# 1　課題 1

状態数が 7 以上でかつ対角成分が全てゼロとなる既約なハイパーリンク行列 $\boldsymbol{H}$ の例を作り, それを書き下しなさい. ただし, 例とする行列 $\boldsymbol{H}$ を正規化して得られる遷移確率行列 $\overline{\boldsymbol{H}}$ が非周期的になるようにすること. さらに, アルゴリズム 1 を C 言語で実装し,$\overline{\boldsymbol{H}}$ によって定まるページランク $\boldsymbol{\pi} = (\pi_i)_{i=1,2,\ldots,N}$ を数値計算し, その結果を出力しなさい. なお, 許容誤差 $\varepsilon$ を $10^{-8}$ 以下に設定すること.

解

$$\boldsymbol{H} = \begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \overline{\boldsymbol{H}} = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

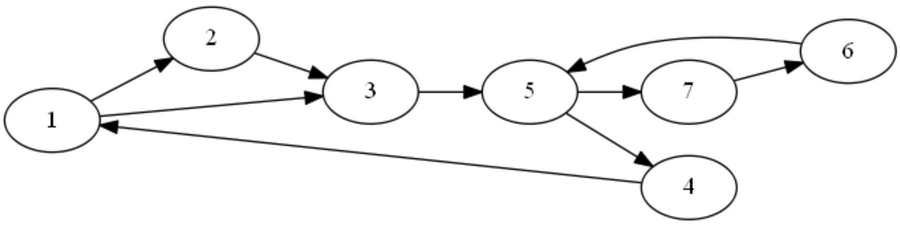


図 1　課題 1 のハイパーリンク行列 $\boldsymbol{H}$ に対応するグラフ

図 1 より,$\boldsymbol{H}$ が既約で, かつ $\overline{\boldsymbol{H}}$ が非周期的であることが確認できる. アルゴリズム 1 を実装したプログラムに行列 $\overline{\boldsymbol{H}}$ を入力として与えたときの出力結果は以下である.(許容誤差 $\varepsilon$ は $10^{-8}$)

$$\boldsymbol{\pi} = (0.133333, 0.066667, 0.133333, 0.133333, 0.266667, 0.133333, 0.133333)$$

# 2　課題 2

行列 $\boldsymbol{F}$ を既約な有限次非負正方行列とし, さらに次式を満たすものとする.

$$\boldsymbol{F}\boldsymbol{e} \leq \boldsymbol{e}, \boldsymbol{F}\boldsymbol{e} \neq \boldsymbol{e} \tag{2.1}$$

式 (2.1) は,$\boldsymbol{F}$ の行和がすべて 1 以下でかつ, 少なくとも 1 つの行和が 1 未満であることを意味している.

さて, 上で導入された行列 $\boldsymbol{F}$ のスペクトル半径は 1 未満となることが知られており,

$$\lim_{k\to\infty} \boldsymbol{F}^k = \boldsymbol{O} \tag{2.2}$$

が成り立つ.

(a) 次式が成り立つことを示しなさい.

$$(\boldsymbol{I}-\boldsymbol{F})^{-1}=\sum_{k=0}^{\infty}\boldsymbol{F}^k \tag{2.3}$$

解

$$(\boldsymbol{I}-\boldsymbol{F})\sum_{k=0}^{n}\boldsymbol{F}^k=\sum_{k=0}^{n}\boldsymbol{F}^k-\sum_{k=1}^{n+1}\boldsymbol{F}^k=\boldsymbol{I}-\boldsymbol{F}^{n+1}\to\boldsymbol{I}(n\to\infty) \tag{2.4}$$

つまり

$$(\boldsymbol{I}-\boldsymbol{F})\sum_{k=0}^{\infty}\boldsymbol{F}^k=\boldsymbol{I} \tag{2.5}$$

同様に

$$(\sum_{k=0}^{n}\boldsymbol{F}^k)(\boldsymbol{I}-\boldsymbol{F})=\sum_{k=0}^{n}\boldsymbol{F}^k-\sum_{k=1}^{n+1}\boldsymbol{F}^k=\boldsymbol{I}-\boldsymbol{F}^{n+1}\to\boldsymbol{I}(n\to\infty) \tag{2.6}$$

$$\therefore(\sum_{k=0}^{\infty}\boldsymbol{F}^k)(\boldsymbol{I}-\boldsymbol{F})=\boldsymbol{I} \tag{2.7}$$

よって

$$\sum_{k=0}^{\infty}\boldsymbol{F}^k=(\boldsymbol{I}-\boldsymbol{F})^{-1}\quad\square \tag{2.8}$$

(b) 任意の $\delta\in(0,1)$ に対して, 以下の不等式を満たすような$\nu\in\mathbb{N}$ が存在することを示しなさい.

$$\boldsymbol{F}^{\nu}\boldsymbol{e}\leq\delta\boldsymbol{e} \tag{2.9}$$

さらに, 任意の自然数 $m$ に対して, 次式が成り立つことを示しなさい.

$$\sum_{k=m\nu}^{\infty}\boldsymbol{F}^k\boldsymbol{e}\leq\frac{\nu\delta^m}{1-\delta}\boldsymbol{e} \tag{2.10}$$

解

$\boldsymbol{F}$ の次数を $M$ とする. $\lim_{k\to\infty}\boldsymbol{F}^k=\boldsymbol{O}$ と $\boldsymbol{F}$ が非負行列であることから,

$$\forall\varepsilon>0,\exists N\in\mathbb{N},\nu\geq N,\nu\in\mathbb{N}\Rightarrow\boldsymbol{F}^{\nu}<\varepsilon\boldsymbol{e}\boldsymbol{e}^{\top} \tag{2.11}$$

