Singular value decomposition (SVD) generalizes matrix diagonalization to non-square matrices. It is powerful because it applies to **any** $m \times n$ matrix.

First we give the recipe:

$$A = U\Sigma V^T. \tag{SVD}$$

Compute AA^T. This product is guaranteed to have m linearly independent eigenvectors by the spectral theorem. Collate those eigenvectors into the columns of a matrix. That matrix is U.

Similarly, compute $A^T A$, which is guaranteed to have n linearly independent eigenvectors that you can stick into V.

Next compute the non-zero eigenvalues λ_i of $A^T A$, take their square roots $\sigma_i = \sqrt{\lambda_i}$, and stick them into the diagonal of Σ.

Now Σ will consist of $\text{rank}(A)$ σ_i's on the diagonal and zeroes elsewhere. Add more zeroes on the diagonal to form a $\min(m, n) \times \min(m, n)$ square matrix. Then add the sufficient amount of rows/columns to form a $m \times n$ matrix - obviously the rows/columns consist entirely of zeroes.

We are almost done. Conventionally, the singular values in Σ should be ordered from greatest to least. This orderering of singular values should correspond to the ordering of eigenvectors in U and V (i.e. $Av_i = \sigma_i u_i$).

Next we look at the intuition.