Homework4

Xiaoma

2023 年 4 月 24 日

1 实验要求

1. 使用 jacobi 方法对随机生成的矩阵 A 进行 SVD 分解
2. 使用 jacobi 方法对鸢尾花数据集进行 PCA

2 算法原理

2.1 主成分分析 (PCA)

2.1.1 PCA 概述

PCA（Principal Component Analysis 是一种常见的数据分析方式，常用于高维数据的降维，可用于提取数据的主要特征分量。

PCA 的主要思想是将 n 维特征映射到 k 维上，这 k 维是全新的正交特征也被称为主成分，是在原有 n 维特征的基础上重新构造出来的 k 维特征。PCA 的工作就是从原始的空间中顺序地找一组相互正交的坐标轴，新的坐标轴的选择与数据本身是密切相关的。其中，第一个新坐标轴选择是原始数据中方差最大的方向，第二个新坐标轴选取是与第一个坐标轴正交的平面中使得方差最大的，第三个轴是与第 1,2 个轴正交的平面中方差最大的。依次类推，可以得到 n 个这样的坐标轴。

2.1.2 PCA 的实现方法

1. 基于特征值分解协方差矩阵实现 PCA 算法

   输入: 数据集 $X = \{x_1, x_2, x_3, \ldots, x_n\}$ ，需要降到 k 维。

   1) 去平均值 (即去中心化)，即每一位特征减去各自的平均值。

   2) 计算协方差矩阵 $\frac{1}{n}XX^T$, 注: 这里除或不除样本数量 n 或 $n-1$, 其实对求出的特征向量没有影响。

   3) 用特征值分解方法求协方差矩阵 $\frac{1}{n}XX^T$ 的特征值与特征向量。

   4) 对特征值从大到小排序，选择其中最大的 k 个。然后将其对应的 k 个特征向量分别作为行向量组成特征向量矩阵 P。

   5) 将数据转换到 k 个特征向量构建的新空间中，即 $Y = PX$。

2. 基于 SVD 分解协方差矩阵实现 PCA 算法

   输入: 数据集 $X = \{x_1, x_2, x_3, \ldots, x_n\}$ ，需要降到 k 维。

   1) 去平均值，即每一位特征减去各自的平均值。

   2) 计算协方差矩阵。

   3) 通过 SVD 计算协方差矩阵的特征值与特征向量。

   4) 对特征值从大到小排序，选择其中最大的 k 个。然后将其对应的 k 个特征向量分别作为列向量组成特征向量矩阵。

   5) 将数据转换到 k 个特征向量构建的新空间中。

2.2 奇异值分解 (SVD)

SVD 概述

奇异值分解 (Singular Value Decomposition，以下简称 SVD) 是在机器学习领域广泛应用的算法，它不光可以用于降维算法中的特征分解，还可以用于推荐系统，以及自然语言处理等领域。是很多机器学习算法的基石。

特征值分解仅适用于提取方阵特征，但在实际应用中，大部分数据对应的矩阵都不是方阵；矩阵可能是有很多 0 的稀疏矩阵，存储量大且浪费空间，这时就需要提取主要特征；奇异值分解是将任意较复杂的矩阵用更小、更简单的 3 个子矩阵的相乘表示，用这 3 个小矩阵来描述大矩阵重要的特性。

SVD 的实现方法

Algorithm 1 basic rSVD

Input: $\mathbf{A} \in \mathbb{R}^{m\times n}$, rank parameter k, power parameter p

Output: $\mathbf{U} \in \mathbb{R}^{m\times k}, \mathbf{S} \in \mathbb{R}^{k\times k}, \mathbf{V} \in \mathbb{R}^{n\times k}$

$\mathbf{\Omega} = \mathrm{randn}(n, k+s)$

$\mathbf{Q} = \mathrm{orth}(\mathbf{A\Omega})$

for $i = 1, 2, \cdots, p$ do

  $\mathbf{G} = \mathrm{orth}\left(\mathbf{A}^{\mathrm{T}}\mathbf{Q}\right)$

  $\mathbf{Q} = \mathrm{orth}(\mathbf{AG})$

end for

$\mathbf{B} = \mathbf{Q}^{\mathrm{T}}\mathbf{A}$

$[\mathbf{U}, \mathbf{S}, \mathbf{V}] = \mathrm{svd}(\mathbf{B})$

$\mathbf{U} = \mathbf{QU}$

$\mathbf{U} = \mathbf{U}(:, 1:k), \mathbf{S} = \mathbf{S}(1:k, 1:k), \mathbf{V} = \mathbf{V}(:, 1:k)$

