



ИПС “Методы линейной алгебры и
анализа данных в экономике”

ФЭН НИУ ВШЭ, Москва

Исмаилов М. А.

Степенной метод / Eigenvalue Power Method (EPM)

Теория Power method для поиска наибольшего собственного значения матрицы
и его приложения.



Постановка задачи

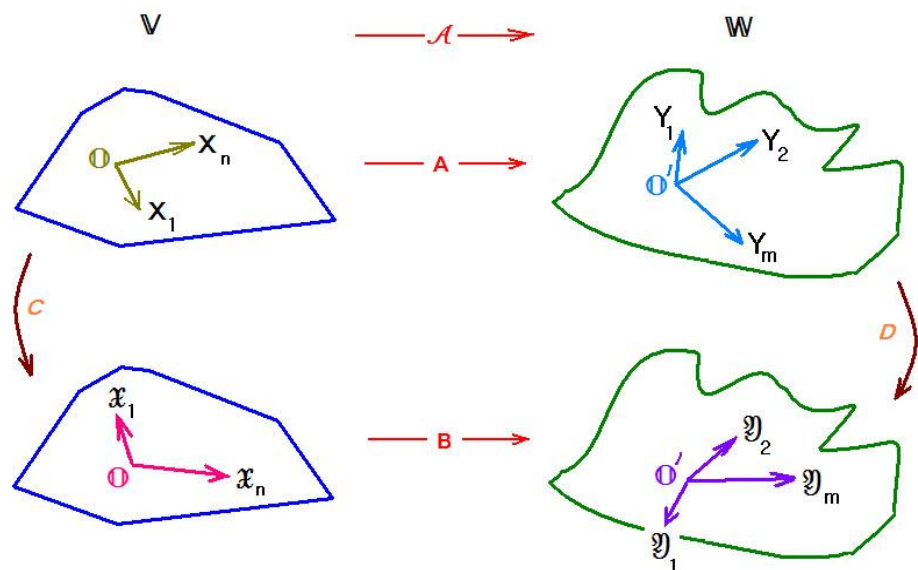
Пусть $A \in M_n(\mathbb{R})$ – квадратная матрица размера $n \times n$.

Собственным значением матрицы A называется число λ такое, что $\exists v \in \mathbb{R}^n : A \cdot v = \lambda \cdot v$. Вектор v в таком случае называют собственным вектором, отвечающим собственному значению λ .

Power Method решает задачу поиска наибольшего собственного значения матрицы.

$$\lambda_1 \lambda_2 \lambda_4 \lambda_5 \lambda_6$$

Мотивировка



Собственные значения матрицы отвечают за растяжение, которому подвергаются векторы при воздействии на них матрицей A в базисе собственных векторов (при линейном отображении).

Это свойство помогает анализировать данные и наблюдать закономерности во многих сферах: экономика, теория графов, машинное обучение и т.д.



Принцип работы алгоритма

Пусть $A \in M_n(\mathbb{R})$ – диагонализуемая квадратная матрица. Алгоритм поиска ее наибольшего собственного значения λ состоит из этапов:

1. Инициализация:

Выбираем случайный вектор $x_0 \in \mathbb{R}^n$; eps – толерантность, с точностью до которой мы ищем λ ; k – максимальное число итераций;

2. Итерация:

На $i+1$ -й итерации поддерживаем x_i – вектор, полученный на прошлой итерации, λ_i – текущее приближение собственного значения.

- умножаем вектор x_i на матрицу A слева \rightarrow получаем x_{i+1} .
- вычисляем λ_{i+1} по формуле $x_{i+1}^T \cdot x_i / (x_i^T \cdot x_i)$
- если $\lambda_{i+1} - \lambda_i < \text{eps}$ или $i + 1 = k$ – прекращаем работу алгоритма
- нормализуем вектор: $x_{i+1} / \|x_{i+1}\|$ чтобы избежать слишком больших / малых значений



Доказательство работы алгоритма

Пусть $A \in M_n(\mathbb{R})$ – матрица диагонализуемого линейного оператора $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ и $\mathcal{P} = \{e_1, \dots, e_n\}$ – базис из собственных векторов, перенумерованный так, что $|\lambda_1| > |\lambda_2| > \dots > |\lambda_n|$, где $\lambda_i = A \cdot e_i$. Рассмотрим итерацию алгоритма:

$$1: \quad x_0 \in \mathbb{R}^n : x_0 = c_1 \cdot e_1 + \dots + c_n \cdot e_n; \text{ Тогда } Ax_0 = \lambda_1 c_1 \cdot e_1 + \dots + \lambda_n c_n \cdot e_n$$

