四川大學
SICHUAN UNIVERSITY

机器学习-第十六章 聚类与降维

教师：胡俊杰 副教授

邮箱：hujunjie@scu.edu.cn

大纲

- K-means聚类
- 奇异值分解（SVD）
- 主成分分析（PCA）
- 总结

无监督学习

- **有监督学习（Supervised learning）**

在一个有监督学习中，数据集具有标签y，目标是找到能够区分各类样本的决策边界

- **无监督学习（ Unsupervised learning）**

在无监督学习中，数据集没有附带标签y，无监督学习应用主要包括聚类、降维、关联规则、推荐系统等方面

K-means聚类

K-均值 (K-means)聚类

- K-means算法是一种常用的聚类算法，其目的在于将无标签数据划分成成不同的组
- K-means算法包含一个迭代过程，其预设k个中心点，将样本划分至k个中心所构成的簇，并计算新的k个中心点，再迭代地划分样本

K-means聚类

K-means算法流程

输入：n个样本的集合X

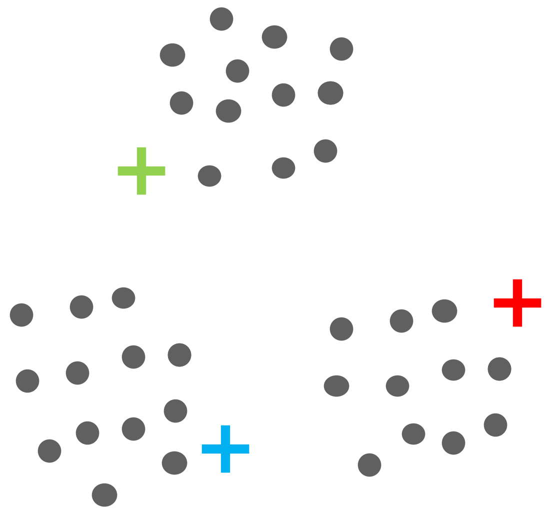
输出：样本集合的聚类结果C

1. 初始化。令$t=0$，随机选择k个样本作为初始聚类中心$m^{(0)}=\left(m_1^{(0)},m_2^{(0)},\dots,m_k^{(0)}\right)$
2. 对样本聚类。对固定的类中心$m^{(t)}=\left(m_1^{(t)},m_2^{(t)},\dots,m_k^{(t)}\right)$，计算每个样本到类中心的距离，将每个样本指派到与其最近的中心的类中，构成聚类结果$C^{(t)}$
3. 计算新的类中心。对聚类结果$C^{(t)}$，计算当前各个类中的样本的均值，作为新的类中心 $m^{(t+1)}=\left(m_1^{(t+1)},m_2^{(t+1)},\dots,m_k^{(t+1)}\right)$
4. 如果已收敛，则输出聚类结果$C^{(t)}$；否则令$t:=t+1$，返回第2步

