

Анализ социальных сетей, лекция 2. Меры центральности

Михаил Пожидаев

16 марта 2024 г.

Матрица

Вид и операции

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$$

Вектором называется столбец или строка матрицы.

Операции

При соблюдении ограничений размерности матрицы можно складывать, умножать между собой, умножать на вектор или число.

Линейные уравнения

Простая и матричная форма

$$\begin{array}{cccccc} a_{11}x_1 & a_{12}x_2 & \dots & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & a_{22}x_2 & \dots & a_{2n}x_n & = & b_2 \\ \vdots & \vdots & \vdots & \vdots & = & \vdots \\ a_{n1}x_1 & a_{n2}x_2 & \dots & a_{nn}x_n & = & b_n \end{array}$$

Матричная форма

$$Ax = b$$

Чудесные значения

Определитель и собственный вектор

Определитель

Определителем квадратной матрицы A называется число, которое стремится к нулю, если матрица вырождается. Обозначается как $\det A$. Если $\det A = 0$, то соответствующая система линейных уравнений будет иметь больше одного решения.

Собственный вектор

Собственным вектором x матрицы A называется вектор, который после умножения на A равен самому себе, умноженному на некоторое число λ ($Ax = \lambda x$).

Для его определения сначала находят λ , которых может быть несколько. Найти их можно через решение уравнения $\det(A - \lambda I) = 0$. Выражение $\det(A - \lambda I)$ принято называть характеристическим многочленом.

Мера центральности

Что такое и для чего?

Мера центральности узла — условная оценка значимости узла, выражаемая численно. В подавляющем большинстве случаев её абсолютное значение не играет роли, но полезны её значения для определения отношения между разными узлами.

Самый простой пример

Самый простой вариант меры центральности — степень связанности узла, которая просто совпадает со степенью вершины.

Степень влияния

Мера центральности в собственных векторах

Дано

Задана сеть (N, E) с матрицей смежности $A = (a_{ij})$, где $i, j \in N$.

Требуется найти

x_i — степень влияния вершины i .

Ограничения

- ▶ $x_i \geq 0$;
- ▶ $a_{ij} = 1$, если i и j смежные, и 0 в противном случае.

Определение

Что считать степенью влияния?

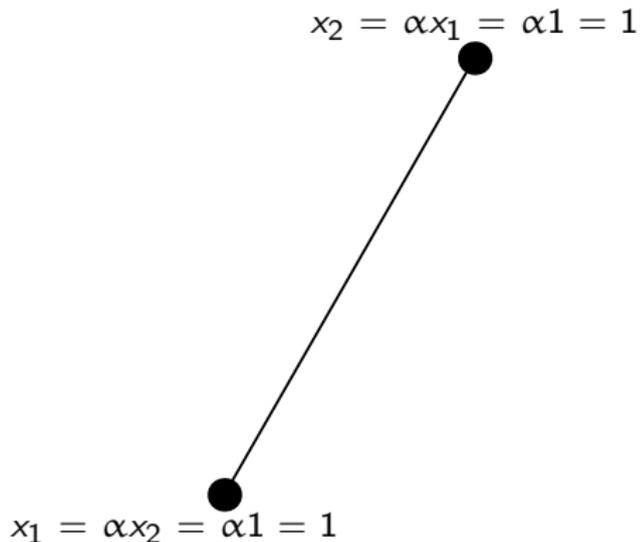
Узел в сети считается тем более «влиятельным», чем более «влиятельны» смежные с ним узлы. Это определение можно записать следующим образом:

$$x_i = \alpha \sum_{j \in N} a_{ij} x_j,$$

где α — это некоторый коэффициент больше нуля, а $i \in N$.

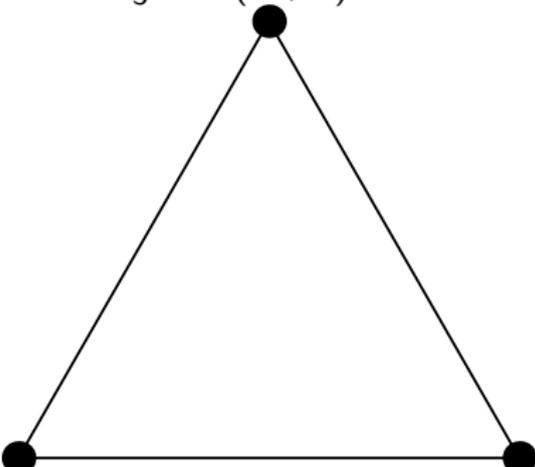
Тривиальные примеры

Как выбрать коэффициент α ?