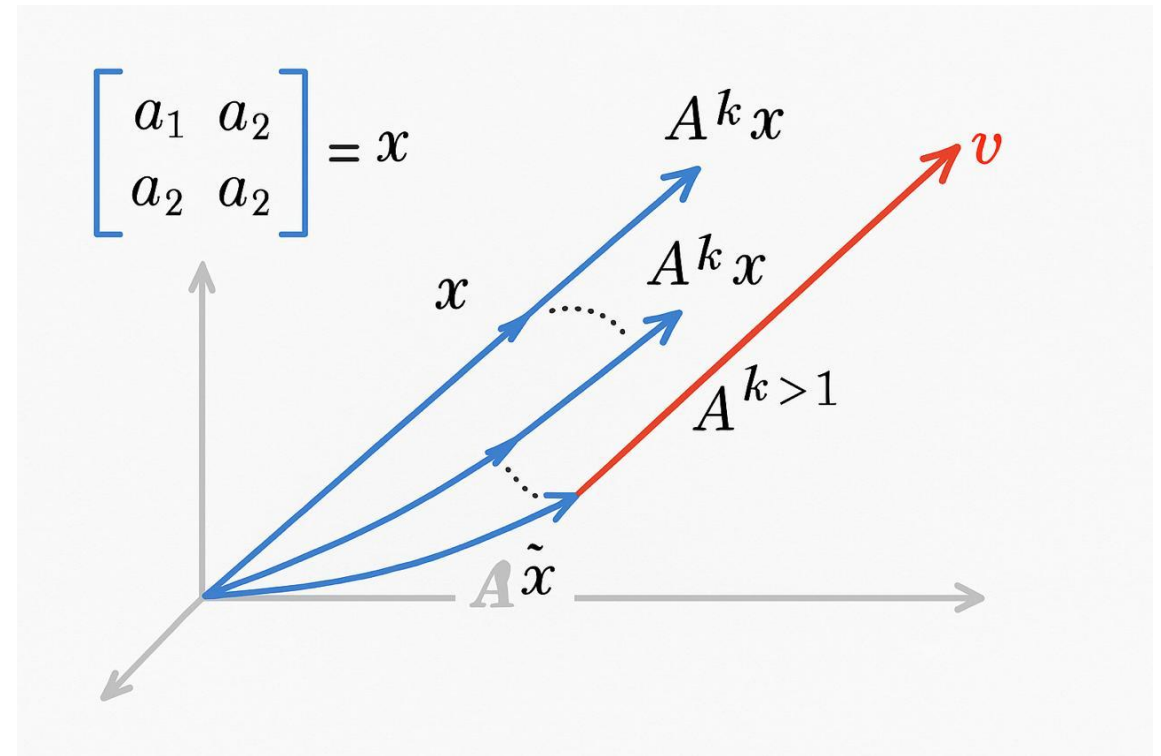
.....

$$i: \quad x_i = \lambda_1^i (c_1 \cdot e_1 + \sum_{k=2}^n (\lambda_k / \lambda_1)^i c_k \cdot e_k)$$

Очевидно, что второе слагаемое в последней скобке стремится к нулю при $i \rightarrow \infty$. Тогда на последнем шаге алгоритма (при котором достигается заданная точность) будем считать, что x_n примерно равно $\lambda_1^i \cdot c_1 \cdot e_1$; Тогда $\lambda_1 = x_n^T \cdot x_{n-1} / (x_{n-1}^T \cdot x_{n-1})$.

Интуиция в работе алгоритма

Каждый раз, когда вектор умножается на матрицу слева (когда к нему применяется линейное отображение) он “выравнивается” в сторону вектора, который отвечает наибольшему собственному значению (т. е. становится коллинеарным вектору e_1 в перенумерованном базисе из собственных векторов).

