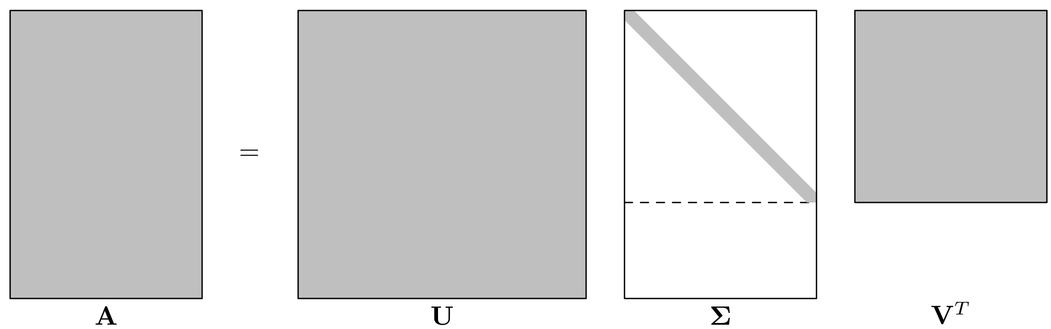
SVD review

Let $\mathbf{A}$ be a rank-ρ matrix in $\mathbb{R}^{m \times n}$ with $m \geq n$. Recall that the *full* SVD of $\mathbf{A}$ takes the form $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, where $\mathbf{U}$ is an $m \times m$ orthonormal matrix (i.e., the columns of $\mathbf{U}$ have unit length and are mutually orthogonal; more concisely, $\mathbf{U}^T\mathbf{U} = \mathbf{I}_m$), $\mathbf{V}$ is an $n \times n$ orthonormal matrix, and $\mathbf{\Sigma}$ is an $m \times n$ diagonal matrix that has nonnegative entries. The columns of $\mathbf{U}$ and $\mathbf{V}$ are called, respectively, the left and right singular vectors of $\mathbf{A}$, and the diagonal entries of $\mathbf{\Sigma}$ are called the singular values of $\mathbf{A}$. In particular, $\mathbf{A}$ has m left singular vectors and n singular values and right singular vectors.



We decompose $\mathbf{U}$ as

$$\mathbf{U} = \begin{bmatrix}\mathbf{u}_1 & \ldots & \mathbf{u}_m\end{bmatrix} = \begin{bmatrix}\mathbf{U}_\rho & \mathbf{U}_\rho^\perp\end{bmatrix},$$

so that $\mathbf{u}_i$ denotes the ith left singular vector of $\mathbf{A}$, and the first ρ left singular vectors of $\mathbf{A}$ constitute the matrix $\mathbf{U}_\rho$, while the remaining left singular vectors constitute $\mathbf{U}_\rho^\perp$. Note that $\mathbf{U}_\rho^T\mathbf{U}_\rho^\perp = \mathbf{0}$. We similarly decompose the matrix of right singular vectors as

$$\mathbf{V} = \begin{bmatrix}\mathbf{v}_1 & \ldots & \mathbf{v}_n\end{bmatrix} = \begin{bmatrix}\mathbf{V}_\rho & \mathbf{V}_\rho^\perp\end{bmatrix},$$

and the matrix of singular values as

$$\mathbf{\Sigma} = \begin{bmatrix}\mathbf{\Sigma}_\rho \\ \mathbf{0}_{m-\rho \times n}\end{bmatrix}.$$

Using this notation, the full SVD of $\mathbf{A}$ has the decomposition

$$\mathbf{A} = \begin{bmatrix}\mathbf{U}_\rho & \mathbf{U}_\rho^\perp\end{bmatrix} \begin{bmatrix}\mathbf{\Sigma}_\rho \\ \mathbf{0}_{m-\rho \times n}\end{bmatrix} \begin{bmatrix}\mathbf{V}_\rho^T \\ (\mathbf{V}_\rho^\perp)^T\end{bmatrix}. \tag{1}$$

The full SVD is useful because in the decomposition $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, the matrices $\mathbf{U}$ and $\mathbf{V}$ are orthonormal, so are invertible, and preserve Euclidean norms of vectors. It also lets you immediately read off orthogonal bases for the four fundamental subspaces associated with $\mathbf{A}$: the kernel/null space (has basis $\mathbf{V}_\rho^\perp$), the column space (has basis $\mathbf{U}_\rho$), the row space (has basis $\mathbf{V}_\rho$), and the cokernel (i.e. the set of vectors so that $\mathbf{x}^T\mathbf{A} = \mathbf{0}$, equivalently the kernel of $\mathbf{A}^T$; this has basis $\mathbf{U}_\rho^\perp$).

However, as you can check by multiplying out equation (1), we can also write $\mathbf{A} = \mathbf{U}_\rho\mathbf{\Sigma}_\rho\mathbf{V}_\rho^T$. This is called the *reduced* SVD, and is a more condensed factorization that is very useful in practice. Now $\mathbf{U}_\rho$ and $\mathbf{V}_\rho$ only contain the singular vectors corresponding to the nonzero singular values of $\mathbf{A}$. Note that if $\mathbf{A}$ is an invertible matrix then the reduced SVD and full SVD are identical.