

Прикладные задачи анализа данных

СПЕКТРАЛЬНАЯ ТЕОРИЯ ГРАФОВ

Дьяконов А.Г.

**Московский государственный университет
имени М.В. Ломоносова (Москва, Россия)**

Спектральная теория графов

Spectral Graph Theory

изучает свойства графов с помощью анализа

- 1) собственных значений,**
- 2) собственных векторов,**
- 3) характеристических полиномов**

матриц, которые связаны с графами:

- 1) матрица сопряжённости,**
- 2) матрица Лапласа,**
- 3) беззнаковая матрица Лапласа.**

**Спектр конечного графа – спектр её матрицы смежности,
спектр Лапласа – спектр матрицы Лапласа графа [дальше].**

Спектры не зависят от нумерации вершин.

Спектр

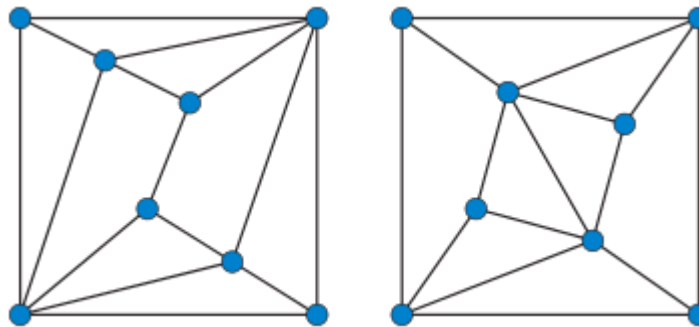
– мультимножество собственных значений.

Графы с одинаковыми спектрами – **изоспектральные**.

Изоспектральные графы не всегда изоморфны: $K_{1,4}$ и $C_4 \cup K_1$.

[Skiena, S. Implementing Discrete Mathematics: Combinatorics and Graph Theory with Mathematica. Reading, MA: Addison-Wesley, p. 85, 1990.]

Ещё пример (из полиэдральных графов).



Теорема. Почти все деревья изоспектральны.

Есть перечень известных изоспектральных графов, см.

<http://mathworld.wolfram.com/CospectralGraphs.html>

Спектр

Матрица сопряжённости неориентированного графа симметричная

\Rightarrow

собственные значения вещественные,

существует базис из ортонормированных собственных векторов.

Зачем нужен «алгебраический» подход к анализу графов.

Инвариант Колен де Вердьера $\mu(G)$ — наибольший коранг $(n - \text{rank}(M))$ среди всех матриц $M \in \mathbb{R}^{n \times n}$:

$$1) M_{ij} = \begin{cases} < 0, & (i, j) \in E, \\ 0 & (i, j) \notin E. \end{cases}$$

2) только одно отрицательное собственное значение (кратности 1),

3) выполняется строгая гипотеза Арнольда

Пример матрицы – матрица смежности

Строгая гипотеза Арнольда:

не существует симметричной матрицы $O^{n \times n} \neq X \in R^{n \times n} : MX = 0,$

$$X_{ij} = 0 \Leftrightarrow \begin{cases} i = j, \\ M_{ij} \neq 0. \end{cases}$$

в монографиях – чуть по-другому

Критерии, связанные с инвариантом.

$\mu \leq 0$ тогда и только тогда, когда нет рёбер

$\mu \leq 1$ тогда и только тогда, когда линейный лес

объединение путей

$\mu \leq 2$ тогда и только тогда, когда внешнепланарный граф

при добавлении вершины и рёбер, которые соединяют текущие вершины с добавленной получаем планарный граф

$\mu \leq 3$ тогда и только тогда, когда планарный граф

$\mu \leq 4$ тогда и только тогда, когда G *linklessly embeddable graph*

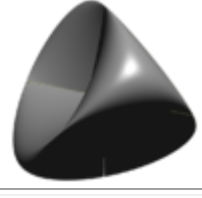
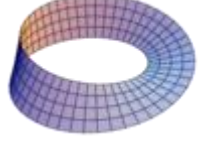
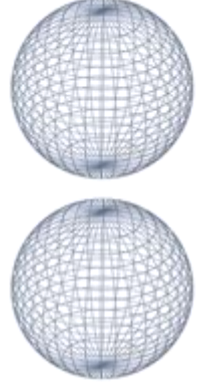
Если вложим в бутылку Клейна, то $\mu \leq 5$

Если вложим в тор, то $\mu \leq 6$

Если вложим в пов-ть с отрицательной характеристикой Эйлера k , то $\mu \leq 4-2k$

off topic – характеристикой Эйлера

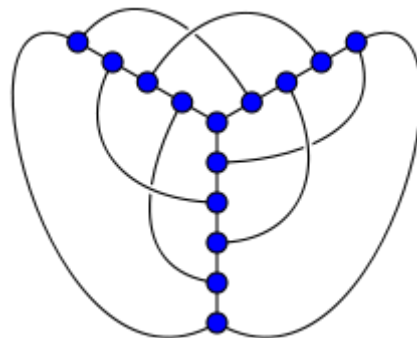
Название	Вид	Эйлерова характеристика
Окружность		0
Круг		1
сфера		2
Тор		0
Двойной тор		-2

Тройной тор		-4
Проективная поверхность		1
Лист Мёбиуса		0
Бутылка Клейна		0
Две сферы		4

Свойства

Любой граф может быть раскрашен в $\mu(G) + 1$ цвет (не доказано?)

