Information Retrieval in High Dimensional Data

Assignment 2

Group 8

Shangyin Gao, Siyuan Liu, Jia Hua, Xianrui Niu, Hanbo Hu

23.11.2017

Task 1

1.1
First we assume s is a normalized vector and s $\in R^n$

$$\mathbf{s} = \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_p \end{bmatrix}$$

From SVD we know that,

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_{11} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_{pp} \end{bmatrix}$$

We should first calculate $s^T \sum \sum^T s$.

$$\begin{aligned}
\mathbf{s^T \Sigma \Sigma^T s} = \mathbf{s^T}(\mathbf{\Sigma \Sigma^T}) &= s^T \begin{bmatrix} \sigma_{11}^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_{pp}^2 \end{bmatrix} s \\
&= \begin{bmatrix} s_1 & s_2 & \dots & s_p \end{bmatrix} \begin{bmatrix} \sigma_{11}^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_{pp}^2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_p \end{bmatrix} \\
&= \begin{bmatrix} s_1\sigma_{11}^2 & s_2\sigma_{22} & \dots & s_p\sigma_{pp}^2 \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_p \end{bmatrix} \\
&= s_1^2\sigma_{11}^2 + s_2^2\sigma_{22}^2 + \ldots + s_p^2\sigma_{pp}^2 \quad (*)
\end{aligned}$$

$$||s|| = 1 \quad \Rightarrow \quad \sqrt{s_1^2 + s_2^2 + \ldots + s_p^2} = 1 \quad \Rightarrow \quad s_1^2 + s_2^2 + \ldots + s_p^2 = 1$$
$$s_1^2 = 1 - \left(s_2^2 + s_3^2 + \ldots + s_p^2\right)$$

$$\begin{aligned}(*) &= \sigma_{11}^2 \left(1 - \left(s_2^2 + s_3^2 + \ldots + s_p^2\right)\right) + s_2^2\sigma_{22}^2 + s_3^2\sigma_{33}^2 + \ldots + s_p^2\sigma_{pp}^2 \\ &= \sigma_{11}^2 + s_2^2\left(\sigma_{22}^2 - \sigma_{11}^2\right) + \ldots + s_p^2\left(\sigma_{pp}^2 - \sigma_{11}^2\right)\end{aligned}$$

We know that the singular values are sorted in a descending manner, i.e. $\sigma_{11} \geq \sigma_{22} \geq \ldots \geq \sigma_{pp}$.

Thus,

$$\sigma_{22}^2 - \sigma_{11}^2 \leq 0 \quad \sigma_{33}^2 - \sigma_{11}^2 \leq 0 \quad \ldots \quad \sigma_{pp}^2 - \sigma_{11}^2 \leq 0$$

If we want to get the vector s to make the equation to be maximal, we should make $s_2 = s_3 = \ldots = s_p = 0 \Rightarrow s_1 = 1$

So we can finally get the vector

$$\hat{\mathbf{s}} = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

1.2

For this part, we should calculate $\frac{1}{N}\ a^T X X^T a$

$$\frac{1}{N} a^T X X^T a = \frac{1}{N} a^T U \Sigma V^T \left(U \Sigma V^T\right)^T a = \frac{1}{N} a^T U \Sigma V^T V \Sigma^T U^T a \quad (1)$$

$$V^T V = I \Rightarrow \quad (1) = \frac{1}{N} a^T U \Sigma \Sigma^T U^T a \quad (2)$$

From definition we know that,

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_{11} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \sigma_{pp} \end{bmatrix}$$

Therefore,

$$\mathbf{\Sigma \Sigma^T} = \begin{bmatrix} \sigma_{11}^2 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \sigma_{pp}^2 \end{bmatrix}$$

$$(2) = \frac{1}{N} a^T U \Sigma \Sigma^T U^T a = \frac{1}{N} a^T \begin{bmatrix} u_1 & u_2 & \dots & u_p \end{bmatrix} \begin{bmatrix} \sigma_{11}^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_{pp}^2 \end{bmatrix} \begin{bmatrix} u_1^T \\ u_2^T \\ \vdots \\ u_p^T \end{bmatrix} a$$

$$= \frac{1}{N} \begin{bmatrix} a^T u_1 & a^T u_2 & \dots & a^T u_p \end{bmatrix} \begin{bmatrix} \sigma_{11}^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_{pp}^2 \end{bmatrix} \begin{bmatrix} u_1^T a \\ u_2^T a \\ \vdots \\ u_p^T a \end{bmatrix} \quad (3)$$

From the Task we know that, a is set to a column of U. Therefore,

$$a = u_i \qquad a^T a = u_i^T u_i = 1 \qquad a^T u_j = u_i^T u_j = 0, j \neq i$$

Therefore,

$$(3) = \frac{1}{N} \begin{bmatrix} u_i^T u_1 & u_i^T u_2 & \dots & u_i^T u_p \end{bmatrix} \begin{bmatrix} \sigma_{11}^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_{pp}^2 \end{bmatrix} \begin{bmatrix} u_1^T u_i \\ u_2^T u_i \\ \vdots \\ u_p^T u_i \end{bmatrix}$$

$$= \frac{1}{N} \begin{bmatrix} 0 & \dots & 0 & u_i^T u_i & 0 & \dots & 0 \end{bmatrix} \begin{bmatrix} \sigma_{11}^2 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_{pp}^2 \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 0 \\ u_i^T u_i \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$= \frac{1}{N} \left[\sigma_{ii}^2 u_i^T u_i + 0 + 0 + \dots + 0 \right] = \frac{1}{N} \left[\sigma_{ii}^2 \times 1 + 0 + 0 + \dots + 0 \right]$$

$$= \frac{1}{N} \sigma_{ii}^2$$

In order to maximize it, we should choose i which can maximize σ, i.e. i=1, since $\sigma_{11} \geq \sigma_{22} \geq \dots \geq \sigma_{pp}$. If i = 1, then $a = u_1$ and

$$\frac{1}{N} \sum_{i=1}^{N} \left(a^T x_i \right)^2 = \frac{1}{N} a^T X X^T a = \frac{1}{N} \sigma_{11}^2 u_1^T u_1 = \frac{1}{N} \sigma_{11}^2 \times 1 = \frac{1}{N} \sigma_{11}^2$$

is maximized.