Geometric Observation

- The image of unit sphere under any $m \times n$ matrix is a \textit{hyperellipse}.
- Give a unit sphere $S$ in $\mathbb{R}^n$, let $AS$ denote the shape after transformation.
- SVD is

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

where $U \in \mathbb{C}^{m \times m}$ and $V \in \mathbb{C}^{n \times n}$ is unitary and $\Sigma \in \mathbb{R}^{m \times n}$ is diagonal.

- \textit{Singular values} are diagonal entries of $\Sigma$, correspond to the principal semiaxes, with entries $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$.

- \textit{Left singular vectors} of $A$ are column vectors of $U$ and are oriented in the directions of the principal semiaxes of $AS$.

- \textit{Right singular vectors} of $A$ are column vectors of $V$ and are the preimages of the principal semiaxes of $AS$.

- $Av_j = \sigma_j u_j$ for $1 \leq j \leq n$.
Two Different Types of SVD

- **Full SVD:** $U \in \mathbb{C}^{m \times m}$, $\Sigma \in \mathbb{R}^{m \times n}$, and $V \in \mathbb{C}^{n \times n}$ is

\[
A = U \Sigma V^* \]

- **Reduced SVD:** $\hat{U} \in \mathbb{C}^{m \times n}$, $\hat{\Sigma} \in \mathbb{R}^{n \times n}$ (assume $m \geq n$)

Furthermore, notice that $A = \min\{m, n\} \sum_{i=1}^{\min\{m, n\}} \sigma_i u_i v_i^*$ so we can keep only entries of $U$ and $V$ corresponding to nonzero $\sigma_i$. 
Two Different Types of SVD

- **Full SVD**: \( U \in \mathbb{C}^{m \times m}, \Sigma \in \mathbb{R}^{m \times n}, \) and \( V \in \mathbb{C}^{n \times n} \) is
  \[
  A = U \Sigma V^*
  \]

- **Reduced SVD**: \( \hat{U} \in \mathbb{C}^{m \times n}, \hat{\Sigma} \in \mathbb{R}^{n \times n} \) (assume \( m \geq n \))
  \[
  A = \hat{U} \hat{\Sigma} V^*
  \]
Two Different Types of SVD

- **Full SVD**: $U \in \mathbb{C}^{m \times m}$, $\Sigma \in \mathbb{R}^{m \times n}$, and $V \in \mathbb{C}^{n \times n}$ is
  \[
  A = U \Sigma V^* \]

- **Reduced SVD**: $\hat{U} \in \mathbb{C}^{m \times n}$, $\hat{\Sigma} \in \mathbb{R}^{n \times n}$ (assume $m \geq n$)
  \[
  A = \hat{U} \hat{\Sigma} V^* \]

- Furthermore, notice that
  \[
  A = \sum_{i=1}^{\min\{m,n\}} \sigma_i u_i v_i^* \]

so we can keep only entries of $U$ and $V$ corresponding to nonzero $\sigma_i$. 
Existence of SVD

(Existence) Every matrix \( A \in \mathbb{C}^{m \times n} \) has an SVD.

Proof: Let \( \sigma_1 = \|A\|_2 \). There exists \( v_1 \in \mathbb{C}^n \) with \( \|v_1\|_2 = 1 \) and \( \|Av_1\|_2 = \sigma_1 \). Let \( U_1 \) and \( V_1 \) be unitary matrices whose first columns are \( u_1 = Av_1/\sigma_1 \) (or any unit-length vector if \( \sigma_1 = 0 \)) and \( v_1 \), respectively. Note that

\[
U_1^*AV_1 = S = \begin{bmatrix}
\sigma_1 & \omega^* \\
0 & B
\end{bmatrix}. \tag{1}
\]

Furthermore, \( \omega = 0 \) because \( \|S\|_2 = \sigma_1 \), and

\[
\left\| \begin{bmatrix}
\sigma_1 & \omega^* \\
0 & B
\end{bmatrix} \begin{bmatrix}
\sigma_1 \\
\omega
\end{bmatrix} \right\|_2 \geq \sigma_1^2 + \omega^*\omega = \sqrt{\sigma_1^2 + \omega^*\omega} \left\| \begin{bmatrix}
\sigma_1 \\
\omega
\end{bmatrix} \right\|_2,
\]

implying that \( \sigma_1 \geq \sqrt{\sigma_1^2 + \omega^*\omega} \) and \( \omega = 0 \).
Existence of SVD Cont’d

We then prove by induction using (1). If \( m = 1 \) or \( n = 1 \), then \( B \) is empty and we have \( A = U_1 S V_1^* \). Otherwise, suppose \( B = U_2 \Sigma_2 V_2^* \), and then

\[
A = U_1 \begin{bmatrix} 1 & 0^* \\ 0 & U_2 \end{bmatrix} \begin{bmatrix} \sigma_1 & 0^* \\ 0 & \Sigma_2 \end{bmatrix} \begin{bmatrix} 1 & 0^* \\ 0 & V_2^* \end{bmatrix} V_1^*,
\]

where \( U \) and \( V \) are unitary.
Uniqueness of SVD

Theorem

(Uniqueness) The singular values \( \{\sigma_j\} \) are uniquely determined. If \( A \) is square and the \( \sigma_j \) are distinct, the left and right singular vectors are uniquely determined \textbf{up to complex signs} (i.e., complex scalar factors of absolute value 1).