Минимальное число пересечений при изображении графа на плоскости $\geq \mu(G) - 3$.



Свойства

Если дополнение графа является линейным лесом, то $\mu(G) \geq |G| - 3$

Если дополнение графа является внешнепланарным графом, то
$$\mu(G) \geq |G| - 4$$

Если дополнение графа G является планарным графом, то
$$\mu(G) \geq |G| - 5$$

Монотонность

Если H получен из G с помощью следующих операций (минорирование):

- 1) удалением изолированных вершин,**
 - 2) удалением рёбер,**
 - 3) сжатием (схлопыванием) рёбер ,**
- тогда $\mu(H) \leq \mu(G)$.**

Для справки

Теорема Робертсона-Сеймура-Томаса

Любое наследуемое свойство графов характеризуется конечным числом запрещенных подграфов.

Наследуемые свойства

планарность, внешнепланарность, вложение в поверхность.

Проблема

вычисление инварианта

Итак, начнём... Граф $G = (V, E)$.

**Чаще – неориентированные простые (без кратных рёбер и петель)
конечные графы.**

иногда – взвешенные.

Матрицы	
сопряжённости	$A \in \{0,1\}^{n \times n} : A_{ij} = 1 \Leftrightarrow (i, j) \in E$
диагональная матрица степеней	$D_{ij} = \begin{cases} \deg(i), & i = j, \\ 0, & i \neq j. \end{cases}$
Распределений (diffusion)	$W = D^{-1}A$
Лапласа	$L = D - A, L = NN^T$
Беззнаковая Лапласа	$Q = D + A, Q = MM^T$
инциденций	$M_{ij} = 1 \Leftrightarrow i \in e_j$
инциденций орграфа	$N_{ij} = \begin{cases} +1, & e_j = (i, *), \\ -1, & e_j = (*, i), \\ 0, & \text{иначе.} \end{cases}$

Напомним...

Собственный вектор (матрицы M) – ненулевой вектор x , для которого справедливо

$$Mx = \lambda x.$$

У симметричных матриц (такие будут у нас)

- **из с.в. можно составить ортонормированный базис (запишем по столбцам в Ψ)**
- **вещественные с.з. (запишем на диагональ Λ)**

$$\Psi^T M \Psi = \Lambda$$

$$M = \Psi \Lambda \Psi^T = \sum_i \lambda_i \psi_i \psi_i^T$$

Отношение Релея –

$$\frac{x^T Mx}{x^T x}$$

Для собственного вектора – $\frac{x^T Mx}{x^T x} = \lambda$.

Теорема. Пусть M – симметричная матрица, тогда максимум отношения Релея равен максимальному собственному значению.

Простое доказательство

$$\frac{\partial f}{\partial x} = \frac{2Mx(x^T x) - 2x(x^T Mx)}{(x^T x)^2} = 0, \quad Mx = \frac{x^T Mx}{(x^T x)} x.$$

Другое доказательство:

взять базис из ортогональных собственных векторов M ,
расписать вектор x , подставить.

Отношение Релея

Кстати, почему максимальное значение всегда существует...

$$\frac{x^T M x}{x^T x}$$

**можно рассматривать только векторы: $\|x\|=1$
(компактное множество)**

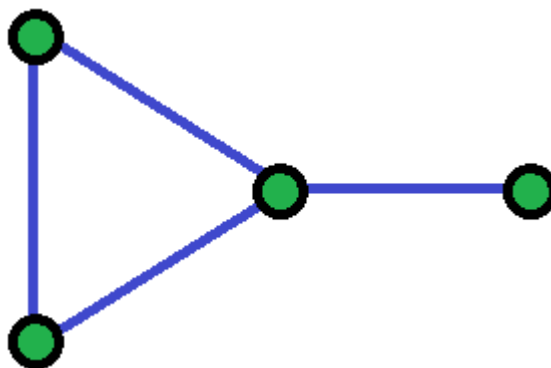
На этом множестве функция непрерывна

Что есть в матрицах...

A_{ij} – число путей из вершины i в вершину j

$$\text{tr}(A^2) = 2 | E |$$

$\text{tr}(A^3) = 6k$, k – число треугольников в графе



	A			
0	1	1	0	0
1	0	1	0	0
1	1	0	1	0
0	0	1	0	0

	A^2			
2	1	1	1	1
1	2	1	1	1
1	1	3	0	0
1	1	0	1	0

	A^3			
2	3	4	1	1
3	2	4	1	1
4	4	2	3	0
1	1	3	0	0

Что есть в матрицах...

