東北大学
産学連携
企画

数理科学 オープンイノベーション セッション

数学者との出会いから
イノベーション創出を狙う

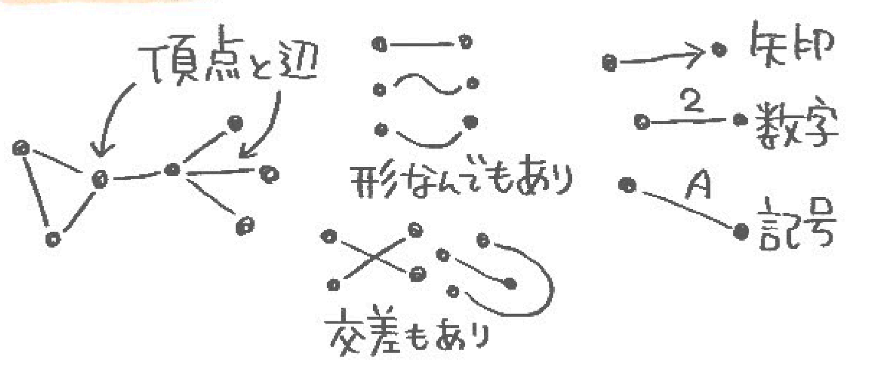
2022.12.13

#02

田中太初さん
東北大学大学院情報科学研究科
純粋・応用数学研究センター

グラフの固有値と固有ベクトル

グラフとは

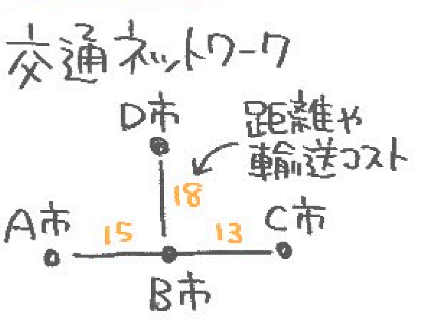


グラフ活用例

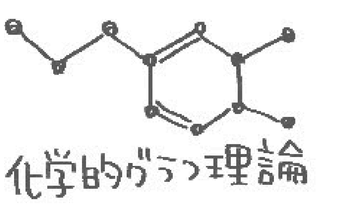
モデル化!

いろいろ使える

例 交通ネットワーク



例 分子構造



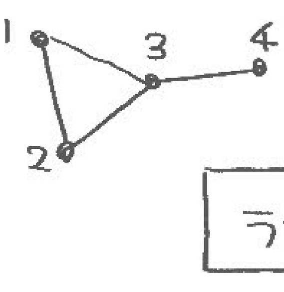
$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

固有ベクトル

固有値

☆ n次正方行列 M の(代数的)固有値は重複度を込めて丁度 n 個

グラフから定まる行列



隣接行列

$$A = \begin{bmatrix} 0&1&1&0 \\ 1&0&1&0 \\ 1&1&0&1 \\ 0&0&1&0 \end{bmatrix}$$

辺の情報が入っている

固有値 2.17009, 0.31111, −1, −1.48119

固有ベクトル
$$\begin{bmatrix} 1.85464 \\ 1.85464 \\ 2.17009 \\ 1 \end{bmatrix}$$

☆ よい性質がいろいろある

☆ 成分が非負
最大固有値の固有ベクトルの成分がどれも0でなくかつ全て同符号

ラプラス行列

$$L = \begin{bmatrix} 2&-1&-1&0 \\ -1&2&-1&0 \\ -1&-1&3&-1 \\ 0&0&-1&1 \end{bmatrix}$$

固有値 4, 3, 1, 0

☆ いろいろなよい性質がある

☆ 必ず0を最小固有値にもつ

推移行列

$$T = \begin{bmatrix} 0&1/2&1/3&0 \\ 1/2&0&1/3&0 \\ 1/2&1/2&0&1 \\ 0&0&1/3&0 \end{bmatrix}$$

