2022秋-计算方法-第四次上机作业说明文档

1 实验目的

实现Jacobi方法求解对称矩阵特征值，应用Jacobi方法实现PCA（主成分分析），并进一步实现矩阵的SVD分解（奇异值分解）。

2 应用问题

2.1 PCA主成分分析

在许多领域我们都需要处理一组大规模的数据，我们希望能够有效地对数据进行压缩而不丢失太多的信息，易于我们进一步的存储或计算。我们首先介绍PCA（主成分分析）降维方法。

我们通常将数据抽象成m个n维数据向量，这些向量可以张成一个欧式空间，一个n维欧式空间的向量可以用n个正交基向量的线性组合表达。一种压缩数据的思路就是对这组向量进行降维，即将这组向量投影到维度更低的子空间上，从而用更少的基向量来刻画这组数据。那么核心问题是如何有效地选择子空间，使得信息的损失最小，也就是投影后的数据尽可能的保持分散。

我们可以用协方差矩阵来刻画数据的分散程度，将m个n维数据向量去中心化后（每个向量减去所有向量的均值）按列排列构成$n\times m$矩阵X，协方差矩阵为$\frac{1}{m}XX^T$。通过对协方差矩阵进行特征值分解，可以得到一组特征向量，将这些特征向量所在的正交方向称为主方向。对应的特征值越大，意味着在这一主方向上包含的信息越多。在PCA降维中，我们将协方差矩阵$\frac{1}{m}XX^T$特征值按从大到小排列，取前k个特征值对应的特征向量为数据的前k个主方向，我们选取这k个特征向量组成我们需要的投影空间，从而实现对这m个n维数据向量进行降维。

PCA有着诸多应用，比如我们可以利用PCA可视化展示一组高维数据的分布。在本次作业中我们将实现这一过程，我们选取这组高维数据的前两个主方向，将这两个方向构成的空间作为我们的视觉平面，将该组数据投影在这一平面上从而实现高维数据的可视化。具体步骤如下：

1. 将去中心化的数据向量按列排列构成$n \times m$矩阵X
2. 计算协方差矩阵$\frac{1}{m}XX^T$
3. 应用Jacobi方法求解$\frac{1}{m}XX^T$的所有特征值
4. 提取最大的两个特征值，计算对应的单位特征向量作为我们视觉平面的两个基$\mathbf{e_1}, \mathbf{e_2}$，计算这组数据在这一平面上的投影，投影坐标的计算$x_i =< \mathbf{x}, \mathbf{e_i} >$，并可视化展示结果

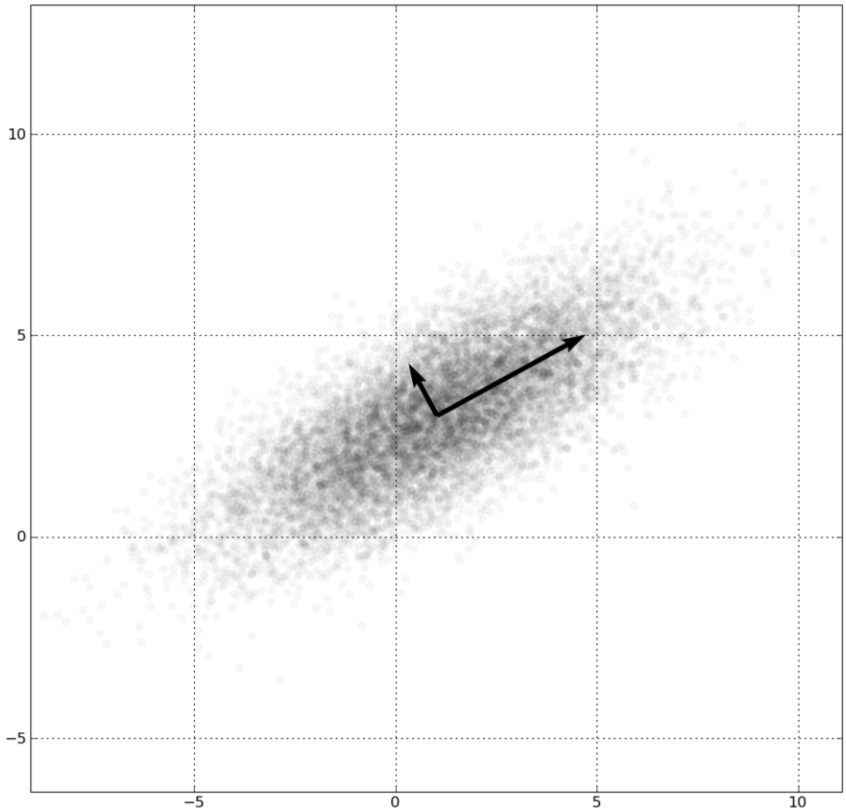


图 1: 一组二维数据的主方向

2.2 SVD分解简介

在PCA中我们实际上找到的是数据矩阵的行向量的各个主方向，对于矩阵的列向量是不是也可以同样进行PCA？是的，我们只需计算$\frac{1}{n}X^TX$的特征值和特征向量即可。我们将$\frac{1}{m}XX^T$各个特征向量按列组成的矩阵记为U，将$\frac{1}{n}X^TX$的特征向量按列组成的矩阵记为V，通过U,V矩阵我们可以对X做这样的分解$X=U\Sigma V^T$，其中Σ 为$n\times m$对角阵，对角元为X^TX特征值的平方根。事实上，对任意一个$n\times m$实矩阵A，我们都可以做这样的分解$A=U\Sigma V^T$，我们将这一分解称为SVD分解（奇异值分解）。可以验证$AA^T=U\Sigma V^TV\Sigma U^T=U\Sigma^2U^T$，$A^TA=V\Sigma U^TU\Sigma V^T=V\Sigma^2V^T$，下面是一个SVD分解的具体例子。

$$\begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} = A = U\Sigma V^T = \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \end{pmatrix} \begin{pmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix}$$

