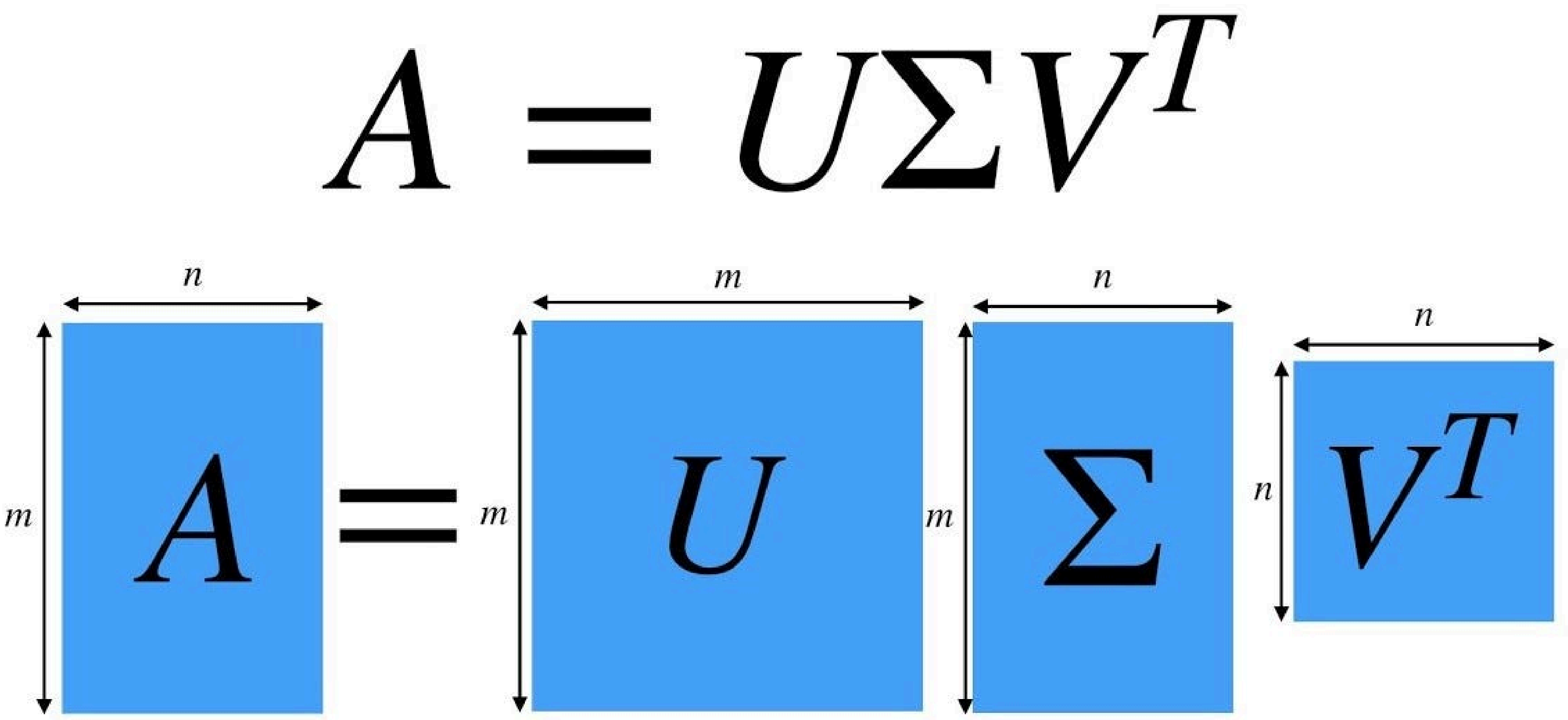
Singular Value Decomposition (SVD)

Singular Value Decomposition is a factorization (breaking down) method in linear algebra, from it's name it decomposes a given matrix into **three other matrices** which gives a way to represent data in terms of **singular values**

Given a **data matrix** A :

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

- U and $V^\intercal$ are both unitary matrices (**Orthogonal**), which means $U^\intercal U = UU^\intercal = \mathbb{I}$
- Σ is a **Diagonal matrix** and hierarchically ordered $\sigma_1 > \sigma_2 \cdots > \sigma_m \geq_0$ which is by importance



$$U = \begin{bmatrix} u_1 & u_2 & \dots & u_n \\ \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

- The columns of U have the same **shape** as the columns in our **data matrix** A, also called the left singular matrix

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_m \end{bmatrix}$$

- Σ called matrix of singular values

$$V^\intercal = \begin{bmatrix} \vdots & \vdots & & \vdots \\ v_1 & v_2 & \dots & v_m \\ \vdots & \vdots & & \vdots \end{bmatrix}^\intercal$$

- $V^\intercal$ also known as the right singular matrix

Interpretation of SVD

The key to understand the **SVD** is viewing the columns of U, Σ and V as representing **concepts** that are hidden in the original matrix A, and in the process of decomposition there will be columns in U and V that correspond to the smallest singular values and in order to get the best approximation of the original matrix, we use the **Reduced SVD** which :

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

- Most of those columns in these decomposed matrices are useless since the one that effect the most comes first (columns are ordered in terms of importance)
- And they will carry zero values on them

Economy SVD eliminate them by taking the first r columns which results in

$$A = \hat{U}\hat{\Sigma}V^\intercal$$

- The $\hat{U}, \hat{\Sigma}$ only took the first m relevant columns getting rid of the **zeros**

Truncated rank SVD this is dimensionality reduction, where it drops the **smallest singular values**

$$A \approx U_r \Sigma_r V_r^\intercal$$

- This saves memory and speeds up the computations for the **SVD**
- It's the backbone of dimensionality reduction methods since it get rid of least relevant columns and only keeps the important columns which reduce the dimensions of the approximated full matrix A

SVD Eigenvectors & Eigenvalues interpretation

The SVD has a nice interpretation with the Eigenvectors & Eigenvalues and the correlation matrix $A^\intercal A$ and $AA^\intercal$:

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

Gives us the **correlation matrix** :

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

With $U_r^T U_r = \mathbb{I}$ the identity matrix

$$A^\intercal A = V\Sigma^2 V^\intercal$$

Multiplying both sides with V :

$$A^\intercal A V = V\Sigma^2$$

This reads as follows :

- The transformation matrix $A^\intercal A$ got V as an **Eigenvectors** matrix
- With Σ_r^2 being the **Eigenvalues** that scales the V_r columns on their span

and calculating $AA^\intercal$ results in :

$$AA^\intercal U = U\Sigma^2$$

- Both of U and V are **Eigenvector** for the correlation matrix with the same **Eigenvalues** Σ^2
- The Σ matrix is the **Square root** of the **Eigenvalues** of both the left and right matrix U, V