

Math 221: LINEAR ALGEBRA

Chapter 8. Orthogonality

§8-6. Singular Value Decomposition

Le Chen¹

Emory University, 2021 Spring

(last updated on 01/25/2021)



Creative Commons License
(CC BY-NC-SA)

¹Slides are adapted from those by Karen Seyffarth from University of Calgary.

Singular Value Decomposition

Examples

Fundamental Subspaces

Applications

Singular Value Decomposition

Examples

Fundamental Subspaces

Applications

Singular Value Decomposition

The diagram illustrates the Singular Value Decomposition (SVD) of a matrix M . It shows the decomposition $M = U \Sigma V^*$ with dimensions and visual representations of each matrix.

M is an $m \times n$ matrix, represented by a 4x4 grid.

U is an $m \times m$ matrix, represented by a 4x4 grid with colored columns (purple, brown, purple, purple).

Σ is an $m \times n$ matrix, represented by a 4x4 grid with blue squares on the diagonal and zeros elsewhere.

V^* is an $n \times n$ matrix, represented by a 4x4 grid with colored rows (green, green, green, green).

The decomposition is shown as:

$$M = U \Sigma V^*$$

Dimensions: $m \times n = m \times m \quad m \times n \quad n \times n$

U is an $m \times m$ matrix, represented by a 4x4 grid with colored columns (purple, brown, purple, purple).

U^* is an $m \times m$ matrix, represented by a 4x4 grid with colored rows (purple, brown, purple, purple).

I_m is an $m \times m$ identity matrix, represented by a 4x4 grid with 1s on the diagonal and 0s elsewhere.

V is an $n \times n$ matrix, represented by a 4x4 grid with colored columns (green, green, green, green).

V^* is an $n \times n$ matrix, represented by a 4x4 grid with colored rows (green, green, green, green).

I_n is an $n \times n$ identity matrix, represented by a 4x4 grid with 1s on the diagonal and 0s elsewhere.

$$U U^* = I_m$$
$$V V^* = I_n$$

Definition

Let A be an $m \times n$ matrix. The **singular values** of A are the square roots of the nonzero eigenvalues of $A^T A$.

Definition

Let A be an $m \times n$ matrix. The **singular values** of A are the square roots of the nonzero eigenvalues of $A^T A$. **Singular Value Decomposition (SVD)** can be thought of as a generalization of orthogonal diagonalization of a symmetric matrix to an arbitrary $m \times n$ matrix.

Definition

Let A be an $m \times n$ matrix. The **singular values** of A are the square roots of the nonzero eigenvalues of $A^T A$. **Singular Value Decomposition (SVD)** can be thought of as a generalization of orthogonal diagonalization of a symmetric matrix to an arbitrary $m \times n$ matrix.

Given an $m \times n$ matrix A , we will see how to express A as a product

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

where

Definition

Let A be an $m \times n$ matrix. The **singular values** of A are the square roots of the nonzero eigenvalues of $A^T A$. **Singular Value Decomposition (SVD)** can be thought of as a generalization of orthogonal diagonalization of a symmetric matrix to an arbitrary $m \times n$ matrix.

Given an $m \times n$ matrix A , we will see how to express A as a product

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

where

- ▶ U is an $m \times m$ orthogonal matrix whose columns are eigenvectors of AA^T .

Definition

Let A be an $m \times n$ matrix. The **singular values** of A are the square roots of the nonzero eigenvalues of $A^T A$. **Singular Value Decomposition (SVD)** can be thought of as a generalization of orthogonal diagonalization of a symmetric matrix to an arbitrary $m \times n$ matrix.

Given an $m \times n$ matrix A , we will see how to express A as a product

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

where

- ▶ U is an $m \times m$ orthogonal matrix whose columns are eigenvectors of AA^T .
- ▶ V is an $n \times n$ orthogonal matrix whose columns are eigenvectors of $A^T A$.

Definition

Let A be an $m \times n$ matrix. The **singular values** of A are the square roots of the nonzero eigenvalues of $A^T A$. **Singular Value Decomposition (SVD)** can be thought of as a generalization of orthogonal diagonalization of a symmetric matrix to an arbitrary $m \times n$ matrix.