K-means聚类

K-means算法流程

1. 初始化。随机选择k个样本作为初始聚类中心

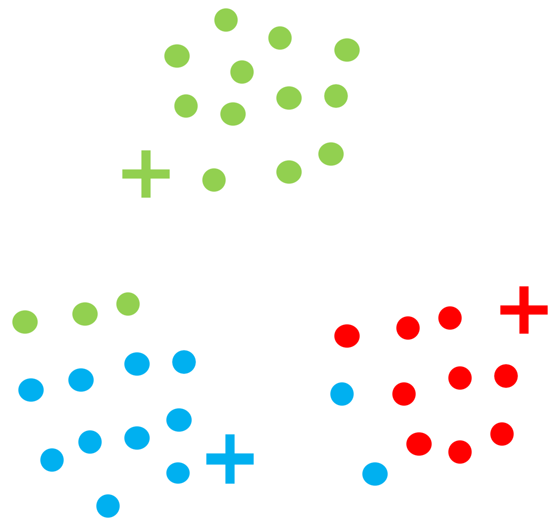


初始化中心

K-means聚类

K-means算法流程

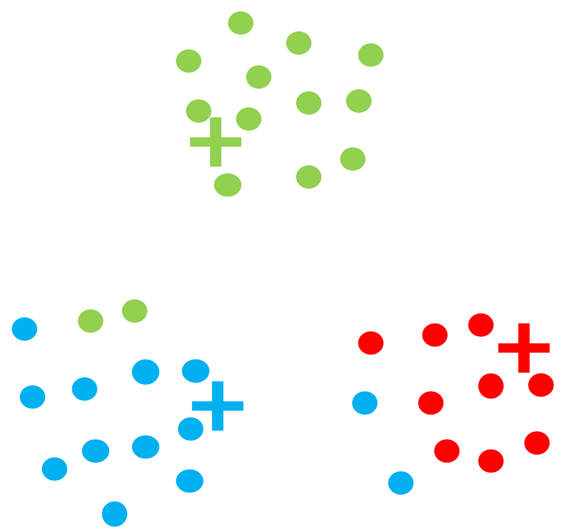
2. 对样本聚类。遍历所有样本，计算每个样本与k个中心之间的距离。将样本指派到与其最近的中心的类中，构成聚类结果$C^{(t)}$



K-means聚类

K-means算法流程

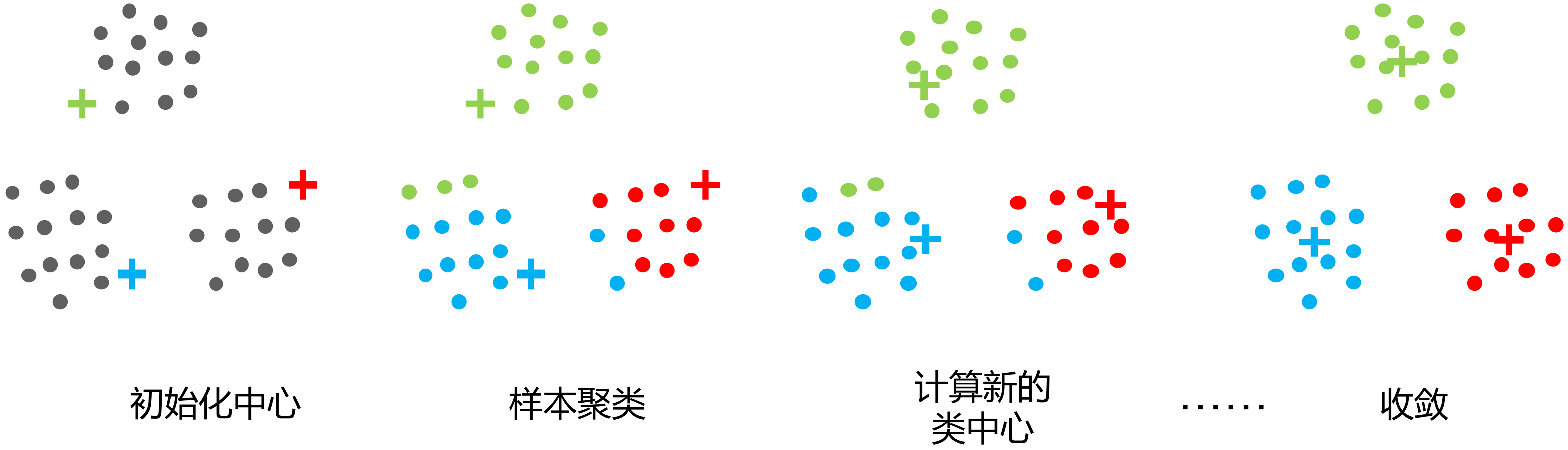
3. 计算新的类中心。对聚类结果$C^{(t)}$，计算当前各个类中的样本的均值，作为新的类中心



迭代更新

K-means聚类

K-means算法流程总结



K-means聚类

K-means的优点

- 计算速度快
- 算法简单，易于理解

K-means的缺点

- 需要预先指定类的数量k，依赖人工经验
- 初始中心的选择直接影响聚类结果

大纲

- K-means聚类
- 奇异值分解（SVD）
- 主成分分析（PCA）
- 总结

奇异值分解

奇异值分解（Singular Value Decomposition, SVD）：将一个非零的$m \times n$实矩阵$A \in R^{m \times n}$表示为以下三个实矩阵的乘积:

$$A = U \Sigma V^T$$

其中U是m阶正交矩阵，V是n阶正交矩阵，Σ是由降序排列的非负对角线元素组成的$m \times n$ 矩形对角矩阵，即

- $UU^T = I$
- $VV^T = I$
- $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \dots, \sigma_p)$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0$
- $p = \min(m, n)$