Теорема Если граф связный (неориентированный) с диаметром d , то существует как минимум $d + 1$ различных с.з. матрицы A (аналогично L, Q).

Доказательство. Пусть $\lambda_1, \dots, \lambda_k$ – все различные с.з., тогда

$$(A - \lambda_1 I) \cdot \dots \cdot (A - \lambda_k I) = 0,$$

поэтому $A^k \in \Lambda(I, A, \dots, A^{k-1})$. Но если диаметр достижим для пары вершин (i, j) , то

$$A_{ij}^t = \begin{cases} 0, & t < d, \\ > 0, & t = d. \end{cases}$$

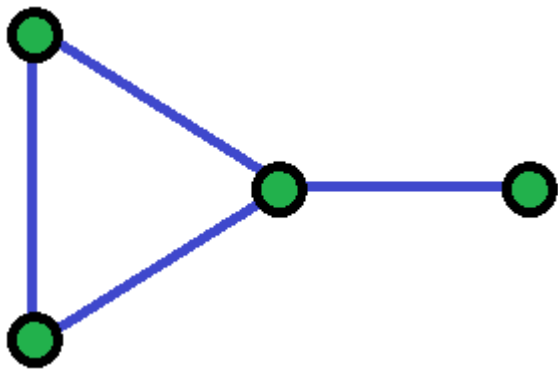
Поэтому $k > d$.

Матрица Лапласа

$$L = D - A$$

Квадратичная форма Лапласа –

$$x^T Lx = \sum_{(i,j)} (x_i - x_j)^2, \quad \mathbf{R}^{|V|} \rightarrow \mathbf{R}.$$



$$[x_1, x_2, x_3, x_4] \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 2 & -1 & 0 \\ -1 & -1 & 3 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} =$$

$$\begin{aligned} & x_1(2x_1 - x_2 - x_3) + x_1(x_1 - x_2) + x_1(x_1 - x_3) + \\ = & x_2(2x_2 - x_1 - x_3) + x_2(x_2 - x_1) + x_2(x_2 - x_3) + \\ = & x_3(3x_3 - x_1 - x_2 - x_4) + x_3(x_3 - x_1) + x_3(x_3 - x_2) + x_3(x_3 - x_4) + \\ & x_4(x_4 - x_3) x_4(x_4 - x_3) \end{aligned}$$

Теорема. Минимальное с.з. матрицы Лапласа = 0

Доказательство:

1 способ) т.к. все с.з. неотрицательны, а матрица вырождена.

2 способ) КФЛ неотрицательна, обращается в ноль. Вспоминаем отношение Релея (по теореме Куранта-Фишера).

Кстати, для беззнаковой матрицы Лапласа

$$x^T Qx = \sum_{(i,j)} (x_i + x_j)^2$$

Собственные значения матрицы Лапласа

Пусть $0 = \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ – с.з. матрицы Лапласа

Теорема. $\lambda_2 = 0 \Leftrightarrow$ граф несвязный

Доказательство.

Если несвязный – в явном виде строятся два ортогональных собственных вектора.

Если связный, то берём вектор ортогональный к константному, в нём есть два различных элемента x_i, x_j , учитывая, что вершины i, j соединяет путь, выражение

$$x^T Lx = \sum_{(i,j)} (x_i - x_j)^2$$

будет положительно. Поэтому это не может быть с.в. с нулевым с.з.

Алгебраическая связность графа

λ_2 называют **алгебраической связностью графа / индексом связности** [Fiedler]

Соответствующий с.в. – вектор Фидлера

Моноotonно не убывает при добавлении рёбер, так как

$$\min_{x^T \tilde{1}=0} \frac{x^T Lx}{x^T x} = \lambda_2.$$

Помним: max – max с.з., min – min с.з.=0.

Важное равенство!

См. также теорему Куранта-Фишера.

Кстати,

$$\min \frac{x^T Lx}{x^T x} = \lambda_1$$

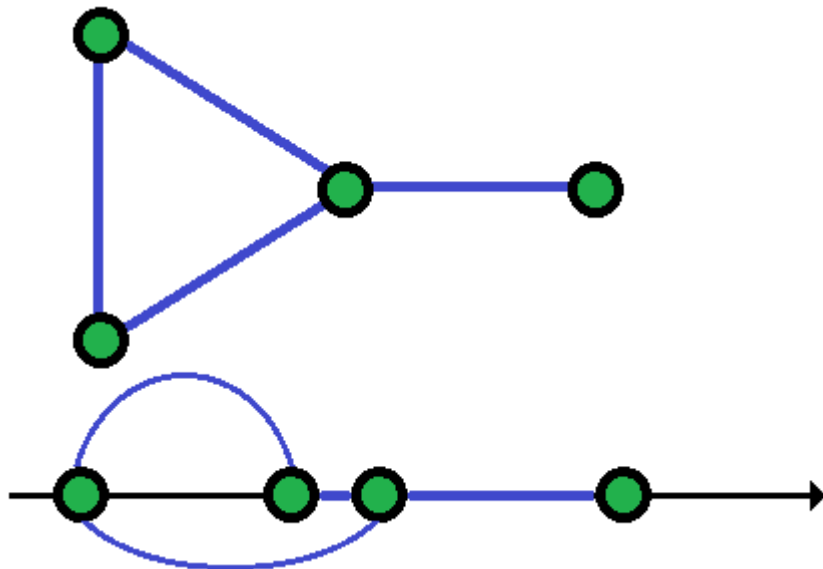
$$\min_{x: x^T \tilde{1}=0} \frac{x^T Lx}{x^T x} = \lambda_2$$

