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

- 実況歓迎

- 発表後追記した部分は緑色で書いた (2019/10/29)

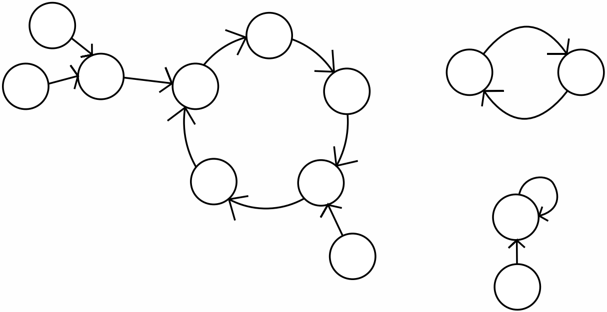
話すこと

- $\underline{n} := \{0, 1, \dots, n-1\}$
- $\underline{n}$ から $\underline{n}$ への写像全体を $\mathrm{Map}(\underline{n}, \underline{n})$ とおく
- $\mathrm{Map}(\underline{n}, \underline{n})$ に一様分布を入れる．つまり，$\mathrm{Map}(\underline{n}, \underline{n})$ 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 $\mathfrak{S}_n$ 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 $\mathfrak{S}_n \subseteq \mathrm{Map}(\underline{n}, \underline{n})$

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:



例題

例題

$f \in \mathrm{Map}(\underline{n}, \underline{n})$ をランダムにとったとき，それが長さ 1 のサイクルを持つ確率は $n \to \infty$ でどうなる？

話すこと

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	$\mathrm{Map}(\underline{n}, \underline{n})$	$\mathfrak{S}_n$
長さ 1 のサイクルを持つ確率	$\to 1 - 1/e \doteq 0.63$	
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	$\sim \frac{e^{-\gamma}}{2} \log n \doteq 0.28 \log n$	$\sim e^{-\gamma} \log n \doteq 0.56 \log n$

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 $f \in \mathrm{Map}(\underline{n}, \underline{n})$ を使った反復列 $(x_n)_{n \in \mathbb{N}}$ すなわち $x_0 = (\text{適当な値}), x_{n+1} = f(x_n)$ 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 (x_0 として採用する適当な値を種という)

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！

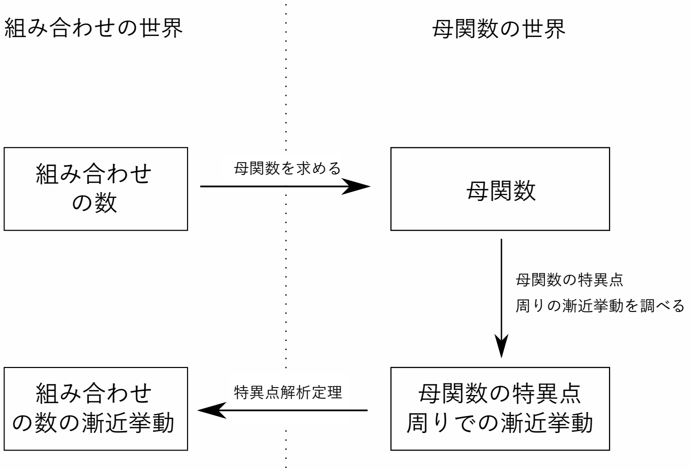
ということを証明する．
実際，「でたらめなアルゴリズム」を選ぶことが $\mathrm{Map}(\underline{n}, \underline{n})$ に一様分布を入れて選ぶことでうまくモデル化できているとすれば，表の左下の結果: $\frac{e^{-\gamma}}{2} \log n$ を使い，たとえば $n = 2^{32}$ のときで，一番短い周期の期待値は 6.2 程度，$n = 2^{64}$ でも 12.4 程度であるとわかる

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

1. 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」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

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 α に対してそのサイズを $|\alpha|$ 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:

