Assignment 2

1

Given $A = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$ the singular value decomposition of the matrix A and $A_k = \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$ its rank-k approximation, we want to express $\|A_k\|_F^2$ in term of the singular values of A.

Let's start from some definitions. Given a matrix M and a vector $\mathbf{w}$

$$\|\mathbf{w}\|_2 = \sqrt{\sum_i w_i^2} \quad \text{is the 2-norm of } \mathbf{w}$$

$$\|M\|_F = \sqrt{\sum_i \sum_j M_{ij}^2} \quad \text{is the Frobenius norm of } M$$

The squares of these values are $\|\mathbf{w}\|_2^2 = \sum_i w_i^2$ and $\|M\|_F^2 = \sum_i \sum_j M_{ij}^2$, respectively. Using these definitions, we can easily state the following relation

$$\|M\|_F^2 = \sum_s \|M_s\|_2^2 \quad \text{where } M_s \text{ is the s-th row of } M$$

Now we are ready to give an answer to the original problem. We can use the last relation with $\|A_k\|_F^2$, obtaining

$$\begin{aligned}
\|A_k\|_F^2 &= \sum_s \|(A_k)_s\|_2^2 && \text{where } (A_k)_s \text{ is a row vector} \\
&= \sum_s (A_k)_s (A_k)_s^T && \text{using the dot product definition} \\
&= \sum_s \left(\sum_{i=1}^{k} \sigma_i (\mathbf{u}_i)_s \mathbf{v}_i^T \sum_{j=1}^{k} \sigma_j (\mathbf{u}_j)_s \mathbf{v}_j \right) && \text{since } (A_k)_s = \sum_{i=1}^{k} \sigma_i (\mathbf{u}_i)_s \mathbf{v}_i^T \\
&= \sum_s \sum_{i=1}^{k} (\sigma_i (\mathbf{u}_i)_s)^2 && \text{since } \mathbf{v}_i^T \mathbf{v}_i = 1 \wedge \mathbf{v}_i^T \mathbf{v}_j|_{i \neq j} = 0 \\
&= \sum_{i=1}^{k} \sigma_i^2 \sum_s (\mathbf{u}_i)_s^2 && \text{exchanging sums} \\
&= \sum_{i=1}^{k} \sigma_i^2 && \text{since } \forall i \, \mathbf{u}_i \text{ is a unit norm vector}
\end{aligned}$$

2

In the case of $\|A - A_k\|_F^2$ we can use the same steps of the previous quesion, taking into consideration that $A - A_k$ is a matrix that can be written as $A - A_k = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T - \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^T = \sum_{i=k+1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T$.

Thus, we obtain

$$\begin{aligned}
\|A - A_k\|_F^2 &= \sum_s \|(A - A_k)_s\|_2^2 \\
&= \sum_s (A - A_k)_s (A - A_k)_s^T \\
&= \sum_s \left(\sum_{i=k+1}^r \sigma_i (\mathbf{u}_i)_s \mathbf{v}_i^T \sum_{j=k+1}^r \sigma_j (\mathbf{u}_j)_s \mathbf{v}_j \right) \\
&= \sum_s \sum_{i=k+1}^r (\sigma_i (\mathbf{u}_i)_s)^2 \\
&= \sum_{i=k+1}^r \sigma_i^2 \sum_s (\mathbf{u}_i)_s^2 \\
&= \sum_{i=k+1}^r \sigma_i^2
\end{aligned}$$

3

Now we have to express $\|A_k\|_2^2$ in term of the singular values of A. Again, let's begin from some definitions. The 2-norm of a matrix M is defined as $\max_{\mathbf{x}:\|\mathbf{x}\|=1} \|M\mathbf{x}\|_2$. Applying it to our case

$$\|A_k\|_2^2 = \max_{\mathbf{x}:\|\mathbf{x}\|=1} \|A_k \mathbf{x}\|_2^2$$

Given a vector $\mathbf{w}$, $\|\mathbf{w}\|_2^2 = \mathbf{w}^T \mathbf{w}$. So we can rewrite the last term as

$$\begin{aligned}
\|A_k\|_2^2 &= \max_{\mathbf{x}:\|\mathbf{x}\|=1} (A_k \mathbf{x})^T A_k \mathbf{x} \\
&= \max_{\mathbf{x}:\|\mathbf{x}\|=1} \mathbf{x}^T A_k^T A_k \mathbf{x}
\end{aligned}$$

where the last equality follow from a trivial property of linear algebra that we can directly check expanding the terms involved.

