

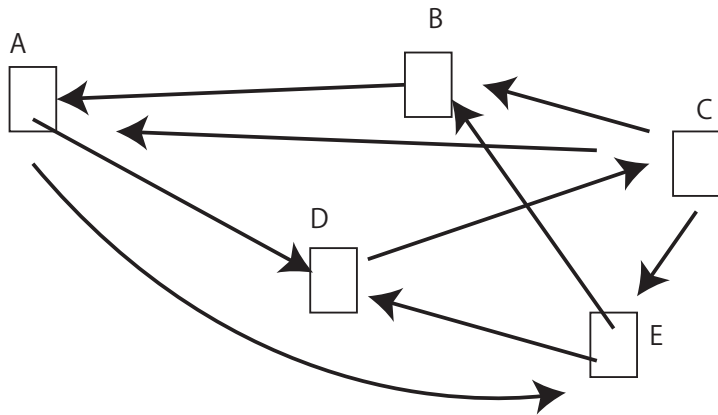
確率行列のペロン-フロベニウスの定理と web 検索

2017.12.18

ウェブ世界を作った定理

- ▶ インフラとしてのウェブ検索技術
- ▶ Google's PageRank
- ▶ 行列に関するペロン-フロベニウスの定理とウェブのリンク構造を結びつけることで検索結果の順位づけに成功
- ▶ 全く違う見かけの概念を結びつける能力

ホームページのリンク関係



数学的には有限グラフとよぶ。

リンクの関係

ホームページ A,B,C,D,E について、リンク関係がわかっている。

1. A から D,E へリンクしている。
2. B から A へリンクしている。
3. C から A,B,E へリンクしている。
4. D から C へリンクしている。
5. E から B,D へリンクしている。

ランキングと連立方程式

1. ホームページ A のランクを x_A という変数で表す。
2. A のランクをリンク先 D, E へ均等配分する。
3. B, C, D, E について繰り返し。

不定形の連立方程式を立てる。全て 0 となる自明な解以外に解は存在するか。

$$x_A = x_B + x_C/3$$

$$x_B = x_C/3 + x_E$$

$$x_C = x_D$$

$$x_D = x_A/2 + x_E/2$$

$$x_E = x_A/2 + x_C/3$$

行列で書き直すと

$$\begin{pmatrix} x_A & x_B & x_C & x_D & x_E \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 & 0 \\ 1/3 & 1/3 & 0 & 0 & 1/3 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \end{pmatrix}$$

確率行列（各行の成分の和が1になる行列）の固有値問題。固有値1に対応して各成分が決まる。

$(x_A, x_B, x_C, x_D, x_E) = (0.52, 0.36, 0.46, 0.46, 0.41)$ この成分が大きいほどランクは高いと解釈する。

確率行列に関するペロン-フロベニウスの定理

$n \times n$ 確率行列 $P = (p_{ij})_{1 \leq i, j \leq n}$ の定義:

- ▶ $p_{ij} \geq 0$
- ▶ $\sum_{j=1}^n p_{ij} = 1$

$$\begin{pmatrix} 0 & 0 & 0 & 1/2 & 1/2 \\ 1 & 0 & 0 & 0 & 0 \\ 1/3 & 1/3 & 0 & 0 & 1/3 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \end{pmatrix}$$

確率行列に関するペロン-フロベニウスの定理

既約な確率行列 P の定義:

- ▶ $P^N = (p_{ij}^{(N)})$ とおく。
- ▶ 任意の $1 \leq i, j \leq n$ について自然数 N が存在し、 $p_{ij}^{(N)} > 0$ が成立。

確率行列に関するペロン-フロベニウスの定理

単位行列は既約ではない。

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

次の行列は既約。

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

確率行列に関するペロン-フロベニウスの定理

既約な確率行列について次が成立:

- ▶ 固有値 1 は単純
- ▶ 固有値 1 に対応する固有ベクトルの成分は全て正
- ▶ 固有値 1 は絶対値最大の固有値

確率行列に関するペロン-フロベニウスの定理

P の固有値 1 に対応する左固有ベクトル p , $p = pP$ の求め方: 行ベクトル v を確率ベクトルとして、次が成立。

$$p = \lim_{N \rightarrow \infty} vP^N$$

繰り返して行列 P をかけることから、べき乗法とよぶ。

初期ベクトル v を固有ベクトルの線形結合によって $v = p + c_2 u_2 + \cdots + c_n u_n$ とおき、次が成立。

$$vP^N = p + c_2 \lambda_2^N u_2 + \cdots + c_n \lambda_n^N u_n$$

$|\lambda_k| < 1$ ($2 \leq k \leq n$) から、 $n \rightarrow \infty$ では p だけが残る。

参考文献

- ▶ Amy N. Langville, Carl D. Meyer, (岩野他訳) 『Google PageRank の数理 —最強検索エンジンのランキング手法を求めて』 (共立出版, 2009)
- ▶ 斎藤正彦 『線形代数演習 (基礎数学 4)』 (東大出版会)

演習

1. 例に挙げた P について固有値を求めよ。 `eigenvalues()` を使う。
2. 固有値 1 についての左固有ベクトルを一つ求めよ。前のスライドで紹介した関係を使う。例として出した固有ベクトルを確率ベクトルとすれば一致する。