Linear Algebra II
Singular Value Decomposition

Ming-Hsuan Kang

Introduction

Singular Value Decomposition (SVD) is a useful factorization for non-square matrix. It is closely related to best-fit subspaces.

Singular Values and Left Singular Vectors

Let A be an $m \times n$ real matrix. Recall that the i-th singular value of A is the i-th largest eigenvalue of A^tA and a subset $\{\vec{v}_1, \cdots, \vec{v}_n\}$ of $\mathbb{R}^n$ forms a set of left singular vectors if it is an corresponding orthonormal eigenbasis of A^tA.

Singular Values of A^t

Lemma

If $\vec{v}$ is an eigenvector of A^tA corresponding the nonzero eigenvalue λ, then $A\vec{v}$ is an eigenvector of AA^t corresponding the eigenvalue λ.

Proof: Since $A^tA\vec{v} = \lambda\vec{v} \neq \vec{0}$, we have $A\vec{v} \neq \vec{0}$. On the other hand, $AA^t(A\vec{v}) = A(A^tA)\vec{v} = A(\lambda\vec{v}) = \lambda A\vec{v}$.

Theorem

The $m \times n$ matrix A and the $n \times m$ matrix A^t have the same set of non-zero singular value (which size is equal to their ranks r).

Recall that A^tA, AA^t, and A all have the same rank. Therefore, the first r singular values of A and A^t are positive and the rest of them equal to zero.

Right Singular Vectors

Theorem

Let $\vec{v}_i$ be the i-th left singular vector of a $m \times n$ matrix A of rank r and λ_i be the i-th singular value of A. Then $\{\vec{u}_1, \cdots, \vec{u}_r\} = \{\frac{1}{\sqrt{\lambda_1}}A\vec{v}_1, \cdots, \frac{1}{\sqrt{\lambda_r}}A\vec{v}_r\}$ is an orthonormal set. Moreover, if we extend this set to an orthonormal basis of $\mathbb{R}^m$, namely $\{\vec{u}_1, \cdots, \vec{u}_r, \vec{u}_{r+1}, \cdots \vec{u}_m\}$. Then the basis forms a set of left singular vectors of A^t.

The set $\{\vec{u}_1, \cdots \vec{u}_m\}$ in the above theorem is called a set of right singular vectors of A induced from $\{v_1, \cdots, v_n\}$.

Proof

From the previous lemma, each of $\{\vec{u}_1, \cdots, \vec{u}_r\}$ is an eigenvector of AA^t corresponding to the eigenvalue λ_i. For $1 \le i, j \le r$,

$$\vec{u}_i^t \vec{u}_j = \frac{1}{\sqrt{\lambda_i \lambda_j}} \vec{v}_i^t A^t A \vec{v}_j = \frac{\lambda_j}{\sqrt{\lambda_i \lambda_j}} \vec{v}_i^t \vec{v}_j = \delta_{ij}.$$

Therefore, $\{\vec{u}_1, \cdots, \vec{u}_r\}$ is an orthonormal set and $\vec{u}_i$ is the i-th left singular vector of A^t. It remains to show that $\vec{u}_{r+1}, \cdots \vec{u}_m$ are all eigenvectors of AA^t corresponding to the zero eigenvalue. Since AA^t is symmetric, all eigenspaces are orthogonal. Especially, its zero eigenspace is orthogonal to the subspace spanned by the first r singular vectors. Therefore, all of $\{\vec{u}_{r+1}, \cdots \vec{u}_m\}$ are contained in the zero eigenspace.

Rectangular Diagonal Matrix

Under the notation in the previous theorem, let $\alpha = \{\vec{v}_1, \cdots, \vec{v}_n\}$ and $\beta = \{\vec{u}_1, \cdots, \vec{u}_m\}$. Then $A(\vec{v}_i) = \sqrt{\lambda_i}\vec{u}_i$ when $i \leq r$ and $A(\vec{v}_i) = \vec{0}$ when $i > r$. Therefore,

$$\mathrm{Rep}_{\alpha,\beta}(L_A) = \begin{pmatrix} \sqrt{\lambda_1} & 0 & 0 & \cdots \\ 0 & \sqrt{\lambda_2} & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}$$

which is a rectangular diagonal matrix, denoted by Σ. Set $V = (\vec{v}_1 \ \cdots \ \vec{v}_n)$ and $U = (\vec{u}_1 \ \cdots \ \vec{u}_m)$. Then we have

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

Singular Value Decomposition

Theorem (SVD for real matrices)

Let A be a real matrix of size $m \times n$. There exists an orthogonal matrix V of size n, an orthogonal matrix U of size m, a rectangular diagonal matrix Σ with non-negative entries of size $m \times n$ such that $A = U\Sigma V^t$.

More precisely, diagonal entries of Σ are square roots of singular values, columns of V are left singular vectors, columns of U are right singular vectors induced from V.

Theorem (SVD for complex matrices)

Let A be a complex matrix of size $m \times n$. There exists a unitary matrix V of size n, a unitary matrix U of size m, a rectangular diagonal matrix Σ with non-negative entries of size $m \times n$ such that $A = U\Sigma V^*$.

