Math 313 Lecture #33
§7.4: The Singular Value Decomposition

We have seen the usefulness of the orthogonal diagonalization of symmetric (square) matrices: $A = PDP^T$.

Is there something similar for non-square matrices? Yes there is.

A Motivating Example. Consider the factorization of the following 3×2 matrix:

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}.$$

We have factored A as a product

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

where U is a 3×3 orthogonal matrix, Σ is a 3×2 matrix whose off diagonal entries are zero, and V is a 2×2 orthogonal matrix.

Can any $m \times n$ matrix A alway be factored as $U\Sigma V^T$?

Theorem 10. For each $A \in \mathbb{R}^{m\times n}$, there exists an orthogonal $m \times m$ matrix U, an $m \times n$ matrix Σ whose off-diagonal entries are all zero and whose diagonal entries satisfy $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$, and an orthogonal $n \times n$ matrix V such that $A = U\Sigma V^T$.

The product $A = U\Sigma V^T$ is called a **singular value decomposition** of A, and the nonnegative diagonal entries σ_i of Σ are called the **singular values** of A.

Example. Find a singular value decomposition of

$$A = \begin{bmatrix} 1 & 0 \\ 2 & 2 \\ 0 & 1 \end{bmatrix}.$$

Step 1. The matrix V is an orthogonal matrix that diagonalizes the symmetric matrix

$$A^TA = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}.$$

The eigenvalues and unit eigenvectors of A^TA are

$$\lambda_1 = 9,\ \vec{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}, \quad \lambda_2 = 1,\ \vec{v}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.$$

Fact: the eigenvalues of A^TA are always real and nonnegative because $A^TA\vec{x} = \lambda\vec{x}$ for $\vec{x} \neq 0$ implies

$$\|A\vec{x}\|^2 = \vec{x}^TA^TA\vec{x} = \lambda\vec{x}^T\vec{x} = \lambda\|\vec{x}\|^2.$$

We have already ordered the eigenvalues in decreasing order.

Since these eigenvectors are orthogonal (they belong to distinct eigenvalues of an symmetric matrix), we have

$$V = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}.$$

Step 2. The singular values of A are $\sigma_1 = \sqrt{\lambda_1} = 3$ and $\sigma_2 = \sqrt{\lambda_2} = 1$, and so

$$\Sigma = \begin{bmatrix} 3 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

Fact: The number of nonzero singular values of A is the rank of A.

Observe that the rank of A is 2 and that A has 2 nonzero singular values.

Step 3. Each nonzero singular value determines a column of U (but not necessarily all of the columns of U).

Since $A = U\Sigma V^T$ and V is orthogonal, we have $AV = U\Sigma$. Passing to the columns of this, we have

$$A\vec{v}_i = \sigma_i \vec{u}_i,$$

which can be solved for $\vec{u}_i$ when $\sigma_i \neq 0$.

Since we have two nonzero singular values, the first two columns of U are

$$\vec{u}_1 = \frac{A\vec{v}_1}{\sigma_1} = \frac{1}{3}\begin{bmatrix} 1 & 0 \\ 2 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \frac{1}{3}\begin{bmatrix} 1/\sqrt{2} \\ 4/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/3\sqrt{2} \\ 4/3\sqrt{2} \\ 1/3\sqrt{2} \end{bmatrix},$$

$$\vec{u}_2 = \frac{A\vec{v}_2}{\sigma_2} = \frac{1}{1}\begin{bmatrix} 1 & 0 \\ 2 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}.$$

Those columns of U determined by the nonzero singular values belong to $\mathrm{Col}(A)$. The remainder of the columns of U are an orthonormal basis for $(\mathrm{Col}\, A)^\perp = \mathrm{Nul}(A^T)$.

Here, the third (and last) column of U is a unit basis vector for $\mathrm{Nul}(A^T)$:

$$A^T = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 2 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1/2 \end{bmatrix} \Rightarrow \vec{u}_3 = \begin{bmatrix} 2/3 \\ -1/3 \\ 2/3 \end{bmatrix}.$$

Then

$$U = \begin{bmatrix} 1/3\sqrt{2} & -1/\sqrt{2} & 2/3 \\ 4/3\sqrt{2} & 0 & -1/3 \\ 1/3\sqrt{2} & 1/\sqrt{2} & 2/3 \end{bmatrix},$$

and we have a singular value decomposition for A:

$$\begin{bmatrix} 1 & 0 \\ 2 & 2 \\ 0 & 1 \end{bmatrix} = A = U\Sigma V^T = \begin{bmatrix} 1/3\sqrt{2} & -1/\sqrt{2} & 2/3 \\ 4/3\sqrt{2} & 0 & -1/3 \\ 1/3\sqrt{2} & 1/\sqrt{2} & 2/3 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^T.$$

A few more observations of the various elements of the procedure for finding a singular value decomposition are in order.

Theorem 9. Let A be an $m \times n$ matrix. Suppose that $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of the symmetric matrix A^TA arranged so that the corresponding eigenvalues of A^TA are $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$. If A has r nonzero singular values, then $\{A\vec{v}_1, A\vec{v}_2, \ldots, A\vec{v}_r\}$ is an orthogonal basis for $\text{Col}(A)$ and $\text{rank}(A) = r$.

