

短いループのできる確率



[https://fujidig.github.io/201812-short-period/
201812-short-period.pdf](https://fujidig.github.io/201812-short-period/201812-short-period.pdf)

でいぐ

2018年12月8日

数物セミナー愛媛談話会

この発表で話すこと

- $\underline{m} = \{0, 1, 2, \dots, m-1\}$ とおく
- \underline{m} から \underline{m} への写像 f が無作為に与えられたときに f による反復列の周期に関する期待値や確率を議論する
- そのような f は m^m 通りあるのでそのそれぞれが等確率であるとする
- m が十分大きいときに興味があるので $m \rightarrow \infty$ の極限を考える

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するということをしない限り)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 を使った反復列によって定められる

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「でたらめな写像」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, 「でたらめな写像」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
- その事実を定量的に述べる一つの議論がこの講演である

Knuthが考えたでたらめな写像

[2]の3.1節より引用.

このアルゴリズムはKnuthが
1959年に考えた.

このアルゴリズムで得られる
写像を使って列を作ると初期
値によっては周期が1や3にな
りえる. つまり疑似乱数とし
ては使い物にならない

(Knuthでさえ、でたらめな写像
が疑似乱数に使えると誤解し
た)

問題1

問題1

写像 $f: \underline{m} \rightarrow \underline{m}$ と $x_0 \in \underline{m}$ が無作為に与えられるときに列 $x_n = f^n(x_0)$ の周期が1になる確率を求めよ.

問題1

写像 $f: \underline{m} \rightarrow \underline{m}$ と $x_0 \in \underline{m}$ が無作為に与えられるときに列 $x_n = f^n(x_0)$ の周期が1になる確率を求めよ.

$x_{i+k} = x_i$ となる $i, k \geq 0$ がある.

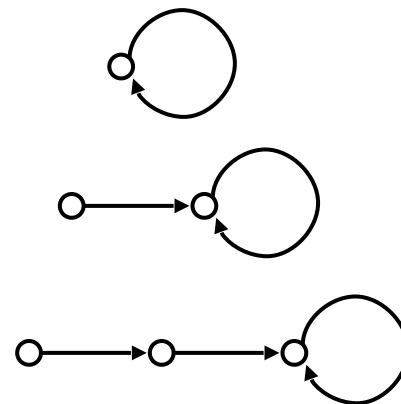
そのような i の最小値を μ , k の最小値を λ とする.

λ が列の周期である.

$$P(\mu = 0, \lambda = 1) = \frac{1}{m}$$

$$P(\mu = 1, \lambda = 1) = \left(1 - \frac{1}{m}\right) \frac{1}{m}$$

$$P(\mu = 2, \lambda = 1) = \left(1 - \frac{1}{m}\right) \left(1 - \frac{2}{m}\right) \frac{1}{m}$$



問題1

写像 $f: \underline{m} \rightarrow \underline{m}$ と $x_0 \in \underline{m}$ が無作為に与えられるときに列 $x_n = f^n(x_0)$ の周期が1になる確率を求めよ.

一般に

$$P(\mu = \mu_0, \lambda = 1) = \left(1 - \frac{1}{m}\right) \left(1 - \frac{2}{m}\right) \dots \left(1 - \frac{\mu_0}{m}\right) \frac{1}{m}$$

よって

$$P(\lambda = 1) = \frac{1}{m} \sum_{i \geq 0} \left(1 - \frac{1}{m}\right) \left(1 - \frac{2}{m}\right) \dots \left(1 - \frac{i}{m}\right)$$

を $Q(m)$ とおく.

問題1

写像 $f: \underline{m} \rightarrow \underline{m}$ と $x_0 \in \underline{m}$ が無作為に与えられるときに列 $x_n = f^n(x_0)$ の周期が1になる確率を求めよ.

$$Q(m) = \sum_{i \geq 0} \left(1 - \frac{1}{m}\right) \left(1 - \frac{2}{m}\right) \dots \left(1 - \frac{i}{m}\right)$$

はRamanujanのQ functionと呼ばれる.

