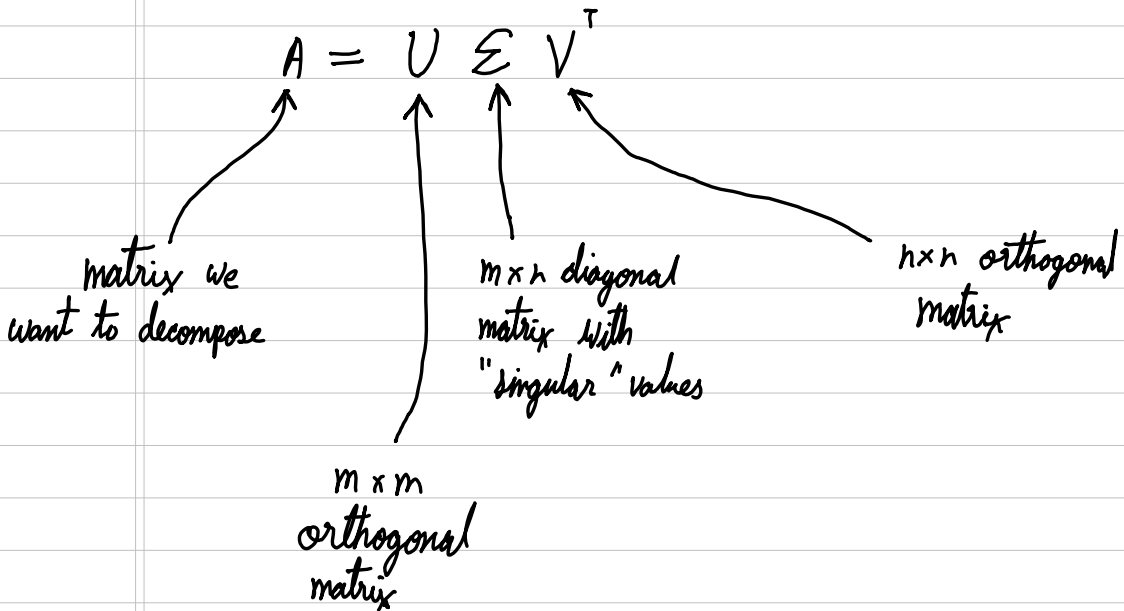


Singular Value Decomposition for Dimensionality Reduction

SVD is a fundamental matrix factorization technique in linear algebra

Any matrix of size $m \times n$ can be split into three matrices



SVD has many applications in data compression, numeric computing and mathematics

$$A = U \Sigma V$$

U_1	U_2	U_3	U_4

σ_1	0	0	0
0	σ_2	0	0
0	0	σ_3	0
0	0	0	σ_4

V_1
V_2
V_3
V_4

$$= \boxed{\sigma_1} \begin{bmatrix} U_1 \end{bmatrix} \boxed{V_1} + \boxed{\sigma_2} \begin{bmatrix} U_2 \end{bmatrix} \boxed{V_2} + \dots$$

the σ are already sorted in ascending value
so

$\sigma_1 \rightarrow$ high
 $\sigma_2 \rightarrow$ medium
 $\sigma_3 \rightarrow$ low
 \vdots
 \vdots

\therefore first term \rightarrow more important
 second term \rightarrow lesser imp.

so we can consider the first k values from m
and the best values with maximum importance

Useful for data compression

PCA using SVD steps

1. Center the data $X_c = X - \bar{X}$
2. Compute SVD of centered data

$$X_c = V \Sigma V^T$$

3. Compute principle components

principle components are given by V

principle scores $Z = X_c \cdot V$ is the transformed data into principle component space

Dimensions can be reduced by selecting k components.

SVD makes PCA easy and efficient when compared to eigenvalue decomposition

SVD is more stable and reliable than eigenvalue decomposition.

For large datasets SVD is faster than computing the full covariance matrix.