Definition 22.3. A triple (U, D, V) such that $A = V D U^\top$, where U and V are orthogonal and D is a diagonal matrix whose entries are nonnegative (it is positive semidefinite) is called a *singular value decomposition (SVD) of* A. If $D = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$, it is customary to assume that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$.

The `Matlab` command for computing an SVD $A = VDU^\top$ of a matrix A is
`[V, D, U] = svd(A)`.

The proof of Theorem 22.5 shows that there are two orthonormal bases $(u_1, \ldots, u_n)$ and $(v_1, \ldots, v_n)$, where $(u_1, \ldots, u_n)$ are eigenvectors of $A^\top A$ and $(v_1, \ldots, v_n)$ are eigenvectors of $AA^\top$. Furthermore, $(u_1, \ldots, u_r)$ is an orthonormal basis of $\mathrm{Im}\, A^\top$, $(u_{r+1}, \ldots, u_n)$ is an orthonormal basis of $\mathrm{Ker}\, A$, $(v_1, \ldots, v_r)$ is an orthonormal basis of $\mathrm{Im}\, A$, and $(v_{r+1}, \ldots, v_n)$ is an orthonormal basis of $\mathrm{Ker}\, A^\top$.

Using a remark made in Chapter 4, if we denote the columns of U by $u_1, \ldots, u_n$ and the columns of V by $v_1, \ldots, v_n$, then we can write

$$A = V D U^\top = \sigma_1 v_1 u_1^\top + \cdots + \sigma_r v_r u_r^\top,$$

with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$. As a consequence, if r is a lot smaller than n (we write $r \ll n$), we see that A can be reconstructed from U and V using a much smaller number of elements. This idea will be used to provide "low-rank" approximations of a matrix. The idea is to keep only the k top singular values for some suitable $k \ll r$ for which $\sigma_{k+1}, \ldots, \sigma_r$ are very small.

Remarks:

(1) In Strang [170] the matrices U, V, D are denoted by $U = Q_2$, $V = Q_1$, and $D = \Sigma$, and an SVD is written as $A = Q_1 \Sigma Q_2^\top$. This has the advantage that Q_1 comes before Q_2 in $A = Q_1 \Sigma Q_2^\top$. This has the disadvantage that A maps the columns of Q_2 (eigenvectors of $A^\top A$) to multiples of the columns of Q_1 (eigenvectors of $A\, A^\top$).

(2) Algorithms for actually computing the SVD of a matrix are presented in Golub and Van Loan [80], Demmel [48], and Trefethen and Bau [176], where the SVD and its applications are also discussed quite extensively.

(3) If A is a symmetric matrix, then in general, there is no SVD $V\Sigma U^\top$ of A with $V = U$. However, if A is positive semidefinite, then the eigenvalues of A are nonnegative, and so the nonzero eigenvalues of A are equal to the singular values of A and SVDs of A are of the form

$$A = V\Sigma V^\top.$$

(4) The SVD also applies to complex matrices. In this case, for every complex $n \times n$ matrix A, there are two unitary matrices U and V and a diagonal matrix D such that

$$A = V D U^*,$$

where D is a diagonal matrix consisting of real entries $\sigma_1, \ldots, \sigma_n$, where $\sigma_1 \geq \cdots \geq \sigma_r$ are the singular values of A, i.e., the positive square roots of the nonzero eigenvalues of $A^* A$ and $A\, A^*$, and $\sigma_{r+1} = \ldots = \sigma_n = 0$.