Fact

$$Q(m) = \sqrt{\frac{\pi m}{2}} - \frac{1}{3} + O(m^{-\frac{1}{2}})$$

問題1

Factより,

$$\rightsquigarrow P(\lambda = 1) \sim \sqrt{\frac{\pi}{2m}} \sim \frac{1.25}{\sqrt{m}}$$

m が十分大きいときには周期が1になる確率は
ほぼ0.

問題2

問題2

写像 $f: \underline{m} \rightarrow \underline{m}$ と $x_0 \in \underline{m}$ が無作為に与えられるときに列 $x_n = f^n(x_0)$ の周期の期待値を求めよ.

問題2

写像 $f: \underline{m} \rightarrow \underline{m}$ と $x_0 \in \underline{m}$ が無作為に与えられるときに列 $x_n = f^n(x_0)$ の周期の期待値を求め一般によ.

$$P(\mu = \mu_0, \lambda = \lambda_0) = \frac{1}{m} \prod_{1 \leq k < \mu_0 + \lambda_0} \left(1 - \frac{k}{m}\right).$$

よって

$$E[\lambda] = \sum_{\substack{1 \leq \lambda_0 \\ 0 \leq \mu_0}} \frac{\lambda_0}{m} \prod_{1 \leq k < \mu_0 + \lambda_0} \left(1 - \frac{k}{m}\right).$$

問題2

写像 $f: \underline{m} \rightarrow \underline{m}$ と $x_0 \in \underline{m}$ が無作為に与えられるときに列 $x_n = f^n(x_0)$ の周期の期待値を求め

$$\mathbb{E}[\lambda] = \sum_{\substack{1 \leq \lambda_0 \\ 0 \leq \mu_0}} \frac{\lambda_0}{m} \prod_{1 \leq k < \mu_0 + \lambda_0} \left(1 - \frac{k}{m}\right)$$

$$= \frac{1}{m} \left\{ 1 + (1+2) \left(1 - \frac{1}{m}\right) + (1+2+3) \left(1 - \frac{1}{m}\right) \left(1 - \frac{2}{m}\right) + \dots \right\}$$

$$= \frac{1}{m} \sum_{n \geq 0} \frac{(n+1)(n+2)}{2} \prod_{1 \leq k < n} \left(1 - \frac{k}{m}\right)$$

問題2

写像 $f: \underline{m} \rightarrow \underline{m}$ と $x_0 \in \underline{m}$ が無作為に与えられるときに列 $x_n = f^n(x_0)$ の周期の期待値を求めよ

Lemma

$$f(a_0, a_1, \dots) = \sum_{n \geq 0} a_n \prod_{k=1}^n \left(1 - \frac{k}{m}\right)$$

なら

$$f(a_0, a_1, \dots) = a_0 + f(a_1, a_2, \dots) - f(a_1, 2a_2, \dots) / m$$

Lemmaに $a_n = \frac{n+1}{2}$ を代入すると

$$E[\lambda] = \frac{1}{m} \sum_{n \geq 0} \frac{(n+1)(n+2)}{2} \prod_{1 \leq k \leq n} \left(1 - \frac{k}{m}\right) = \frac{1 + Q(m)}{2}.$$

問題2

写像 $f: \underline{m} \rightarrow \underline{m}$ と $x_0 \in \underline{m}$ が無作為に与えられるときに列 $x_n = f^n(x_0)$ の周期の期待値を求めよ.

$$E[\lambda] = \frac{1 + Q(m)}{2} \sim \sqrt{\frac{\pi m}{8}} + \frac{1}{3}.$$

つまり周期 λ の期待値は \sqrt{m} に比例する程度

問題3

問題3

写像 $f: \underline{m} \rightarrow \underline{m}$ が無作為に与えられるとき, f が不動点を持つ確率を求めよ.

問題3

写像 $f: \underline{m} \rightarrow \underline{m}$ が無作為に与えられるとき, f が不動点を持つ確率を求めよ.

不動点とは $f(x) = x$ となる $x \in \underline{m}$ のこと.

すべての x で $f(x) \neq x$ となる f の個数は $(m-1)^m$

一方, すべての f の個数は m^m .