奇异值分解

子空间

若S是向量空间V的非空子集，且S满足以下条件：

- 对任意实数a，若$x \in S$，则$ax \in S$
- 若$x \in S$且$y \in S$，则$x + y \in S$

则S称为V的子空间

设$v_1, v_2, \dots, v_n$为向量空间V的向量，则其线性组合

$$a_1v_1 + a_2v_2 + \cdots + a_nv_n$$

构成V的子空间，称为$v_1, v_2, \dots, v_n$张成（span）的子空间

如果满足以下两个条件：

1. $v_1, v_2, \dots, v_n$线性无关
2. $v_1, v_2, \dots, v_n$张成向量空间V

则$v_1, v_2, \dots, v_n$构成向量空间V的基（base），基的个数即向量空间的维数

奇异值分解

行空间

- 设A为$m \times n$矩阵，则由A的行向量张成的R^n的子空间，称为A的行空间
- 由A的列向量张成的R^m的子空间，称为A的列空间

零空间

- 设A为$m \times n$矩阵，令$N(A)$为齐次方程组$Ax = 0$的所有解的集合，则$N(A)$为R^n的一个子空间，称为A的零空间（null space），即

$$N(A) = \{x \in R^n | Ax = 0\}$$

正交补

- 设Y为R^n的子空间，R^n中与Y中的每一向量正交的向量集合记为$Y^{\perp}$，即

$$Y^{\perp} = \{x \in R^n | x^T y = 0, \forall y \in Y\}$$

集合$Y^{\perp}$称为Y的正交补

奇异值分解

$A \in R^{m\times n}$

$Ax = 0$

- 从列的角度理解 $\begin{array}{c} \left[\begin{matrix} | & | & \dots & | \end{matrix}\right] \\ \begin{matrix} \text{column 1} & & \text{column n} \end{matrix} \end{array} \begin{bmatrix} x_1 \\ x_2 \\ \cdots \\ x_n \end{bmatrix} = x_1(\text{column1}) + \cdots + x_n(\text{column}\ n) = 0$
- 从行的角度理解，A的行向量与x正交

奇异值分解

$A \in R^{m\times n}$

$$Ax = 0 \quad \longleftrightarrow \quad \begin{bmatrix} -row_1 - \\ -row_2 - \\ \cdots \\ -row_m - \end{bmatrix} x = \begin{bmatrix} 0 \\ 0 \\ \cdots \\ 0 \end{bmatrix}$$

- 矩阵的A的行空间与零空间正交

$$A^T y = 0 \quad \longleftrightarrow \quad \begin{bmatrix} -column_1 - \\ -column_1 - \\ \cdots \\ -column_n - \end{bmatrix} y = \begin{bmatrix} 0 \\ 0 \\ \cdots \\ 0 \end{bmatrix}$$

- 矩阵的A的列空间与**左零空间**正交

注：$A^T y = 0$与$y^T A = 0^T$含义一样

奇异值分解

四个基本子空间：行空间、列空间、零空间、左零空间

$m \times n$矩阵A

行空间的维数为r

零空间的维数为$n - r$

n

列空间的维数为r

左零空间的维数为$m - r$

m

- 零空间是行空间的正交补
- 左零空间是列空间的正交补

奇异值分解

$A = U\sum V^T$

- $A \in R^{m\times n}, U \in R^{m\times m}, V \in R^{n\times n}, \sum \in R^{m\times n}, UU^T = I, VV^T = I$

首先考虑A^TA，其为n阶实对称矩阵，且特征值非负（证明如下）

假设λ是A^TA的特征值，x是λ对应的特征向量

$\|Ax\|^2 = (Ax)^T(Ax) = x^TA^TAx = \lambda x^Tx = \lambda\|x\|^2$,其中$\lambda = \frac{\|Ax\|^2}{\|x\|^2} \geq 0$