これは十分大きな自然数$\nu$を取れば $\boldsymbol{F}^{\nu}$の各成分は$\varepsilon$以下の非負の値になることを言っている.
$\varepsilon$は任意であるので, $\varepsilon=\delta/M$ とすると,

$$\boldsymbol{F}^{\nu}\boldsymbol{e}\leq\frac{\delta}{M}\boldsymbol{e}\boldsymbol{e}^{\top}\boldsymbol{e}=\frac{\delta}{M}\boldsymbol{e}M=\delta\boldsymbol{e}\quad\square \tag{2.12}$$

続いて, 後半を証明する.自然数 $n'$ を $n \le n'\nu - 1$ を満たすようにとると

$$\sum_{k=m\nu}^{n} \boldsymbol{F}^k \boldsymbol{e} \le \sum_{k=m\nu}^{n'\nu-1} \boldsymbol{F}^k \boldsymbol{e}\ [\because n \le n'\nu - 1, \boldsymbol{F}^k \boldsymbol{e} \ge 0] \tag{2.13}$$

$$= \boldsymbol{F}^{m\nu}\boldsymbol{e} + \boldsymbol{F}^{m\nu+1}\boldsymbol{e} + \cdots + \boldsymbol{F}^{n'\nu-1}\boldsymbol{e} \tag{2.14}$$

$$\le \boldsymbol{F}^{m\nu}\boldsymbol{e} + \boldsymbol{F}^{m\nu}\boldsymbol{e} + \cdots + \boldsymbol{F}^{m\nu}\boldsymbol{e} + \boldsymbol{F}^{(m+1)\nu}\boldsymbol{e} + \cdots + \boldsymbol{F}^{(n'-1)\nu}\boldsymbol{e} \tag{2.15}$$

$$[\because \boldsymbol{F}\boldsymbol{e} \le \boldsymbol{e}\ \text{より}\ i < j \Rightarrow \boldsymbol{F}^i\boldsymbol{e} \ge \boldsymbol{F}^j\boldsymbol{e}]$$

$$= \nu\boldsymbol{F}^{m\nu}\boldsymbol{e} + \nu\boldsymbol{F}^{(m+1)\nu}\boldsymbol{e} + \cdots + \nu\boldsymbol{F}^{(n'-1)\nu}\boldsymbol{e} \tag{2.16}$$

$$\le \nu\delta^m\boldsymbol{e} + \nu\delta^{m+1}\boldsymbol{e} + \cdots + \nu\delta^{n'-1}\boldsymbol{e} \tag{2.17}$$

$$[\because \boldsymbol{F}^{\nu}\boldsymbol{e} \le \delta\boldsymbol{e}\ \text{より}\ \boldsymbol{F}^{i\nu}\boldsymbol{e} \le \delta^i\boldsymbol{e}]$$

$$= \nu\delta^m \frac{1-\delta^{n'-m}}{1-\delta}\boldsymbol{e} \tag{2.18}$$

$$\to \nu\delta^m \frac{1}{1-\delta}\boldsymbol{e}\ (n \to \infty)\ [\because n \to \infty \Rightarrow n' \to \infty] \tag{2.19}$$

よって

$$\sum_{k=m\nu}^{\infty} \boldsymbol{F}^k \boldsymbol{e} \le \frac{\nu\delta^m}{1-\delta}\boldsymbol{e} \quad \square \tag{2.20}$$

(c) 任意の $\varepsilon \in (0,1)$ に対して, $N(\varepsilon)$ を

$$N(\varepsilon) = \left\lceil \nu \frac{\log\{\varepsilon\delta(1-\delta)/\nu\}}{\log\delta} \right\rceil \tag{2.21}$$

と定義する. ただし, $\lceil\cdot\rceil$ は天井関数を表す. このとき, 式 (2.10) を用いて, 次式が成り立つことを示しなさい.

$$\sum_{k=n+1}^{\infty} \boldsymbol{F}^k \boldsymbol{e} \le \varepsilon\boldsymbol{e}, \forall n \ge N(\varepsilon) \tag{2.22}$$

解

$\boldsymbol{F}$ が非負行列であるので $\forall k \in \mathbb{N}$ に対して $\boldsymbol{F}^k\boldsymbol{e} \ge 0$ である.よって

$$\sum_{k=n+1}^{\infty} \boldsymbol{F}^k \boldsymbol{e} \le \sum_{k=n}^{\infty} \boldsymbol{F}^k \boldsymbol{e} \tag{2.23}$$

$$= \sum_{k=\frac{n}{\nu}\nu}^{\infty} \boldsymbol{F}^k \boldsymbol{e} \tag{2.24}$$

$$\le \sum_{k=[\frac{n}{\nu}]\nu}^{\infty} \boldsymbol{F}^k \boldsymbol{e} \tag{2.25}$$

$$\le \frac{\nu\delta^{[\frac{n}{\nu}]}}{1-\delta}\boldsymbol{e}\ [\because (2.10)\ \text{式}] \tag{2.26}$$