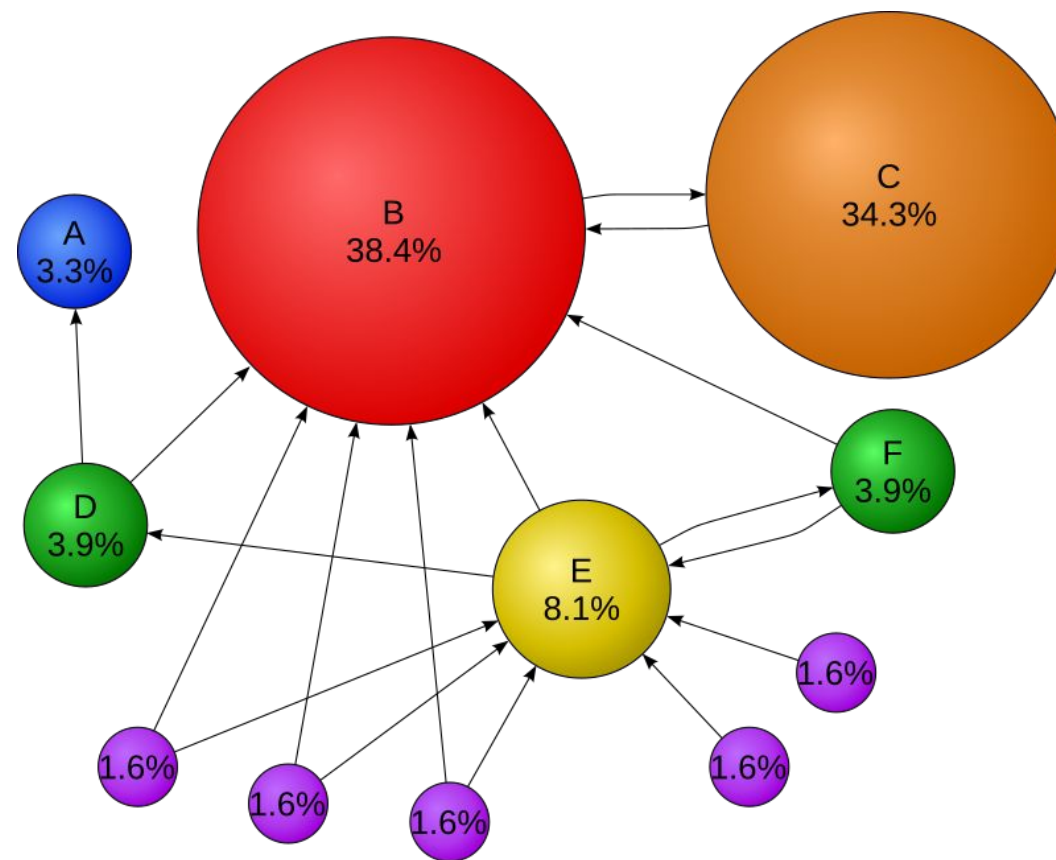


Применение

Google – PageRank

Один из самых известных примеров применения алгоритма – ранжирование веб-страниц по важности.

PageRank – один из индикаторов, который помогает Google сортировать веб-страницы. Сеть страниц при поисковом запросе образует граф, вершины которого – вероятность перехода на страницу. Поиском наибольшего собственного значения система (каким то образом) оценивает важность каждой страницы.



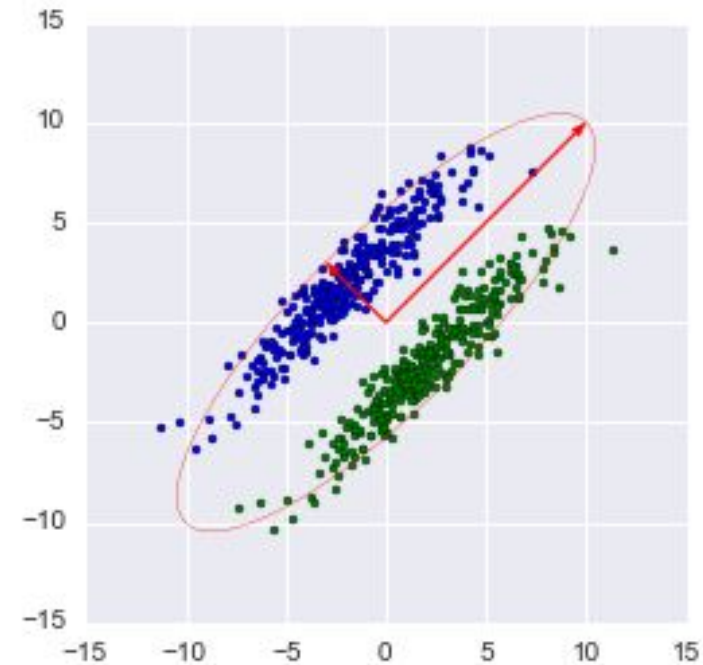
Применение

РСА – метод главных компонент.

Не менее яркий пример применения степенного метода для снижения размерности данных в экономических моделях.

Степенной метод в РСА находит самое важное направление в данных — ось, вдоль которой точки разбросаны сильнее всего.

Многokrатно уточняя это направление, алгоритм выделяет главную компоненту, которая сохраняет максимальную информацию о данных, позволяя уменьшить их размерность без существенной потери информации.





Заключение

Eigen Power Method – это полезный алгоритм который много где применяется.