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

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 ϵ 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 ⓪ と書く．
- ⓪ のみからなるラベル付きクラスを $\textsc{Unit}$ と書く．
- $\{\epsilon\}$, $\textsc{Unit}$ 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 $\sigma \in \mathfrak{S}_n$ 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

$$\sigma_0 \longrightarrow \sigma_1 \longrightarrow \cdots \longrightarrow \sigma_{n-1}$$

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 $\textsc{Perm}$ 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

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

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 $\textsc{Urn}$ で書く:

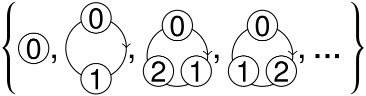
$$\textsc{Urn} = \left\{ \epsilon, ⓪, \boxed{⓪ \quad ①}, \boxed{\begin{array}{cc} ⓪ & ① \\ \multicolumn{2}{c}{②} \end{array}}, \boxed{\begin{array}{cc} ⓪ & ① \\ ② & ③ \end{array}}, \boxed{\begin{array}{cc} ⓪ & ① \\ ② & ③ \\ \multicolumn{2}{c}{④} \end{array}}, \dots \right\}$$

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，$\textsc{Urn}$ の指数型母関数は

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

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 $\textsc{Cyc}$ で表す:

$\textsc{Cyc} =$ 

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，$\textsc{Cyc}$ の指数型母関数は

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

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．

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

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削減とは，ラベルの相対的な順序は保ったまま，ラベルの集合を $\underline{n}$ にしたものをいう．

$$7 \rightarrow 3 \rightarrow 9 \rightarrow 2 \quad \xRightarrow{\text{削減}} \quad 2 \rightarrow 1 \rightarrow 3 \rightarrow 0$$

ラベル付きオブジェクトの積

定義 (積)

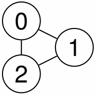
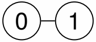
よくラベル付けされたオブジェクト β, γ に対して,

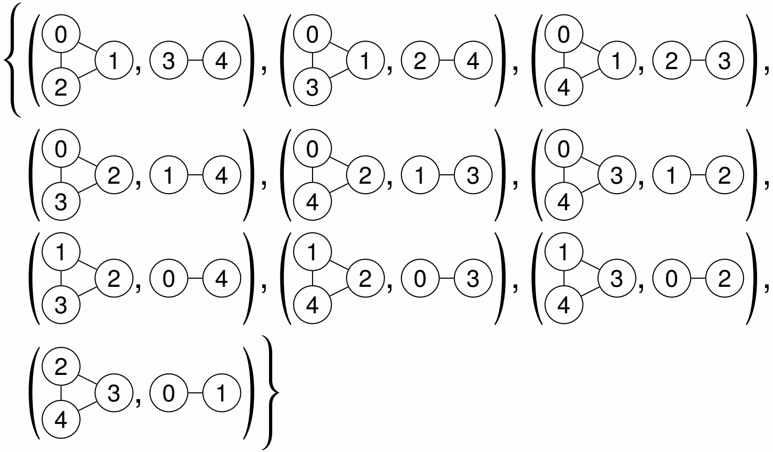
$$\beta \star \gamma = \{(\beta', \gamma') \mid (\beta', \gamma') \text{ はよくラベル付けされていて } \beta \text{ は } \beta' \text{ の削減かつ } \gamma \text{ は } \gamma' \text{ の削減}\}$$

とおく.

$\beta \star \gamma$ の各元はラベル付きオブジェクトの順序対だが, これもラベル付きオブジェクトとみなす. (β', γ') のラベル全体といったら, β' のラベル全体と γ' のラベル全体の和集合と考える.