ここで, $n \geq N(\varepsilon)$ の条件より

$$n \geq \left\lceil \nu \frac{\log\{\varepsilon\delta(1-\delta)/\nu\}}{\log\delta} \right\rceil \tag{2.27}$$

$$n \geq \nu \frac{\log\{\varepsilon\delta(1-\delta)/\nu\}}{\log\delta} \tag{2.28}$$

$$\frac{n}{\nu} \geq \frac{\log\{\varepsilon\delta(1-\delta)/\nu\}}{\log\delta} \tag{2.29}$$

$$\frac{n}{\nu}\log\delta \leq \log\varepsilon + \log(1-\delta) + \log\delta - \log\nu \tag{2.30}$$

$$\log\nu + \left(\frac{n}{\nu} - 1\right)\log\delta - \log(1-\delta) \leq \log\varepsilon \tag{2.31}$$

$$\log\nu + \left[\frac{n}{\nu}\right]\log\delta - \log(1-\delta) \leq \log\varepsilon \quad \left[\because \frac{n}{\nu} - 1 \leq \left[\frac{n}{\nu}\right], \log\delta < 0\right] \tag{2.32}$$

$$\log\frac{\nu\delta^{\left[\frac{n}{\nu}\right]}}{1-\delta} \leq \log\varepsilon \tag{2.33}$$

$$\frac{\nu\delta^{\left[\frac{n}{\nu}\right]}}{1-\delta} \leq \varepsilon \tag{2.34}$$

よって

$$\frac{\nu\delta^{\left[\frac{n}{\nu}\right]}}{1-\delta}\boldsymbol{e} \leq \varepsilon\boldsymbol{e} \tag{2.35}$$

以上から

$$\sum_{k=n+1}^{\infty} \boldsymbol{F}^k\boldsymbol{e} \leq \varepsilon\boldsymbol{e} \quad \square \tag{2.36}$$

(d) 式 (2.22) は次式と等価である.

$$\left\| \sum_{k=0}^{n} \boldsymbol{F}^k - (\boldsymbol{I} - \boldsymbol{F})^{-1} \right\|_{\infty} \leq \varepsilon, \forall n \geq N(\varepsilon) \tag{2.37}$$

ただし, $\|\cdot\|_\infty$は $\infty$-ノルムを表す.

式 (2.37) から次のことがわかる.任意の$\varepsilon \in (0,1)$ に対して, $\boldsymbol{I}, \boldsymbol{F}, \boldsymbol{F}^2, \ldots, \boldsymbol{F}^{N(\varepsilon)}$を順に足し上げた結果は, $(\boldsymbol{I} - \boldsymbol{F})^{-1}$の近似とみなすことができ, その誤差は $\infty$-ノルムで測ると$\varepsilon$未満になる.この事実に基づいた $(\boldsymbol{I} - \boldsymbol{F})^{-1}$の近似計算手順 (アルゴリズムAと呼ぶ) を書き下しなさい (アルゴリズム 1 の記述を参考にすること).

解

アルゴリズム A

入力 : 確率行列 $\boldsymbol{F}, \varepsilon \in (0,1), \delta \in (0,1)$

変数 : $\boldsymbol{X} = (X_{i,j})_{i,j\in\mathbb{S}}, \boldsymbol{Y} = (Y_{i,j})_{i,j\in\mathbb{S}}, \boldsymbol{Z} = (Z_{i,j})_{i,j\in\mathbb{S}}, \boldsymbol{A} = (A_{i,j})_{i,j\in\mathbb{S}}, \nu, N(\varepsilon)$

(i)$\boldsymbol{X} := \boldsymbol{F}, \boldsymbol{Y} := \boldsymbol{I}, \boldsymbol{Z} := \boldsymbol{I}, \nu := 0$

(ii)$\boldsymbol{Ye} \leq \delta \boldsymbol{e}$ となるまで, 以下の 3 つの操作を上から順に繰り返し実行する.

$$\boldsymbol{A} := \boldsymbol{Y}$$
$$\boldsymbol{Y} := \boldsymbol{AX}$$
$$\nu := \nu + 1$$

(iii)$N(\varepsilon) := \left\lceil \nu \frac{\log\{\varepsilon\delta(1-\delta)/\nu\}}{\log \delta} \right\rceil$
(iv)$\boldsymbol{Y} := \boldsymbol{I}$
(v)以下の 3 つの操作を上から順に $N(\varepsilon)$ 回繰り返し実行する.

$$\boldsymbol{A} := \boldsymbol{Y}$$
$$\boldsymbol{Y} := \boldsymbol{AX}$$
$$\boldsymbol{Z} := \boldsymbol{Z} + \boldsymbol{Y}$$

(vi)$\boldsymbol{Z}$ を結果として出力する.
このアルゴリズムにおける $\boldsymbol{Z}$ が $(\boldsymbol{I} - \boldsymbol{F})^{-1}$の近似となる.