Given an $m \times n$ matrix A , we will see how to express A as a product

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

where

- ▶ U is an $m \times m$ orthogonal matrix whose columns are eigenvectors of AA^T .
- ▶ V is an $n \times n$ orthogonal matrix whose columns are eigenvectors of $A^T A$.
- ▶ Σ is an $m \times n$ matrix whose only nonzero values lie on its main diagonal, and are the square roots of the eigenvalues of both AA^T and $A^T A$.

Theorem

If A is an $m \times n$ matrix, then $A^T A$ and AA^T have the same nonzero eigenvalues.

Theorem

If A is an $m \times n$ matrix, then $A^T A$ and AA^T have the same nonzero eigenvalues.

Proof.

Suppose A is an $m \times n$ matrix, and suppose that λ is a nonzero eigenvalue of $A^T A$. Then there exists a nonzero vector $\vec{x} \in \mathbb{R}^n$ such that

$$(A^T A)\vec{x} = \lambda\vec{x}. \quad (1)$$

Multiplying both sides of this equation by A :

$$\begin{aligned} A(A^T A)\vec{x} &= A\lambda\vec{x} \\ (AA^T)(A\vec{x}) &= \lambda(A\vec{x}). \end{aligned}$$

Since $\lambda \neq 0$ and $\vec{x} \neq \vec{0}_n$, $\lambda\vec{x} \neq \vec{0}_n$, and thus by equation (1), $(A^T A)\vec{x} \neq \vec{0}_n$; thus $A^T(A\vec{x}) \neq \vec{0}_n$, implying that $A\vec{x} \neq \vec{0}_m$.

Therefore $A\vec{x}$ is an eigenvector of AA^T corresponding to eigenvalue λ . An analogous argument can be used to show that every nonzero eigenvalue of AA^T is an eigenvalue of $A^T A$, thus completing the proof. ■

Singular Value Decomposition

Examples

Fundamental Subspaces

Applications

Examples

Example

Let $A = \begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix}$. Then

$$AA^T = \begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ -1 & 1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 11 & 5 \\ 5 & 11 \end{bmatrix}.$$

$$A^T A = \begin{bmatrix} 1 & 3 \\ -1 & 1 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 10 & 2 & 6 \\ 2 & 2 & -2 \\ 6 & -2 & 10 \end{bmatrix}.$$

Example (continued)

Since AA^T is 2×2 while $A^T A$ is 3×3 , and AA^T and $A^T A$ have the same nonzero eigenvalues, compute $c_{AA^T}(x)$ (because it's easier to compute than $c_{A^T A}(x)$).

$$\begin{aligned}c_{AA^T}(x) &= \det(xI - AA^T) = \begin{vmatrix} x - 11 & -5 \\ -5 & x - 11 \end{vmatrix} \\ &= (x - 11)^2 - 25 \\ &= x^2 - 22x + 121 - 25 \\ &= x^2 - 22x + 96 \\ &= (x - 16)(x - 6).\end{aligned}$$

Therefore, the eigenvalues of AA^T are $\lambda_1 = 16$ and $\lambda_2 = 6$.

Example (continued)

The eigenvalues of $A^T A$ are $\lambda_1 = 16$, $\lambda_2 = 6$, and $\lambda_3 = 0$, and the singular values of A are $\sigma_1 = \sqrt{16} = 4$ and $\sigma_2 = \sqrt{6}$. By convention, we list the eigenvalues (and corresponding singular values) in nonincreasing order (i.e., from largest to smallest).

Example (continued)

The eigenvalues of $A^T A$ are $\lambda_1 = 16$, $\lambda_2 = 6$, and $\lambda_3 = 0$, and the singular values of A are $\sigma_1 = \sqrt{16} = 4$ and $\sigma_2 = \sqrt{6}$. By convention, we list the eigenvalues (and corresponding singular values) in nonincreasing order (i.e., from largest to smallest).

To find the matrix V , find eigenvectors for $A^T A$. Since the eigenvalues of AA^T are distinct, the corresponding eigenvectors are orthogonal, and we need only normalize them.