$$\min_{x: x^T \tilde{1}=0, x^T \psi_2=0} \frac{x^T Lx}{x^T x} = \lambda_3$$

и т.д.

Что получится если \min заменить на argmin ?

Проблема вложения графа [Hall, 70]



Вложить граф в прямую:

$$x^T Lx = \sum_{(i,j)} (x_i - x_j)^2 \rightarrow \min_x,$$

где $x = (x_1, \dots, x_n)$ – координаты наших вершин.

Избежать очевидного константного решения:

$$\tilde{1}^T x = 0,$$

учесть масштаб:

$$\|x\| = 1.$$

Решение – собственный вектор, соответствующий второму по величине с.з. матрицы Лапласа.

Проблема вложения графа [Hall, 70]

Теперь вкладываем в плоскость:

$$\sum_{(i,j) \in E} \|(x_i, y_i) - (x_j, y_j)\|^2 = \sum_{(i,j) \in E} (x_i - x_j)^2 + \sum_{(i,j) \in E} (y_i - y_j)^2 \rightarrow \min$$

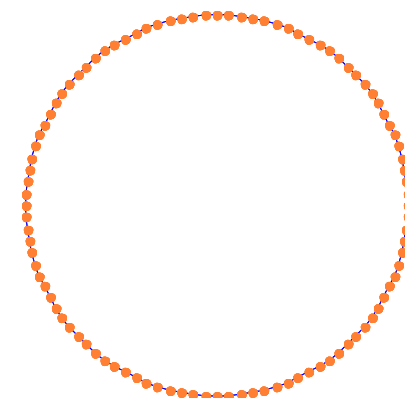
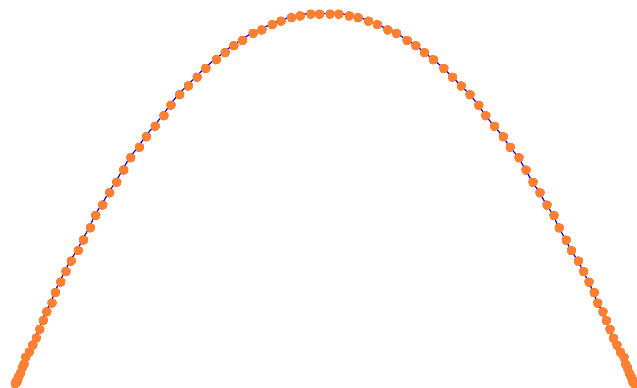
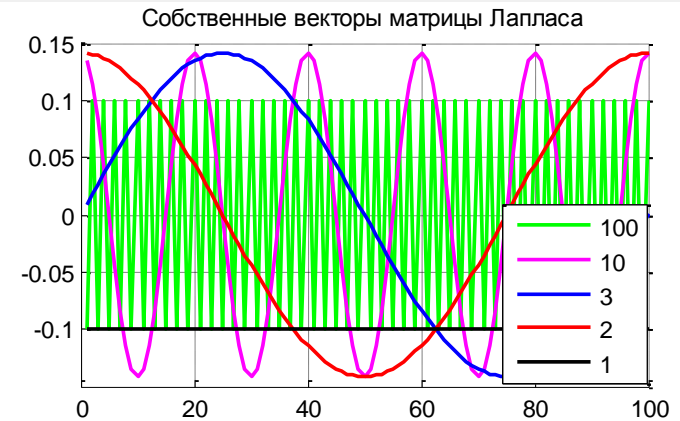
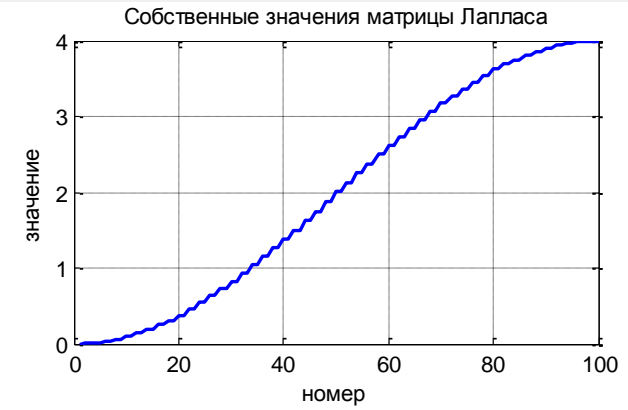
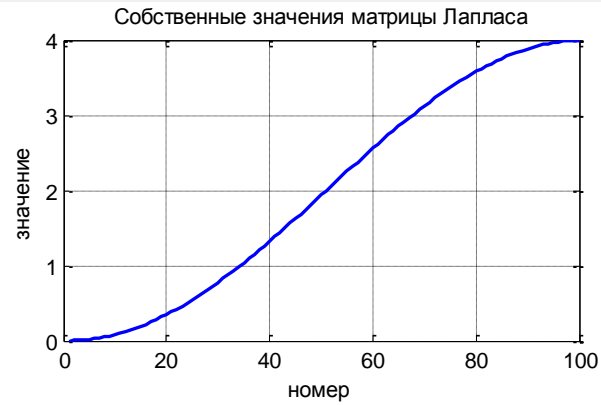
при условии