Proof. Because $\vec{v}_i$ and $\lambda_j\vec{v}_j$ are orthogonal when $i \neq j$, we have

$$(A\vec{v}_i)^T(A\vec{v}_j) = \vec{v}_i^T A^T A\vec{v}_j = \vec{v}_i(\lambda\vec{v}_j) = 0.$$

This means that $\{A\vec{v}_1, A\vec{v}_2, \ldots, A\vec{v}_n\}$ is an orthogonal set.

Furthermore, since

$$\|A\vec{v}_i\|^2 = (A\vec{v}_i)^T A\vec{v}_i = \vec{v}_i^T A^T A\vec{v}_i = \vec{v}_i^T(\lambda\vec{v}_i) = \lambda_i(\vec{v}_i^T\vec{v}_i) = \lambda_i\|\vec{v}_i\|^2 = \lambda_i,$$

then each $\lambda_i \geq 0$ and $A\vec{v}_i \neq \vec{0}$ only when $\lambda_i \neq 0$.

Thus the orthogonal set of nonzero vectors $\{A\vec{v}_1, A\vec{v}_2, \ldots, A\vec{v}_r\}$ is linearly independent.

Now we have that $\text{Span}\{A\vec{v}_1, A\vec{v}_2, \ldots, A\vec{v}_r\}$ is a subspace of $\text{Col}(A)$.

To get the other inclusion, we notice for $\vec{y} \in \text{Col}(A)$ that there is

$$\vec{x} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n \in \mathbb{R}^m$$

such that

$$\begin{aligned}\vec{y} = A\vec{x} &= c_1A\vec{v}_1 + c_2A\vec{v}_2 + \cdots + c_rA\vec{v}_r + c_{r+1}A\vec{v}_{r+1} + \cdots + c_nA\vec{v}_n \\ &= c_1A\vec{v}_1 + c_2A\vec{v}_2 + \cdots + c_rA\vec{v}_r + \vec{0}.\end{aligned}$$

Thus $\vec{y}$ belongs to $\text{Span}\{A\vec{v}_1, A\vec{v}_2, \ldots, A\vec{v}_r\}$.

Since $\{A\vec{v}_1, A\vec{v}_2, \ldots, A\vec{v}_r\}$ is a linearly independent spanning set for $\text{Col}(A)$, the dimension of $\text{Col}(A)$ is r, and so $\text{rank}(A) = r$. □

Example. Find a singular value decomposition of

$$A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}.$$

Step 1. The matrix V is an orthogonal matrix that diagonalizes the symmetric matrix

$$A^TA = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}.$$

The eigenvalues of A^TA are $\lambda_1 = 360$, $\lambda_2 = 90$, and $\lambda_3 = 0$, and corresponding unit eigenvectors are

$$\vec{v}_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}, \; \vec{v}_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}, \; \vec{v}_3 = \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix}.$$

These are the columns of V.

Step 2. The singular values of A are $\sigma_1 = \sqrt{\lambda_1} = 6\sqrt{10}$, $\sigma_2 = \sqrt{\lambda_2} = 3\sqrt{10}$, and $\sigma_3 = \sqrt{\lambda_3} = 0$.

The matrix Σ (of the same size as A) is

$$\Sigma = \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}.$$

Step 3. Construct the matrix U.

The $r = \text{rank}(A)$ nonzero singular values of A determine the first r columns of U.

From $AV = U\Sigma$, we have

$$\vec{u}_1 = \frac{A\vec{v}_1}{\sigma_1} = \frac{1}{6\sqrt{10}} \begin{bmatrix} 18 \\ 6 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix},$$
$$\vec{u}_2 = \frac{A\vec{v}_2}{\sigma_2} = \frac{1}{3\sqrt{10}} \begin{bmatrix} 3 \\ -9 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{10} \\ -3\sqrt{10} \end{bmatrix}.$$

Do we need to find more columns of U? No, because the vectors $\vec{u}_1$ and $\vec{v}_2$ form an orthonormal basis for $\mathbb{R}^2$ (and the dimension of $\text{Nul}(A^T)$ is 0 too).

A singular value decomposition of A is

$$\begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix} \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 & 0 \end{bmatrix} \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}^T.$$

Is this singular value decomposition of A unique? No because replacing $\vec{v}_1$ with $-\vec{v}_1$ and $\vec{u}_1$ with $-\vec{u}_1$ gives another singular value decomposition of A.

Just some other things to recongize: for $m \times n$ matrix A with rank $r \geq 1$, in a singular value decomposition $A = U\Sigma V^T$ we have

(i) the columns $\vec{v}_1, \ldots, \vec{v}_r$ of V form an orthonormal basis for $\text{Col}(A^T) = \text{Row}(A)$,
(ii) the columns $\vec{v}_{r+1}, \ldots, \vec{v}_n$ of V form an orthonormal basis for $\text{Nul}(A)$ when $r < n$,
(iii) the columns $\vec{u}_1, \ldots, \vec{u}_r$ of U form an orthonormal basis for $\text{Col}(A)$,
(iv) the columns $\vec{u}_{r+1}, \ldots, \vec{u}_m$ of U form an orthonormal basis for $\text{Nul}(A^T)$ when $r < m$.