2 Extra Credit Challenge (5 Points)

Suppose we have a matrix $\mathbf{A} \in \mathbb{R}^{n \times d}$ with SVD $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^T$, where $\mathbf{U} \in \mathbb{R}^{n \times r}$, $\mathbf{D} \in \mathbb{R}^{r \times r}$, $\mathbf{V} \in \mathbb{R}^{d \times r}$.

(a) (1 point) Show that

$$\mathbf{A} = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T \tag{5}$$

proof: 4th property of SVD

$$A = UDV^T = U \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_r \\ & & & \vdots \end{bmatrix} V^T$$

$$= U \left(\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \ddots \end{bmatrix} + \begin{bmatrix} 0 & & \\ & \sigma_2 & \\ & & \ddots \end{bmatrix} + \cdots + \begin{bmatrix} 0 & & & \\ & 0 & & \\ & & \ddots & \\ & & & \sigma_r \end{bmatrix} \right) V^T$$

$$= \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \cdots + \sigma_r u_r v_r^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

(b) (1 point) Show that

$$\mathbf{u}_i = \frac{1}{\sigma_i}\mathbf{A}\mathbf{v}_i \tag{6}$$

In particular, the components of $\mathbf{u}_i$ represent the size of the projection of the rows of $\mathbf{A}$ onto $\mathbf{v}_i$ (scaled by σ_i).

proof: Since $A = UDV^T$ and

$V^TV = I \iff V^T = V^{-1} \iff (V^T)^{-1} = V$

we have $A(V^T)^{-1} = UD$ that is $AV = UD$

in matrix form:

$$A\begin{bmatrix} v_1 & v_2 & \cdots & v_r \end{bmatrix} = \begin{bmatrix} u_1 & u_2 & \cdots & u_r \end{bmatrix}\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}$$

compute matrix, we have.

$$\begin{cases} Av_1 = u_1\sigma_1 = \sigma_1 u_1 \\ Av_2 = u_2\sigma_2 = \sigma_2 u_2 \\ \vdots \\ Av_r = u_r\sigma_r = \sigma_r u_r \end{cases}, \text{ that is } \begin{cases} u_1 = \frac{1}{\sigma_1}Av_1 \\ u_2 = \frac{1}{\sigma_2}Av_2 \\ \vdots \\ u_r = \frac{1}{\sigma_r}Av_r \end{cases}$$

$$\Rightarrow u_i = \frac{1}{\sigma_i}Av_i \quad (1 \leq i \leq r)$$

(c) (1 point) One way of finding a reduced rank approximation of $\mathbf{A}$ is by hard-setting all but the k largest σ_i to 0. This approximation is called the truncated SVD, and by (a) we see it can be written as

$$\mathbf{A}_k := \sum_{i=1}^{k} \sigma_i \mathbf{u}_i \mathbf{v}_i^T \tag{7}$$

From (a), we see the truncated SVD can also be written as $\mathbf{A}_k = \mathbf{U}_k \mathbf{D}_k \mathbf{V}_k^T$, where $\mathbf{U}_k \in \mathbb{R}^{n\times k}$, $\mathbf{V} \in \mathbb{R}^{d\times k}$ are the first k columns of $\mathbf{U}, \mathbf{V}$, and $\mathbf{D} \in \mathbb{R}^{k\times k}$ has the first k singular values.

Show that the rows of $\mathbf{A}_k$ are the projections of the rows of $\mathbf{A}$ onto the subspace of V_k spanned by the first k right singular vectors.

Hint: Recall that the projection of a vector $\mathbf{a}$ onto a subspace spanned by $\mathbf{v}_1, \ldots, \mathbf{v}_k$ where the $\mathbf{v}_i$ are pairwise orthogonal is given by the sum of projections of $\mathbf{a}$ onto the individual $\mathbf{v}_i$.

The projection of the rows of A onto subspace of U_k spanned by the first K right svectors, can be written as:

$\begin{bmatrix} a_1 \\ | \\ a_n \end{bmatrix}$ onto subspace of V_K spanned by $(v_1, -, v_K)$.

From Hint, we know that the projection of each row of A onto a subspace spanned by $v_1, -, v_K$ is given by the sum of projection of each row of A on to the individual V_i, that is

$\sum_{i=1}^{K} A v_i v_i^T$ (definition from 1.2 (2))

From part (b) we have $A v_i = \sigma_i u_i$

so that $\sum_{i=1}^{K} A v_i v_i^T = \sum_{i=1}^{K} \sigma_i u_i v_i^T = A_K$.

(d) (2 points) The Frobenius norm of a matrix $\mathbf{M} \in \mathbb{R}^{n \times m}$ is defined as

$$||\mathbf{M}||_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{m} M_{ij}^2} \tag{8}$$

Show that

$$\mathbf{A}_k = \underset{\text{rank}(\mathbf{B})=k}{\arg\min} ||\mathbf{A} - \mathbf{B}||_F \tag{9}$$

where the arg min is taken over matrices of rank k.

Hint: Use the fact that V_k is the best-fit k-dimensional subspace for the rows of $\mathbf{A}$.

The key is use L_2 norm derive Frobenious.
L_2 — vector, Frobenius — matrix.

$$\underset{\text{rank}(B)=k}{\arg} ||A-B||_F = \arg\min \sum_{i=1}^{n}\sum_{j=1}^{m} (A-B)_{ij}$$

$$= \arg\min \sum_{i=1}^{n} \left(\sum_{j=1}^{m} (A-B)_{ij} \right)$$

$$= \arg\min \sum_{i=1}^{n} || (A-B)_{i\cdot} ||_2^2$$

$$= \sum_{i=1}^{n} \arg\min || (A_i - B_i) ||_2^2$$

By Hint and part (c).

$$= \sum_{i=1}^{n} V_k. \quad = A_k.$$