设矩阵A的秩为r，则A^TA的秩也是r，由于A^TA是对称矩阵，A^TA的秩等于正特征值的个数，因此

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_r > 0 \qquad \lambda_{r+1} = \lambda_{r+2} = \cdots = \lambda_n = 0$$

令$\sigma_j = \sqrt{\lambda_j}$,对应地

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 \qquad \sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_n = 0$$

注：σ_j即为奇异值

奇异值分解

令$v_1,..,v_r$为A^TA正特征值对应的向量， $v_{r+1},..,v_n$为0特征值对应的向量

$V=[V_1,V_2]$，其中$V_1=[v_1,..,v_r]$ $V_2=[v_{r+1},..,v_n]$

令 $\Sigma_1=\begin{bmatrix}\sigma_1 & & & \\ & \sigma_2 & & \\ & & \cdots & \\ & & & \sigma_r\end{bmatrix}$ $\Sigma=\begin{bmatrix}\Sigma_1 & 0\\ 0 & 0\end{bmatrix}$, $\Sigma\in R^{m\times n}$

$A=\boxed{[U_1,U_2]}\begin{bmatrix}\Sigma_1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}V_1^T\\ V_2^T\end{bmatrix}$ 其中$U=[U_1,U_2]$， $U\in R^{m\times m}$，且U为正交矩阵

?

奇异值分解

$$A = [U_1, U_2]\begin{bmatrix}\Sigma_1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}V_1^T\\ V_2^T\end{bmatrix}$$

$$= U_1\Sigma_1 V_1^T$$

- 当$U_1\Sigma_1 = AV_1$时，上式成立

  即要求$u_j = \boxed{\frac{1}{\sigma_j}Av_j}, j = 1,2,\dots,r$　　$U_1 = [u_1 \quad u_2 \quad \dots \quad u_r]$

  已知

  $u_i, u_j \in U_1$，　$u_i^T u_j = \frac{1}{\sigma_i}v_i^T A^T \frac{1}{\sigma_j}Av_j = \frac{1}{\sigma_i\sigma_j}v_i^T(A^T Av_j) = \frac{1}{\sigma_i\sigma_j}\sigma_j^2 v_i^T v_j = 0$

  $u_1, u_2, \dots, u_r$构成A的列空间的一组标准正交基（因为$\frac{1}{\sigma_j}Av_j$是对A的列向量的线性组合）

- 如何确定剩下的U_2？

奇异值分解

$$A = [U_1, U_2]\begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix}$$

$U_1 = [u_1, u_2, \dots, u_r]$构成$A$的列空间的一组标准正交基

要求$U = [U_1, U_2]$为正交矩阵，即求U_1 的正交补

因为$U_1 = [u_1, u_2, \dots, u_r]$张成$A$的列空间，可知$U_2$即$A$的左零空间（左零空间是列空间的正交补）

即求$A^T y = 0$的解，进而构建$U_2 = [u_{r+1}, u_{r+2}, \dots, u_m]$

至此，已实现$A = [U_1, U_2]\begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} V_1^T \\ V_2^T \end{bmatrix} = U\Sigma V^T$

- 以上也称为矩阵的完全奇异值分解（full singular value decomposition）

奇异值分解

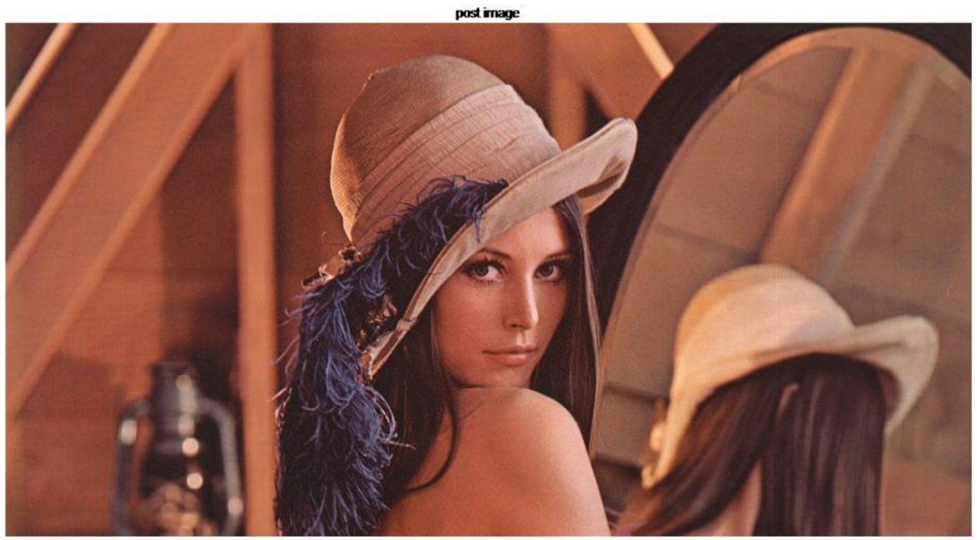
- 在矩阵的奇异值分解中，只取最大的k个奇异值对应的部分，则得到矩阵的截断奇异值分解（truncated singular value decomposition）