Geometric argument: If the lengths of semiaxes of a hyperellipse are distinct, then the semiaxes themselves are determined by the geometry up to signs.
Uniqueness of SVD Cont’d

Algebraic argument: Based on 2-norm and prove by induction. Consider the case where the $\sigma_j$ are distinct. The 2-norm is unique, so is $\sigma_1$. If $v_1$ is not unique up to sign, then the orthonormal bases of these vectors are right singular vectors of $A$, implying that $\sigma_1$ is not a simple singular value.

Once $\sigma_1$, $u_1$, and $v_1$ are determined, the remainder of SVD is determined by the space orthogonal to $v_1$. Because $v_1$ is unique up to sign, the orthogonal subspace is uniquely defined. Then prove by induction.
Uniqueness of SVD Cont’d

Algebraic argument: Based on 2-norm and prove by induction. Consider the case where the $\sigma_j$ are distinct. The 2-norm is unique, so is $\sigma_1$. If $v_1$ is not unique up to sign, then the orthonormal bases of these vectors are right singular vectors of $A$, implying that $\sigma_1$ is not a simple singular value.

Once $\sigma_1$, $u_1$, and $v_1$ are determined, the remainder of SVD is determined by the space orthogonal to $v_1$. Because $v_1$ is unique up to sign, the orthogonal subspace is uniquely defined. Then prove by induction.

● Question: What if we change the sign of a singular vector?
Uniqueness of SVD Cont’d

Algebraic argument: Based on 2-norm and prove by induction. Consider the case where the $\sigma_j$ are distinct. The 2-norm is unique, so is $\sigma_1$. If $v_1$ is not unique up to sign, then the orthonormal bases of these vectors are right singular vectors of $A$, implying that $\sigma_1$ is not a simple singular value.

Once $\sigma_1$, $u_1$, and $v_1$ are determined, the remainder of SVD is determined by the space orthogonal to $v_1$. Because $v_1$ is unique up to sign, the orthogonal subspace is uniquely defined. Then prove by induction.

- Question: What if we change the sign of a singular vector?

- Question: What if $\sigma_i$ is not distinct?
SVD vs. Eigenvalue Decomposition

- Eigenvalue decomposition of nondefective matrix $A$ is $A = X\Lambda X^{-1}$
SVD vs. Eigenvalue Decomposition

- Eigenvalue decomposition of nondefective matrix $A$ is $A = X \Lambda X^{-1}$

- Differences between SVD and Eigenvalue Decomposition
  - Not every matrix has eigenvalue decomposition, but every matrix has singular value decomposition
  - Eigenvalues may not always be real numbers, but singular values are always non-negative real numbers
  - Eigenvectors are not always orthogonal to each other (orthogonal for symmetric matrices), but left (or right) singular vectors are orthogonal to each other

Similarities
- Singular values of $A$ are square roots of eigenvalues of $A^*A$ and $AA^*$, and their eigenvectors are left and right singular vectors, respectively
- Singular values of hermitian matrices are absolute values of eigenvalues, and eigenvectors are singular vectors (up to complex signs)
- This relationship can be used to compute singular values by hand
SVD vs. Eigenvalue Decomposition

- Eigenvalue decomposition of nondefective matrix $A$ is $A = X\Lambda X^{-1}$

- Differences between SVD and Eigenvalue Decomposition
  - Not every matrix has eigenvalue decomposition, but every matrix has singular value decomposition
  - Eigenvalues may not always be real numbers, but singular values are always non-negative real numbers
  - Eigenvectors are not always orthogonal to each other (orthogonal for symmetric matrices), but left (or right) singular vectors are orthogonal to each other

- Similarities
  - Singular values of $A$ are square roots of eigenvalues of $AA^*$ and $A^*A$, and their eigenvectors are left and right singular vectors, respectively
  - Singular values of hermitian matrices are absolute values of eigenvalues, and eigenvectors are singular vectors (up to complex signs)
  - This relationship can be used to compute singular values by hand
Matrix Properties via SVD

- Let $r$ be number of nonzero singular values of $A \in \mathbb{C}^{m \times n}$
  - $\text{rank}(A)$ is $r$
  - $\text{range}(A) = \langle u_1, u_2, \ldots, u_r \rangle$
  - $\text{null}(A) = \langle v_{r+1}, v_{r+2}, \ldots, v_n \rangle$

- 2-norm and Frobenius norm
  - $\|A\|_2 = \sigma_1$ and $\|A\|_F = \sqrt{\sum_i \sigma_i^2}$

- Determinant of matrix
  - For $A \in \mathbb{C}^{m \times m}$, $|\det(A)| = \prod_{i=1}^m \sigma_i$
Matrix Properties via SVD

- Let \( r \) be number of nonzero singular values of \( A \in \mathbb{C}^{m \times n} \)
  - \( \text{rank}(A) \) is \( r \)
  - \( \text{range}(A) = \langle u_1, u_2, \ldots, u_r \rangle \)
  - \( \text{null}(A) = \langle v_{r+1}, v_{r+2}, \ldots, v_n \rangle \)

- 2-norm and Frobenius norm
  - \( \|A\|_2 = \sigma_1 \) and \( \|A\|_F = \sqrt{\sum_i \sigma_i^2} \)

- Determinant of matrix
  - For \( A \in \mathbb{C}^{m \times m} \), \( |\det(A)| = \prod_{i=1}^{m} \sigma_i \)

- However, SVD may not be the most efficient way in solving problems
- Algorithms for SVD are similar to those for eigenvalue decomposition and we will discuss them later in the semester