

Численные методы, осень 2022

Задание 2 [SVD-разложение и его применения]

Всего баллов: 40 Срок сдачи: 21 октября

РЕКОМЕНДОВАННАЯ ЛИТЕРАТУРА

- Лекции 4-5 из [1]
- Лекция 2 из [2]
- StackExchange: в чём смысл собственных векторов, собственных чисел и анализа главных компонент

УПРАЖНЕНИЯ

- (10) В этом упражнении мы познакомимся с тремя основными алгоритмами вычисления сингулярного разложения, доступными в Python: `numpy.linalg.svd`, `scipy.sparse.linalg.svds` и `sklearn.utils.extmath.randomized_svd`.
 - Создайте матрицу A размера $n \times n$ ($n = 2000$) со случайными элементами из стандартного нормального распределения.
 - С помощью этих трёх алгоритмов (и теоремы Эккарта-Янга?) аппроксимируйте матрицу A матрицами ранга 2. Вы получите матрицы A_{svd} , A_{svds} и A_{rsvd} . Измерьте времена выполнения.
 - Вычислите нормы отклонений: $\|A - A_{\text{svd}}\|_F$ $\|A - A_{\text{svds}}\|_F$ $\|A - A_{\text{rsvd}}\|_F$
Объясните результат.
- (5) Пусть матрица A размером $m \times n$ имеет сингулярное разложение $A = U\Sigma V^T$. Выразите через U , Σ и V сингулярные разложения следующих матриц:
 - (i) $(A^T A)^{-1}$
 - (ii) $(A^T A)^{-1} A^T$
 - (iii) $A (A^T A)^{-1}$
 - (iv) $A (A^T A)^{-1} A^T$

3. (10) Рассмотрим матрицу:

$$\begin{bmatrix} -2 & 11 \\ -10 & 5 \end{bmatrix}$$

- Выпишите сингулярные числа, а также левые и правые сингулярные векторы матрицы A . Так как SVD-разложение не единственно, найдите то разложение, в котором матрицы U и V имеют наименьшее количество отрицательных элементов.
- Нарисуйте единичный круг и его образ под действием оператора A . Изобразите сингулярные векторы и отметьте их координаты.
- Чему равна спектральная норма и норма Фробениуса матрицы A ?
- Найдите A^{-1} с помощью SVD-разложения.
- Найдите собственные числа λ_1 и λ_2 матрицы A .

4. (5) В файле `A.pru` находится матрица A размера $n \times n$. Определите наилучшее приближение следующего вида, в котором переменные разделяются: $A_{ij} \approx h_i \eta_j$. Чему равна относительная ошибка такой аппроксимации?

$$\delta_{\text{err}} = \frac{\sqrt{\sum_{ij} (A_{ij} - h_i \eta_j)^2}}{\sqrt{\sum_{ij} A_{ij}^2}}$$

Сколько понадобится слагаемых, чтобы приближение стало точным?

$$A_{ij} = \sum_{\alpha=1}^K h_{\alpha i} \eta_{\alpha j} \quad K = ?$$

5. (10) В этом упражнении мы применим SVD-разложение к задаче уменьшения размерности. Начнём с загрузки датасета:

```
from sklearn.datasets import load_digits
digits = load_digits()
A = digits.data
y = digits.target
```

Каждая строка массива A состоит из 64 чисел с плавающей точкой, которые задают черно-белое изображение 8×8 . Это изображение цифры. Сама цифра (метка данных) указана в массиве y .

- Изучите датасет. Постройте изображения нескольких цифр, например, 0, 3, 7
- Отнормируйте датасет A .
- С помощью SVD-разложения спроецируйте датасет, сделайте его из $N \times 64$ -мерного $N \times 2$ -мерным. Постройте двумерный `scatter_plot`, окрасьте точки, соответствующие разным цифрам, в разные цвета.

-
- [1] L. Trefethen and D. Bau, *Numerical Linear Algebra*, Other Titles in Applied Mathematics (SIAM, 1997).
- [2] E. Tyrtyshnikov, *A Brief Introduction to Numerical Analysis* (Birkhäuser Boston, 2012).