设A为$m \times n$实矩阵，秩为r，对于$0 < k < r$，称$U_k \sum_k V_k^T$为矩阵A的截断奇异值分解

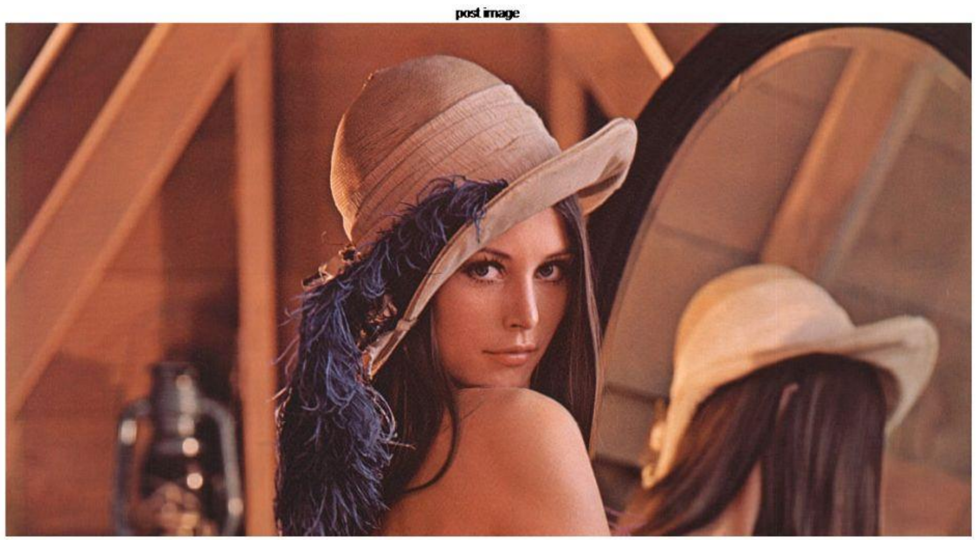
$$A \approx U_k \sum_k V_k^T$$

其中U_k是$m \times k$矩阵，对应完全奇异值分解中U的前k列。V_k是$n \times k$矩阵，对应V 的前k 列。 $\sum_k$ 是k阶对角矩阵，由$\sum$的前k个对角线元素构成。截断奇异值分解对应有损压缩。

奇异值分解



原始图像



处理后的图像

$m = 575$， $n = 1081$

原始图像维度$575 \times 1081 \times 3 = 1{,}864{,}725$

设$k = 150$，则经过截断奇异值分解后的矩阵及维度：

$U_{m\times k} = 575 \times 150$，$\sum_{k\times k} = 150 \times 150$，$V_{k\times n}^{T} = 1081 \times 150$

原始图像经过压缩后的维度：$3 \times (575 \times 150 + 150 \times 150 + 1081 \times 150) =$
$812{,}700$