(e) つづいて, 以下の再帰式で生成される列 $\{\boldsymbol{V_n}; n \in \mathbb{Z}_+\}$ を考える.

$$\boldsymbol{U_0} = \boldsymbol{F},\ \boldsymbol{V_0} = \boldsymbol{I} \tag{2.38}$$
$$\boldsymbol{U_n} = \boldsymbol{U_{n-1}^2}, \qquad n \geq 1 \tag{2.39}$$
$$\boldsymbol{V_n} = (\boldsymbol{I} + \boldsymbol{U_{n-1}})\boldsymbol{V_{n-1}}, \quad n \geq 1 \tag{2.40}$$

次式が成り立つことを示しなさい.

$$\lim_{n\to\infty} \boldsymbol{V_n} = (\boldsymbol{I} - \boldsymbol{F})^{-1} \tag{2.41}$$

解
$\boldsymbol{U_n}$ についての再帰式 (2.39) より,

$$\boldsymbol{U_n} = \boldsymbol{U_{n-1}}^2 = \boldsymbol{U_{n-2}}^{2^2} = \cdots = \boldsymbol{U_0}^{2^n} = \boldsymbol{F}^{2^n} \tag{2.42}$$

よって,$\boldsymbol{U_n} = \boldsymbol{F}^{2^n}$である.
次 に $\boldsymbol{V_n}$ について

$$\boldsymbol{V_1} = (\boldsymbol{I} + \boldsymbol{U_0})\boldsymbol{V_0} = \boldsymbol{I} + \boldsymbol{F} \tag{2.43}$$
$$\begin{aligned}\boldsymbol{V_2} = (\boldsymbol{I} + \boldsymbol{U_1})\boldsymbol{V_1} &= (\boldsymbol{I} + \boldsymbol{F}^2)(\boldsymbol{I} + \boldsymbol{F}) \\ &= \boldsymbol{I} + \boldsymbol{F} + \boldsymbol{F}^2 + \boldsymbol{F}^3\end{aligned} \tag{2.44}$$
$$\begin{aligned}\boldsymbol{V_3} = (\boldsymbol{I} + \boldsymbol{U_2})\boldsymbol{V_2} &= (\boldsymbol{I} + \boldsymbol{F}^4)(\boldsymbol{I} + \boldsymbol{F} + \boldsymbol{F}^2 + \boldsymbol{F}^3) \\ &= \boldsymbol{I} + \boldsymbol{F} + \boldsymbol{F}^2 + \cdots + \boldsymbol{F}^7\end{aligned} \tag{2.45}$$

$\boldsymbol{V_1}, \boldsymbol{V_2}, \boldsymbol{V_3}$ から, $\boldsymbol{V_n} = \sum_{k=0}^{2^n-1} \boldsymbol{F}^k$であることが推測できる.$\boldsymbol{V_n} = \sum_{k=0}^{2^n-1} \boldsymbol{F}^k$を $n$ についての数学的帰納法により証明する.

(i)$n=0$ のとき,

$$(\text{左辺}) = \boldsymbol{V_0} = \boldsymbol{I} \tag{2.46}$$

$$(\text{右辺}) = \sum_{k=0}^{2^0-1} \boldsymbol{F}^k = \sum_{k=0}^{0} \boldsymbol{F}^k = \boldsymbol{I} \tag{2.47}$$

となり, 成り立つ.

(ii)$n=m$ のとき, $\boldsymbol{V_m} = \sum_{k=0}^{2^m-1} \boldsymbol{F}^k$が成り立つと仮定すると,

$n=m+1$ のとき

$$\begin{aligned}
\boldsymbol{V_{m+1}} &= (\boldsymbol{I} + \boldsymbol{U_m})\boldsymbol{V_m} && (2.48)\\
&= (\boldsymbol{I} + \boldsymbol{F}^{2^m}) \sum_{k=0}^{2^m-1} \boldsymbol{F}^k && (2.49)\\
&= \sum_{k=0}^{2^m-1} \boldsymbol{F}^k + \sum_{k=0}^{2^m-1} \boldsymbol{F}^{2^m+k} && (2.50)\\
&= \sum_{k=0}^{2^m-1} \boldsymbol{F}^k + \sum_{k=2^m}^{2^{m+1}-1} \boldsymbol{F}^k && (2.51)\\
&= \sum_{k=0}^{2^{m+1}-1} \boldsymbol{F}^k && (2.52)
\end{aligned}$$

よって, $n=m+1$ のときも成り立つ.

したがって, $\boldsymbol{V_n} = \sum_{k=0}^{2^n-1} \boldsymbol{F}^k$であり, $\lim_{n\to\infty} \boldsymbol{V_n} = \sum_{k=0}^{\infty} \boldsymbol{F}^k = (\boldsymbol{I} - \boldsymbol{F})^{-1}$ [$\because$ (2.3) 式] $\square$

(f) 再帰式 (2.38)(2.39)(2.40) を用いると,$(\boldsymbol{I} - \boldsymbol{F})^{-1}$を近似的に計算できるが, その際, 繰り返し計算の停止基準を設ける必要がある.そこで, 任意の$\varepsilon \in (0,1)$ に対して,

$$\left\| \boldsymbol{V_n} - (\boldsymbol{I} - \boldsymbol{F})^{-1} \right\|_\infty \le \varepsilon \tag{2.53}$$

を満たす最小の自然数を$\tilde{N}(\varepsilon)$ とし, $\boldsymbol{V_{\tilde{N}(\varepsilon)}}$ を $(\boldsymbol{I} - \boldsymbol{F})^{-1}$の近似として定めることにする.

以上の設定のもと, $\boldsymbol{V_{\tilde{N}(\varepsilon)}}$ の計算手順 (アルゴリズムBと呼ぶ) 書き下しなさい (アルゴリズム 1 の記述を参考にすること).