よって

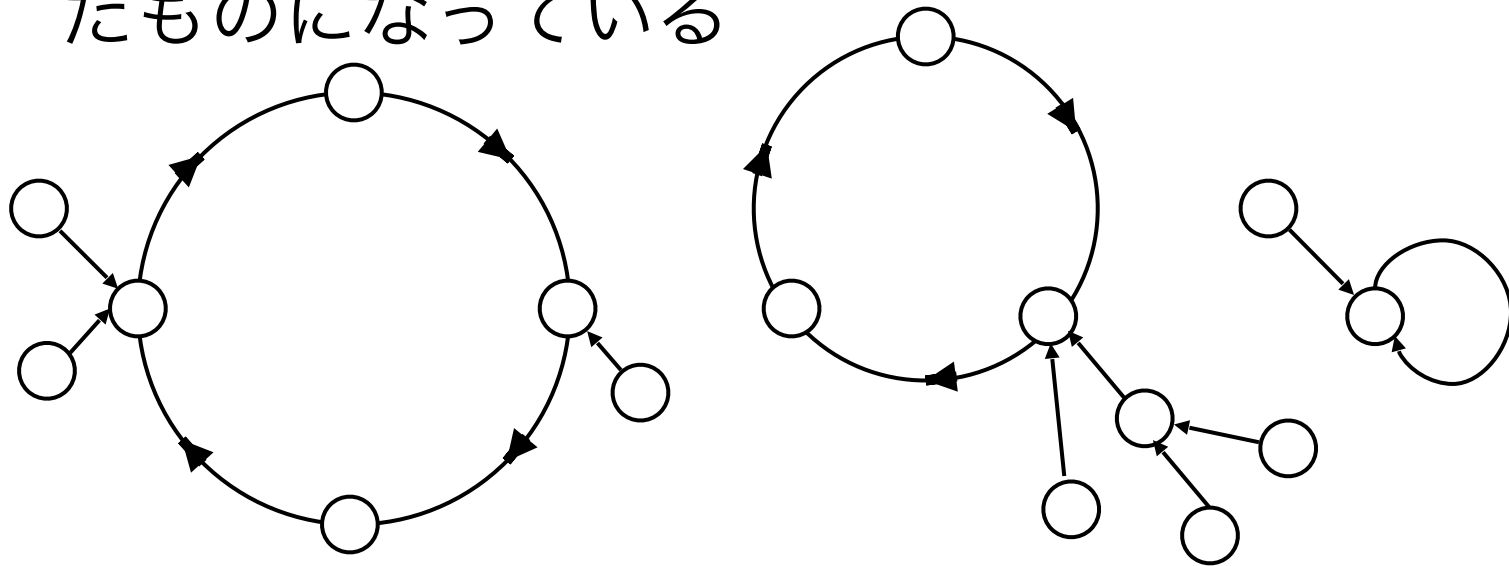
$$\begin{aligned} P(\text{不動点を持つ}) &= 1 - \frac{(m-1)^m}{m^m} \\ &= 1 - \left(1 - \frac{1}{m}\right)^m \end{aligned}$$

$m \rightarrow \infty$ の極限は $1 - e^{-1} = 0.632\dots$ である.

問題4

...の前に観察

- $0, 1, 2, \dots, m-1$ を頂点として $f(i) = j$ なら i から j への有向辺を伸ばしたグラフを考える
- このグラフはいくつかのサイクルに木を生やしたものになっている



問題4

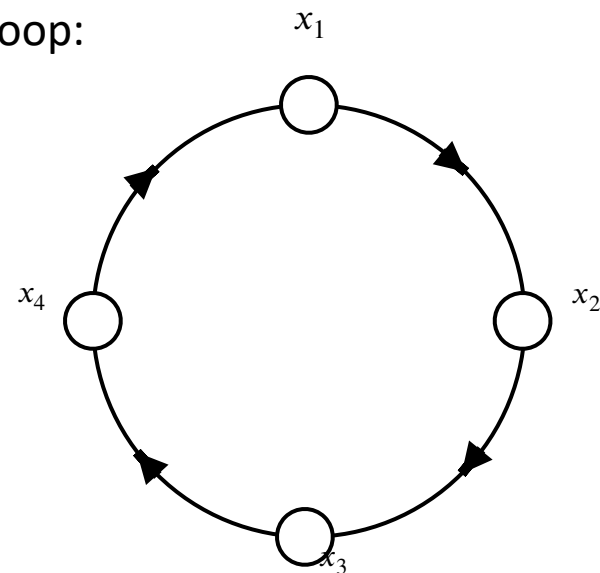
問題3を一般化しよう.