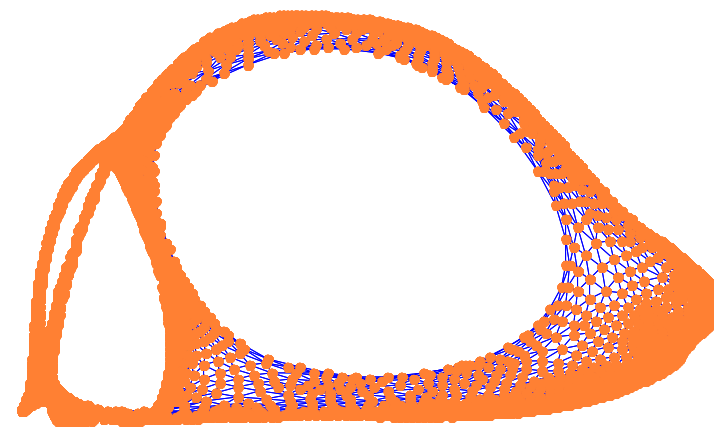
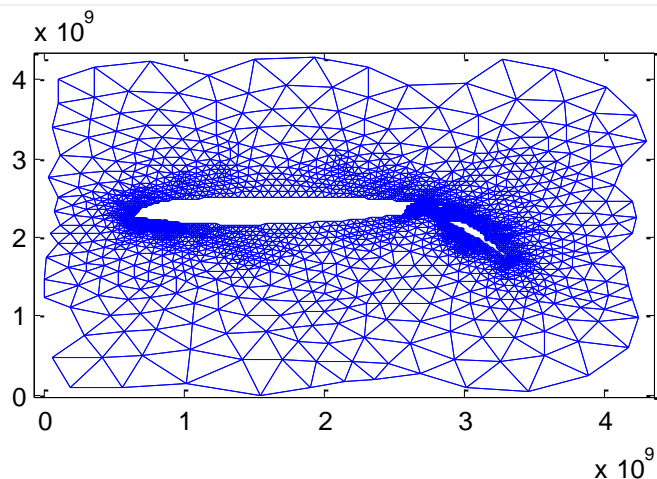
$$\sum_{i \in V} (x_i, y_i) = (0, 0).$$

Если добавить условие ортогональности x и y , то получим, что решение – с.в., соответствующие второму и третьему с.з. матрицы Лапласа.

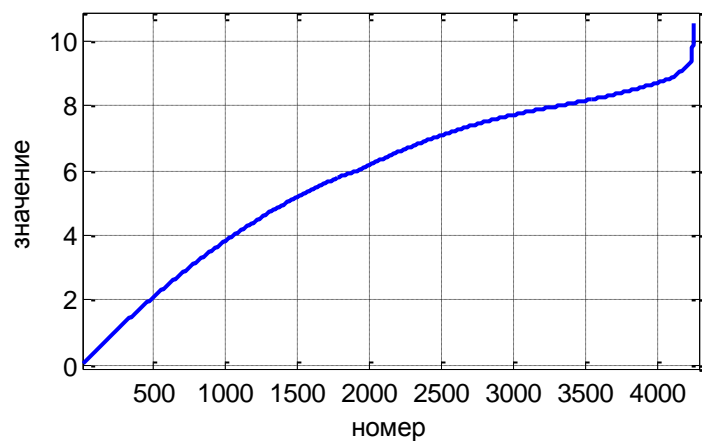
Вот почему визуализация графа по с.в.!

Сейчас будут картинки... откуда берутся синусоиды?

