7.4 The Singular Value Decomposition

Let A be an $m \times n$ matrix. Then A^TA is symmetric and can be orthogonally diagonalized. Let $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ be an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of A^TA, and let $\lambda_1, \ldots, \lambda_n$ be the associated eigenvalues of A^TA. Then for $1 \le i \le n$,

$$\|A\mathbf{v}_i\|^2 = (A\mathbf{v}_i)^T A\mathbf{v}_i = \mathbf{v}_i^T A^T A\mathbf{v}_i = \mathbf{v}_i^T(\lambda_i \mathbf{v}_i) = \lambda_i. \tag{1}$$

So the eigenvalues of A^TA are all nonnegative. By renumbering, if necessary, we may assume that the eigenvalues are arranged such that

$$\lambda_1 \ge \cdots \ge \lambda_n \ge 0.$$

The *singular values* of A are the square roots of the eigenvalues of A^TA, denoted by $\sigma_1, \ldots, \sigma_n$. By (1), the singular values of A are the lengths of the vectors $A\mathbf{v}_1, \ldots, A\mathbf{v}_n$.

Example 1. *Find the singular values of* $A = \begin{bmatrix} -3 & 0 \\ 0 & 0 \end{bmatrix}$ *and* $B = \begin{bmatrix} 3 & 0 \\ 8 & 3 \end{bmatrix}$

Theorem 9. *Suppose* $\{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ *is an orthogonal basis of* $\mathbb{R}^n$ *consisting of eigenvectors of* A^TA*, arranged so that the corresponding eigenvalues of* A^TA *satisfy* $\lambda_1 \ge \cdots \ge \lambda_n$*, and suppose* A *has* r *nonzero singular values. Then* $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$ *is an orthogonal basis for* Col A*, and* rank $A = r$*.*

Theorem 10 (The Singular Value Decomposition)**.** *Let* A *be an* $m \times n$ *matrix with rank* r*. Then there exists an* $m \times n$ *matrix* $\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$ *for which the diagonal entries in the* $r \times r$ *diagonal matrix* D *are the first* r *singular values of* A*,* $\sigma_1 \ge \cdots \ge \sigma_r > 0$*, and there exists an* $m \times n$ *orthogonal matrix* U *and an* $n \times n$ *orthogonal matrix* V *such that*

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

Any factorization $A = U\Sigma V^T$, with U and V orthogonal, $\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$, and positive diagonal entries in D, is called a *singular value decomposition* (or SVD) of A. The columns of U in such a decomposition are called *left singular vectors* of A, and the columns of V are called *right singular vectors* of A.

Proof. Let $\lambda_1 \ge \cdots \ge \lambda_n$ be the eigenvalues of A^TA for eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$, respectively. Let $\sigma_i = \sqrt{\lambda_i}$ for $1 \le i \le n$. Then by the previous theorem, $\{A\mathbf{v}_1, \ldots, A\mathbf{v}_r\}$ is an orthogonal basis for Col A. Normalize $A\mathbf{v}_i$ to obtain basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$, where

$$\mathbf{u}_i = \frac{1}{\|A\mathbf{v}_i\|} A\mathbf{v}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$$

and hence

$$A\mathbf{v}_i = \sigma_i \mathbf{u}_i, \qquad 1 \le i \le r. \tag{2}$$

Now extend $\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$ to an orthonormal basis $\{\mathbf{u}_1, \ldots, \mathbf{u}_m\}$ of $\mathbb{R}^m$, and let

$$U = [\mathbf{u}_1 \quad \cdots \quad \mathbf{u}_m] \quad \text{and} \quad V = [\mathbf{v}_1 \quad \cdots \quad \mathbf{v}_n].$$

Note that U and V are orthogonal matrices. Also, from (2),

$$AV = [A\mathbf{v}_1 \quad \cdots \quad A\mathbf{v}_r \quad \mathbf{0} \quad \cdots \quad \mathbf{0}] = [\sigma_1\mathbf{u}_1 \quad \cdots \quad \sigma_r\mathbf{u}_r \quad \mathbf{0} \quad \cdots \quad \mathbf{0}].$$

Let D be the diagonal matrix with diagonal entries $\sigma_1 \ldots, \sigma_r$, and let $\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$. Then

$$\begin{aligned} U\Sigma &= [\mathbf{u}_1 \quad \cdots \mathbf{u}_m] \left[\begin{array}{cccc|c} \sigma_1 & & & 0 & \\ & \sigma_2 & & & 0 \\ & & \ddots & & \\ 0 & & & \sigma_r & \\ \hline & & 0 & & 0 \end{array}\right] \\ &= [\sigma_1\mathbf{u}_1 \quad \cdots \quad \sigma_r\mathbf{u}_r \quad \mathbf{0} \quad \cdots \quad \mathbf{0}] \\ &= AV. \end{aligned}$$

Since V is an orthogonal matrix, $U\Sigma V^T = AVV^T = A$. □

To construct a singular value decomposition of a matrix A:

1. Find an orthogonal diagonalization of $A^T A$.
2. Set up V and Σ.
3. Construct U.

Example 2. *Find an SVD of* $\begin{bmatrix} 7 & 1 \\ 5 & 5 \\ 0 & 0 \end{bmatrix}$.

Theorem (IMT (concluded))**.** *Let A be an $n \times n$ matrix. Then the following are each equivalent to the statement that A is an invertible matrix.*

21. $(\text{Col } A)^\perp = \{\mathbf{0}\}$.

22. $(\text{Nul } A)^\perp = \mathbb{R}^n$.

23. Row $A = \mathbb{R}^n$.

24. A has n nonzero singular values.

Example 3. *When Σ contains rows or columns of zeros, a more compact decomposition of A is possible. Let $r = \text{rank } A$, and partition U and V into submatrices whose first blocks contain r columns:*

$$\begin{aligned} U &= [U_r \quad U_{m-r}], \qquad \textit{where } U_r = [\mathbf{u}_1 \quad \cdots \mathbf{u}_r] \\ V &= [V_r \quad V_{n-r}], \qquad \textit{where } V_r = [\mathbf{v}_1 \quad \cdots \mathbf{v}_r]. \end{aligned}$$

Then U_r is $m \times r$ and V_r is $n \times r$. Then partitioned matrix multiplication shows that

$$A = [U_r \quad U_{m-r}] \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_r^T \\ V_{n-r}^T \end{bmatrix} = U_r D V_r^T. \tag{3}$$

This factorization of A is called a reduced singular value decomposition *of A. Since the diagonal entries in D are nonzero, D is invertible. The following matrix is called the* pseudo-inverse*, or the* Moore-Penrose inverse*, of A:*

$$A^+ = V_r D^{-1} U_r^T.$$