Advanced Linear Algebra – Week 9

The Singular Value Decomposition

This week we will learn about:

- The singular value decomposition (SVD),
- Orthogonality of the fundamental matrix subspaces, and
- How the SVD relates to other matrix decompositions,

Extra reading and watching:

- Section 2.3.1 and 2.3.2 in the textbook
- Lecture videos 34, 35, 36, and 37 on YouTube
- Singular value decomposition at Wikipedia
- Fundamental Theorem of Linear Algebra at Wikipedia

Extra textbook problems:

⋆ 2.3.1, 2.3.4(a,b,c,f,g,i)

⋆⋆ 2.3.3, 2.3.5, 2.3.7

⋆⋆⋆ 2.3.14, 2.3.20

☠ 2.3.26

If the Schur decomposition theorem from last week was "big", then the upcoming theorem is "super-mega-gigantic". The singular value decomposition is possibly the biggest and most widely-used theorem in all of linear algebra (and is my personal favourite), so we're going to spend some time focusing on it.

Theorem 9.1 — Singular Value Decomposition (SVD)

Suppose $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, and $A \in \mathcal{M}_{m,n}(\mathbb{F})$. Then there exist unitary matrices $U \in \mathcal{M}_m(\mathbb{F})$ and $V \in \mathcal{M}_n(\mathbb{F})$ and a diagonal matrix $\Sigma \in \mathcal{M}_{m,n}(\mathbb{R})$ with non-negative entries such that

Furthermore, the diagonal entries of Σ (called the **singular values** of A) are the non-negative square roots of the eigenvalues of A^*A.

Let's compare how this decomposition theorem is good and bad compared to our previous decomposition theorems:

- Good:
- Good:
- Kinda good, kinda bad:

Proof. Consider the matrix A^*A and assume that $m \geq n$...

■

Some notes about the SVD are in order:

- The singular values of A are exactly the square roots of the eigenvalues of A^*A. Alternatively...

- Even though the singular values are uniquely determined by A, the diagonal matrix Σ isn't.

- The unitary matrices U and V are often not uniquely determined by A. Example:

Example. *Let's find the singular values of a matrix.*

To compute a full singular value decomposition (not just the singular values), we again leech off of diagonalization. Notice that

Similarly, the columns of U are eigenvectors of AA^*, but a slightly quicker (and slightly more correct) way to compute the columns of U is to notice that

Example. *Compute a singular value decomposition of the matrix*

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

Before delving into what makes the singular value decomposition so useful, it is worth noting that if $A \in \mathcal{M}_{m,n}(\mathbb{F})$ has singular value decomposition $A = U\Sigma V^*$ then A^T and A^* have singular value decompositions

In particular,

Geometric Interpretation

Recall that we think of unitary matrices as arbitrary-dimensional rotations and/or reflections. Using this intuition gives the singular value decomposition a simple geometric interpretation. Specifically, it says that every matrix $A = U\Sigma V^* \in \mathcal{M}_{m,n}(\mathbb{F})$ acts as a linear transformation from $\mathbb{F}^n$ to $\mathbb{F}^m$ in the following way:

- First,
- Then,
- Finally,

Let's illustrate this geometric interpretation in the $m = n = 2$ case:

In particular, it is worth keeping track not only of how the linear transformation changes a unit square grid on $\mathbb{R}^2$ into a parallelogram grid, but also how it transforms...

Furthermore, the two radii of the ellipse are exactly

In higher dimensions, linear transformations send (hyper-)ellipsoids to (hyper-)ellipsoids. For example, the matrix

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

from earlier deforms the unit sphere as follows:

The fact that the unit sphere is turned into a 2D ellipse by this matrix corresponds to the fact that...

In fact, the first two left singular vectors $\mathbf{u}_1$ and $\mathbf{u}_2$ (which point in the directions of the major and minor axes of the ellipse) form an orthonormal basis of the range.