Example (continued)

The eigenvalues of $A^T A$ are $\lambda_1 = 16$, $\lambda_2 = 6$, and $\lambda_3 = 0$, and the singular values of A are $\sigma_1 = \sqrt{16} = 4$ and $\sigma_2 = \sqrt{6}$. By convention, we list the eigenvalues (and corresponding singular values) in nonincreasing order (i.e., from largest to smallest).

To find the matrix V , find eigenvectors for $A^T A$. Since the eigenvalues of AA^T are distinct, the corresponding eigenvectors are orthogonal, and we need only normalize them.

$\lambda_1 = 16$: solve $(16I - A^T A)\vec{y}_1 = \vec{0}$.

$$\left[\begin{array}{ccc|c} 6 & -2 & -6 & 0 \\ -2 & 14 & 2 & 0 \\ -6 & 2 & 6 & 0 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right], \text{ so } \vec{y}_1 = \begin{bmatrix} t \\ 0 \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, t \in \mathbb{R}.$$

Example (continued)

The eigenvalues of $A^T A$ are $\lambda_1 = 16$, $\lambda_2 = 6$, and $\lambda_3 = 0$, and the singular values of A are $\sigma_1 = \sqrt{16} = 4$ and $\sigma_2 = \sqrt{6}$. By convention, we list the eigenvalues (and corresponding singular values) in nonincreasing order (i.e., from largest to smallest).

To find the matrix V , find eigenvectors for $A^T A$. Since the eigenvalues of AA^T are distinct, the corresponding eigenvectors are orthogonal, and we need only normalize them.

$\lambda_1 = 16$: solve $(16I - A^T A)\vec{y}_1 = \vec{0}$.

$$\left[\begin{array}{ccc|c} 6 & -2 & -6 & 0 \\ -2 & 14 & 2 & 0 \\ -6 & 2 & 6 & 0 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right], \text{ so } \vec{y}_1 = \begin{bmatrix} t \\ 0 \\ t \end{bmatrix} = t \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, t \in \mathbb{R}.$$

$\lambda_2 = 6$: solve $(6I - A^T A)\vec{y}_2 = \vec{0}$.

$$\left[\begin{array}{ccc|c} -4 & -2 & -6 & 0 \\ -2 & 4 & 2 & 0 \\ -6 & 2 & -4 & 0 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right], \text{ so } \vec{y}_2 = \begin{bmatrix} -s \\ -s \\ s \end{bmatrix} = s \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, s \in \mathbb{R}.$$

Example (continued)

$\lambda_3 = 0$: solve $(-A^T A)\vec{y}_3 = \vec{0}$.

$$\left[\begin{array}{ccc|c} -10 & -2 & -6 & 0 \\ -2 & -2 & 2 & 0 \\ -6 & 2 & -10 & 0 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right], \text{ so } \vec{y}_3 = \begin{bmatrix} -r \\ 2r \\ r \end{bmatrix} = r \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, r \in \mathbb{R}.$$

Example (continued)

$\lambda_3 = 0$: solve $(-A^T A)\vec{y}_3 = \vec{0}$.

$$\left[\begin{array}{ccc|c} -10 & -2 & -6 & 0 \\ -2 & -2 & 2 & 0 \\ -6 & 2 & -10 & 0 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right], \text{ so } \vec{y}_3 = \begin{bmatrix} -r \\ 2r \\ r \end{bmatrix} = r \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, r \in \mathbb{R}.$$

Let

$$\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \vec{v}_2 = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, \vec{v}_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}.$$

Then

$$V = \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{3} & -\sqrt{2} & -1 \\ 0 & -\sqrt{2} & 2 \\ \sqrt{3} & \sqrt{2} & 1 \end{bmatrix}.$$

Example (continued)

$\lambda_3 = 0$: solve $(-A^T A)\vec{y}_3 = \vec{0}$.

$$\left[\begin{array}{ccc|c} -10 & -2 & -6 & 0 \\ -2 & -2 & 2 & 0 \\ -6 & 2 & -10 & 0 \end{array} \right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right], \text{ so } \vec{y}_3 = \begin{bmatrix} -r \\ 2r \\ r \end{bmatrix} = r \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}, r \in \mathbb{R}.$$

