习题1.4的参考答案

1. 主成分对样本的重构

标准化样本矩阵$\mathbf{X} \in \mathbb{R}^{N\times p}$的SVD分解为：

$$\mathbf{X} = \mathbf{U}\Sigma\mathbf{V}^{\top}, \tag{0.1}$$

其中，$\mathbf{U} \in \mathbb{R}^{N\times p}$是左特征向量, $\Sigma \in \mathbb{R}^{p\times p}$是一个对角矩阵， $\mathbf{V} \in \mathbb{R}^{p\times p}$是右特征向量。 可以验证，$V_k = \alpha_k$（$V_k$是$\mathbf{V}$的第$k$列， α_k是第k个主成分对应的权重）。

如果保留前k个主成分，那么样本主成分（也就是主成分得分）可以表示为：

$$\mathbf{Z} = \mathbf{X}\mathbf{V}^{(k)} = \mathbf{U}^{(k)}\Sigma^{(k)} \in \mathbb{R}^{N\times k} \tag{0.2}$$

其中 $\mathbf{V}^{(k)}$ 是$\mathbf{V}$的前k列, $\mathbf{U}^{(k)}$是$\mathbf{U}$的前k列, $\Sigma^{(k)}$ 是Σ前$k\times k$子矩阵。 如果希望从主成分重构$\mathbf{X}$，则可以得到

$$\mathbf{X}^{(k)} = \mathbf{U}^{(k)}\Sigma^{(k)}(\mathbf{V}^{(k)})^{\top} \tag{0.3}$$

Why?

用主成分$\mathbf{Z}^{(k)}$重构$\mathbf{X}$的含义是希望寻找一个$\mathbf{B} \in \mathbb{R}^{k\times N}$，使得

$$\mathbf{B} = \arg\min_{\mathbf{B}} \left\|\mathbf{X} - \mathbf{Z}\mathbf{B}\right\|_F^2 \tag{0.4}$$

可以求得$\mathbf{B} = (\mathbf{V}^{(k)})^{\top}$.

2. 习题1.4的求解

题目解析：这个题目是想要我们证明，求解下面的问题，与求解主成分分析是等价的：

$$\begin{aligned} &\min_{\mathbf{L}} \quad \|\mathbf{X} - \mathbf{L}\|_F \\ &\text{s.t.} \quad \operatorname{rank}(\mathbf{L}) \leqslant k \end{aligned}$$

这里的$\mathbf{L}$相当于是对$\mathbf{X}$的重构，同时，$\mathbf{L}$是一个rank不大于k的低秩矩阵。 联系到主成分分析的问题，我们希望验证$\mathbf{L} = \mathbf{X}^{(k)}$.

Proof. 首先，注意到

$$\|\mathbf{X} - \mathbf{X}^{(k)}\|_F^2 = \|\sum_{i=k+1}^{N} \sigma_i U_i V_i^\top\|_F^2 = \sum_{k+1}^{N} \sigma_i^2 \tag{0.5}$$

其中σ_i是Σ第(i,i)个元素， U_i是矩阵$\mathbf{U}$的第i列， V_i是矩阵$\mathbf{V}$的第i列.

接下来，我们希望验证，对于任意的rank为k的矩阵$\mathbf{L}$，都成立

$$\|\mathbf{X} - \mathbf{X}^{(k)}\|_F^2 = \sum_{i=k+1}^{N} \sigma_i^2 \leq \|\mathbf{X} - \mathbf{L}\|_F^2 \tag{0.6}$$

这里插播一个矩阵分析中的结论： 对于任意矩阵$\mathbf{A}, \mathbf{B} \in \mathbb{R}^{m \times n}$ $(n \leq m)$, 成立

$$\sigma_{i+j-1}(\mathbf{A} + \mathbf{B}) \leq \sigma_i(\mathbf{A}) + \sigma_j(\mathbf{B}) \tag{0.7}$$

其中，$1 \leq i, j \leq n$， $i + j - 1 \leq n$。 感兴趣的同学可以看*Topics in Matrix Analysis*这本书的定理3.3.16的详细证明。 注意这是一个经典结论，希望大家掌握。

回到本证明中，这里我们令$\mathbf{A} = \mathbf{X} - \mathbf{L}$， $\mathbf{B} = \mathbf{L}$。可知$\sigma_{k+1}(\mathbf{B}) = 0$. 则可得$\sigma_{i+k}(\mathbf{X}) \leq \sigma_i(\mathbf{X} - \mathbf{L})$. 因此成立

$$\|\mathbf{X} - \mathbf{X}^{(k)}\|_F^2 = \sum_{i=k+1}^{N} \sigma_i^2 \leq \sum_{i=1}^{N} \sigma_i^2(\mathbf{X} - \mathbf{L}) = \|\mathbf{X} - \mathbf{L}\|_F^2 \tag{0.8}$$

结论得证。

$\square$