Lecture 17: Singular Value Decomposition

0.1 Constructing the Full SVD

Let

$$A \in \mathbb{C}^{m \times n}$$

have rank r.

Since A^*A is Hermitian and positive semidefinite, it has an orthonormal basis of eigenvectors

$$v_1, \ldots, v_n$$

with corresponding eigenvalues

$$\lambda_1 \geq \cdots \geq \lambda_r > 0 = \lambda_{r+1} = \cdots = \lambda_n.$$

Define

$$\sigma_j = \sqrt{\lambda_j}, \qquad j = 1, \ldots, r.$$

For each $j = 1, \ldots, r$, define

$$\boxed{u_j = \frac{Av_j}{\sigma_j}.}$$

Then

$$Av_j = \sigma_j u_j.$$

As shown previously,

$$u_1, \ldots, u_r$$

are orthonormal.

Moreover, each u_j lies in $\mathcal{C}(A)$. Since there are $r = \text{rank}(A)$ such vectors, they form an orthonormal basis for $\mathcal{C}(A)$:

$$\boxed{\mathcal{C}(A) = \text{span}\{u_1, \ldots, u_r\}.}$$

For $j = r+1, \ldots, n$, we have

$$\lambda_j = 0.$$

Therefore,

$$A^*Av_j = 0.$$

Since

$$v_j^* A^* A v_j = \|Av_j\|_2^2,$$

we obtain

$$\boxed{Av_j = 0, \qquad j = r+1, \ldots, n.}$$

Thus

$$v_{r+1}, \ldots, v_n$$

form an orthonormal basis for $\mathcal{N}(A)$.

Now there is an important point.

The vectors

$$u_1, \ldots, u_r$$

do not necessarily form a basis for all of $\mathbb{C}^m$ when $r < m$.

To construct the *full* SVD, choose additional vectors

$$u_{r+1}, \ldots, u_m$$

so that

$$u_1, \ldots, u_m$$

form an orthonormal basis of $\mathbb{C}^m$.

Equivalently, we may choose

$$u_{r+1}, \ldots, u_m$$

to be an orthonormal basis for

$$\mathcal{C}(A)^\perp = \mathcal{N}(A^*).$$

These additional vectors are not obtained from the formula

$$u_j = \frac{Av_j}{\sigma_j},$$

since the corresponding singular values are zero.

Now define

$$U = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_m \\ | & & | \end{pmatrix} \in \mathbb{C}^{m \times m},$$

and

$$V = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{pmatrix} \in \mathbb{C}^{n \times n}.$$

Both U and V are unitary.

Let

$$\Sigma \in \mathbb{R}^{m \times n}$$

be the rectangular diagonal matrix

$$\Sigma_{jj} = \begin{cases} \sigma_j, & j \le r, \\ 0, & j > r. \end{cases}$$

Then

$$\begin{aligned} AV &= \begin{pmatrix} | & & & & | \\ Av_1 & \cdots & Av_r & Av_{r+1} & \cdots \\ | & & & & | \end{pmatrix} \\ &= \begin{pmatrix} | & & & & | \\ \sigma_1 u_1 & \cdots & \sigma_r u_r & 0 & \cdots \\ | & & & & | \end{pmatrix} \\ &= U\Sigma. \end{aligned}$$

Since V is unitary,

$$VV^* = I_n.$$

Therefore,

$$\boxed{A = AVV^* = U\Sigma V^*.}$$

Suppose

$$A \in \mathbb{C}^{m\times n}$$

has rank r, and let

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

be a full singular value decomposition.

Since $\operatorname{rank}(A) = r$, exactly r singular values are positive:

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$

Write

$$U_r = \begin{pmatrix} | & & | \\ u_1 & \cdots & u_r \\ | & & | \end{pmatrix} \in \mathbb{C}^{m\times r},$$

and

$$V_r = \begin{pmatrix} | & & | \\ v_1 & \cdots & v_r \\ | & & | \end{pmatrix} \in \mathbb{C}^{n\times r}.$$

Also define

$$\Sigma_r = \operatorname{diag}(\sigma_1, \ldots, \sigma_r) \in \mathbb{R}^{r\times r}.$$

The columns associated with zero singular values do not contribute to the product $U\Sigma V^*$. Therefore,

$$\boxed{A = U_r\Sigma_r V_r^*.}$$

This is called the *compact singular value decomposition.*

Theorem 0.1 (Compact SVD)**.** *If*

$$A \in \mathbb{C}^{m\times n}$$

has rank r, then

$$\boxed{A = U_r\Sigma_r V_r^*,}$$

where U_r and V_r have orthonormal columns and

$$\Sigma_r = \operatorname{diag}(\sigma_1, \ldots, \sigma_r)$$

contains exactly the nonzero singular values of A.

The compact SVD keeps only the directions on which A acts nontrivially.

0.2 SVD as a Sum of Rank-One Matrices

The compact SVD

$$A = U_r \Sigma_r V_r^*$$

can also be written as

$$\boxed{A = \sum_{j=1}^{r} \sigma_j u_j v_j^*.}$$

Indeed,

$$U_r \Sigma_r = \begin{pmatrix} | & & | \\ \sigma_1 u_1 & \cdots & \sigma_r u_r \\ | & & | \end{pmatrix},$$

and multiplying by V_r^* gives

$$A = \sigma_1 u_1 v_1^* + \cdots + \sigma_r u_r v_r^*.$$

Each matrix

$$u_j v_j^*$$

has rank one.