ラベル付きオブジェクトの積

例: $\beta =$ , $\gamma =$  に対して

$\beta \star \gamma =$ 

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

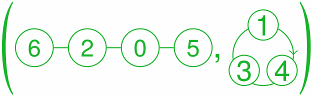
ラベル付きクラスの積

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

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

で定める．

例:



$\in \textsc{Perm} \star \textsc{Cyc}$

ラベル付きクラスの積と指数型母関数

定理

$\mathcal{B}, \mathcal{C}$ をラベル付きクラスとする．その積を $\mathcal{A} = \mathcal{B} \star \mathcal{C}$ として，$\mathcal{A}, \mathcal{B}, \mathcal{C}$ 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,

$$\textsc{Sequence}_k(\mathcal{B}) = \underbrace{\mathcal{B} \star \mathcal{B} \star \cdots \star \mathcal{B}}_{k \text{ 個}}$$

と定義し,

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

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

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

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

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 $\mathcal{B}$ に対して,

$$\textsc{Set}_k(\mathcal{B}) = \textsc{Sequence}_k(\mathcal{B})/\sim$$

と定義する．ここに $\sim$ は並び替えて等しくなるという同値関係.

ただし $\textsc{Sequence}_k(\mathcal{B})/\sim$ の元 (同値類) もラベル付きオブジェクトとみなす．その際，同値類のラベル集合とは代表元のラベル集合とする.

集合クラス

定義 (集合クラス)

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

と定義する.

集合クラス

サイズ k の $\textsc{Sequence}(\mathcal{B})$ 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

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

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

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 $\mathcal{B}$ に対して,

$$\textsc{Cycle}_k(\mathcal{B}) = \textsc{Sequence}_k(\mathcal{B})/\sim$$

と定義する．ここに $\sim$ 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

$$\textsc{Cycle}(\mathcal{B}) = \bigcup_{k \geq 1} \textsc{Cycle}_k(\mathcal{B})$$

と定義する．

サイクルクラス

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

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

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:

$$\begin{aligned} \textsc{Perm} &= \textsc{Sequence}(\textsc{Unit}) \\ \textsc{Urn} &= \textsc{Set}(\textsc{Unit}) \\ \textsc{Cyc} &= \textsc{Cycle}(\textsc{Unit}) \end{aligned}$$

置換のクラス

置換全体のクラス $\textsc{Perm}$ を思い出そう．その指数型母関数は

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

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

$$\textsc{Perm} \simeq \textsc{Set}(\textsc{Cycle}(\textsc{Unit}))$$

が成り立つ．右辺の指数型母関数を計算してみると

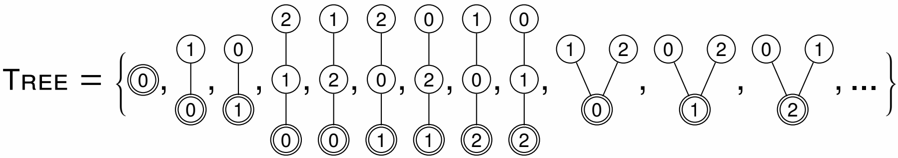
$$\exp\left(\log \frac{1}{1-z}\right) = \frac{1}{1-z}$$

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

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

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:

$$\textsc{Tree} \simeq \textsc{Unit} \star \textsc{Set}(\textsc{Tree})$$

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 $\textsc{Func}$ 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，

$$\textsc{Func} \simeq \textsc{Set}(\textsc{Component})$$
$$\textsc{Component} = \textsc{Cycle}(\textsc{Tree})$$

が成立．したがって $\textsc{Func}, \textsc{Component}$ の指数型母関数を F, K とすれば，

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

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	$\mathcal{A} = \mathcal{B} \star \mathcal{C}$	$A(z) = B(z)C(z)$
列	$\mathcal{A} = \textsc{Sequence}(\mathcal{B})$	$A(z) = \frac{1}{1-B(z)}$
集合	$\mathcal{A} = \textsc{Set}(\mathcal{B})$	$A(z) = \exp B(z)$
サイクル	$\mathcal{A} = \textsc{Cycle}(\mathcal{B})$	$A(z) = \log \frac{1}{1-B(z)}$