解

アルゴリズム B

入力 : 確率行列 $\boldsymbol{F}, \varepsilon \in (0,1), \delta \in (0,1)$

変数 : $\boldsymbol{X} = (X_{i,j})_{i,j\in\mathbb{S}}, \boldsymbol{Y} = (Y_{i,j})_{i,j\in\mathbb{S}}, \boldsymbol{A} = (A_{i,j})_{i,j\in\mathbb{S}}, \boldsymbol{B} = (B_{i,j})_{i,j\in\mathbb{S}}, \nu, N(\varepsilon), \tilde{N}(\varepsilon)$

(i)$\boldsymbol{X} := \boldsymbol{F}, \boldsymbol{Y} := \boldsymbol{I}, \nu := 0$

(ii)$\boldsymbol{Y}\boldsymbol{e} \le \delta \boldsymbol{e}$ となるまで, 以下の 3 つの操作を上から順に繰り返し実行する.

$$\begin{aligned}
\boldsymbol{A} &:= \boldsymbol{Y}\\
\boldsymbol{Y} &:= \boldsymbol{A}\boldsymbol{X}\\
\nu &:= \nu + 1
\end{aligned}$$

(iii)$N(\varepsilon) := \left\lceil \nu \frac{\log\{\varepsilon\delta(1-\delta)/\nu\}}{\log\delta} \right\rceil, \tilde{N}(\varepsilon) = \lceil \log_2\{N(\varepsilon)+1\} \rceil$
(iv)$\boldsymbol{Y} := \boldsymbol{I}$
(v)以下の 4 つの操作を上から順に$\tilde{N}(\varepsilon)$ 回繰り返し実行する.

$$\begin{aligned}\boldsymbol{A} &:= \boldsymbol{X}\\ \boldsymbol{B} &:= \boldsymbol{Y}\\ \boldsymbol{X} &:= \boldsymbol{A}^2\\ \boldsymbol{Y} &:= (\boldsymbol{I}+\boldsymbol{A})\boldsymbol{B}\end{aligned}$$

(vi)$\boldsymbol{Y}$ を結果として出力する.
このアルゴリズムにおける $\boldsymbol{Y}$ が $(\boldsymbol{I}-\boldsymbol{F})^{-1}$の近似となる.

(g) 許容誤差 $\varepsilon$ を同一にしたとき, アルゴリズム A と B の計算量について, 適宜数式を用いて考察しなさい.

解

入力する確率行列のサイズは $M$ とする. また, 許容誤差が同じとき, アルゴリズム A と B の手順 (ii) の繰り返し回数は同じであり, その回数を $l$ とする. このとき, アルゴリズム A と B の計算量はそれぞれ以下のようになる.

アルゴリズム A:
(i)$O(3M^2+1) = O(M^2)$
(ii)1 回の繰り返しにあたり, ループの停止判断に $O(M^2)$, 行列 $\boldsymbol{A}, \boldsymbol{Y}$, および$\nu$の値更新にそれぞれ $O(M^2), O(M^3), O(1)$ の計算量がかかる. 手順 (ii) の繰り返し回数は $l$ であるので, 計算量の合計は $O(lM^3)$
(iii)行列のサイズに依存しないので, $O(1)$
(iv)$O(M^2)$
(v)1 回の繰り返しにあたり, 行列 $\boldsymbol{A}, \boldsymbol{Y}$ および $\boldsymbol{Z}$ の値更新にそれぞれ $O(M^2), O(M^3), O(M^2)$ の計算量がかかる. 手順 (v) の繰り返し回数は $N(\varepsilon)$ であるので, 計算量の合計は $O\left(N(\varepsilon)M^3\right)$

以上により, アルゴリズム A の計算量は

$$O(M^2) + O(lM^3) + O(1) + O(M^2) + O\left(N(\varepsilon)M^3\right) = O(lM^3) + O\left(N(\varepsilon)M^3\right)$$

アルゴリズム B:
(i)$O(2M^2+1) = O(M^2)$
(ii)1 回の繰り返しにあたり, ループの停止判断に $O(M^2)$, 行列 $\boldsymbol{A}, \boldsymbol{Y}$, および$\nu$の値更新にそれぞれ $O(M^2), O(M^3), O(1)$ の計算量がかかる. 手順 (ii) の繰り返し回数は $l$ であるので, 計算量の合計は $O(lM^3)$
(iii)$N(\varepsilon), \tilde{N}(\varepsilon)$ ともに行列のサイズに依存しないので, $O(1)$
(iv)$O(M^2)$
(v)1 回の繰り返しにあたり, 行列 $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{X}$ および $\boldsymbol{Y}$ の値更新にそれぞれ $O(M^2), O(M^2), O(M^3), O(M^3)$ の計算量がかかる. 手順 (v) の繰り返し回数は $\tilde{N}(\varepsilon)$ であるので, 計算量の合計は $O\left(\tilde{N}(\varepsilon)M^3\right)$

以上により, アルゴリズム B の計算量は

$$O(M^2) + O(lM^3) + O(1) + O(M^2) + O\left(\tilde{N}(\varepsilon)M^3\right) = O(lM^3) + O\left(\tilde{N}(\varepsilon)M^3\right)$$

ここで,$\tilde{N}(\varepsilon) = \lceil \log_2\{N(\varepsilon)+1\} \rceil$ であるが, 誤差は高々 1 であるので$\tilde{N}(\varepsilon) = \log_2\{N(\varepsilon)+1\}$ と近似してみなすことができる.よって, アルゴリズムA の計算量が $O(lM^3) + O\left(N(\varepsilon)M^3\right)$ であるのに対し,B の計算量は $O(lM^3) + O(M^3 \log_2\{N(\varepsilon)+1\})$ であり, 第二項が $N(\varepsilon)$ の log のオーダーであるので, アルゴリズム B は A と比較して効率的なアルゴリズムであることがわかる.