3 实验结果 (详见程序生成的 md 文件)

A

0.894135	0.787075	0.0657304
0.957633	0.965759	0.390754
0.190948	0.506086	0.659409
0.436956	0.800031	0.736373

(a) 矩阵A

AAT

1.42329	1.64206	0.612404	1.06878
1.64206	2.00244	0.929282	1.47882
0.612404	0.929282	0.727404	0.973891
1.06878	1.47882	0.973891	1.37323

(b) 矩阵$A * A^T$

U

0.499848	-0.460479	-0.415391	0.502919
0.629718	0.770804	0.605195	-0.414927
0.325482	-0.0749708	0.353568	0.643525
0.497663	-0.43382	-0.579809	-0.400977

(c) 矩阵U

Sigma

2.22595	0	0
0	0.754982	0
0	0	0.0385276
0	0	0

(d) 矩阵Σ

VT

0.597307	0.702818	0.386356
-0.516541	0.715917	-0.469732
0.394548	-0.664099	0.635063

(e) 矩阵V^T

```
[1.42328494 1.64206236 0.61240415 1.06878414]
[1.64206236 2.0024401  0.92928192 1.47882132]
[0.61240415 0.92928192 0.72740441 0.97389135]
[1.06878414 1.47882132 0.97389135 1.37322534]
```

(f) $A * A^T$-numpy 计算结果

```
2.225954491437638
0.7549814881659662
0.03852714111468048
```

(g) Σ-numpy 计算结果

Sum_N_Diag

11.0326

4.40888

0.220458

0.0836733

0.0380921

0.0205735

0.00701574

1.012e-05

3.93372e-06

1.09825e-06

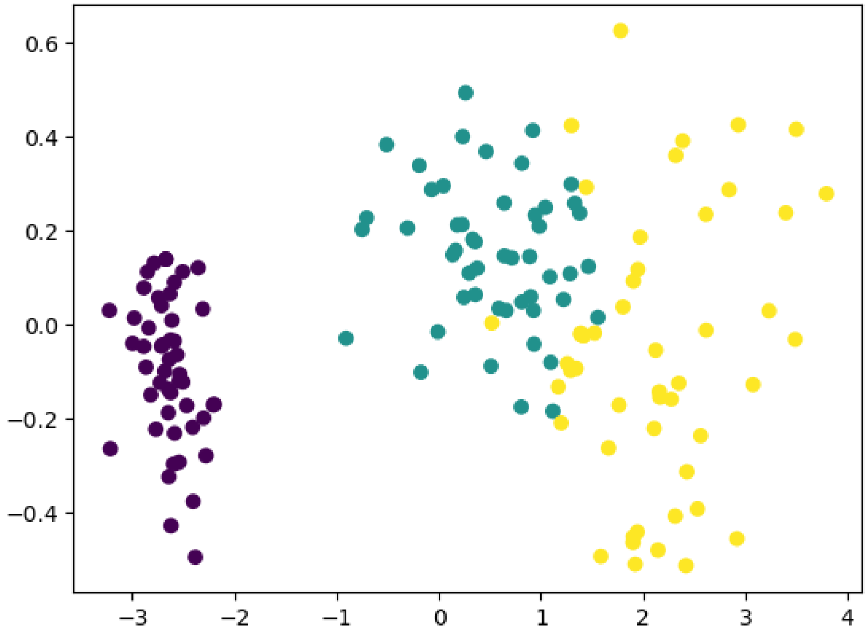
3.94345e-09

迭代过程

1/m XXT

0.681122	-0.0390067	1.26519	0.513458
-0.0390067	0.186751	-0.319568	-0.117195
1.26519	-0.319568	3.09242	1.28774
0.513458	-0.117195	1.28774	0.578532

矩阵$\frac{1}{m}X * X^T$



PCA 分布

4 实验分析

1. jacobi 方法迭代过程中矩阵非对角元素的平方和呈下降趋势。通过对结果比较可知，求得的特征值是对称矩阵特征值的近似值。

2. 通过观察 PCA 的结果，可以发现降维后仍能有效的区分不同标签对应的数据。