MATH-UA 140 - Linear Algebra

Lecture 26 : Singular Value Decomposition (SVD)

This lecture covers one of the most important topics in linear algebra from the point of view of applications and modern numerical methods. We will show that ANY matrix A (not necessarily square) can be written as

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

where U and V are orthogonal matrices, and Σ is diagonal with positive entries.

Why is that possible for any matrix? We will see that in this lecture. At this point, observe that U and V are different orthogonal matrices, unlike the decomposition $A = Q \Lambda Q^T$ for symmetric matrices.

I) SVD : Theory

Let us rewrite $A = U \Sigma V^T$ by multiplying by V on the right:

$$AV = U\Sigma$$

If A is an $m \times n$ matrix, V is an $n \times n$ matrix, U an $m \times m$ matrix, and Σ a $m \times n$ matrix.

V is orthogonal, so its columns are an orthonormal basis of $\mathbb{R}^n$.

U is orthogonal, so its columns are an orthonormal basis of $\mathbb{R}^m$.

Let r be the rank of A. To within a reshuffling of the columns of V, we can say that the first r columns are in the row space of A. The remaining $n-r$ columns are in the nullspace of A.

For the first r columns of V, $AV = U\Sigma$ means that A takes the r orthonormal columns $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r$ and turns them into an orthogonal basis $\sigma_1\vec{u}_1, \sigma_2\vec{u}_2, \ldots, \sigma_r\vec{u}_r$ of its column space:

$$A\begin{bmatrix}\vec{v}_1 & \cdots & \vec{v}_r\end{bmatrix} = \begin{bmatrix}\vec{u}_1 & \cdots & \vec{u}_r\end{bmatrix}\begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_r\end{bmatrix} = \begin{bmatrix}\sigma_1\vec{u}_1 & \cdots & \sigma_r\vec{u}_r\end{bmatrix}$$

where all the $\sigma_i \neq 0$ since the $\vec{v}_i$ are not in the nullspace of A.

To get the full $AV = U\Sigma$ form with V and U square matrices, we need to include $n-r$ orthonormal vectors $\vec{v}_{r+1}, \ldots, \vec{v}_n$ in the nullspace of A, $m-r$ orthonormal vectors $\vec{u}_{r+1}, \ldots, \vec{u}_m$ in the left nullspace of A, and $m-r$ zero rows and $n-r$ zero columns in Σ:

$$A\begin{bmatrix}\vec{v}_1 & \cdots & \vec{v}_r & \vec{v}_{r+1} & \cdots & \vec{v}_n\end{bmatrix} = \begin{bmatrix}\vec{u}_1 & \cdots & \vec{u}_r & \vec{u}_{r+1} & \cdots & \vec{u}_m\end{bmatrix}\begin{bmatrix}\sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & 0 & \\ & & & & \ddots & \\ & & & & & 0\end{bmatrix}$$

The full SVD then reads

$$AV = U\Sigma \iff A = U\Sigma V^T = \vec{u}_1\sigma_1\vec{v}_1^T + \cdots + \vec{u}_r\sigma_r\vec{v}_r^T$$

The vectors $\vec{v}_1, \dots, \vec{v}_r$ are called singular vectors. They can be viewed as generalized eigenvectors.

The values $\sigma_1, \dots, \sigma_r$ are all strictly positive, and are called singular values. They can be viewed as generalized eigenvalues.

II] Computing an SVD

We now know what the SVD of a matrix means. How would we go about computing it? Let us derive expressions for the matrices V and U, and for the singular values σ_i.

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

To get rid of U in this expression, let us multiply by A^T on the left:

$$A^T A = V \Sigma^T \underbrace{U^T U}_{I} \Sigma V^T = V \underbrace{\begin{bmatrix} \sigma_1^2 & & & & \\ & \ddots & & & \\ & & \sigma_r^2 & & \\ & & & 0 & \\ & & & & \ddots \; 0 \end{bmatrix}}_{n \times n \text{ diagonal matrix}} V^T$$

$A^T A$ is symmetric positive semidefinite, and we recognize here its diagonalization. We conclude that V is the eigenvector matrix of $A^T A$, and the entries σ_i of Σ are the square roots of the eigenvalues of $A^T A$.

more precisely: the matrix with orthonormal eigenvectors

To get an expression for U, we multiply $A = U \Sigma V^T$ by A^T on the right:

$$A A^T = U \Sigma \underbrace{V^T V}_{=I} \Sigma^T U^T = U \overbrace{\begin{bmatrix} \sigma_1^2 & & & & \\ & \ddots & & & \\ & & \sigma_r^2 & & \\ & & & 0 & \\ & & & & \ddots \; 0 \end{bmatrix}}^{m \times m \text{ matrix}} U^T$$

We conclude that U is the eigenvector matrix of AA^T, where the eigenvectors are constructed to be orthonormal.

III) Examples

* Let us compute the SVD of $A = \begin{bmatrix} 1 & 1 \\ -3 & 3 \end{bmatrix}$.

• $A^T A = \begin{bmatrix} 1 & -3 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ -3 & 3 \end{bmatrix} = \begin{bmatrix} 10 & -8 \\ -8 & 10 \end{bmatrix}$

Its eigenvalues are 2 and 18. The corresponding eigenvectors are $\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ and $\vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$. Thus,

$$V = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \text{ and } \Sigma = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 3\sqrt{2} \end{bmatrix}$$

• $AA^T = \begin{bmatrix} 1 & 1 \\ -3 & 3 \end{bmatrix} \begin{bmatrix} 1 & -3 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 18 \end{bmatrix}$

As expected, the eigenvalues are 2 and 18. The orthonormal eigenvectors are $\vec{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\vec{u}_2 = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$ so that $A\vec{v}_1 = \sqrt{2}\,\vec{u}_1$ and $A\vec{v}_2 = 3\sqrt{2}\,\vec{u}_2$

so $U = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ and the SVD for A is

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 3\sqrt{2} \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

* Let us compute the SVD of $A = \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$

This is a rank 1 matrix, so very little computation is required:

- An orthonormal basis for the row space is $\vec{v}_1 = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}$
- An orthonormal basis for the nullspace is $\vec{v}_2 = \begin{bmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$
- $A^T A = \begin{bmatrix} 2 & 1 \\ 4 & 2 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 10 & 20 \end{bmatrix}$

so the eigenvalues of $A^T A$ are 0 and 25

The singular values of A are 0 and 5

- An orthonormal basis for the column space is $\vec{u}_1 = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix}$
- An orthonormal basis for the left nullspace is $\vec{u}_2 = \begin{bmatrix} -\frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix}$

We can therefore write

$$\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{5}} & -\frac{1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$$