## 3　課題 3

(a) 少なくとも一つのぶら下がりノードを含むような状態数が 5 以上のハイパーリンク行列 $\boldsymbol{H}$ の例を作り, 対応する確率行列 $\boldsymbol{S}$ とGoogle行列 $\boldsymbol{G}$ を書き下しなさい.

解

$$\boldsymbol{H} = \begin{pmatrix} 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \boldsymbol{S} = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 1 \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix}$$

$$\boldsymbol{G} = \alpha\boldsymbol{S} + (1-\alpha)\boldsymbol{E} = \begin{pmatrix} \frac{1-\alpha}{5} & \frac{2+3\alpha}{10} & \frac{1-\alpha}{5} & \frac{2+3\alpha}{10} & \frac{1-\alpha}{5} \\ \frac{1-\alpha}{5} & \frac{1-\alpha}{5} & \frac{1+4\alpha}{5} & \frac{1-\alpha}{5} & \frac{1-\alpha}{5} \\ \frac{2+3\alpha}{10} & \frac{1-\alpha}{5} & \frac{1-\alpha}{5} & \frac{1-\alpha}{5} & \frac{2+3\alpha}{10} \\ \frac{1-\alpha}{5} & \frac{1-\alpha}{5} & \frac{1-\alpha}{5} & \frac{1-\alpha}{5} & \frac{1+4\alpha}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix}$$

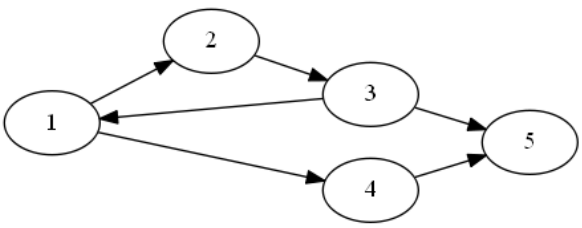


図 2　課題 3 のハイパーリンク行列 $\boldsymbol{H}$ に対応するグラフ

図 2 より, ノード 5 がぶら下がりノードになっていることが確認できる.

(b) (a) で書き下した Google 行列が定める PageRank $\boldsymbol{\gamma} = (\gamma_i)_{i=1,2,\ldots,N}$ は $\boldsymbol{F} = \alpha \boldsymbol{S}$ とおいてアルゴリズム A あるいは B を実行すれば求まる (式 (3.1) および (3.2) 参照). そこで, 実際に, アルゴリズム A および B を用いた $\boldsymbol{\gamma}$ の数値計算プログラムを実装し,

$$\alpha = 0.01, 0.1, 0.5, 0.85, 0.99$$

に対する $\boldsymbol{\gamma}$ の計算結果をそれぞれ出力しなさい.

$$\boldsymbol{\gamma} = \frac{(1-\alpha)\boldsymbol{e}^\top}{N}(\boldsymbol{I} - \alpha \boldsymbol{S})^{-1} \tag{3.1}$$

$$(\boldsymbol{I} - \alpha \boldsymbol{S})^{-1} = \sum_{n=0}^{\infty} (\alpha \boldsymbol{S})^n \tag{3.2}$$

解

アルゴリズム A を用いて実装した数値計算プログラムに行列 $\boldsymbol{S}$ を入力として与えたときの出力結果は以下である.

$\alpha = 0.010000$ のとき $\boldsymbol{\gamma} = (0.199405, 0.199400, 0.200397, 0.199400, 0.201399)$
$\alpha = 0.100000$ のとき $\boldsymbol{\gamma} = (0.194461, 0.194000, 0.203677, 0.194000, 0.213861)$
$\alpha = 0.500000$ のとき $\boldsymbol{\gamma} = (0.179592, 0.171429, 0.212245, 0.171429, 0.265306)$
$\alpha = 0.850000$ のとき $\boldsymbol{\gamma} = (0.172422, 0.154986, 0.213446, 0.154986, 0.304160)$
$\alpha = 0.990000$ のとき $\boldsymbol{\gamma} = (0.170349, 0.149322, 0.212828, 0.149322, 0.318178)$

アルゴリズム B を用いて実装した数値計算プログラムに行列 $\boldsymbol{S}$ を入力として与えたときの出力結果は以下である.

$\alpha = 0.010000$ のとき $\boldsymbol{\gamma} = (0.199405, 0.199400, 0.200397, 0.199400, 0.201399)$
$\alpha = 0.100000$ のとき $\boldsymbol{\gamma} = (0.194461, 0.194000, 0.203677, 0.194000, 0.213861)$
$\alpha = 0.500000$ のとき $\boldsymbol{\gamma} = (0.179592, 0.171429, 0.212245, 0.171429, 0.265306)$
$\alpha = 0.850000$ のとき $\boldsymbol{\gamma} = (0.172422, 0.154986, 0.213446, 0.154986, 0.304160)$
$\alpha = 0.990000$ のとき $\boldsymbol{\gamma} = (0.170349, 0.149322, 0.212828, 0.149322, 0.318178)$

(c) (b)と同じ設定の下で数値実験を行い, $\alpha$を横軸に, アルゴリズムAの $N(\varepsilon)$ およびアルゴリズムBの$\tilde{N}(\varepsilon)$ を縦軸に取ったグラフを作成しなさい.ただし, $\varepsilon$は $10^{-8}$の値で固定し, $\alpha$の値の刻み幅は, グラフの線が十分なめらかになるように, 十分小さく取ること.

