Obs. Seja $r = n \leq m$. Então a matriz

$$C = A^T A \text{ é S.P.D.}$$

Denotando os autovalores de C, em ordem decrescente como λ_i^C, $i = 1, \ldots, n$ então note que:

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

já que a matriz U é ortogonal.

Dessa forma, a matriz $\Sigma^T \Sigma$ é diagonal com elementos $\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2$, e V^T é uma matriz que garante uma "transformação semelhante" com C. Além disso,

$$\sigma_i = \sqrt{\lambda_i^C} \quad \text{e} \quad \|A\|_2 = \sigma_1$$

Teorema (Melhor aproximação de posto inferior de uma matriz)
A melhor aproximação de posto r, denotada como A_r, de uma matriz A, que admite a decomposição

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

e a relação

~~seguinte~~ $\|A - A_r\|_2 = \cancel{\min \|A - A\|} = \sigma_{r+1}$ é minimizada, é a matriz

$$A_r = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

onde u_i e v_i são os i-ésimos vetores colunas de U e V, respectivamente.

Obs. Esse Teorema é aplicado para a construção de uma decomposição SVD "truncada", que faz a "redução dimensional" de um determinado problema.