大纲

- **K-means聚类**
- **奇异值分解（SVD）**
- **主成分分析（PCA）**
- **总结**

主成分分析

- 主成分分析（Principal Component Analysis, PCA）是一种常用的无监督学习方法，其利用正交变换把由线性相关变量表示的观测数据转换为少数几个由线性无关变量表示的数据，线性无关的变量称为主成分
- 主成分的个数（如k）通常小于原始变量（如m）的个数，所以主成分分析属于降维方法

$$X_{m\times n} \xrightarrow{\text{PCA}} Y_{k\times n}$$

- n个样本
- 每个样本m维

- n个样本
- 每个样本k维

主成分分析

核心思想

- 主成分分析中，首先对给定数据进行规范化，使得数据每一变量的平均值为0，方差为1
- 之后对数据进行正交变换，原来由线性相关变量表示的数据，通过正交变换变成由若干个线性无关的新变量表示的数据
- 新变量是可能的正交变换中变量的方差（信息保存）最大的，方差表示在新变量上信息的大小
- 可以把数据由少数主成分表示，实现对数据的降维

主成分分析

- 在数据总体（population)上进行的主成分分析称为总体主成分分析
- 在有限样本上进行的主成分分析称为样本主成分分析
- 总体主成分分析是样本主成分分析的基础

主成分分析

- 假设$\boldsymbol{x} = (x_1, x_2, \dots, x_m)^T$是$m$维随机变量，其均值向量是$\boldsymbol{u}$

$$\boldsymbol{u} = E(\boldsymbol{x}) = (u_1, u_2, \dots, u_m)^T$$

- 协方差矩阵为Σ

$$\Sigma = \operatorname{cov}(\boldsymbol{x}, \boldsymbol{x}) = E[(\boldsymbol{x} - \boldsymbol{u})(\boldsymbol{x} - \boldsymbol{u})^T]$$

- 考虑m维随机变量$\boldsymbol{x}$到m维随机变量$\boldsymbol{y} = (y_1, y_2, \dots, y_m)^T$的线性变换

$$y_i = \boldsymbol{\alpha}_i^T \boldsymbol{x} = \alpha_{1i} x_1 + \alpha_{2i} x_2 + \cdots + \alpha_{mi} x_m$$

其中$\boldsymbol{\alpha}_i^T = (\alpha_{1i}, \alpha_{2i}, \dots, \alpha_{mi}), i = 1,2, \dots, m$

- y_i性质如下

$$E(y_i) = \boldsymbol{\alpha}_i^T \boldsymbol{u}, i = 1,2, \dots, m$$

$$\operatorname{var}(y_i) = E\left[(\boldsymbol{\alpha}_i^T \boldsymbol{x} - \boldsymbol{\alpha}_i^T \boldsymbol{u})(\boldsymbol{\alpha}_i^T \boldsymbol{x} - \boldsymbol{\alpha}_i^T \boldsymbol{u})^T\right] = \boldsymbol{\alpha}_i^T \Sigma \boldsymbol{\alpha}_i$$

$$\operatorname{cov}(y_i, y_j) = E\left[(\boldsymbol{\alpha}_i^T \boldsymbol{x} - \boldsymbol{\alpha}_i^T \boldsymbol{u})(\boldsymbol{\alpha}_j^T \boldsymbol{x} - \boldsymbol{\alpha}_j^T \boldsymbol{u})^T\right] = \boldsymbol{\alpha}_i^T \Sigma \boldsymbol{\alpha}_j$$

主成分分析

总体主成分 $$y_i = \boldsymbol{\alpha}_i^T \boldsymbol{x} = \alpha_{1i}x_1 + \alpha_{2i}x_2 + \cdots + \alpha_{mi}x_m$$

- $y_1, y_2, \dots, y_m$为$\boldsymbol{x}$的第1主成分、第2主成分、…、第m主成分，当满足以下条件时：
  - 系数向量$\boldsymbol{\alpha}_i^T$是单位向量，即$\boldsymbol{\alpha}_i^T \boldsymbol{\alpha}_i = 1, i = 1,2,\dots,m, \boldsymbol{\alpha}_i^T \boldsymbol{\alpha}_j = 0\ (i \neq j)$
  - 变量y_i与y_j互不相关，即$\mathrm{cov}(y_i, y_j) = 0, (i \neq j)$
  - 变量y_1是$\boldsymbol{x}$的所有线性变换中方差最大的； y_2是与y_1不相关的$\boldsymbol{x}$的所有线性变换中方差最大的；一般地， y_i是与$y_1, y_2, \dots, y_{i-1}$都不相关的$\boldsymbol{x}$的所有线性变换中方差最大的

