Topics on Linear Algebra

Based on MIT 18.06sc and 18.065

Tianyu Du

June 12, 2020

1 Projection onto Subspaces

2 Singular Value Decomposition

Decomposition Let $A \in \mathbb{R}^{m\times n}$, suppose $m > n$, then A can be written as

$$A = U\Sigma V^T \tag{1}$$

where U is a $m \times m$ orthonormal matrix with **left singular vectors** as its columns, Σ is a $m \times n$ orthonormal matrix with **singular values** on its diagonal, and V is a $n \times n$ matrix with **right singular vectors** as its columns. Note that Σ is constructed by stacking a $n \times n$ diagonal matrix $diag(\sigma_1, \sigma_2, \cdots, \sigma_n)$ with a zero matrix of size $(m - n) \times n$.

$$U = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_m] \tag{2}$$

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \sigma_n \\ 0 & \cdots & \ddots & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \tag{3}$$

$$V = [\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_m] \tag{4}$$

Singular Values and Singular Vectors Like solving $A\mathbf{x} = \lambda\mathbf{x}$ for eigenvalues/vectors, here we wish to identify $r = rank(A)$ triples of $(\mathbf{v}_i, \sigma_i, \mathbf{u}_i)$ such that $(\mathbf{v}_i)$ and $(\mathbf{u}_i)$ are orthonormal. Moreover, these singular values/vectors need to satisfy

$$A\mathbf{v}_i = \sigma_i \mathbf{u}_i \quad \forall i \in \{1, 2, \cdots, r\} \tag{5}$$

$$A\mathbf{v}_j = 0 \quad \forall j \in \{r+1, r+2, \cdots, n\} \quad (\dagger) \tag{6}$$

Finding Singular Values and Vectors Suppose $A = U\Sigma V^T$,

$$A^T A = (U\Sigma V^T)^T U\Sigma V^T \tag{7}$$

$$= V\Sigma^T U^T U\Sigma V^T \tag{8}$$

$$= V\Sigma^T \Sigma V^T \tag{9}$$

$$= V diag(\sigma_1^2, \sigma_2^2, \cdots, \sigma_n^2) V^T \tag{10}$$

Because $A^T A$ is symmetric and positive semidefinite, all of it's eigenvalues are non-negative. Moreover, $A^T A$ admits the eigenvalue decomposition $Q\Lambda Q^T$. Therefore, $V = Q$ and $\sigma_i = \sqrt{\lambda_i}$.

Similarly, $AA^T = U\Sigma\Sigma^T U^T$, therefore, U consists of eigenvectors of AA^T.

Note that $rank(A^T A) = rank(A) = r$, $A^T A \in \mathbb{R}^{n\times n}$ has $n - r$ eigenvectors corresponding to $\lambda = 0$. Let $\{\mathbf{v}_1, \cdots, \mathbf{v}_r\}$ denote eigenvectors of $A^T A$ with $\lambda > 0$, and $\{\mathbf{v}_{r+1}, \cdots, \mathbf{v}_n\}$ are eigenvectors with zero eigenvalues.

Similarly, let $\{\mathbf{u}_1, \cdots, \mathbf{u}_r\}$ and $\{\mathbf{u}_{r+1}, \cdots, \mathbf{u}_m\}$ denote eigenvectors of AA^T corresponding to positive and zero eigenvalues.

As a result, the representation in (†) can be written as (p.s. probably this argument only works for $rank(A) = n$?)

$$A[\mathbf{v}_1, \cdots, \mathbf{v}_r, \cdots, \mathbf{v}_n] = [\mathbf{u}_1, \cdots, \mathbf{u}_r, \cdots, \mathbf{u}_n, \cdots, \mathbf{u}_m] \begin{bmatrix} \sigma_1 & \cdots & 0 \\ 0 & \ddots & 0 \\ 0 & \cdots & \sigma_n \\ 0 & \cdots & 0 \end{bmatrix} \tag{11}$$

$$\implies AV = U\Sigma \tag{12}$$

$$\implies AVV^T = U\Sigma V^T \tag{13}$$

$$\implies A = U\Sigma V^T \tag{14}$$

Which gives us the singular value decomposition of A.

Geometric Interpretation The singular value decomposition implies any linear transformation A can be decomposed into a sequence of rotation, stretching, and rotation operations.

3 Graph Clustering