f の l -loopとは \underline{m} の相異なる元

x_1, x_2, \dots, x_l であって

$f(x_1) = x_2, \dots, f(x_l) = x_1$ をみたすもののこと.

4-loop:



写像 $f: \underline{m} \rightarrow \underline{m}$ が無作為に与えられるとき, n 以下のloopができる確率を求めよ.

問題4

写像 $f: \underline{m} \rightarrow \underline{m}$ が無作為に与えられるとき, n 以下のloopができる確率を求めよ.

かんたんのため, $n = 2$ で解く.

Fact (包除原理)

I を有限全順序集合, 各 $i \in I$ に対し A_i を事象とする. このとき

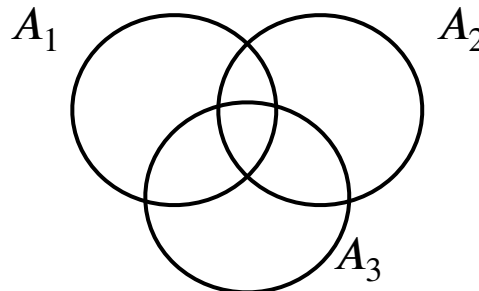
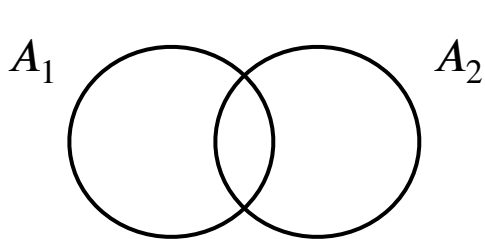
$$P\left(\bigcup_{i \in I} A_i\right) = \sum_{k=1}^{|I|} (-1)^{k-1} \sum_{i_1 < \dots < i_k} P(A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k})$$

問題4

写像 $f: \underline{m} \rightarrow \underline{m}$ が無作為に与えられるとき, n 以下のloopができる確率を求めよ.

包除原理は $|I| = 2, 3$ のときは次の式を表す:

- $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$
- $P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3)$



問題4

写像 $f: \underline{m} \rightarrow \underline{m}$ が無作為に与えられるとき, n 以下のloopができる確率を求めよ.

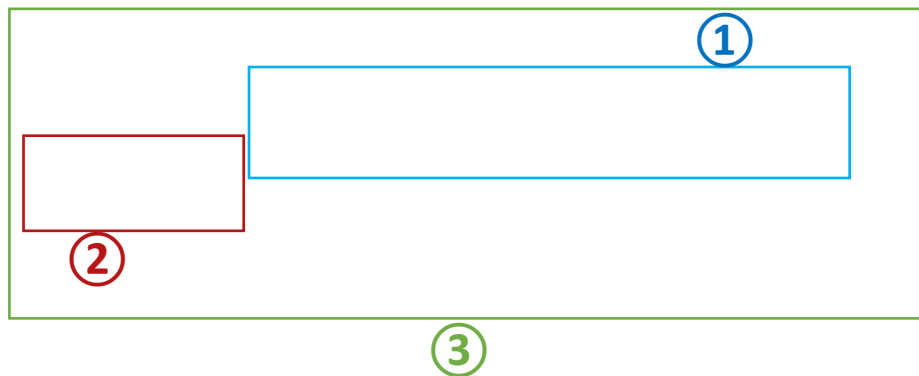
- $[\underline{m}]^2 = \left\{ \{x, y\} \subset \underline{m} \mid x \neq y \right\}$ とおく
- $x \in \underline{m}$ に対し $A_x = \left\{ f: \underline{m} \rightarrow \underline{m} \mid x \text{ は } f \text{ の不動点} \right\}$
- $\{x, y\} \in [\underline{m}]^2$ に対し
 $A_{\{x, y\}} = \left\{ f: \underline{m} \rightarrow \underline{m} \mid \{x, y\} \text{ は } f \text{ の } 2\text{-loop} \right\}$
と定める.

