Singular value decomposition (SVD)

Zecheng Zhang

September 9, 2022

This section will discuss singular value decomposition (SVD) of a matrix $A \in \mathbb{R}^{m\times n}$. Geometric understanding of SVD (in two-dimensional space): A maps a unit circle to an ellipse. Singular values are the length of the principle semiaxes of the ellipse. the left singular vectors are unit vectors oriented in the direction of the principle semiaxes.

1 Construction

The first singular value is defined as:

$$\sigma_1 = \sup_{\|v\|=1} \|Av\|.$$

Remark 1. The first singular value is well defined, i.e., such a $v_1 \in \mathbb{R}^n$ always exists. Non-rigorous argument: the function : $v \to \|Av\|$ is continuous and with a compact domain.

Now one can find $u_1 \in \mathbb{R}^m$ with $\|u_1\| = 1$ such that $Av_1 = \sigma_1 u_1$.

One can follow the definition of the first singular value and define the second singular value as,

$$\sigma_2 = \sup_{\|v\|=1, v\perp v_1} \|Av\|.$$

The remark 1 implies that such a v_2 always exists and let us denote it as v_2. In addition, we can find $u_2 \in \mathbb{R}^m$ with $\|u_2\| = 1$ such that $Av_2 = \sigma_2 u_2$.

Remark 2. $\sigma_2 \leq \sigma_1$ because v_2 is taken from a smaller subspace $\{v_1\}^\perp \subset \mathbb{R}^n$.

Theorem 1.1. u_1 and u_2 which are defined above are orthogonal.

The theorem implies that $u_1 \perp u_2$. Repeat the process, one can find a unit vector $v_3 \in W_2 = \{v_1, v_2\}^\perp$ such that it admits

$$\sigma_3 = \sup_{\|v\|=1, v\in V_2} \|Av\|.$$

In addition, one can find a unit vector u_3 such that $Av_3 = \sigma_3 u_3$. One can show that $\{u_1, u_2, u_3\}$ are orthogonal.

Remark 3. Let us define $W_p = \{v_1, v_2, ..., v_p\}^\perp$. If $\sup_{v\in W_p} \|Av\| = 0$, or, $Av_{p+1} = 0$, we can make u_{p+1} (nonzero if possible) to be any vector which is orthogonal to $\{u_1, ..., u_p\}$. If u_{p+1} has to be zero, $span\{u_1, ..., u_p\} = \mathbb{R}^m$

Theorem 1.2. $rank(A)$ equals to the number of nonzero singular values.

Proof. Let us assume $\{\sigma_1, ..., \sigma_p\}$ are all nonzero but $\sigma_{p+1} = 0$. Let $V_p = \{u_1, ..., u_p\}$ be the singular vector corresponding to $\sigma_1, ...\sigma_p$. We claim that $V_p \subset row(A)$. For $v_i \in V_p$, we have,

$$\begin{aligned}
&Av_i = \sigma_i u_i \\
\Rightarrow &u_i^t A v_i v_i^t = \sigma_i u_i^t u_i v_i^t \\
\Rightarrow &\frac{1}{\sigma_i} u_i^t A = v_i^t.
\end{aligned}$$

This implies that $v_i \in row(A)$, the claim is proved. By theorem in the last section, $null(A) = row(A)^\perp \subset V_p^\perp$. Now, for $v \in V_p^\perp$, we have $Av = 0$, otherwise contradicts with the definition of V_p. As a result, $V_p^\perp \subset null(A)$, i.e., $V_p^\perp = null(A)$. By the rank theorem, $dim(V_p) = rank(A)$. $\square$

Repeat the process for n times, we then can construct an orthonormal matrix $V = [v_1, ..., v_n] \in \mathbb{R}^{n\times n}$, another matrix with orthonormal columns $U = [u_1, ..., u_n] \in \mathbb{R}^{m\times n}$ up to some 0 columns, and a diagonal matrix $\Sigma \in \mathbb{R}^{n\times n}$ with diagonal entries being $\sigma_1, ..., \sigma_n$. Recall the matrix multiplication, we have,

$$AV = U\Sigma.$$

Full SVD: make U matrix orthonormal when $m > n$. One can append an additional $m - n$ orthonormal columns to fulfill this goal. Σ should change as well so that the product $AV = U\Sigma$ still holds. To do this, one can append $m - n$ zero rows to the bottom of Σ. As a result, we have $AV = U\Sigma$ where $V \in \mathbb{R}^{n\times n}$, $U \in \mathbb{R}^{m\times m}$ and $\Sigma \in \mathbb{R}^{m\times n}$. Since V is orthonormal, we have:

$$A = U\Sigma V^{-1}.$$

2 Symmetric matrix

Let S be a real $n \times n$ matrix. S is symmetric if $S = S^t$.

Theorem 2.1. All eigenvalues of S are real.

Theorem 2.2. S has n linearly independent eigenvectors.

Theorem 2.3 (Spetral theorem)**.** If S is real symmetric, then $S = QDQ^t$ for Q orthogonal and D diagonal.