In fact, SVD is a very nature decomposition. For an $m \times n$ matrix A, there are two square matrices associated to A, namely AA^t and A^tA. These two matrices are both positive semi-definite and they have the same set of positive eigenvalues. The SVD of A just consists of an orthonormal eigenbasis of A^tA, an orthonormal eigenbasis of AA^t, and the square roots of those positive eigenvalues.

Example

Let $A = \begin{pmatrix} -3 & 4 & 2 \\ 2 & -2 & 0 \end{pmatrix}$. Then $A^tA = \begin{pmatrix} 13 & -16 & -6 \\ -16 & 20 & 8 \\ -6 & 8 & 4 \end{pmatrix}$ which has eigenvectors $\frac{1}{3\sqrt{5}}\begin{pmatrix} 4 \\ -5 \\ -2 \end{pmatrix}$, $\frac{1}{3}\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$, and $\frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}$ corresponding the eigenvalues 36, 1, and 0 respectively. Next, we have

$$A\begin{pmatrix} 4 \\ -5 \\ -2 \end{pmatrix} = \begin{pmatrix} -36 \\ 18 \end{pmatrix} \quad \text{and} \quad A\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Finally, we obtain the SVD of A as

$$A = \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 6 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \frac{4}{3\sqrt{5}} & -\frac{\sqrt{5}}{3} & -\frac{2}{3\sqrt{5}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{2}{\sqrt{5}} \\ -\frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{pmatrix}$$

Remark. We also have

$$A = \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \frac{4}{3\sqrt{5}} & -\frac{\sqrt{5}}{3} & -\frac{2}{3\sqrt{5}} \\ \frac{1}{\sqrt{5}} & 0 & \frac{2}{\sqrt{5}} \end{pmatrix}.$$

Example

Let $A = \begin{pmatrix} -3 & 2 \\ 4 & -2 \\ 2 & 0 \end{pmatrix}$. Then $A^tA = \begin{pmatrix} 29 & -14 \\ -14 & 8 \end{pmatrix}$ which has eigenvectors $\frac{1}{\sqrt{5}}\begin{pmatrix} -2 \\ 1 \end{pmatrix}$ and $\frac{1}{\sqrt{5}}\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ corresponding the eigenvalues 36 and 1 respectively. Next, we have

$$A\begin{pmatrix} -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 8 \\ -10 \\ -4 \end{pmatrix} \quad \text{and} \quad A\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}.$$

Third, we find a vector orthogonal to the above two vectors, namely, $\begin{pmatrix} 2 \\ 2 \\ -1 \end{pmatrix}$. Finally, we obtain the SVD of A as

$$A = \begin{pmatrix} \frac{4}{3\sqrt{5}} & \frac{1}{\sqrt{5}} & \frac{2}{3} \\ -\frac{\sqrt{5}}{3} & 0 & \frac{2}{3} \\ -\frac{2}{3\sqrt{5}} & \frac{2}{\sqrt{5}} & -\frac{1}{3} \end{pmatrix} \begin{pmatrix} 6 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}$$

Remark. We also have

$$A = \begin{pmatrix} \frac{4}{3\sqrt{5}} & \frac{1}{\sqrt{5}} \\ -\frac{\sqrt{5}}{3} & 0 \\ -\frac{2}{3\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 6 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{pmatrix}.$$

Rank-one Matrix Factorization

For the SVD of A, we have

$$A = \begin{pmatrix} \vec{u}_1 & \cdots & \vec{u}_m \end{pmatrix} \begin{pmatrix} \sqrt{\lambda_1} & 0 & 0 & \cdots \\ 0 & \sqrt{\lambda_2} & 0 & \cdots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix} \begin{pmatrix} \vec{v}_1^t \\ \vdots \\ \vec{v}_n^t \end{pmatrix}$$

$$= \begin{pmatrix} \sqrt{\lambda_1}\vec{u}_1 & \cdots & \sqrt{\lambda_r}\vec{u}_r & \vec{0} & \cdots & \vec{0} \end{pmatrix} \begin{pmatrix} \vec{v}_1^t \\ \vdots \\ \vec{v}_n^t \end{pmatrix} = \sum_{i=1}^{r} \sqrt{\lambda_i}\vec{u}_i\vec{v}_i^t.$$

Note that each $\sqrt{\lambda_i}\vec{u}_i\vec{v}_i^t$ is an $m \times n$ matrix of rank one. The meaning of this factorization will be discussed in the next lecture.

Compact SVD

From the rank-one matrix factorization of A, we have

$$A = \sum_{i=1}^{r} \sqrt{\lambda_i}\vec{u}_i\vec{v}_i^t = \begin{pmatrix} \vec{u}_1 & \cdots & \vec{u}_r \end{pmatrix} \begin{pmatrix} \sqrt{\lambda_1} & 0 & \cdots & 0 \\ 0 & \sqrt{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sqrt{\lambda_r} \end{pmatrix} \begin{pmatrix} \vec{v}_1^t \\ \vdots \\ \vec{v}_r^t \end{pmatrix}$$

Theorem (Compact SVD)

Let A be a real matrix of size $m \times n$ of rank r. There exist an $n \times r$ matrix V with orthonormal columns, an $m \times r$ matrix U with orthonormal columns, and a square diagonal matrix Σ with positive entries of size r such that $A = U\Sigma V^t$.