Singular Value Decomposition (SVD)

Consider a Matrix $A \in \mathbb{R}^{m \times n}$
We can decompose the matrix into U, Σ and V such that

$$A = U\Sigma V^T$$

where U and V are orthogonal matrices
($UU^T = I$ and $VV^T = I$)

And Σ has only Diagonal Entries

$$\Sigma_{m\times n} = \begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & \sigma_R & 0 & \dots \\ 0 & 0 & \dots & 0 & 0 & \vdots \\ \vdots & \vdots & \dots & \vdots & \dots & \ddots \end{bmatrix}_{m\times n}$$

The diagonal entries are the eigen values of the matrix A arranged in descending order
$\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \geq \sigma_R$
where R is the rank of the matrix A

It can be seen that

$$\begin{aligned} A &= \sum_{i=1}^{R} \sigma_i \mathbf{u_i} \mathbf{v_i^T} \\ A &= \sigma_1 \mathbf{u_1} \mathbf{v_1^T} + \sigma_2 \mathbf{u_2} \mathbf{v_2^T} + \dots + \sigma_R \mathbf{u_R} \mathbf{v_R^T} \end{aligned}$$

where $\mathbf{u_i}$ and $\mathbf{v_i}$ are the i^{th} columns of matrices U and V respectively

Low Rank Approximation

Consider the following problem.
$A \in \mathbb{R}^{m\times n},\ \text{rank}(A) = R$
Find $A_k \in \mathbb{R}^{m\times n},\ \text{rank}(A_k) = k$ such that
$||A - A_k||$ is minimized over some norm

As an optimization problem, we find A_k (for some norm)

$$\begin{aligned} \min_{A_k} \quad & ||A - A_k|| \\ \text{such that} \quad & \text{rank}(A_k) \leq k \end{aligned}$$

SVD as a low rank Approximation

Eckart–Young–Mirsky Theorem:

Consider $A \in \mathbb{R}^{m \times n}$, $\text{rank}(A) = R$

Suppose $A = U \Sigma V$ where $UU^T = I_m$, $VV^T = I_n$
and $\Sigma_{R \times R}$ is a Diagonal Matrix containing the non-zero singular values of A in a non-decreasing order.

The, the solution of the Optimization Problem,

$$\begin{aligned} \text{minimize over } A_k \qquad ||A - A_k|| \qquad & \text{subject to } \text{rank}(A_k) \leq k \\ \text{is given by } A_k \ = \sum_{i=1}^{k} \sigma_i \mathbf{u_i} \mathbf{v_i^T} & \end{aligned}$$

where $\mathbf{u_i}$ and $\mathbf{v_i}$ are the i^{th} columns of U and V

(Math Stack Exchange) [Link]

Proof : (For Spectral Norm)

$$\begin{aligned} A_k \quad &= \quad \sum_{i=1}^{k} \sigma_i \mathbf{u_i} \mathbf{v_i^T} \\ \implies \quad ||A - A_k|| \quad &= \quad \left|\left| \sum_{i=k+1}^{R} \sigma_i \mathbf{u_i} \mathbf{v_i^T} \right|\right|_2 \quad = \quad \sigma_{k+1} \end{aligned}$$

Consider some Matrix B such that $\text{rank}(B) \leq k$
$\text{rank}(B) \leq k \implies \dim \mathcal{N}(B) \geq n - k$
Suppose $V_{k+1} = [v_1 \ v_2 \ \dots \ v_{k+1}]$ and $\mathcal{R}(V_{k+1})$ be the range space of V_k
$\dim \mathcal{N}(B) + \dim \mathcal{R}(V_{k+1}) \geq n - k + k + 1 = n + 1$

$$\implies \qquad \exists\, x \in \mathcal{N}(B) \cap \mathcal{R}(V_{k+1}) \qquad \|x\|_2 = 1$$

$$\implies \quad x = \gamma_1 v_1 \;+\; \gamma_2 v_2 \;+\; \ldots + \gamma_{k+1} v_{k+1}, \quad \sum_{i=1}^{k+1} \gamma_i^2 = 1$$

By definition,

$$\begin{aligned} \|A-B\|_2^2 \, \|x\|_2^2 \quad &\geq \quad \|(A-B)x\|_2^2 \\ \implies \quad \|A-B\|_2^2 \quad &\geq \quad \|(A-B)x\|_2^2 \end{aligned}$$

We know that $Bx = 0$ since $x \in \mathcal{N}(B)$

$$\implies \; \|(A-B)x\|_2^2 \;\geq\; \|Ax\|_2^2 \;=\; \sum_{i=1}^{k+1} \sigma_i^2 \gamma_i^2 \;\geq\; \sigma_{k+1}^2$$

This implies

$$\|A-B\|_2^2 \;\geq\; \sigma_{k+1}^2 \quad \forall\, B$$

In other words,

$$\|A-B\|_2 \;\geq\; \|A-A_k\|_2 \quad \forall\, B$$

Hence Proved.

Proof : (For Frobenius Norm)

$$||A - A_k||_F = ||U\Sigma V^T - A_k||_F = ||\Sigma - U^T A_k V||_F$$

Suppose $N = U^T A_k V$

$$||\Sigma - N||_F^2 = \sum_{i,\,j} |\Sigma_{ij} - N_{ij}|^2 = \sum_{i=1}^{k} |\sigma_i - N_{ii}|^2 + \sum_{i>k} |N_{ii}|^2 + \sum_{i \neq j} |N_{ij}|^2$$

The minimal is achieved when $N_{ii} = \sigma_i \;\forall\, i \leq k$ AND
All the non-diagonal terms of N and
all the diagonals with $i > k$ are equal to 0

This implies

$$A_k \;=\; \sum_{i=1}^{k} \sigma_i \mathbf{u_i} \mathbf{v_i^T}$$

Hence Proved.