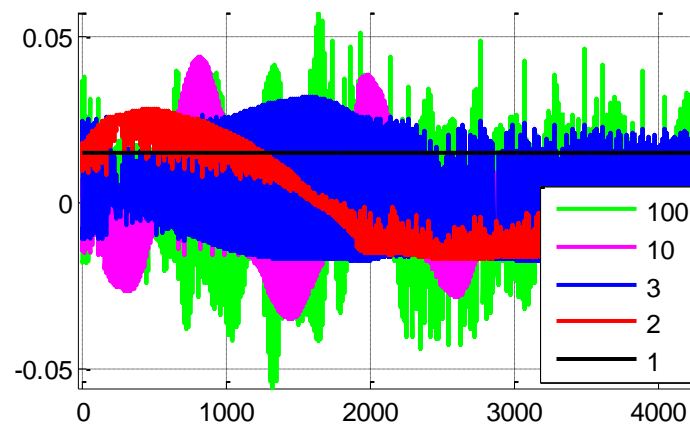




Собственные значения матрицы Лапласа

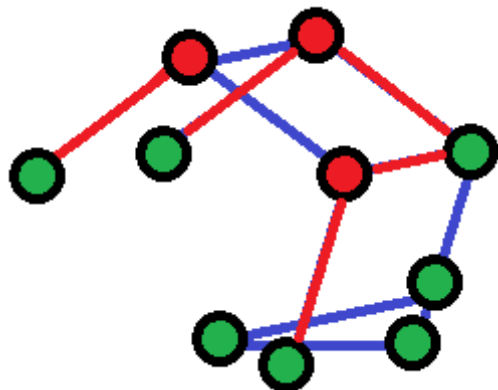


Собственные векторы матрицы Лапласа



[Hal70] K. M. Hall. An r -dimensional quadratic placement algorithm. *Management Science*, 17:219-229, 1970.

Разбиение графа



Рёберная граница –

$$\partial S = \{(i, j) \in E \mid i \in S, j \notin S\}.$$

Число Чигера (изопериметрическое число) – $h(G) = \min_{0 < |S| \leq n/2} \frac{|\partial S|}{|S|}.$

Оценивает, есть ли в графе "узкое горло".

Разбиение графа

Теорема. $h(G) \geq \lambda_2(1 - s)/2$, где $s = |S| / |V|$.

Если λ_2 – большое с.з., то граф «сильно связан».

Неравенство Чигера [Wiki, без доказательства]

В k -регулярном графе $(k - \lambda_2)/2 \leq h(G) \leq \sqrt{2k(k - \lambda_2)}$

Часто называют одним из основных результатов в СТК.

Теорема. $h(G) \geq \lambda_2(1-s)/2$, где $s = |S|/|V|$.

Доказательство. Известно, что

$$\min_{x^T \tilde{1} = 0} \frac{x^T Lx}{x^T x} = \lambda_2.$$

Поэтому для любого вектора x ортогонального к $\tilde{1}$ выполняется

$$x^T Lx \geq \lambda_2 x^T x.$$

Если $x = x_S - s\tilde{1}$, где x_S – характеристический вектор множества S (поправка x_S до ортогональности к $\tilde{1}$), то

$$x^T Lx = \sum_{(i,j) \in E} (x_i - x_j) = |\delta S|$$

и $x^T \tilde{1} = 0$.

Из

$$x^T x = |S|(1-s)^2 + (|V| - |S|)s^2 = |S|(1-s)$$

следует утверждение теоремы.

Применение в комбинаторике

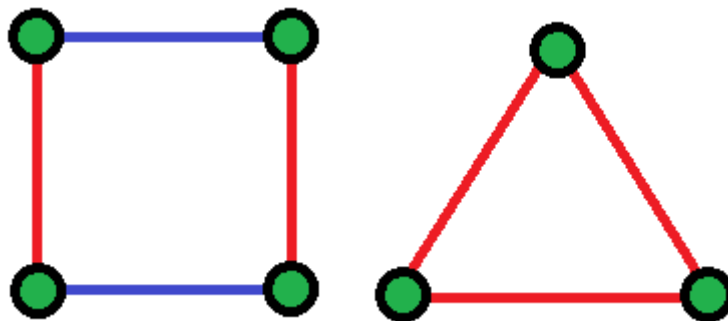
Теорема. $h(G) \geq \lambda_2(1-s)/2$, где $s = |S|/|V|$.

Следствие (можно показать, зная спектр гиперкуба), что для любого подмножества вершин $S : |S| \leq 2^{n-1}$ справедливо $|\partial S| \geq |S|$ (это простое некомбинаторное доказательство).

Теорема [без доказательства]. В неориентированном мультиграфе число остовных деревьев равно

$$\det(L + J / n^2) = \mu_2 \cdot \dots \cdot \mu_n / n.$$

Теорема [без доказательства]. $|V| = n = 2k$, с.з. Лапласа $0 \leq \mu_1 \leq \dots \leq \mu_n$, если $\mu_n \leq 2\mu_2$, то в графе есть совершенное соответствие (подмножество рёбер такое, что любая вершина инцидентна только одному ребру множества).



Теорема [без доказательства]. Кратность нуля как с.з. (неориентированного графа) равна числу компонент связности.

Как быть с двудольностью

Спектр Лапласа не распознаёт двудольность: $K_{1,3}$ и $K_1 \cup K_3$

Теорема [без доказательства]. Кратность нуля как с.з. (неориентированного графа) беззнакового Лапласа равна числу компонент двудольности.

Теорема [без доказательства]. Граф двудольный тогда и только тогда, когда спектр Лапласа равен "беззнаковому" спектру Лапласа.

