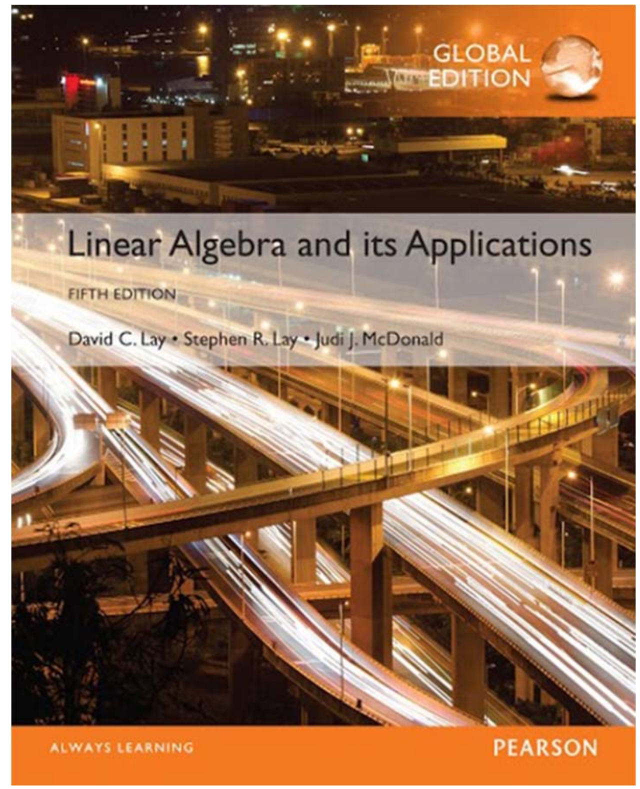
7 Symmetric Matrices and Quadratic Forms

7.4

THE SINGULAR VALUE DECOMPOSITION



PEARSON

THE SINGULAR VALUES OF AN $m \times n$ MATRIX

- Let A be an $m \times n$ matrix.
  - A^TA is symmetric and can be orthogonally diagonalized.
- Let $\{v_1, \ldots, v_n\}$ is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of A^TA and $\lambda_1 \geq \cdots \geq \lambda_n$ be the associated eigenvalues of A^TA.

$$\begin{aligned} \|A\mathbf{v}_i\|^2 &= (A\mathbf{v}_i)^T A\mathbf{v}_i = \mathbf{v}_i^T A^T A\mathbf{v}_i \\ &= \mathbf{v}_i^T(\lambda_i \mathbf{v}_i) && \text{Since } \mathbf{v}_i \text{ is an eigenvector of } A^TA \\ &= \lambda_i && \text{Since } \mathbf{v}_i \text{ is a unit vector} \end{aligned}$$

- Hence, eigenvalues of A^TA are nonnegative.

THE SINGULAR VALUES OF AN $m \times n$ MATRIX

- **Singular values of A:**
  - Square roots of the eigenvalues of A^TA

- **Theorem 9** Suppose $\{v_1, \ldots, v_n\}$ is an orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of A^TA, arranged so that the corresponding eigenvalues of A^TA satisfy $\lambda_1 \geq \cdots \geq \lambda_n$, and suppose A has r nonzero singular values. Then $\{Av_1, \ldots, Av_r\}$ is an orthogonal basis for Col A, and rank $A = r$.

THE SINGULAR VALUES OF AN $m \times n$ MATRIX

- **Proof** Because v_i and $\lambda_j v_j$ are orthogonal for $i \neq j$,
$$(Av_i)^T(Av_j) = v_i^T A^T A v_j = v_i^T(\lambda_j v_j) = 0$$

- Thus $\{Av_1, \ldots, Av_n\}$ is an orthogonal set.

- Since the lengths of the vectors $Av_1, \ldots, Av_n$ are the singular values of A, and since there are r nonzero singular values, $Av_i \neq 0$ if and only if $1 \leq i \leq r$.

- So $Av_1, \ldots, Av_r$ are linearly independent vectors, and they are in Col A.

- Finally, for any y in Col A—say, y = Ax—we can write $x = c_1v_1 + \cdots + c_nv_n$, and

$$
\begin{aligned}
y &= Ax \\
&= c_1Av_1 + \cdots + c_rAv_r + c_{r+1}Av_{r+1} + \cdots \\
&\quad + c_nAv_n
\end{aligned}
$$

THE SINGULAR VALUES OF AN $m \times n$ MATRIX

$$= c_1 A v_1 + \cdots + c_r A v_r + 0 + \cdots + 0$$

- Thus **y** is in Span$\{Av_1, \ldots, Av_r\}$, which shows that

$$\{Av_1, \ldots, Av_r\}$$

is an (orthogonal) basis for Col A.

- Hence:

$$\text{rank } A = \dim \text{Col } A = r.$$

THE SINGULAR VALUE DECOMPOSITION

- **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 Σ as:

$$\Sigma = \begin{array}{ll} \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} & \leftarrow m - r \text{ rows} \\ \quad\quad\ \uparrow n - r \text{ columns} & \end{array} \tag{3}$$

for which the diagonal entries in D are the first r singular values of A, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, and there exist an $m \times m$ orthogonal matrix U and an $n \times n$ orthogonal matrix V such that

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

THE SINGULAR VALUE DECOMPOSITION

- Any factorization $A = U\Sigma V^T$, with U and V orthogonal, Σ as in (3), 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 λ_i and v_i be as in Theorem 9, so that $\{Av_1, \ldots, A\mathrm{v}_\mathrm{r}\}$ is an orthogonal basis for Col A.