これらを利用して functional グラフの指数型母関数を求めた．

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．

二つの定義

定義

原点周りで正則な複素関数 f について，原点でのテイラー展開の z^n の係数を $[z^n]f(z)$ と表す．

定義

二つの数列 $(a_n), (b_n)$ について $a_n \sim b_n \ (n \to \infty)$ とは $a_n/b_n \to 1 \ (n \to \infty)$ となることを意味する．数列ではなく二つの関数のある点に近づけたときの場合についても同様に定義．

特異点解析

定理 (特異点解析; Flajolet & Odlyzko)

$f(z)$ を領域

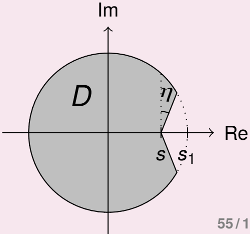
$$D = \{z \in \mathbb{C} \mid |z| < s_1, |\arg(z - s)| > \frac{\pi}{2} - \eta, z \neq s\}$$

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

Tree の指数型母関数 $T(z)$ は収束半径 $1/e$ であり $1/e$ に特異点がある．特異点解析定理に載せるために $1/e$ 周りでの $T(z)$ の漸近挙動を得たい．

$T(z) = ze^{T(z)}$ なので $z = \varphi(T(z))$ where $\varphi(t) = te^{-t}$ である．ここに $\varphi(t)$ の $t = 1$ 周りでのテイラー展開は

$$\varphi(t) = \frac{1}{e} - \frac{(t-1)^2}{2e} + \dots .$$

よって $z = \varphi(t)$ のとき

$$z \fallingdotseq \frac{1}{e}\left(1 - \frac{(t-1)^2}{2}\right) \quad \therefore 1 - ez \fallingdotseq \frac{(t-1)^2}{2}$$

$$\therefore t \fallingdotseq 1 - \sqrt{2}\sqrt{1 - ez}$$

木関数の漸近展開

この議論を厳密化＆精密化すれば次を得る．

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 $z \to e^{-1}$ としたとき

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

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 χ 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

χ 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	$\mathcal{A} = \mathcal{B} \star \mathcal{C}$	$A(z,u) = B(z,u)C(z,u)$
列	$\mathcal{A} = \textsc{Sequence}(\mathcal{B})$	$A(z,u) = \frac{1}{1-B(z,u)}$
集合	$\mathcal{A} = \textsc{Set}(\mathcal{B})$	$A(z,u) = \exp B(z,u)$
サイクル	$\mathcal{A} = \textsc{Cycle}(\mathcal{B})$	$A(z,u) = \log \frac{1}{1-B(z,u)}$

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)
6. 一番短い周期の期待値
7. 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	$\mathrm{Map}(\underline{n}, \underline{n})$	$\mathfrak{S}_n$
長さ 1 のサイクルを持つ確率	$\to 1 - 1/e \doteq 0.63$	
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	$\sim \frac{e^{-\gamma}}{2} \log n \doteq 0.28 \log n$	$\sim e^{-\gamma} \log n \doteq 0.56 \log n$

連結成分の個数の期待値

連結成分の個数をパラメータとする Fᴜɴᴄ 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

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

連結成分の個数の期待値

木関数の近似

$T(z) = 1 - \sqrt{2}\sqrt{1-ez} + O(1-ez) \ \ (z \to 1/e)$ 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

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

連結成分の個数の期待値

よって連結成分の個数の期待値は

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

となる.

連結成分の個数の期待値

量	$\mathrm{Map}(\underline{n}, \underline{n})$	$\mathfrak{S}_n$
長さ 1 のサイクルを持つ確率	$\to 1 - 1/e \doteq 0.63$	
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	$\sim \frac{e^{-\gamma}}{2} \log n \doteq 0.28 \log n$	$\sim e^{-\gamma} \log n \doteq 0.56 \log n$

ランダムな点からのサイクルの長さの期待値

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.