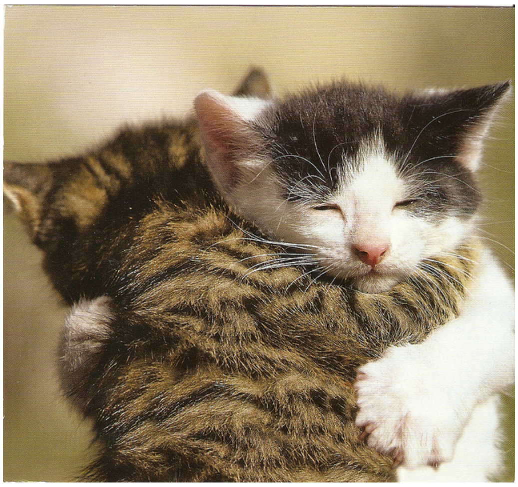
SVD分解的具体步骤如下

1. 对实矩阵A,应用Jacobi方法法计算AA^T的特征值$\lambda_1\lambda_2...\lambda_n$，满足$\lambda_1>=\lambda_2>=...>=\lambda_n$
2. 计算AA^T各个特征值对应的单位特征向量$u_1,u_2,...,u_n$按列组成正交矩阵U
3. 计算$\Sigma=diag(\sqrt{\lambda_1},\sqrt{\lambda_2},...,\sqrt{\lambda_n})$

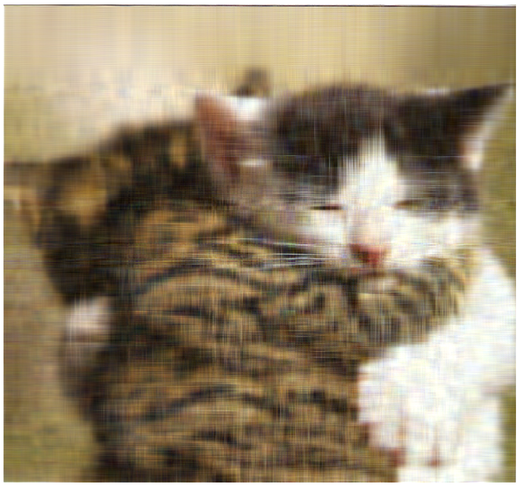
4. 根据公式$V\Sigma = A^TU$计算V，具体计算步骤如下：对上面步骤得到的每一个非零的λ_i，计算单位特征向量$v_i = \frac{A^T u_i}{\sqrt{\lambda_i}}$，将得到的$v_i$扩充成一组完整的正交基按列排列得到正交矩阵$V$

注：SVD分解存在无需计算AA^T更高效的迭代算法，因其较为复杂不在这里介绍，感兴趣的同学可以自行查阅相关资料。

这里我们从PCA出发得到SVD分解，不过事实上SVD有着更为广泛的意义。它是计算方法中比特征值分解更为常用的矩阵分解，在机器学习、计算机图形学等领域有着许多重要应用。SVD能直接应用于数据压缩，下面是应用SVD压缩图像的例子



(a) 原始图像



(b) SVD降维后图像

图 2: 图像压缩

3 实验要求

(a) 自行生成一个4×3的随机矩阵A，应用Jacobi方法求解矩阵AA^T的特征值，计算矩阵A的SVD分解。要求A的每个元素均为$[0, 1]$区间内的随机数。

(b) 对iris（鸢尾花）数据集进行PCA，iris数据集包含150条数据，从提供的文件读取，每条数据有4个属性值和一个标签（标签取值为0，1，2）。要求对这150个4维数据进行PCA，可视化展示这些数据在前两个主方向上的分布，其中不同标签的数据需用不同的颜色或形状加以区分。

要求**作业里所有特征值求解均按照课本上Jacobi方法的流程进行**，控制精度$e = 10^{-6}$。
程序实现完毕后，应撰写实验报告。实验报告中应包含如下内容：

1. 标题、学号、姓名。

2. 实验结果。

   (a) 输出AA^T矩阵Jacobi方法的结果，并输出在迭代过程中，每一次更新的矩阵上非对角元素的平方和。输出矩阵A的SVD分解$A = U\Sigma V^T$中各矩阵U，Σ，V^T。需要注明随机矩阵A的初始值（即生成的随机矩阵是什么）。

   (b) 输出计算得到的协方差矩阵$\frac{1}{m}XX^T$，可视化展示这组数据在前两个主方向上投影的结果。

3. 结果分析：
   (a) 分析 Jacobi方法迭代过程中矩阵非对角元素的平方和是否呈下降趋势。分析求得的特征值是否为对称矩阵特征值的近似值（即计算$det(AA^T - \lambda_i I)$。如果算法无法运行，试分析原因。
   (b) 描述你得到的可视化结果。

4 提交要求

4.1 提交方式

请提交源代码和实验报告。新建目录，并以“A组-HW4-学号-姓名”方式命名，该目录下应包含如下内容：

- src\ （文件夹，存放你的源代码）
- report.pdf （你的实验报告）

将该文件夹以压缩包方式（压缩包命名方式为“A组-HW4-学号-姓名.zip”）
发送到课程邮箱computation_22_1@163.com(周三周五班)，**邮件标题**以同样方式命名。
请严格按照命名方式要求提交，不要交错邮箱，否则可能漏记成绩。

4.2 截止时间

在10月30日23:59分前提交。 **若有特殊情况请向助教说明。**

5 参考

【机器学习】降维——PCA（非常详细）