Let

$$\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \vec{v}_2 = \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}, \vec{v}_3 = \frac{1}{\sqrt{6}} \begin{bmatrix} -1 \\ 2 \\ 1 \end{bmatrix}.$$

Then

$$V = \frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{3} & -\sqrt{2} & -1 \\ 0 & -\sqrt{2} & 2 \\ \sqrt{3} & \sqrt{2} & 1 \end{bmatrix}.$$

Also,

$$\Sigma = \begin{bmatrix} 4 & 0 & 0 \\ 0 & \sqrt{6} & 0 \end{bmatrix},$$

and we use A , V^T , and Σ to find U .

Example (continued)

Since V is orthogonal and $A = U\Sigma V^T$, it follows that $AV = U\Sigma$. Let $V = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix}$, and let $U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 \end{bmatrix}$, where \vec{u}_1 and \vec{u}_2 are the two columns of U .

Example (continued)

Since V is orthogonal and $A = U\Sigma V^T$, it follows that $AV = U\Sigma$. Let $V = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix}$, and let $U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 \end{bmatrix}$, where \vec{u}_1 and \vec{u}_2 are the two columns of U . Then we have

$$\begin{aligned} A \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} &= \begin{bmatrix} \vec{u}_1 & \vec{u}_2 \end{bmatrix} \Sigma \\ \begin{bmatrix} A\vec{v}_1 & A\vec{v}_2 & A\vec{v}_3 \end{bmatrix} &= \begin{bmatrix} \sigma_1\vec{u}_1 + 0\vec{u}_2 & 0\vec{u}_1 + \sigma_2\vec{u}_2 & 0\vec{u}_1 + 0\vec{u}_2 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1\vec{u}_1 & \sigma_2\vec{u}_2 & \vec{0} \end{bmatrix} \end{aligned}$$

which implies that $A\vec{v}_1 = \sigma_1\vec{u}_1 = 4\vec{u}_1$ and $A\vec{v}_2 = \sigma_2\vec{u}_2 = \sqrt{6}\vec{u}_2$.

Example (continued)

Since V is orthogonal and $A = U\Sigma V^T$, it follows that $AV = U\Sigma$. Let $V = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix}$, and let $U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 \end{bmatrix}$, where \vec{u}_1 and \vec{u}_2 are the two columns of U . Then we have

$$\begin{aligned} A \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} &= \begin{bmatrix} \vec{u}_1 & \vec{u}_2 \end{bmatrix} \Sigma \\ \begin{bmatrix} A\vec{v}_1 & A\vec{v}_2 & A\vec{v}_3 \end{bmatrix} &= \begin{bmatrix} \sigma_1\vec{u}_1 + 0\vec{u}_2 & 0\vec{u}_1 + \sigma_2\vec{u}_2 & 0\vec{u}_1 + 0\vec{u}_2 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1\vec{u}_1 & \sigma_2\vec{u}_2 & \vec{0} \end{bmatrix} \end{aligned}$$

which implies that $A\vec{v}_1 = \sigma_1\vec{u}_1 = 4\vec{u}_1$ and $A\vec{v}_2 = \sigma_2\vec{u}_2 = \sqrt{6}\vec{u}_2$. Thus,

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

and

$$\vec{u}_2 = \frac{1}{\sqrt{6}}A\vec{v}_2 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix} \frac{1}{\sqrt{3}} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = \frac{1}{3\sqrt{2}} \begin{bmatrix} 3 \\ -3 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

Example (continued)

Therefore,

$$U = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix},$$

and

$$\begin{aligned} A &= \begin{bmatrix} 1 & -1 & 3 \\ 3 & 1 & 1 \end{bmatrix} \\ &= \left(\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \right) \begin{bmatrix} 4 & 0 & 0 \\ 0 & \sqrt{6} & 0 \end{bmatrix} \left(\frac{1}{\sqrt{6}} \begin{bmatrix} \sqrt{3} & 0 & \sqrt{3} \\ -\sqrt{2} & -\sqrt{2} & \sqrt{2} \\ -1 & 2 & 1 \end{bmatrix} \right). \end{aligned}$$

Problem

Find an SVD for $A = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$.