Тривиальные примеры

Как выбрать коэффициент α ?



$x_3 = \alpha(1 + 1) = 1$

$x_1 = \alpha(1 + 1) = 1$

$x_2 = \alpha(1 + 1) = 1$

Матричный вид

Получение собственных векторов

Коэффициент α заменим на коэффициент $\lambda = \frac{1}{\alpha}$. Таким образом, выражение меры центральности узла i можно записать в следующем виде:

$$x_i = \frac{1}{\lambda} \sum_{j \in N} a_{ij} x_j$$

После перехода к векторной форме записи получим выражение:

$$\lambda x = A x$$

Получившееся выражение является выражением собственного вектора x для линейного оператора A с собственным значением λ .

Алгоритм

Влияние в собственных векторах

1. Находим главное собственное значение матрицы смежности (нахождение корней характеристического уравнения $\det(A - \lambda I)$).
2. Подставляем найденное значение в уравнение собственного вектора и получаем систему линейных уравнений.
3. Заменяем любое уравнение уравнением нормировки.
4. Решаем полученную систему и находим все x_i .

Напоминание!

Система $(A - \lambda I)x = 0$ будет иметь ненулевое решение только при вырожденной матрице.

Центральность по Кацу

Всем немного своей центральности

Каждый узел получает некоторую величину собственной центральности, обозначенную через β . Уравнение центральности по Кацу имеет вид:

$$x_i = \alpha \sum_{j \in N} (a_{ij} x_j) + \beta$$

Если уравнение переписать в матричном виде, то получим систему:

$$x = \alpha Ax + \beta \mathbf{1}$$

Через $\mathbf{1}$ обозначен единичный вектор.

Выражение x

Уравнение с x в правой части

$$x = \alpha Ax + \beta 1$$

$$x - \alpha Ax = \beta 1$$

$$x = \beta(I - \alpha A)^{-1}1$$

I — единичная матрица.

α и β

Какие выбрать для аргументов значения?

Спасает требование нормализации!

β можно исключить, потому что нас не интересуют абсолютные значения.

α задаёт соотношение в смеси своей «влиятельности» и «влиятельности» в собственных векторах. При росте α $\det(I - \alpha A)$ будет стремиться к нулю, из-за чего невозможно вычисление обратной величины.

$$\det(I - \alpha' A) = 0$$

$$\det\left(\frac{1}{\alpha'} I - A\right) = 0$$

PageRank

Ранжирование в поисковой выдаче Google

Плохо для Интернета!

В Интернете нужны ориентированные графы и предотвращение передачи влияния от сайтов-агрегаторов.

$$x_i = \alpha \sum_{j \in N} \frac{a_{ij} x_j}{d_j} + \beta$$

d_j — исходящая степень вершины j .

Матричный вид

$$x = \alpha A d^{-1} x + \beta \mathbf{1}$$

Математическая запись

Общий вид центральности PageRank

$$x = (I - \alpha A d^{-1})^{-1} \mathbf{1}$$

Полагаем β равной 1. α Google использует равным 0,85.

Любопытный факт!

При $\beta = 0$, $\alpha = 1$ и x_i равной степени вершины i формула вырождается в тождество, что показывает не такое уж плохое качество центральности в виде степени связности.

HITS

Hyperlink induced topic search

Хабы и авторитеты в публикациях

Обозначим через x_i источники информации, а через y_i агрегаторы информации.

Получим в результате систему матричных уравнений:

$$x_i = \alpha \sum_{j \in N} a_{ij} y_j$$

$$y_i = \beta \sum_{j \in N} a_{ji} x_j$$

Решение HITS

Выражение для новой системы

Матричный вид :

$$x = \alpha Ay$$

$$y = \beta A^T x$$

Подставляя x и y , получим новую систему:

$$AA^T x = \lambda x$$

$$A^T A y = \lambda y$$

$$\lambda = \frac{1}{\alpha\beta}$$

Мера близости

Анализ расстояний между вершинами

$$l_i = \frac{\sum_{j \in N} d_{ij}}{n_i},$$

где d_{ij} — расстояние между вершинами i и j , а n_i — количество вершин, достигаемых из вершины i .

Степень близости вершины $c_i = \frac{1}{l_i}$.

Недостаток!

Применение степени близости затруднено, если в сети более одной компоненты связности.

Спасибо за внимание!

Всё о курсе: <https://marigostra.ru/materials/networks.html>

E-mail: msp@luwrain.org

Канал в Телеграм: <https://t.me/MarigostraRu>