Thus every rank-r matrix can be written as the sum of r rank-one matrices:

$$\boxed{A = \sigma_1 u_1 v_1^* + \cdots + \sigma_r u_r v_r^*.}$$

This representation will be particularly useful when we study low-rank approximation.

0.3 The Four Fundamental Subspaces and the SVD

The SVD gives orthonormal bases for all four fundamental subspaces of a matrix.

Let

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

be a full SVD of a rank-r matrix.

Recall that for $j = 1, \ldots, r$,

$$Av_j = \sigma_j u_j.$$

Since $\sigma_j > 0$, the vectors

$$u_1, \ldots, u_r$$

span the column space of A. Therefore,

$$\boxed{\mathcal{C}(A) = \text{span}\{u_1, \ldots, u_r\}.}$$

For $j > r$,

$$Av_j = 0.$$

Hence

$$v_{r+1}, \ldots, v_n$$

form an orthonormal basis for the null space:

$$\boxed{\mathcal{N}(A) = \text{span}\{v_{r+1}, \ldots, v_n\}.}$$

Similarly,

$$A^* u_j = \sigma_j v_j, \qquad j = 1, \ldots, r,$$

so

$$\boxed{\mathcal{C}(A^*) = \text{span}\{v_1, \ldots, v_r\}.}$$

Finally,

$$A^* u_j = 0, \qquad j = r+1, \ldots, m,$$

and therefore

$$\boxed{\mathcal{N}(A^*) = \text{span}\{u_{r+1}, \ldots, u_m\}.}$$

Thus the SVD organizes the four fundamental subspaces as follows:

$$\boxed{\begin{array}{c} \mathcal{C}(A) = \text{span}\{u_1, \ldots, u_r\}, \\ \mathcal{N}(A^*) = \text{span}\{u_{r+1}, \ldots, u_m\}, \\ \mathcal{C}(A^*) = \text{span}\{v_1, \ldots, v_r\}, \\ \mathcal{N}(A) = \text{span}\{v_{r+1}, \ldots, v_n\}. \end{array}}$$

Notice that this also makes the familiar orthogonal decompositions transparent:

$$\mathbb{C}^m = \mathcal{C}(A) \oplus \mathcal{N}(A^*),$$

and

$$\mathbb{C}^n = \mathcal{C}(A^*) \oplus \mathcal{N}(A).$$

0.4 Example: Computing an SVD by Hand

Consider

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

We will compute both its full and compact SVD.

First,

$$A^* A = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

The eigenvalues of $A^* A$ are

$$4, \qquad 2, \qquad 0.$$

An orthonormal set of corresponding eigenvectors is

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \qquad v_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

Therefore, the singular values are

$$\boxed{\sigma_1 = 2, \qquad \sigma_2 = \sqrt{2}, \qquad \sigma_3 = 0.}$$

Now compute the left singular vectors.

For v_1,

$$Av_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix},$$

so

$$u_1 = \frac{Av_1}{\sigma_1} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

For v_2,

$$Av_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 2 \end{pmatrix} = \begin{pmatrix} 0 \\ \sqrt{2} \end{pmatrix}.$$

Hence

$$u_2 = \frac{Av_2}{\sigma_2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Thus

$$U = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$V = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix},$$

and

$$\Sigma = \begin{pmatrix} 2 & 0 & 0 \\ 0 & \sqrt{2} & 0 \end{pmatrix}.$$

Therefore the full SVD is

$$\boxed{A = U\Sigma V^*.}$$

Since

$$\operatorname{rank}(A) = 2,$$

the compact SVD keeps only the first two right singular vectors:

$$U_r = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

$$\Sigma_r = \begin{pmatrix} 2 & 0 \\ 0 & \sqrt{2} \end{pmatrix},$$

and

$$V_r = \begin{pmatrix} 1 & 0 \\ 0 & 1/\sqrt{2} \\ 0 & 1/\sqrt{2} \end{pmatrix}.$$

Thus

$$\boxed{A = U_r \Sigma_r V_r^*.}$$

The rank-one expansion is

$$A = 2u_1 v_1^* + \sqrt{2}\, u_2 v_2^*.$$

Explicitly,

$$A = 2\begin{pmatrix}1\\0\end{pmatrix}\begin{pmatrix}1 & 0 & 0\end{pmatrix} + \sqrt{2}\begin{pmatrix}0\\1\end{pmatrix}\frac{1}{\sqrt{2}}\begin{pmatrix}0 & 1 & 1\end{pmatrix}.$$

Finally, the four fundamental subspaces are immediately visible from the SVD:

$$\mathcal{C}(A) = \operatorname{span}\{u_1, u_2\} = \mathbb{R}^2,$$

$$\mathcal{N}(A^*) = \{0\},$$

$$\mathcal{C}(A^*) = \operatorname{span}\{v_1, v_2\},$$

and

$$\boxed{\mathcal{N}(A) = \operatorname{span}\{v_3\} = \operatorname{span}\left\{\begin{pmatrix}0\\1\\-1\end{pmatrix}\right\}.}$$