問題4

写像 $f: \underline{m} \rightarrow \underline{m}$ が無作為に与えられるとき, n 以下のloopができる確率を求めよ.

すると

$$\begin{aligned}
 P(\exists 1\text{-loop} \vee \exists 2\text{-loop}) &= P\left(\bigcup_{t \in \underline{m} \cup [\underline{m}]^2} A_t\right) = \sum_{k=1}^{|\underline{m}|} (-1)^{k-1} \sum_{\substack{t_1 < \dots < t_k \\ t_i \in \underline{m} \cup [\underline{m}]^2}} P(A_{t_1} \cap A_{t_2} \cap \dots \cap A_{t_k}) \\
 &= \sum_{k=1} (-1)^{k-1} \sum_{\substack{r+s=k \\ r+2s \leq m}} \sum_{\substack{x_1 < \dots < x_r \in \underline{m} \\ p_1 < \dots < p_s \in [\underline{m}]^2}} P(A_{x_1} \cap \dots \cap A_{x_r} \cap A_{p_1} \cap \dots \cap A_{p_s})
 \end{aligned}$$



問題4

写像 $f: \underline{m} \rightarrow \underline{m}$ が無作為に与えられるとき, n 以下のloopができる確率を求めよ.

① について計算. 固定された

$x_1, \dots, x_r, p_1, \dots, p_s$ について

$$\begin{aligned} & \sum_{\substack{x_1 < \dots < x_r \in \underline{m} \\ p_1 < \dots < p_s \in \underline{[m]}^2}} P(A_{x_1} \cap \dots \cap A_{x_r} \cap A_{p_1} \cap \dots \cap A_{p_s}) \\ &= \frac{m^{m-r-2s}}{m^m} = \frac{1}{m^{r+2s}} \end{aligned}$$

問題4

写像 $f: \underline{m} \rightarrow \underline{m}$ が無作為に与えられるとき, n 以下のloopができる確率を求めよ.

②で走らせているものの個数を数える.

$$\# \left\{ \left((x_1, \dots, x_r), (p_1, \dots, p_s) \right) \in \underline{m}^r \times \left([\underline{m}]^2 \right)^s \left| \begin{array}{l} x_1 < \dots < x_r, p_1 < \dots < p_s, \\ \{x_1\}, \dots, \{x_r\}, p_1, \dots, p_s \text{ は pairwise disjoint} \end{array} \right. \right\}$$

$$= C(m, r) \frac{P(m-r, 2s)}{(2!)^s s!}$$

ここに $C(-, -)$ は組み合わせの数, $P(-, -)$ は順列の数を表す記号.

よって, ①と②とから③が計算される.

$$\sum_{\substack{x_1 < \dots < x_r \in \underline{m} \\ p_1 < \dots < p_s \in [\underline{m}]^2}} P(A_{x_1} \cap \dots \cap A_{x_r} \cap A_{p_1} \cap \dots \cap A_{p_s}) = C(m, r) \frac{P(m-r, 2s)}{2^s s!} \frac{1}{m^{r+2s}}$$

$$= \frac{1}{2^s r! s!} \left(1 - \frac{1}{m}\right) \left(1 - \frac{2}{m}\right) \dots \left(1 - \frac{r+2s-1}{m}\right)$$

問題4

写像 $f: \underline{m} \rightarrow \underline{m}$ が無作為に与えられるとき, n 以下のloopができる確率を求めよ.

