Existence and Computation of the SVD

• Recall the Singular Value Decomposition:

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

where \( A \) is \( m \times n \), \( U \) is \( m \times m \) orthogonal, \( \Sigma \) is \( m \times n \) diagonal, and \( V \) is \( n \times n \) orthogonal.

• The actual computation of the singular value decomposition is tedious and rarely done by hand.

• Like most linear algebra computations you use computers for all but the most simple problems.

• For example, in Matlab you can compute the SVD with the statement

\[
[U,\text{Sigma},V] = \text{svd}(A)
\]

• The Matlab algorithm for computing the SVD is quite complicated.

• Here is a simple procedure that can be used in principle.

• Suppose \( m \geq n \). Recall that

\[ A = U\Sigma V^T \quad \text{and} \quad A^T = V\Sigma^T U^T. \]
Thus

\[ A^T A = V \Sigma^T U^T U \Sigma V^T = V \Sigma^T \Sigma V^T \quad (1) \]

and

\[ A A^T = U \Sigma V^T V \Sigma^T U^T = U \Sigma \Sigma^T U^T. \quad (2) \]

- The matrix \( \Sigma^T \Sigma \) is \( n \times n \) and diagonal. The diagonal entries are the squares of the singular values. Thus equation (1) implies that the \( v_i \), the right eigenvectors, are the eigenvectors of \( A A^T \), and the singular values are the square roots of the eigenvalues of \( A A^T \).

- Similarly, the matrix \( \Sigma \Sigma^T \) is a diagonal \( m \times m \) matrix, with the first \( n \) diagonal entries again being the squares of the singular values. The remaining diagonal entries are zero. Hence the \( u_i \), the left singular vectors, are the eigenvectors of \( A A^T \). Of course, if \( m \geq n \) only the first \( n \) right singular vectors contribute to the computation of \( A \).

- In this approach, computing the SVD amounts to computing the orthogonal transforms to diagonal form of the symmetric matrices \( A^T A \) and \( A A^T \).
• Finally, I’ll describe the textbook proof which can be used in principle to compute the SVD of small matrices by hand.

• So suppose $A$ is $m \times n$. Then $A^T A$ is symmetric and can be orthogonally diagonalized. Suppose

\[
\{v_1, v_2, \ldots, v_n\}
\]

is a set of orthonormal eigenvectors of $A^T A$, with corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$.

• The eigenvalues of $A^T A$ are all non-negative because we saw in class that any matrix of the form $A^T A$ is positive semidefinite.

• However, this can also be seen directly:

\[
0 \leq \|Av_i\|^2
= (Av_i)^T Av_i
= v_i^T A^T A v_i
= v_i^T (\lambda_i v_i)
= \lambda_i (v_i^T v_i)
= \lambda_i
\]

• We may assume that the eigenvalues of $A^T A$ are ordered such that

\[
\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0.
\]

• We define the singular values of $A$ to be the square roots of the $\lambda_i$:

\[
\sigma_i = \sqrt{\lambda_i} = \|Av_i\|.
\]
Theorem 9 in the textbook states: Suppose that \( \{v_1, \ldots, v_n\} \) is an orthonormal basis of \( \mathbb{R}^n \) consisting of eigenvectors of \( A^T A \), arranged so that the corresponding eigenvalues of \( A^T A \) satisfy

\[
\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0,
\]

and suppose that \( A^T A \) has \( r \) nonzero singular values. Then the set

\[
B = \{ A v_1, A v_2, \ldots, A v_r \}
\]

is an orthogonal basis for \( \text{Col} A \), and

\[
\text{rank} A = r.
\]

We first show that \( B \) is an orthogonal set. Since \( v_i \) and \( v_j \) are orthogonal we get for \( i \neq j \) that

\[
(A v_i)^T A(v_j) = v_i^T A^T A v_j = v_i^T (\lambda_j v_j) = 0.
\]

Next, notice that since \( \| A v_i \| = \sigma_i \) and there are \( r \) non-zero singular values we see that \( A v_i \) is non-zero only if and only if \( i \leq r \).

Since the vectors in \( B \) are non-zero and orthogonal they are linearly independent.

The set \( B \) is also a spanning set of \( \text{Col}(A) \). To see this suppose \( y \) is in the column space of \( A \). Then

\[
y = A x
\]
for some \( \mathbf{x} \) which we can write as.

\[
\mathbf{x} = \sum_{i=1}^{n} c_i \mathbf{v}_i.
\]

Hence

\[
\mathbf{y} = A\mathbf{x} = \sum_{i=1}^{n} c_i \mathbf{v}_i = \sum_{i=1}^{r} c_i (A\mathbf{v}_i).
\]

- Hence \( B \) is a basis of the column space of \( A \), and the rank of \( A \)—the dimension of the column space—equals \( r \).
- This completes the proof of Theorem 9.
- Next, normalize the orthogonal vectors \( A\mathbf{v}_i \) to get the orthonormal vectors

\[
\mathbf{u}_i = \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|} = \frac{A\mathbf{v}_i}{\sigma_i}.
\]

- Expand the orthonormal set

\[
\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}
\]

to an orthonormal basis

\[
\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r, \mathbf{u}_{r+1}, \ldots, \mathbf{u}_m\}
\]

- Now define

\[
U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \ldots & \mathbf{u}_m \end{bmatrix}.
\]
\[ V = [v_1 \ v_2 \ \ldots \ v_n], \]

and

\[
\Sigma = \begin{bmatrix}
\sigma_1 & 0 & \ldots & 0 \\
0 & \sigma_2 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \sigma_n \\
0 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 0
\end{bmatrix}
\]

(4)

- Then \( U \) is \( m \times m \) orthogonal, \( V \) is \( n \times n \) orthogonal, and \( \Sigma \) is \( m \times n \) diagonal as required for the SVD. Moreover,

\[
U \Sigma = [u_1 \ u_2 \ \ldots \ u_m] \begin{bmatrix}
\sigma_1 & 0 & \ldots & 0 \\
0 & \sigma_2 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & \sigma_n \\
0 & 0 & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
0 & 0 & \ldots & 0
\end{bmatrix}
\]

\[
= [\sigma_1 u_1 \ \sigma_2 u_2 \ \ldots \sigma_r u_2 \ 0 \ \ldots \ 0]
\]

\[
= [A v_1 \ A v_2 \ \ldots \ A v_r u_2 \ 0 \ \ldots \ 0]
\]

\[
= AV
\]
• Since $V$ is orthogonal, right-multiplying with $V^T$ on both sides of

$$U\Sigma = AV$$

gives the Singular Value Decomposition

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