Матрица смежности

Пошли другие обозначения

$$A \sim \text{с.з. } \lambda_1 \geq \dots \geq \lambda_n$$

$$L = kI - A \sim \text{с.з. } 0 = \mu_1 \leq \mu_2 \leq \dots \leq \mu_n$$

Теорема $d_{\text{avr}} \leq \lambda_1 \leq d_{\text{max}}$ ■

Доказательство

$$\lambda_1 = \max_x \frac{x^T A x}{x^T x} \geq \frac{\tilde{1}^T A \tilde{1}}{\tilde{1}^T \tilde{1}} = \frac{\sum A_{ij}}{n} = \frac{\sum \text{deg}(i)}{n}$$

Пусть v – собственный вектор, соответствующий μ_1 с i -м максимальным элементом (можно считать ненулевым), тогда

$$\lambda_1 = \frac{(Av)_i}{v_i} = \frac{\sum_{j:(i,j) \in E} v_j}{v_i} \leq \sum_{j:(i,j) \in E} \frac{v_j}{v_i} \leq \sum_{j:(i,j) \in E} 1 = \text{deg}(i) \leq d_{\text{max}} \quad \blacksquare$$

Теорема $d_{\text{avr}} \leq \lambda_1 \leq d_{\text{max}}$ ■

Замечание Если удалить вершину с наименьшей степенью, то средняя степень d_{avr} неубывает, а λ_1 невозрастает, т.е. не смотря на оценку они ведут себя по-разному!

$$\lambda_1 = \max_x \frac{x^T A x}{x^T x} \geq \max_y \frac{\begin{pmatrix} y \\ 0 \end{pmatrix}^T A \begin{pmatrix} y \\ 0 \end{pmatrix}}{\begin{pmatrix} y \\ 0 \end{pmatrix}^T \begin{pmatrix} y \\ 0 \end{pmatrix}}$$

Замечание Из этой схемы доказательства понятно, что

$$\lambda_{\min}(A) \leq \lambda_{\min}(A') \leq \lambda_{\max}(A') \leq \lambda_{\max}(A),$$

где A' – подматрица A , образованная строками и столбцами из $\{i_1, \dots, i_k\}$

Следствие. Граф раскрашиваем в $d_{\max} + 1$ цвет (очевидно). Граф раскрашиваем в $\lfloor \lambda_1 \rfloor + 1$ цвет. По индукции. Оценка точна! **ДЗ Почему?**

Замечание [без доказательства]. Хроматическое число

$$\begin{aligned} &\geq \frac{\lambda_n}{\lambda_n - d_{\text{avr}}} \\ &\geq 1 + \frac{\lambda_1}{-\lambda_n} \end{aligned}$$

[через оценку сумм с.з. для блочковых матриц]

Напоминалка

Граф 2-раскрашиваем \Leftrightarrow двудольный

Граф 3-раскрашиваем – NP-полная задача

Граф планарный \Rightarrow 4-раскрашиваем

Лемма. Если в конечном графе $\lambda_1 = d_{\max}$, то он d_{\max} -регулярный.

Доказательство. У нас при доказательстве было неравенство

$$\lambda_1 = \frac{(Av)_i}{v_i} = \frac{\sum_{j:(i,j) \in E} v_j}{v_i} \leq \sum_{j:(i,j) \in E} \frac{v_j}{v_i} \leq \sum_{j:(i,j) \in E} 1 = \deg(i) \leq d_{\max}$$

Теперь – это равенство:

$$\frac{v_j}{v_i} = 1, (i, j) \in E,$$

**не только у i -й вершины максимальная степень, но и у всех соседей.
Из связности графа \Rightarrow у всех вершин в графе максимальная степень.**

Для DM. Спектр пополнять другими характеристиками графа.

Теорема (Фробениуса-Перрона) [без доказательства].

Пусть граф связный и взвешенный, тогда

1) $\lambda_1 \geq -\lambda_n$ [они все вещественные, пока не больше]

2) $\lambda_1 > \lambda_2$

3) для λ_1 есть положительный собственный вектор

[Можно добавить доказательство + вспомогательная лемма]

Теорема (Ф-П для Лапласианов) [без доказательства].

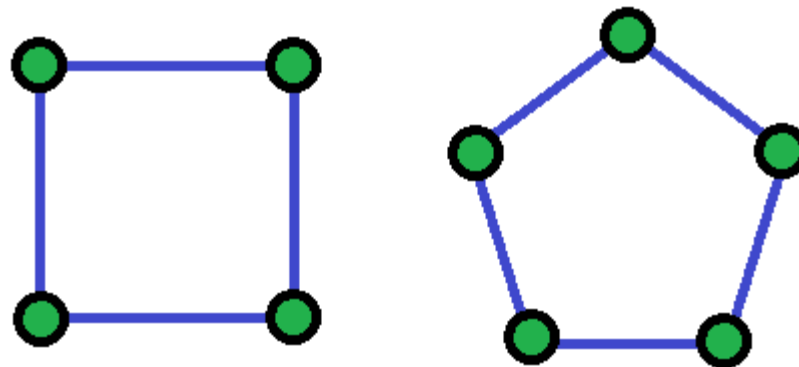
Пусть матрица M имеет неположительные недиагональные элементы, граф ненулевых недиагональных элементов связан.