**解**

$\alpha$の刻み幅を 0.001 として, $N(\varepsilon)$ と$\tilde{N}(\varepsilon)$ をプロットしたグラフが図 3 である.

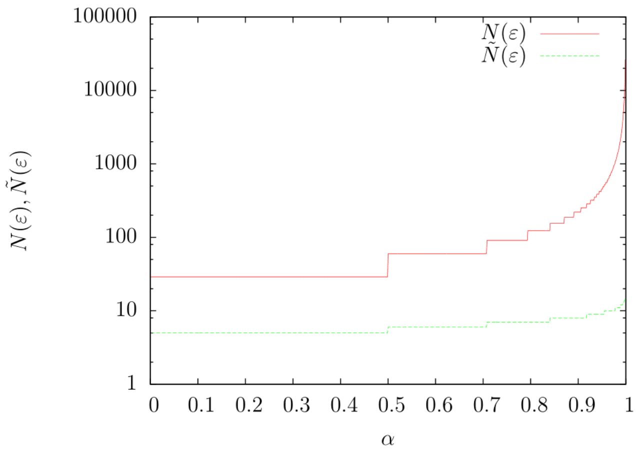


図 3　$\alpha$と $N(\varepsilon)$ の関係 (赤線), $\alpha$と$\tilde{N}(\varepsilon)$ の関係 (緑線)

なお, 図 3 のグラフは縦軸が対数軸の片対数グラフであることに注意されたい.

図 3 より, $N(\varepsilon)$ は$\alpha$が 1 に近づいていくにつれ, 値が 5 桁まで増えた一方で, $\tilde{N}(\varepsilon)$ は$\alpha$が 1 に近づいても 2 桁に収まっていることがわかる. $N(\varepsilon)$ と$\tilde{N}(\varepsilon)$ は それぞれアルゴリズム A と B の手順 (v)の繰り返し回数であり, $N(\varepsilon)$ に比べ$\tilde{N}(\varepsilon)$ がはるかに小さいので, 課題 2 の (g) でも述べたようにアルゴリズム B のほうがアルゴリズム Aと比較して効率的である.

## 4　課題 4

(a) Google 行列 $\boldsymbol{G}$ が定める PageRank, すなわち, $\boldsymbol{G}$ の定常分布ベクトル $\boldsymbol{\gamma} = (\gamma_i)_{i=1,2,\dots,N}$が式 (4.1) で与えられることを示しなさい.

$$\boldsymbol{\gamma} = \frac{(1-\alpha)\boldsymbol{e}^\top}{N}(\boldsymbol{I} - \alpha\boldsymbol{S})^{-1} \tag{4.1}$$

**解**

$\boldsymbol{\gamma}$ は $\boldsymbol{G}$ の定常分布ベクトルであるので,

$$\boldsymbol{\gamma}\boldsymbol{G} = \boldsymbol{\gamma} \tag{4.2}$$

が成り立つ. この式を変形し,

$$\boldsymbol{\gamma}\left\{\alpha\boldsymbol{S}+(1-\alpha)\frac{\boldsymbol{e}\boldsymbol{e}^\top}{N}\right\}=\boldsymbol{\gamma} \tag{4.3}$$

$$\boldsymbol{\gamma}(\boldsymbol{I}-\alpha\boldsymbol{S})=\frac{1-\alpha}{N}\boldsymbol{\gamma}\boldsymbol{e}\boldsymbol{e}^\top \tag{4.4}$$

$$\boldsymbol{\gamma}(\boldsymbol{I}-\alpha\boldsymbol{S})=\frac{1-\alpha}{N}\boldsymbol{e}^\top\ [\because \boldsymbol{\gamma}\text{ の成分の総和は }1] \tag{4.5}$$

$$\boldsymbol{\gamma}=\frac{(1-\alpha)\boldsymbol{e}^\top}{N}(\boldsymbol{I}-\alpha\boldsymbol{S})^{-1}\quad\square \tag{4.6}$$

(b) 確率行列 $\boldsymbol{S}$ の全ての固有値 (ただし重複を許す) が $\{1,\lambda_2,\lambda_3,\dots,\lambda_N\}$ であるとすると,Google行列 $\boldsymbol{G}$ の全ての固有値は,

$$\{1,\alpha\lambda_2,\alpha\lambda_3,\dots,\alpha\lambda_N\} \tag{4.7}$$

となることを示しなさい.

解

任意の $N$ 次確率行列 $\boldsymbol{A}$ において, 1 以外の固有値に属する固有ベクトルはその成分の総和が 0 であること $\cdots(*)$ を示す.

今, $\boldsymbol{x}$ が $\boldsymbol{A}$ の固有値$\lambda(\neq 1)$ に属する固有ベクトルとする.このとき

$$\boldsymbol{x}\boldsymbol{A}=\lambda\boldsymbol{x} \tag{4.8}$$

ここで, $\boldsymbol{d}=(1,1,\dots,1)$ とすると, $\boldsymbol{A}$ の各行の成分の総和が 1 であることから

$$\boldsymbol{d}\boldsymbol{A}^\top=\boldsymbol{d} \tag{4.9}$$

(4.8) 式の両辺と $\boldsymbol{d}$ の内積をとると

$$(\boldsymbol{d},\boldsymbol{x}\boldsymbol{A})=(\boldsymbol{d},\lambda\boldsymbol{x}) \tag{4.10}$$

$$(\text{左辺})=\boldsymbol{d}(\boldsymbol{x}\boldsymbol{A})^\top=\boldsymbol{d}\boldsymbol{A}^\top\boldsymbol{x}^\top=(\boldsymbol{d}\boldsymbol{A}^\top,\boldsymbol{x})=(\boldsymbol{d},\boldsymbol{x}) \tag{4.11}$$

$$(\text{右辺})=\lambda(\boldsymbol{d},\boldsymbol{x}) \tag{4.12}$$

$$\therefore (1-\lambda)(\boldsymbol{d},\boldsymbol{x})=0 \tag{4.13}$$

$\lambda\neq 1$ なので, $(\boldsymbol{d},\boldsymbol{x})=0$, すなわち $\sum_{j=1}^{N}x_j=0$ である. 以上より $(*)$ が示された.

