低秩矩阵分解(Low-rank Matrix Factorization)

1. 背景介绍

矩阵分解(Matrix Factorization, MF)是数据科学领域一项重要的技术，其关键在于分析数据中存在的潜在结构，可通过发现这种结构获取数据的压缩表示.矩阵分解通过将一个原矩阵分解为两个低秩矩阵的乘积或一个低秩矩阵与一个稀疏矩阵之和的形式，为降维、聚类、矩阵补全等提供了一种统一的方法.

矩阵分解模型主要包括以下几类：

1. 基本矩阵分解(Basic Matrix Factorization)；
2. 非负矩阵分解(Non-negative Matrix Factorization)；
3. 正交非负矩阵分解(Orthogonal non-negative Matrix Factorization)

本次主要讨论基本矩阵分解模型.

1. 模型描述
   1. **Low-rank matrix approximation**

给定一个秩为的矩阵，欲求其最优秩近似矩阵，该问题可形式化为

其中，矩阵Frobenius范数. 选取F范数为假设数据受到独立同分布的高斯噪声污染.

* 1. **Basic Matrix Factorization**

给定样本矩阵. 希望找到低秩矩阵，最小化

其中，为正则项，通常选择范数或范数.

若令，可得：

因此，为一低秩矩阵.

* 1. **Robust principal component analysis**

给定样本矩阵. 经典的PCA研究希望找到低秩矩阵，使得和之间的差异最小. 得以下优化问题：

其中为所求子空间的目标维数. 该模型即问题2.1的重述.

经典的PCA假设数据的噪声是高斯的，对于大的噪声或严重离群点，PCA会受影响而无法正常工作. 为改善此缺陷提出了鲁棒PCA(Robust principal component analysis, RPCA)算法.

鲁棒主成分分析算法认为，一个数据矩阵既包含结构信息，也包含噪声. 因此我们可将该矩阵分解为两个矩阵相加，一个是低秩的（由内部结构导致各行或各列线性相关），另一个是稀疏的（数据含有稀疏的噪声）. 即如下双目标优化问题：

引入折中因子，将双目标优化问题转为单目标优化问题：

其中，矩阵0-范数为矩阵中非0元素的个数.

将NP难问题作凸松弛后，得：

其中，矩阵核范数，矩阵1,1范数. 容易证明矩阵的核范数可用其奇异值表示，即.

1. 模型求解
   * 1. **SVD分解**

SVD(Singular Value Decomposition)，奇异值分解. 对任意矩阵，可将其唯一分解为：

其中为正交矩阵，只有对角位上有非零元素，且从大到小排列，这些对角元素称为奇异值.

利用奇异值分解的性质：

选取前个奇异值，的前列，的前行近似描述矩阵，即

则

就是式(2.1)的最优解，此即Eckart-Young-Mirsky定理.

* + 1. **梯度下降法**

假设模型存在高斯噪声，选取范数和范数，将式(2.2)实例化得：

y利用如下公式：

计算得：

利用规则：

计算对偏导数：

由此，可通过梯度下降法交替更新.

* + 1. **阈值迭代算法**

将原样本矩阵按行零均值化，得矩阵

则协方差矩阵可表示为

SVD(Singular Value Decomposition)，奇异值分解. 对任意矩阵，可将其唯一分解为：

其中为正交矩阵，只有对角位上有非零元素，且从大到小排列，这些对角元素称为奇异值.

奇异值分解的性质如下：

1. 为正交矩阵，即；
2. ，，即是的特征向量，是的特征向量，奇异值的平方是和的特征值；
3. 选取前个奇异值，的前列，的前行近似描述矩阵，即

根据，我们只需将做SVD分解，并选取右奇异矩阵的前列.

**算法描述**

1. 求零均值化矩阵;
2. 对作奇异值分解;
3. 选取右奇异矩阵的前列.
4. 模型推广

对原样本矩阵作主成分分析后，得到新样本矩阵和变换矩阵，希望根据变换矩阵将原样本矩阵分解为标准正交基的线性组合.

假设，则有

其中，.

由于的行向量相互正交，因此可通过下式求解.

则有

或对做SVD分解，得左奇异矩阵奇异值矩阵和右奇异矩阵. 由奇异值分解性质有

可根据上式将原样本矩阵近似表示为个正交基的线性组合.