主成分分析

- 定理：设$\boldsymbol{x}$是m维随机变量，Σ是$\boldsymbol{x}$的协方差矩阵，Σ的特征值分别是$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$，特征值对应的单位特征向量分别是$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \ldots, \boldsymbol{\alpha}_m$，则$\boldsymbol{x}$的第$k$个主成分是

$$y_k = \boldsymbol{\alpha}_k^T \boldsymbol{x} = \alpha_{1k} x_1 + \alpha_{2k} x_2 + \cdots + \alpha_{mk} x_m, k = 1,2, \ldots, m$$

$\boldsymbol{x}$的第k主成分的方差即协方差矩阵Σ的第k个特征值

$$\mathrm{var}(y_k) = \boldsymbol{\alpha}_k^T \Sigma \boldsymbol{\alpha}_k = \lambda_k$$

主成分分析

证明：采用拉格朗日乘子法求出主成分

首先求$\boldsymbol{x}$的第一主成分$y_1=\boldsymbol{\alpha}_1^T\boldsymbol{x}$，即求系数向量$\boldsymbol{\alpha}_1$。由主成分定义可知，$\boldsymbol{\alpha}_1$是在$\boldsymbol{\alpha}_1^T\boldsymbol{\alpha}_1=1$条件下，使得方差$\mathrm{var}(\boldsymbol{\alpha}_1^T\boldsymbol{x})=\boldsymbol{\alpha}_1^T\Sigma\boldsymbol{\alpha}_1$达到最大的，即求解带约束的优化问题：

$$\max_{\boldsymbol{\alpha}_1}\boldsymbol{\alpha}_1^T\Sigma\boldsymbol{\alpha}_1$$

$$s.t.\ \boldsymbol{\alpha}_1^T\boldsymbol{\alpha}_1=1$$

定义拉格朗日函数$\boldsymbol{\alpha}_1^T\Sigma\boldsymbol{\alpha}_1-\lambda(\boldsymbol{\alpha}_1^T\boldsymbol{\alpha}_1-1)$，其中$\lambda$是拉格朗日乘子。对$\boldsymbol{\alpha}_1$求导，并令其为0，得

$$\Sigma\boldsymbol{\alpha}_1-\lambda\boldsymbol{\alpha}_1=0$$

即λ是Σ的特征值， $\boldsymbol{\alpha}_1$是对应的单位特征向量。假设$\boldsymbol{\alpha}_1$是Σ最大特征值λ_1对应的单位特征向量，则$\boldsymbol{\alpha}_1$与λ_1为最优化问题的解，$\boldsymbol{\alpha}_1^T\boldsymbol{x}$构成第一主成分，其方差为协方差矩阵的最大特征值

主成分分析

求第二主成分即求解以下带约束的优化问题：

$$\max_{\boldsymbol{\alpha}_2} \boldsymbol{\alpha}_2^T \Sigma \boldsymbol{\alpha}_2$$

$$s.t. \boldsymbol{\alpha}_2^T \boldsymbol{\alpha}_2 = 1, \operatorname{cov}(y_1, y_2) = \boldsymbol{\alpha}_1^T \Sigma \boldsymbol{\alpha}_2 = 0$$

注意到$\boldsymbol{\alpha}_1^T \Sigma \boldsymbol{\alpha}_2 = \boldsymbol{\alpha}_2^T \Sigma \boldsymbol{\alpha}_1 = \boldsymbol{\alpha}_2^T \lambda_1 \boldsymbol{\alpha}_1 = \lambda_1 \boldsymbol{\alpha}_2^T \boldsymbol{\alpha}_1 = \lambda_1 \boldsymbol{\alpha}_1^T \boldsymbol{\alpha}_2$

(协方差矩阵对称)