Now we can use the following result: if M is a symmetric matrix, $\max_{\mathbf{x}:\|\mathbf{x}\|=1} \mathbf{x}^T M\mathbf{x} = \lambda_1$, where λ_1 is the largest eigenvalue of M. Since $A_k^T A_k$ is clearly symmetric and λ_1 is its largest eigenvalue, we obtain

$$\|A_k\|_2^2 = \lambda_1 = \sigma_1^2$$

4

Lastly, we can repeat the same considerations done previously for $\|A - A_k\|_2^2$. In order to simplify the notation, we'll use $C = A - A_k$, and it's important to recall the decomposition $C = \sum_{i=k+1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T$.

$$\begin{aligned}
\|A - A_k\|_2^2 &= \|C\|_2^2 \\
&= \max_{\mathbf{x}:\|\mathbf{x}\|=1} \|C\mathbf{x}\|_2^2 \\
&= \max_{\mathbf{x}:\|\mathbf{x}\|=1} (C\mathbf{x})^T C\mathbf{x} \\
&= \max_{\mathbf{x}:\|\mathbf{x}\|=1} \mathbf{x}^T C^T C\mathbf{x} \\
&= \sigma_{k+1}^2
\end{aligned}$$

Assignment 3

1

Relying on what we did, we can now prove that $\sigma_k \leq \|A\|_F / \sqrt{k}$. We already know that $\|A\|_F^2 = \sum_{i=1}^{r} \sigma_i^2$. All the elements of the original inequality are greater than zero, hence we can rewrite it and square both sides as follows

$$\begin{aligned}
\|A\|_F &\geq \sigma_k \sqrt{k} \\
\|A\|_F^2 &\geq \sigma_k^2 k \\
\sum_{i=1}^{r} \sigma_i^2 &\geq \sigma_k^2 k
\end{aligned}$$

Now we have to prove the last inequality. Of course

$$\sum_{i=1}^{r} \sigma_i^2 \geq \sum_{i=1}^{k} \sigma_i^2$$

for all $k \leq r$. Since $\sigma_1^2 \geq \sigma_2^2 \geq ... \geq \sigma_r^2$, it holds that $\sigma_{k-1}^2 + \sigma_k^2 \geq \sigma_k^2 + \sigma_k^2 \geq 2\sigma_k^2$. Thus

$$\sum_{i=1}^{k} \sigma_i^2 \geq \sigma_k^2 k$$

and this concludes the proof.

2

The last step is to prove that there exists a matrix B of rank at most k such that $\|A - B\|_2 \leq \|A\|_F / \sqrt{k}$. Setting $B = A_k$, we have proved that $\|A - A_k\|_2^2 = \sigma_{k+1}^2$. Again, we rewrite the inequality

$$\begin{aligned} \|A\|_F &\geq \|A - A_k\|_2 \sqrt{k} \\ \|A\|_F^2 &\geq \|A - A_k\|_2^2 \, k \\ \sum_{i=1}^{r} \sigma_i^2 &\geq \sigma_{k+1}^2 k \end{aligned}$$

It is sufficient to use the previous result, by which $\sum_{i=1}^{r} \sigma_i^2 \geq \sigma_k^2 k$. As noted before, $\sigma_k^2 \geq \sigma_{k+1}^2$. Thus, $\sum_{i=1}^{r} \sigma_i^2 \geq \sigma_{k+1}^2 k$ holds.

Assignment 1

Here's the link to Google Colab notebook:
Image LRA