よって

$$\begin{aligned}
 & P(\exists 1\text{-loop} \vee \exists 2\text{-loop}) \\
 &= \sum_{k=1} (-1)^{k-1} \sum_{\substack{r+s=k \\ r+2s \leq m}} (\textcircled{3}) \\
 &= \sum_{k=1} (-1)^{k-1} \sum_{\substack{r+s=k \\ r+2s \leq m}} \frac{1}{2^s r! s!} \left(1 - \frac{1}{m}\right) \left(1 - \frac{2}{m}\right) \dots \left(1 - \frac{r+2s-1}{m}\right) \\
 &= \sum_{\substack{r+s \geq 1 \\ r+2s \leq m}} \frac{(-1)^{r+s-1}}{2^s r! s!} \left(1 - \frac{1}{m}\right) \left(1 - \frac{2}{m}\right) \dots \left(1 - \frac{r+2s-1}{m}\right)
 \end{aligned}$$

問題4

写像 $f: \underline{m} \rightarrow \underline{m}$ が無作為に与えられるとき, n 以下のloopができる確率を求めよ.

$$P(\exists 1\text{-loop} \vee \exists 2\text{-loop})$$

$$= \sum_{\substack{r+s \geq 1 \\ r+2s \leq m}} \frac{(-1)^{r+s-1}}{2^s r! s!} \left(1 - \frac{1}{m}\right) \left(1 - \frac{2}{m}\right) \dots \left(1 - \frac{r+2s-1}{m}\right)$$

$m \rightarrow \infty$ とすると

$$\begin{aligned} \sum_{r+s \geq 1} \frac{(-1)^{r+s-1}}{2^s r! s!} &= 1 - \sum_{r \geq 0} \frac{(-1)^r}{r!} \sum_{s \geq 0} \frac{(-1)^s}{2^s s!} \\ &= 1 - e^{-1} e^{-\frac{1}{2}} \\ &= 1 - e^{-\frac{3}{2}} \end{aligned}$$

(ルベークの収束定理とフビニの定理を使った)

問題4

写像 $f: \underline{m} \rightarrow \underline{m}$ が無作為に与えられるとき, n 以下のloopができる確率を求めよ.

一般の n では

$$\begin{aligned} P_n &:= \lim_{m \rightarrow \infty} P\left(\{f: \underline{m} \rightarrow \underline{m} \mid f \text{ が } n \text{ 以下のloopを持つ}\}\right) \\ &= 1 - e^{-H_n} \left(H_n = 1 + \frac{1}{2} + \dots + \frac{1}{n} \right) \end{aligned}$$

となる.

$$P_1 = 0.632\dots$$

$$P_2 = 0.776\dots$$

$$P_3 = 0.840\dots$$

$$P_{10} = 0.946\dots$$

$$P_{100} = 0.994\dots$$

問題5

問題5

写像 $f: \underline{m} \rightarrow \underline{m}$ が無作為に与えられるとき, f の一番短いloopの長さの期待値は？

(未解決？)

具体例

- $m = 2^{24} = 16777216$ として f は次のものを採用

- $$f(x) = \left\lfloor \frac{(ax + b) \bmod 2^{64}}{2^{40}} \right\rfloor$$

where $a = 10208877246009069551$, $b = 17024059796969606487$

- このときループの長さは

4051, 2523, 259, 197, 165, 140, 139, 2

となった (これで全て)

なお, $\sqrt{\frac{\pi m}{8}} = 2566. \dots$

参考文献

- [1] Donald E. Knuth “The Art of Computer Programming, Volume 1: Fundamental Algorithms” Addison-Wesley Professional, 1997.
- [2] Donald E. Knuth “The Art of Computer Programming, Volume 2: Seminumerical Algorithms” Addison-Wesley Professional, 2014.

本発表は主に[2]の3.1節の演習問題に依っている.

まとめ

- (問題1) 写像と初期値を無作為に選んだとき、周期が1になる確率は約 $\sqrt{\frac{\pi}{2m}}$
- (問題2) 周期の期待値は約 $\sqrt{\frac{\pi m}{8}}$
- (問題3) 写像を無作為に選んだとき不動点ができる確率は
約 $1 - e^{-1} = 0.632\dots$
- (問題4) n 以下のloopができる確率は約 $1 - e^{-H_n}$ ($H_n = 1 + \frac{1}{2} + \dots + \frac{1}{n}$)
- (問題5) 一番短いloopの長さの期待値は？
- “The random numbers should not be generated with a method chosen at random.” [Knuth '2] 3.1節)
- ご清聴ありがとうございました

