Advanced Linear Algebra

Alvin Lin

January 2019 - May 2019

Singular Value Decomposition

Recall that the diagonalizaton of A can be written as:

$$A = \begin{bmatrix} \vec{v_1} & \dots & \vec{v_n} \end{bmatrix} \begin{bmatrix} \lambda_1 & \vdots \\ \dots & \lambda_n \end{bmatrix} \begin{bmatrix} \vec{v_1} & \dots & \vec{v_n} \end{bmatrix}^{-1}$$

If A is symmetric, we can orthogonally diagonalize it as $A = PDP^T$.

$$A = \begin{bmatrix} \hat{v_1} & \dots & \hat{v_n} \end{bmatrix} \begin{bmatrix} \lambda_1 & \vdots \\ \dots & \lambda_n \end{bmatrix} \begin{bmatrix} \dots & \vec{v_1}^T & \dots \\ \dots & \vec{v_n}^T & \dots \end{bmatrix}^{-1}$$

Recall from the definition of singular values of A:

$$(A^T A)\hat{v_i} = \lambda_i \hat{v_i} \quad (\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n)$$

We define the singular values $\sigma_i = \sqrt{\lambda_i} = \|A\hat{v_i}\|$.

$$\begin{aligned} \|A\hat{v_i}\|^2 &= (A\hat{v_i})^T(A\hat{v_i}) \\ &= \hat{v_i}^T A^T A \hat{v_i} \\ &= \lambda(\hat{v_i}^T \hat{v_i}) = \lambda \end{aligned}$$

Singular Value Decomposition

$$A = U\Sigma V^T$$

$$[A] = \begin{bmatrix} \hat{u_1} & \dots & \hat{u_n} \end{bmatrix} \begin{bmatrix} \sigma_1 & \vdots \\ \dots & \sigma_n \end{bmatrix} \begin{bmatrix} \dots & \hat{v_1}^T & \dots \\ \dots & \hat{v_n}^T & \dots \end{bmatrix}$$

1. Find the eigenvalues λ_i and eigenvectors $\hat{v}_i$ of $A^T A$.
2. Let $\sigma_i = \sqrt{\lambda_i}$ and normalize $\hat{v}_i$.
3. Let $\hat{u}_i = \frac{1}{\sigma_i} A\hat{v}_i$

Why does this work?

$$\begin{aligned}
AV &= U\Sigma \\
A\begin{bmatrix}\hat{v}_1 & \dots & \hat{v}_n\end{bmatrix} &\stackrel{?}{=} \begin{bmatrix}\hat{u}_1 & \dots & \hat{u}_n\end{bmatrix}\Sigma \\
\begin{bmatrix}A\hat{v}_1 & \dots & A\hat{v}_n\end{bmatrix} &\stackrel{?}{=} \begin{bmatrix}\sigma_1\hat{u}_1 & \dots & \sigma_n\hat{u}_n\end{bmatrix} \\
\begin{bmatrix}\sigma_1\hat{u}_1 & \dots & \sigma_n\hat{u}_n\end{bmatrix} &= \begin{bmatrix}\sigma_1\hat{u}_1 & \dots & \sigma_n\hat{u}_n\end{bmatrix}
\end{aligned}$$

Observations:

- V is orthogonal because the columns are eigenvectors of the symmetric matrix $A^T A$.
- U is orthogonal because the columns $\hat{u}_i$ and $\hat{u}_j$ are orthogonal.

$$\begin{aligned}
\hat{u}_i \cdot \hat{u}_j &= \left(\frac{1}{\sigma_i} A\hat{v}_i\right) \cdot \left(\frac{1}{\sigma_j} A\hat{v}_j\right) \\
&= \frac{1}{\sigma_i\sigma_j}(A\hat{v}_j)^T(A\hat{v}_i) \\
&= \frac{1}{\sigma_i\sigma_j}\hat{v}_j{}^T(A^T A\hat{v}_i) \\
&= \frac{\lambda_i}{\sigma_i\sigma_j}\hat{v}_j{}^T\hat{v}_i \\
&= 0
\end{aligned}$$

In general, $A = U\Sigma V^T$ where

- A is an $m \times n$ matrix
- U is an $m \times m$ orthogonal matrix
- Σ is an $m \times n$ matrix
- V^T is an $n \times n$ orthogonal matrix

Example

Find the singular value decomposition of $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

$$A = U\Sigma V^T$$

$$A^T A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\lambda_1 = 2 \quad \sigma_1 = \sqrt{2} \quad \hat{v_1} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

$$\lambda_2 = 1 \quad \sigma_2 = \sqrt{1} \quad \hat{v_2} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\lambda_3 = 0 \quad \sigma_3 = 0 \quad \hat{v_3} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

$$V = \begin{bmatrix} \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ 0 & 1 & 1 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\hat{u_1} = \frac{1}{\sigma_1} A\hat{v_1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

$$\hat{u_2} = \frac{1}{\sigma_2} A\hat{v_2} = \frac{1}{1} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}$$

Example

Find the singular value decomposition of $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = U\Sigma V^T$$

$$A^T A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\lambda_1 = 3 \quad \sigma_1 = \sqrt{3} \quad \hat{v_1} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\lambda_2 = 1 \quad \sigma_1 = \sqrt{1} \quad \hat{v_2} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$$V = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\hat{u_1} = \frac{1}{\sigma_1} A\hat{v_1} = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}$$

$$\hat{u_2} = \frac{1}{\sigma_2} A\hat{v_2} = \frac{1}{\sqrt{1}} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

The question now is, how do we find a $\hat{u_3}$ such that it is an orthonormal vector to both $\hat{u_1}$ and $\hat{u_2}$. Let

$$\left\{ \begin{bmatrix} \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}, \begin{bmatrix} 0 \\ -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$$

be a basis for $\mathbb{R}^3$ and use the Gram-Schmidt process to get $\hat{u_3}$ as

$$\begin{aligned}
\hat{u_3} &= \hat{e_3} - \text{proj}_W \hat{e_3} \quad W = span(\hat{u_1}, \hat{u_2}) \\
&= \hat{e_3} - \left(\frac{\hat{u_1} \cdot \hat{e_3}}{\hat{u_1} \cdot \hat{u_1}}\right)\hat{u_1} - \left(\frac{\hat{u_2} \cdot \hat{e_3}}{\hat{u_2} \cdot \hat{u_2}}\right)\hat{u_2} = \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} \\
U &= \begin{bmatrix} \frac{2}{\sqrt{6}} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix}
\end{aligned}$$

Projection form of a symmetrix matrix

$$A = PDP^T = \lambda_1 \hat{v_1}\hat{v_1}^T + \cdots + \lambda_n \hat{v_n}\hat{v_n}^T$$

This can be analogously done for a singular value decomposition.

$$A = U\Sigma V^T = \sigma_1 \hat{u_1}\hat{v_1}^T + \cdots + \sigma_r \hat{u_r}\hat{v_r}^T \quad (\sigma_1 \geq \cdots \geq \sigma_r, \sigma_{r+1} = \cdots = \sigma_n = 0)$$

You can find all my notes at `http://omgimanerd.tech/notes`. If you have any questions, comments, or concerns, please contact me at alvin@omgimanerd.tech