各列の合計が1

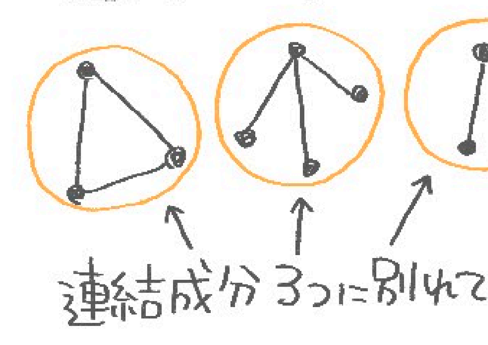
固有値 1, 0.22871, −0.5, −0.72871

☆ 必ず1を最大固有値にもつ

☆ 成分が非負
最大固有値1の固有ベクトルの成分がどれも0でなくかつ全て同符号

グラフの固有値からわかること

固有値0の重複度
=
連結成分の個数

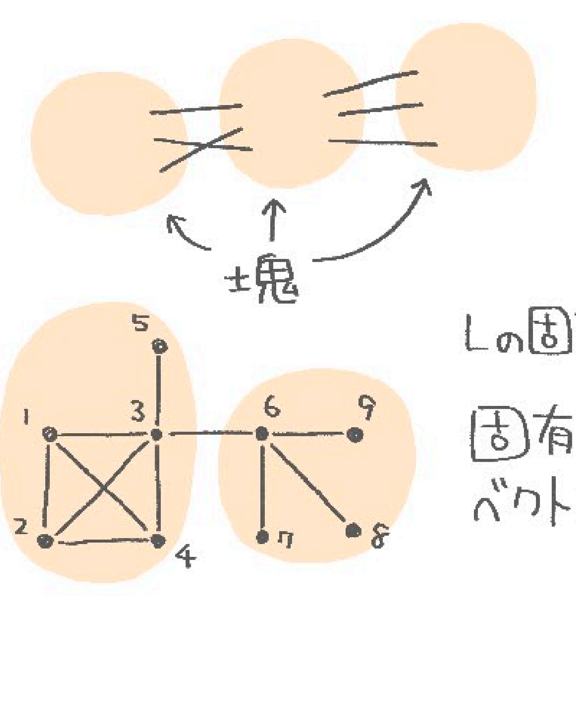


発展

スペクトラルクラスタリング

使いやすい

はやい!

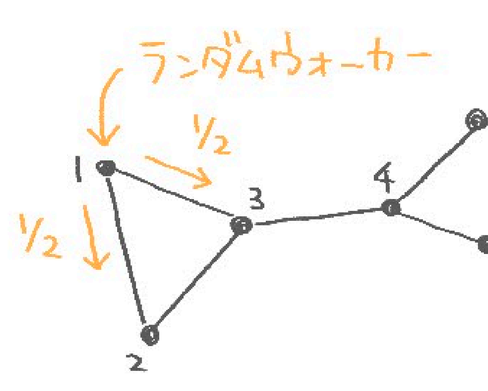


Lの固有値を小さい方から3つ使う

Lの固有値: 0, 0.32718, 1, …

固有ベクトル
$$\begin{bmatrix} -0.78600 \\ -0.78600 \\ -0.52884 \\ -0.78600 \\ -0.78600 \\ 0.67282 \\ \vdots \end{bmatrix}$$

ランダムウォーク



$$T = \begin{bmatrix} 0&1/2&1/3&0&0&0 \\ 1/2&0&1/3&0&0&0 \\ 1/2&1/2&0&1/3&0&0 \\ 0&0&1/3&0&1&1 \\ 0&0&0&1/3&0&0 \\ 0&0&0&1/3&0&0 \end{bmatrix}$$

固有値 1, 0.73741, …, −0.90407

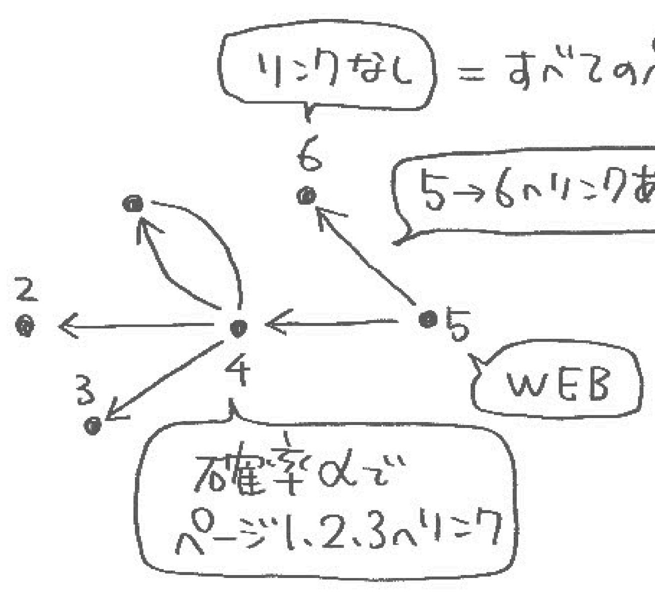
固有ベクトル
$$\begin{bmatrix} 2 \\ 2 \\ 3 \\ 3 \\ 1 \\ 1 \end{bmatrix}$$

☆ 1以外の固有値の絶対値が小さいほど収束が早い!

例 検索アルゴリズムでの活用
Google Page Rank (~2019)

Googleはこれを活用し世界最大の企業に

リンクなし = すべてのページへリンクと解釈



$$T = \begin{bmatrix} 0&1/6&1/6&1/3&0&1/6 \\ 0&1/6&1/6&1/3&0&1/6 \\ 0&1/6&1/6&1/3&0&1/6 \\ 1&1/6&1/6&0&1/2&1/6 \\ 0&1/6&1/6&0&0&1/6 \\ 0&1/6&1/6&0&1/2&1/6 \end{bmatrix}$$

$0 < \alpha < 1$ に対して

$$R = \frac{1-\alpha}{6}J + \alpha T$$

↑ 成分が全て1の正方行列

☆ ページkのランク v_k の意味
- 多くのページからリンクされると↑
- ランクの大きいページからリンクで↑
- リンク数の少ないページからリンクで↑

☆ 成分が非負
最大固有値1の固有ベクトル v の成分がどれも0でなくかつ全て同符号

☆ αの選び方がポイント
※Googleでは $\alpha = 0.85$ を使用

リンクをクリックする確率

良いグラフとは何か?