ランダムな点からのサイクルの長さの期待値

$\varphi \in \textsc{Func}_n, i \in \underline{n}$ に対して

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

とおく．この量に関する指数型母関数を

$$\xi(z) = \sum_{n \geq 0} \Bigg(\sum_{\substack{\varphi \in \textsc{Func}_n \\ i \in \underline{n}}} \mathrm{NCycPts}(\varphi, i) \Bigg) \frac{z^n}{n!}$$

とする．

ランダムな点からのサイクルの長さの期待値

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．ξ を求めよう．

ランダムな点からのサイクルの長さの期待値

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:

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

ランダムな点からのサイクルの長さの期待値

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

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

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

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

ランダムな点からのサイクルの長さの期待値

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:

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

ランダムな点からのサイクルの長さの期待値

$g(z)$ と $\textsc{Func}$ 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．

$$\begin{aligned}
\because\ & n![z^n]g(z)\frac{1}{1-T(z)} \\
&= n!\sum_{k+l=n}([z^k]g(z))([z^l]\frac{1}{1-T(z)}) \\
&= n!\sum_{k+l=n}\Bigg(\sum_{\substack{\varphi\ \text{は単成分の写像} \\ |\varphi|=k}}\mathrm{NCycPts}(\varphi)\Bigg)\cdot k\cdot\frac{1}{k!}\cdot\frac{l^l}{l!} \\
&= \sum_{k+l=n}\binom{n}{k}\Bigg(\sum_{\substack{\varphi\ \text{は単成分の写像} \\ |\varphi|=k}}\mathrm{NCycPts}(\varphi)\Bigg)\cdot k\cdot l^l \\
&= \sum_{\substack{\varphi\in\textsc{Func}_n \\ i\in\underline{n}}}\mathrm{NCycPts}(\varphi,i)
\end{aligned}$$

ランダムな点からのサイクルの長さの期待値

この $\xi(z) = (zT'(z))/(1 - T(z))^3$ 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

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

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

ランダムな点からのサイクルの長さの期待値

量	$\mathrm{Map}(\underline{n}, \underline{n})$	$\mathfrak{S}_n$
長さ 1 のサイクルを持つ確率	$\to 1 - 1/e \doteq 0.63$	
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	$\sim \frac{e^{-\gamma}}{2} \log n \doteq 0.28 \log n$	$\sim e^{-\gamma} \log n \doteq 0.56 \log n$

一番長いサイクルの長さの期待値

次に一番長いサイクルの長さの期待値を考察する．
一般に，$\varphi \in \textsc{Func}_n$ と $i \in \underline{n}$ で決まるパラメータを $\xi(\varphi, i)$ としよう．このとき

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

と母関数

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

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

一番長いサイクルの長さの期待値

このとき，ラベル付きクラス

$$\textsc{Func}^{[k]} = \{\varphi \in \textsc{Func} \mid \xi^{\max}(\varphi) \leq k\}$$

を考えて，その指数型母関数を $F^{[k]}(z)$ とすると

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

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:

$$\mathbf{E}[X] = \sum_{k \geq 1} k\mathbf{P}(X = k) = \sum_{k \geq 1} \mathbf{P}(X \geq k) = \sum_{k \geq 1} (1 - \mathbf{P}(X < k))$$

一番長いサイクルの長さの期待値

さて，$\xi(\varphi, i)$ が NCycPts(φ, i)，つまり φ 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 $I_k(z) = \sum_{j=1}^{k} \frac{u^j}{j}$ である．
よって，

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

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

一番長いサイクルの長さの期待値

$T(z)$ 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 $\mathrm{Re}\, x > 0$ 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:

$$\begin{aligned}
\Xi(z) &= \frac{1}{1-T(z)} \sum_{k\geq 1} \Bigl(1 - \exp(-\sum_{j\geq k} \frac{e^{-jx}}{j})\Bigr) \\
&\sim \frac{1}{1-T(z)} \sum_{k\geq 1} \Bigl(1 - \exp(-\int_{kx}^{\infty} \frac{e^{-v}}{v} dv)\Bigr) \\
&\sim \frac{1}{(1-T(z))x} \int_0^{x\infty} \Bigl(1 - \exp(-\int_u^{\infty} \frac{e^{-v}}{v} dv)\Bigr) du \\
&\sim \frac{1}{(1-T(z))^2} \int_0^{\infty} \Bigl(1 - \exp(-\int_u^{\infty} \frac{e^{-v}}{v} dv)\Bigr) du
\end{aligned}$$