This same type of argument works in general and leads to the following theorem:

Theorem 9.2 — Bases of the Fundamental Subspaces

Let $A \in \mathcal{M}_{m,n}$ be a matrix with rank r and singular value decomposition $A = U\Sigma V^*$, where

$$U = \left[\, \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_m \,\right] \quad \text{and} \quad V = \left[\, \mathbf{v}_1 \mid \mathbf{v}_2 \mid \cdots \mid \mathbf{v}_n \,\right].$$

Then

a) $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ is an orthonormal basis of range(A),

b) $\{\mathbf{u}_{r+1}, \mathbf{u}_{r+2}, \ldots, \mathbf{u}_m\}$ is an orthonormal basis of null(A^*),

c) $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_r\}$ is an orthonormal basis of range(A^*), and

d) $\{\mathbf{v}_{r+1}, \mathbf{v}_{r+2}, \ldots, \mathbf{v}_n\}$ is an orthonormal basis of null(A).

Proof. Let's compute $A\mathbf{v}_j$:

■

Corollary 9.3 — Orthogonality of the Fundamental Subspaces

Suppose $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, and $A \in \mathcal{M}_{m,n}(\mathbb{F})$. Then

a) range(A) is orthogonal to null(A^*), and

b) null(A) is orthogonal to range(A^*).

In this corollary, when we say that one subspace is orthogonal to another, we mean that

Example. *Compute a singular value decomposition of the matrix*

$$A = \begin{bmatrix} 1 & 1 & 1 & -1 \\ 0 & 1 & 1 & 0 \\ -1 & 1 & 1 & 1 \end{bmatrix},$$

and use it to construct bases of the four fundamental subspaces of A.

Relationship With Other Matrix Decompositions

We now make sure that we really understand where the SVD fits into our world of matrix decompositions. For example, one way of rephrasing the singular value decomposition is as saying that we can always write a rank-r matrix as a sum of r rank-1 matrices in a very special way:

Theorem 9.4 — Orthogonal Rank-One Sum Decomposition

Suppose $\mathbb{F} = \mathbb{R}$ or $\mathbb{F} = \mathbb{C}$, and $A \in \mathcal{M}_{m,n}(\mathbb{F})$ is a matrix with $\text{rank}(A) = r$. Then there exist orthonormal sets of vectors $\{\mathbf{u}_i\}_{i=1}^r \subset \mathbb{F}^m$ and $\{\mathbf{v}_i\}_{i=1}^r \subset \mathbb{F}^n$ such that

where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ are the non-zero singular values of A.

- This formulation is sometimes useful because...

- In fields other than $\mathbb{R}$ and $\mathbb{C}$, ...

Proof. For simplicity, we again assume that $m \leq n$ throughout this proof, and then we just do block matrix multiplication in the singular value decompositon:

■

In fact the singular value decomposition and the orthogonal rank-one sum decomposition are "equivalent" in the sense that you can prove one to quickly prove the other, and vice-versa. Sometimes they are both just called the singular value decomposition.

Example. *Compute an orthogonal rank-one sum decomposition of the matrix*

$$A = \begin{bmatrix} 1 & 1 & 1 & -1 \\ 0 & 1 & 1 & 0 \\ -1 & 1 & 1 & 1 \end{bmatrix}.$$

Similarly, the singular value decomposition is also "essentially equivalent" to the polar decomposition:

In the opposite direction,

If $A \in \mathcal{M}_n$ is positive semidefinite, then the singular value decomposition coincides exactly with the spectral decomposition:

A slight generalization of this type of argument leads to the following theorem:

Theorem 9.5 — Singular Values of Normal Matrices

Suppose $A \in \mathcal{M}_n$ is a normal matrix. Then the singular values of A are the absolute values of its eigenvalues.

Proof. Since A is normal, we can use the spectral decomposition to write $A = UDU^*$, where U is unitary and D is diagonal...

■

To see that the above theorem does not hold for non-normal matrices, consider the following example:

Example. *Compute the eigenvalues and singular values of the matrix*

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