一般のグラフ × 「良い」グラフ

固有値、固有ベクトルは近似計算

汚い　まとまらない

固有値、固有ベクトルはしばしば明示的に求まる

良い性質が使える

代数的手法が使える

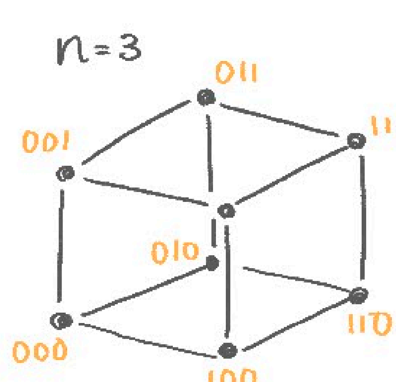
対称性をたくさんもつ

活用できる

☆ 種々の代数的手法を駆使!

例 超立方体 Q_n

良いグラフ!



- 頂点:長さnのビット列 (0000, …, 1111…, 1010…等)
- 2つのビット列は1ビットだけ異なるときに辺で結ぶ
- 異なる固有値は $n, n-2, \ldots, n-2i, \ldots, 2-n, -n$
  重複度 $\binom{n}{i}$

例 誤り訂正符号 = Q_n の頂点の集まり

↑ 符号語

誤り訂正能力を指定した上でどれだけの符号語を取れるか?

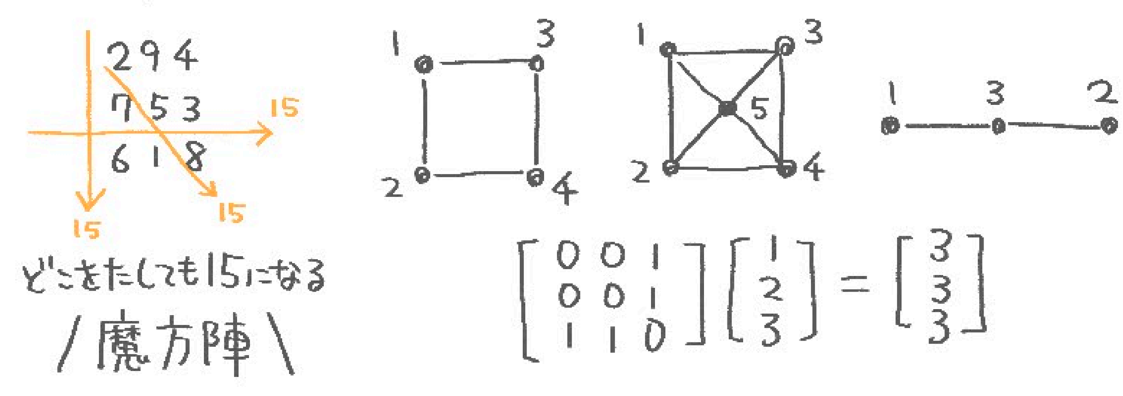
$n=3, d=3$ → 2個まで

$n=16, d=6$ → 256個まで

☆ 異なる符号語は少なくとも d ビット異なる

例 グラフの距離魔ラベリング (distance magic labeling)

各頂点に対し隣接した頂点のラベルの和が一定



$$\begin{bmatrix} 0&0&1 \\ 0&0&1 \\ 1&1&0 \end{bmatrix}\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \\ 3 \end{bmatrix}$$

☆ 超立方体 Q_n が距離魔ラベリングを持つなら n は4で割って2余る

ディスカッション

世の中の安定した自然現象はよい行列になる?

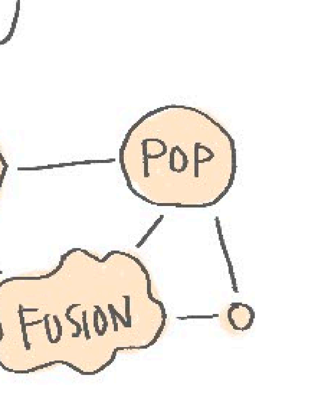
よい形のグラフは自然界にはまれな存在

自然界のものを大事なもの/でないものに分ければ数学的に捉える?

スペクトラルクラスタリング = 塊をみつけてくる 単純化してモデル化

固有値が虚数になるものは現実世界にある?

人のつながり信頼度に応用できる?



WEB広告のどれが刺さったのか分析への活用も?

スペクトラルクラスタリング = 教師なし機械学習で使われている
業務で使っていた!

WEBサイトリニューアル時 広告性質 × WEB性質 ひもづけられないか?

感性の定量化
いろんなジャンルで分岐のネットワーク

グラフ化する中で固有ベクトルで塊を見出せる

圧縮→単純化してカンタンに自然現象をとらえられないか?

ページランクを人のつながりに活用できないか?

いろんな社会リソースをつなぐ

グラフをどうつくるかが重要

どう抽象化するのか…

辺に重みをつけてもいい