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 $\sqrt{n}$ の比例係数は

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

一番長いサイクルの長さの期待値

量	$\mathrm{Map}(\underline{n}, \underline{n})$	$\mathfrak{S}_n$
長さ 1 のサイクルを持つ確率	$\to 1 - 1/e \doteq 0.63$	
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	$\sim \frac{e^{-\gamma}}{2} \log n \doteq 0.28 \log n$	$\sim e^{-\gamma} \log n \doteq 0.56 \log n$

一番短い周期の期待値

a_n を φ が $\mathrm{Map}(\underline{n}, \underline{n})$ 全体を走るときの φ 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:

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

なぜならば

$$\textsc{Set}(\textsc{Cycle}_{\geq k}(\textsc{Tree})) - \{\varnothing\} \Leftrightarrow \exp \left(\sum_{j \geq k} \frac{T(z)^j}{j} \right) - 1.$$

一番短い周期の期待値

ここで

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

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
$\mathrm{Re}(x) > 0$ で $A(e^{-x})$ の $x \to 0$ の挙動を調べればよい．

一番短い周期の期待値

関数 $f : (0, \infty) \to \mathbb{C}$ のメリン変換は次で定義される:

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

一番短い周期の期待値

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

の挙動を調べる．まず内側の和を積分で近似すると

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

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

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

一番短い周期の期待値

計算をがちゃがちゃやれば $\mathcal{M}\tilde{A}(s)$ は

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

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

一番短い周期の期待値

$s = 1$ 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

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

を得る．ここに $\asymp$ は差が正則関数の意味.

一番短い周期の期待値

定義 (fundamental strip)

実数 $a < b$ について $\langle a, b\rangle = \{s \in \mathbb{C} \mid a < \mathrm{Re}(s) < b\}$ 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．

一番短い周期の期待値

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 $\mathcal{M}f(s)$ は fundamental strip $\langle \alpha, \beta \rangle$ を持つとする．ある $\alpha' < \alpha$ について，$\mathcal{M}f(s)$ 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，$\alpha' \le \mathrm{Re}(s) \le \eta$ で $|s| \to \infty$ としたとき $\mathcal{M}f(s) = O(|s|^{-r})$ 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

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

ならば $x \to +0$ で

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

一番短い周期の期待値

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，

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

一番短い周期の期待値

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

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

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.

一番短い周期の期待値

量	$\mathrm{Map}(\underline{n}, \underline{n})$	$\mathfrak{S}_n$
長さ 1 のサイクルを持つ確率	$\to 1 - 1/e \doteq 0.63$	
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	$\sim \frac{e^{-\gamma}}{2} \log n \doteq 0.28 \log n$ ✔	$\sim e^{-\gamma} \log n \doteq 0.56 \log n$

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)
6. 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:
  1. 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	母関数	漸近挙動
連結成分の個数	$\frac{1}{1-T(z)} \log \frac{1}{1-T(z)}$	$\frac{1}{2} \log n$
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	$\frac{zT'(z)}{(1-T(z))^3}$	$\sqrt{\pi n/8}$
一番長いサイクルの長さ	$\sum_{k\geq 0} \left(\frac{1}{1-T(z)} - \exp \sum_{j=1}^{k} \frac{T(z)^j}{j}\right)$	$c_1 \sqrt{n}$
一番短いサイクルの長さ	$\sum_{k\geq 1} \left(-1 + \exp \sum_{j\geq k} \frac{T(z)^j}{j}\right)$	$\frac{e^{-\gamma}}{2} \log n$

※ 母関数の列を表に追加した

参考文献

[1] 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．web で PDF が公開されている．(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．