次に, 本題を証明する.

まず, ペロン = フロベニウスの定理より, 既約な確率行列ならば最大固有値 1 の重複度は 1 である.行列 $\boldsymbol{S}$ は既約であるので, $\lambda_2,\lambda_3,\dots,\lambda_N$は 1 ではない.

各固有ベクトル$\lambda_i(i=2,3,\dots,N)$ に属する固有ベクトルを $\boldsymbol{x_i}$($\lambda_i\neq 1$ なので $(*)$ より $\boldsymbol{x_i}$ の成分の総和は 0) とすると

$$\boldsymbol{x_i}\boldsymbol{S}=\lambda_i\boldsymbol{x_i} \tag{4.14}$$

$\boldsymbol{x_i}$ と $\boldsymbol{G}$ の積をとると

$$\boldsymbol{x_i G} = \boldsymbol{x_i}\left(\alpha \boldsymbol{S} + \frac{1-\alpha}{N}\boldsymbol{e}\boldsymbol{e}^\top\right) \tag{4.15}$$

$$= \alpha \boldsymbol{x_i S} + \frac{1-\alpha}{N}\boldsymbol{x_i e e}^\top \tag{4.16}$$

$$= \alpha\lambda_i \boldsymbol{x_i}\ [\because (4.14)\ 式, \boldsymbol{x_i e} = 0] \tag{4.17}$$

これは, $\alpha\lambda_i$が $\boldsymbol{G}$ の固有値であることを示している.よって, $\alpha\lambda_2, \alpha\lambda_3, \ldots, \alpha\lambda_N$は $\boldsymbol{G}$ の固有値である.
また, $\boldsymbol{\gamma}$ は $\boldsymbol{G}$ の定常分布ベクトルなので $\boldsymbol{\gamma G} = \boldsymbol{\gamma}$ となり, 1 は $\boldsymbol{G}$ の固有値である.
以上から, $\{1, \alpha\lambda_2, \alpha\lambda_3, \ldots, \alpha\lambda_N\}$ は $\boldsymbol{G}$ の固有値である.

最後に, $\boldsymbol{G}$ の固有値が $\{1, \alpha\lambda_2, \alpha\lambda_3, \ldots, \alpha\lambda_N\}$ のみであることを示す.
ここで, $\boldsymbol{G}$ が $1, \alpha\lambda_2, \alpha\lambda_3, \ldots, \alpha\lambda_N$とは異なる固有値$\lambda'$をもつと仮定する.(背理法)
$\lambda'$に属する固有ベクトルを $\boldsymbol{y}$ ($\lambda' \neq 1$ なので $(*)$ より $\sum_{j=1}^{N} y_j = 0$) とすると

$$\boldsymbol{yG} = \lambda' \boldsymbol{y} \tag{4.18}$$

$$\boldsymbol{y}\left(\alpha \boldsymbol{S} + \frac{1-\alpha}{N}\boldsymbol{e}\boldsymbol{e}^\top\right) = \lambda' \boldsymbol{y} \tag{4.19}$$

$$\alpha \boldsymbol{yS} + \frac{1-\alpha}{N}\boldsymbol{yee}^\top = \lambda' \boldsymbol{y} \tag{4.20}$$

$$\alpha \boldsymbol{yS} = \lambda' \boldsymbol{y}\ [\because \boldsymbol{ye} = 0] \tag{4.21}$$

$$\boldsymbol{yS} = \frac{\lambda'}{\alpha}\boldsymbol{y} \tag{4.22}$$

つまり, $\frac{\lambda'}{\alpha}$は $\boldsymbol{S}$ の固有値である.
しかし, $\lambda' \neq \alpha\lambda_2, \alpha\lambda_3, \ldots, \alpha\lambda_N$なので$\frac{\lambda'}{\alpha} \neq \lambda_2, \lambda_3, \ldots, \lambda_N$である.また, $\frac{\lambda'}{\alpha} = 1$ とすると, ペロン = フロベニウスの定理より最大固有値 1 に属する固有ベクトルは非負または非正であるので, $\sum_{j=1}^{N} y_j = 0$ より, $y_1 = y_2 = \cdots = y_N = 0$ となる.これは固有ベクトルの定義に反するので, $\frac{\lambda'}{\alpha} \neq 1$ である.
以上から, $\frac{\lambda'}{\alpha} \neq 1, \lambda_2, \lambda_3, \ldots, \lambda_N$であるが, これは $\boldsymbol{S}$ の固有値が $\{1, \lambda_2, \lambda_3, \ldots, \lambda_N\}$ であることに矛盾する.
したがって, $\boldsymbol{G}$ の固有値は $\{1, \alpha\lambda_2, \alpha\lambda_3, \ldots, \alpha\lambda_N\}$ のみである. □