Пусть λ_1 – наименьшее с.з. с с.в. v^1 . Тогда можно выбрать v^1 положительным и λ_1 имеет кратность 1.

Теорема Граф двудольный тогда и только тогда, когда для любого с.з. λ величина $(-\lambda)$ тоже является с.з.

Связный граф с наибольшим с.з. λ двудольный тогда и только тогда, когда $(-\lambda)$ тоже является с.з.

Сильно регулярный граф – простой, ориентированный, без петель, существуют параметры (n, k, k_1, k_2) такие, что $|V| = n$, $\forall i \deg(i) = k$, $\forall (i, j) \in E \deg(i, j) = k_1$, $\forall (i, j) \notin E \deg(i, j) = k_2$.



$(4, 2, 0, 2)$, $(5, 2, 0, 1)$

$\deg(i, j) = k$ – вершины i, j имеют k общих соседей.

Теорема Для простого нетривиального (не полного и не пустого) * графа порядка n следующие утверждения эквивалентны:

1) граф (n, k, k_1, k_2) -сильно регулярный

2) $A^2 = (k_1 - k_2)A + (k - k_2)I + k_2J$ для некоторых вещественных k, k_1, k_2

3) есть два с.з. с с.в. ортогональными к $\tilde{1}$

Доказательство Первые два утв. очевидно эквивалентны. Пусть верно второе и v – с.в. с с.з. λ , тогда

$$A^2v = (k_1 - k_2)Av + (k - k_2)Iv + k_2Jv$$

$$\lambda^2v = (k_1 - k_2)\lambda v + (k - k_2)v + k_2\left(\sum v_i\right)v$$

Для вектора ортогонального к $\tilde{1}$ –

$$\lambda^2 = (k_1 - k_2)\lambda + (k - k_2)$$

Здесь два разных решения.

Если верно третье утв. и соответствующие с.з. λ, λ' , то

$$(A - \lambda I)(A - \lambda' I) = sJ$$

для некоторого J , поэтому $A^2 \in \Lambda(A, I, J)$.

Теорема.

Граф с одним с.з. – без рёбер

Связный граф с двумя с.з. – полный

Связный регулярный граф с 3 с.з. – строго регулярный

Связный регулярный граф с 4 с.з. – "walk-regular" (для любого $k \geq 2$ число путей через вершину длины k не зависит от вершины)

МАТЕМАТИЧЕСКИЕ ОСНОВЫ СТГ

Теорема (Куранта-Фишера) / Min-max theorem

Пусть A – симметричная матрица с с.з. $\lambda_1 \geq \dots \geq \lambda_n$, тогда

$$\lambda_k = \max_{\substack{S \subseteq \mathbf{R}^n \\ \dim(S)=k}} \min_{x \in S} \frac{x^T A x}{x^T x} = \min_{\substack{T \subseteq \mathbf{R}^n \\ \dim(T)=n-k+1}} \max_{x \in T} \frac{x^T A x}{x^T x}.$$

Следствие. Если A – симметричная матрица с с.з. $\alpha_1 \geq \dots \geq \alpha_n$, матрица B получена из неё удалением i -й строки и i -го столбца, её с.з. $\beta_1 \geq \dots \geq \beta_n$, тогда

$$\alpha_1 \geq \beta_1 \geq \dots \geq \alpha_n \geq \beta_n.$$

тут ошибка;)

МАТЕМАТИЧЕСКИЕ ОСНОВЫ СТГ

$L \succ 0$, если L – неотрицательно(!) определённая матрица

$G \succ H$, если $L_G \succ L_H$, если $L_G - L_H \succ 0$

Лемма. Если $G \succ cH$, то $\mu_k(G) \geq c\mu_k(H)$ для всех k (здесь умножение на c – умножение весов графа).

Доказательство очевидно из

$$\lambda_k(G) = \max_{\substack{S \subseteq \mathbf{R}^n \\ \dim(S)=k}} \min_{x \in S} \frac{x^T L_G x}{x^T x} \geq c \max_{\substack{S \subseteq \mathbf{R}^n \\ \dim(S)=k}} \min_{x \in S} \frac{x^T L_H x}{x^T x} \geq c \lambda_k(H).$$

Аналогична монотонность при добавлении рёбер и увеличении отдельных весов.

МАТЕМАТИЧЕСКИЕ ОСНОВЫ СТГ

Теорема об аппроксимации

Для любого $\varepsilon > 0$ существует $d > 0$, что **для всех** достаточно больших n существует d -регулярный граф G :

$$(1 + \varepsilon)G \succ K_n \succ (1/(1 + \varepsilon))G.$$

Полные графы аппроксимируются графами с малым числом рёбер!