SVD applications: rank, column, row, and null spaces

Rank: the rank of a matrix is equal to:

- number of linearly independent columns
- number of linearly independent rows

(Remarkably, these are always the same!).

For an $m \times n$ matrix, the rank must be less than or equal to $\min(m, n)$. The rank can be thought of as the dimensionality of the vector space spanned by its rows or its columns.

Lastly, the rank of $A$ is equal to the number of non-zero singular values!

Consider the SVD of a matrix $A$ that has rank $k$: 

$$A = USV^\top$$

Column space: Since $A$ is rank $k$, the first $k$ left singular vectors, $\{\vec{u}_1, \ldots, \vec{u}_k\}$ (the columns of $U$), provide an orthonormal basis for the column space of $A$.

Row space: Similarly, the first $k$ right singular vectors, $\{\vec{v}_1, \ldots, \vec{v}_k\}$ (the columns of $V$, or the rows of $V^\top$), provide an orthonormal basis for the row space of $A$.

Null space: The last right singular vectors, $\{\vec{v}_{k+1}, \ldots, \vec{v}_n\}$ (the last columns of $V$, or the last rows of $V^\top$), provide an orthonormal basis for the null space of $A$.

Let’s prove this last one, just to see what such a proof looks like.

First, consider a vector $\vec{x}$ that can be expressed as a linear combination of the last $n - k$ columns of $V$:

$$\vec{x} = \sum_{i=k+1}^{n} w_i \vec{v}_i,$$

for some real-valued weights $\{w_i\}$. To show that $\vec{x}$ lives in the null space of $A$, we need to show that $A\vec{x} = 0$. Let’s go ahead and do this now. (It isn’t that hard, and this gives the flavor of what a lot of proofs in linear algebra look like)
\[ A\vec{x} = A \left( \sum_{i=k+1}^{n} w_i \vec{v}_i \right) \quad \text{(by definition of } \vec{x}) \quad (1) \]
\[ = \sum_{i=k+1}^{n} w_i (A\vec{v}_i) . \quad \text{(by definition of linearity)} \quad (2) \]

Now let’s look at any one of the terms in this sum:
\[ A\vec{v}_i = (USV^\top)\vec{v}_i = US(V^\top\vec{v}_i) = US\vec{e}_i, \quad (3) \]
where \( \vec{e}_i \) is the “identity” basis vector consisting of all 0’s except for a single 1 in the \( i \)'th row. This follows from the fact that \( \vec{v}_i \) is orthogonal to every row of \( V^\top \) except the \( i \)'th row, which gives \( \vec{v}_i \cdot \vec{v}_i = 1 \) because \( \vec{v}_i \) is a unit vector.

Now, because \( i \) in the sum only ranged over \( k+1 \) to \( n \), then when we multiply \( \vec{e}_i \) by \( S \) (which has non-zeros along the diagonal only up to the \( k \)'th row / column), we get zero:
\[ S\vec{e}_i = 0 \quad \text{for } i > k. \]
Thus
\[ US\vec{e}_i = 0 \]
which means that the entire sum
\[ \sum_{i=k+1}^{n} US\vec{e}_i = 0. \]
So this shows that \( A\vec{x} = 0 \) for any vector \( \vec{x} \) that lives in the subspace spanned by the last \( n - k \) columns of \( V \), meaning it lies in the null space. This is of course equivalent to showing that the last \( n - k \) columns of \( V \) provide an (orthonormal) basis for the null space!

2 Positive semidefinite matrix

Positive semi-definite (PSD) matrix is a matrix that has all eigenvalues \( \geq 0 \), or equivalently, a matrix \( A \) for which \( \vec{x}^\top A\vec{x} \geq 0 \) for any vector \( \vec{x} \).

To generate an \( n \times n \) positive semi-definite matrix, we can take any matrix \( X \) that has \( n \) columns and let \( A = X^\top X \).

3 Relationship between SVD and eigenvector decomposition

Definition: An eigenvector of a square matrix \( A \) is defined as a vector satisfying the equation
\[ A\vec{x} = \lambda\vec{x}, \]
and λ is the corresponding eigenvalue. In other words, an eigenvector of A is any vector that, when multiplied by A, comes back as itself scaled by λ.

**Spectral theorem:** If a matrix A is symmetric and positive semi-definite, then the SVD also an eigendecomposition, that is, a decomposition in terms of an orthonormal basis of eigenvectors:

\[ A = U \Sigma U^\top, \]

where the columns of U are eigenvectors and the diagonal entries \( \{s_i\} \) of \( S \) are the eigenvalues. Note that for such matrices, \( U = V \), meaning the left and right singular vectors are identical.

Exercise: prove to yourself that: \( A\vec{u}_i = s_i \vec{u}_i \)

**SVD of matrix times its transpose.** In class we showed that if \( A = U \Sigma V^\top \), then \( A^\top A \) (which it turns out, is symmetric and PSD) has the singular value decomposition (which is also an eigendecomposition): \( A^\top A = V \Sigma^2 V^\top \). Test yourself by deriving the SVD of \( AA^\top \).

## 4 Linearity and Linear Systems

Linear system is a kind of mapping \( f(\vec{x}) \rightarrow \vec{y} \) that has the following two properties:

1. homogeneity ("scalar multiplication"):
   \[ f(ax) = af(x) \]

2. additivity:
   \[ f(\vec{x}_1 + \vec{x}_2) = f(\vec{x}_1) + f(\vec{x}_2) \]

Of course we can combine these two properties into a single requirement and say: \( f \) is a linear function if and only if it obeys the principal of superposition:

\[ f(a\vec{x}_1 + b\vec{x}_2) = af(\vec{x}_1) + bf(\vec{x}_2) \]

**General rule:** we can write any linear function in terms of a matrix operation:

\[ f(\vec{x}) = A\vec{x} \]

for some matrix \( A \).

**Question:** is the function \( f(x) = ax + b \) a linear function? Why or why not?