定义拉格朗日函数 $\boldsymbol{\alpha}_2^T \Sigma \boldsymbol{\alpha}_2 - \lambda(\boldsymbol{\alpha}_2^T \boldsymbol{\alpha}_2 - 1) - \phi \boldsymbol{\alpha}_2^T \boldsymbol{\alpha}_1$，其中$\lambda$和$\phi$是拉格朗日乘子，对$\boldsymbol{\alpha}_2$求导，并令其为0，得

$$2\Sigma \boldsymbol{\alpha}_2 - 2\lambda \boldsymbol{\alpha}_2 - \phi \boldsymbol{\alpha}_1 = 0$$

$$\Big\downarrow \text{左乘}\boldsymbol{\alpha}_1^T$$

$$2\underbrace{\boxed{\boldsymbol{\alpha}_1^T \Sigma \boldsymbol{\alpha}_2}}_{}^{0} - 2\lambda \overbrace{\boxed{\boldsymbol{\alpha}_1^T \boldsymbol{\alpha}_2}}^{0} - \phi \boldsymbol{\alpha}_1^T \boldsymbol{\alpha}_1 = 0$$

可知$\phi = 0$，因此$\Sigma \boldsymbol{\alpha}_2 - \lambda \boldsymbol{\alpha}_2 = 0$

λ是Σ的特征值， $\boldsymbol{\alpha}_2$是对应的单位特征向量。假设$\boldsymbol{\alpha}_2$是Σ的第二大特征值λ_2对应的单位特征向量，则$\boldsymbol{\alpha}_2$与λ_2构成以上优化问题的解。于是$\boldsymbol{\alpha}_2^T \boldsymbol{x}$构成第二主成分，其方差为协方差矩阵的第二大特征值

主成分分析

- 一般地，系数向量$\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_m$分别是$\Sigma$的第1直至第$m$个单位特征向量， $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m \geq 0$为对应的特征值。第$k$个主成分的方差为$\Sigma$的第$k$大特征值：

$$\mathrm{var}(y_k) = \mathrm{var}\left(\boldsymbol{\alpha}_k^T \boldsymbol{x}\right) = \lambda_k, k = 1,2, \dots, m$$

- 主成分分析的主要目的在于降维，因此一般选择$k(k \ll m)$个主成分（线性无关变量）来代替m个原有变量（线性相关变量），从而简化问题，并能保留原有变量的大部分信息

主成分分析

样本主成分分析

- 假设对m维随机变量$\boldsymbol{x} = (x_1, x_2, \dots, x_m)^T$进行$n$次独立观测，得到$\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$个观测样本，用样本矩阵$X$表示

$$X = [\boldsymbol{x}_1\ \boldsymbol{x}_2\ \ \dots \boldsymbol{x}_n] = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{bmatrix}$$

- 样本主成分分析即先规范化处理X（各变量均值为0，方差为1），计算样本的协方差矩阵Σ，并对Σ进行特征值分解，得到前k个特征值对应的单位特征向量$\boldsymbol{\alpha}_i, i = 1,2, \dots, k$
- 计算n个样本的k个主成分

$$Y = A^T X$$

$$A_{m\times k} = [\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2, \dots, \boldsymbol{\alpha}_k] \quad X_{m\times n} = [\boldsymbol{x}_1\ \boldsymbol{x}_2\ \ \dots \boldsymbol{x}_n] \quad Y_{k\times n} = [\boldsymbol{y}_1\ \boldsymbol{y}_2\ \ \dots \boldsymbol{y}_n]$$

主成分分析

数据矩阵的奇异值分解算法

- 对于规范化后的$m \times n$样本矩阵X，构造新的$n \times m$矩阵X'

$$X' = \frac{1}{\sqrt{n-1}} X^T$$

可知${X'}^T X' = \left(\frac{1}{\sqrt{n-1}} X^T\right)^T \left(\frac{1}{\sqrt{n-1}} X^T\right) = \frac{1}{n-1} X X^T$，即${X'}^T X'$为$X$的协方差矩阵，主成分分析转化为求$\Sigma = {X'}^T X'$的特征值和对应的特征向量

- 假设X'的截断奇异值分解为$X' = U \Sigma V^T$，则V的列向量就是$\Sigma = {X'}^T X'$的单位特征向量，即X的主成分可通过X'的奇异值分解实现

大纲

- K-means聚类
- 奇异值分解（SVD）
- 主成分分析（PCA）
- 总结

总结

考试时间地点
6月17日（17周周二） 16:45-18:45，综合楼B407

考试题型：名词解释、填空、问答题

谢 谢！