THE SINGULAR VALUE DECOMPOSITION

- Normalize each Av_i to obtain an orthonormal basis $\{u_1, \ldots, u_r\}$, where

$$u_i = \frac{1}{\|Av_i\|} Av_i = \frac{1}{\sigma_i} Av_i$$

- And

$$Av_i = \sigma_i u_i \qquad (1 \leq i \leq r) \tag{4}$$

- Now extend $\{u_1, \ldots, u_r\}$ to an orthonormal basis $\{u_1, \ldots, u_m\}$ of $\mathbb{R}^m$, and let

$$U = [u_1 \;\; u_2 \;\; \ldots \;\; u_m] \quad \text{and} \quad V = [v_1 \;\; v_2 \;\; \ldots \;\; v_n]$$

- By construction, U and V are orthogonal matrices.

THE SINGULAR VALUE DECOMPOSITION

- Also, from (4),

  $AV = [Av_1 \ \ldots \ Av_r \ \ 0 \ \ldots \ 0] = [\sigma_1 u_1 \ldots \sigma_r u_r \ \ 0 \ldots 0]$
- Let D be the diagonal matrix with diagonal entries $\sigma_1, \ldots, \sigma_r$, and let Σ be as in (3) above. Then

$$U\Sigma = [\,\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_m\,] \left[\begin{array}{cccc|c} \sigma_1 & & & 0 & \\ & \sigma_2 & & & 0 \\ & & \ddots & & \\ 0 & & & \sigma_r & \\ \hline & & 0 & & 0 \end{array}\right]$$

$$= [\,\sigma_1\mathbf{u}_1 \quad \cdots \quad \sigma_r\mathbf{u}_r \quad \mathbf{0} \quad \cdots \quad \mathbf{0}\,]$$

$$= AV$$

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

THE SINGULAR VALUE DECOMPOSITION

- **Example 3** Use the results of Examples 1 and 2 to construct a singular value decomposition of $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$.

- **Solution** A construction can be divided into three steps.
- ***Step 1. Find an orthogonal diagonalization of A^TA***. That is, find the eigenvalues of A^TA and a corresponding orthonormal set of eigenvectors. If A had only two columns, the calculations could be done by hand. Larger matrices usually require a matrix program. However, for the matrix A here, the eigendata for A^TA are provided in Example 2.

THE SINGULAR VALUE DECOMPOSITION

We find eigenvalues and their corresponding (orthonormal) eigenvectors of A^TA.

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

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

$$\mathbf{v}_1 = \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix}, \quad \mathbf{v}_2 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}, \quad \mathbf{v}_3 = \begin{bmatrix} 2/3 \\ -2/3 \\ 1/3 \end{bmatrix}$$

THE SINGULAR VALUE DECOMPOSITION

- ***Step 2*. *Set up V and Σ*.** Arrange the eigenvalues of A^TA in decreasing order. In Example 1, the eigenvalues are already listed in decreasing order: 360, 90, and 0. The corresponding unit eigenvectors, $\mathbf{v}_1$, $\mathbf{v}_2$, and $\mathbf{v}_3$, are the right singular vectors of A. Using Example 1, construct

$$V = [v_1 \ \ v_2 \ \ v_3] = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}$$

- The square roots of the eigenvalues are the singular values:

$$\sigma_1 = 6\sqrt{10}, \qquad \sigma_2 = 3\sqrt{10}, \qquad \sigma_3 = 0$$

THE SINGULAR VALUE DECOMPOSITION

- The nonzero singular values are the diagonal entries of *D*. The matrix Σ is the same size as *A*, with *D* in its upper left corner and with 0's elsewhere.

$$D = \begin{bmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \end{bmatrix}, \qquad \Sigma = [D \;\; 0] = \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix}$$

- ***Step 3. Construct U***. When *A* has rank *r*, the first *r* columns of *U* are the normalized vectors obtained from $Av_1, \ldots, Av_r$. In this example, A has two nonzero singular values, so rank A = 2. Recall from equation (2) and the paragraph before Example 2 that $\|Av_1\| = \sigma_1$ and $\|Av_2\| = \sigma_2$.

THE SINGULAR VALUE DECOMPOSITION

- Thus

$$u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{6\sqrt{10}} \begin{bmatrix} 18 \\ 6 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}$$

$$u_2 = \frac{1}{\sigma_2} A v_2 = \frac{1}{3\sqrt{10}} \begin{bmatrix} 3 \\ -9 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$$

- Note that $\{\mathbf{u}_1, \mathbf{u}_2\}$ is already a basis for $\mathbb{R}^2$. Thus no additional vectors are needed for U, and $U = [\mathbf{u}_1 \;\; \mathbf{u}_2]$. The singular value decomposition of A is

$$A = \begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3/\sqrt{10} \end{bmatrix} \begin{bmatrix} 6/\sqrt{10} & 0 & 0 \\ 0 & 3/\sqrt{10} & 0 \end{bmatrix} \begin{bmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}$$

THE SINGULAR VALUE DECOMPOSITION

- **Theorem: The Invertible Matrix Theorem (concluded)**
- Let A be an $n \times n$ matrix. Then the following statements are each equivalent to the statement that A is an invertible matrix.

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

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

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

x. A has n nonzero singular values.