Solution

Since A is 3×1 , $A^T A$ is a 1×1 matrix whose eigenvalues are easier to find than the eigenvalues of the 3×3 matrix AA^T .

$$A^T A = \begin{bmatrix} -1 & 2 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 9 \end{bmatrix}.$$

Thus $A^T A$ has eigenvalue $\lambda_1 = 9$, and the eigenvalues of AA^T are $\lambda_1 = 9$, $\lambda_2 = 0$, and $\lambda_3 = 0$. Furthermore, A has only one singular value, $\sigma_1 = 3$.

Problem

Find an SVD for $A = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}$.

Solution

Since A is 3×1 , $A^T A$ is a 1×1 matrix whose eigenvalues are easier to find than the eigenvalues of the 3×3 matrix AA^T .

$$A^T A = \begin{bmatrix} -1 & 2 & 2 \end{bmatrix} \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 9 \end{bmatrix}.$$

Thus $A^T A$ has eigenvalue $\lambda_1 = 9$, and the eigenvalues of AA^T are $\lambda_1 = 9$, $\lambda_2 = 0$, and $\lambda_3 = 0$. Furthermore, A has only one singular value, $\sigma_1 = 3$.

To find the matrix V , find an eigenvector for $A^T A$ and normalize it. In this case, finding a unit eigenvector is trivial: $\vec{v}_1 = \begin{bmatrix} 1 \end{bmatrix}$, and

$$V = \begin{bmatrix} 1 \end{bmatrix}.$$

Solution (continued)

Also, $\Sigma = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$, and we use A , V^T , and Σ to find U .

Solution (continued)

Also, $\Sigma = \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix}$, and we use A , V^T , and Σ to find U .

Now $AV = U\Sigma$, with $V = \begin{bmatrix} \vec{v}_1 \end{bmatrix}$, and $U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{bmatrix}$, where \vec{u}_1 , \vec{u}_2 , and \vec{u}_3 are the columns of U . Thus

$$\begin{aligned} A \begin{bmatrix} \vec{v}_1 \end{bmatrix} &= \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \vec{u}_3 \end{bmatrix} \Sigma \\ \begin{bmatrix} A\vec{v}_1 \end{bmatrix} &= \begin{bmatrix} \sigma_1\vec{u}_1 + 0\vec{u}_2 + 0\vec{u}_3 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1\vec{u}_1 \end{bmatrix} \end{aligned}$$

This gives us $A\vec{v}_1 = \sigma_1\vec{u}_1 = 3\vec{u}_1$, so

$$\vec{u}_1 = \frac{1}{3}A\vec{v}_1 = \frac{1}{3} \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} \begin{bmatrix} 1 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}.$$

Solution (continued)

The vectors \vec{u}_2 and \vec{u}_3 are eigenvectors of AA^T corresponding to the eigenvalue $\lambda_2 = \lambda_3 = 0$. Instead of solving the system $(0I - AA^T)\vec{x} = \vec{0}$ and then using the Gram-Schmidt orthogonalization algorithm on the resulting set of two basic eigenvectors, the following approach may be used.

Find vectors \vec{u}_2 and \vec{u}_3 by first extending $\{\vec{u}_1\}$ to a basis of \mathbb{R}^3 , then using the Gram-Schmidt algorithm to orthogonalize the basis, and finally normalizing the vectors.

Solution (continued)

The vectors \vec{u}_2 and \vec{u}_3 are eigenvectors of AA^T corresponding to the eigenvalue $\lambda_2 = \lambda_3 = 0$. Instead of solving the system $(0I - AA^T)\vec{x} = \vec{0}$ and then using the Gram-Schmidt orthogonalization algorithm on the resulting set of two basic eigenvectors, the following approach may be used.

Find vectors \vec{u}_2 and \vec{u}_3 by first extending $\{\vec{u}_1\}$ to a basis of \mathbb{R}^3 , then using the Gram-Schmidt algorithm to orthogonalize the basis, and finally normalizing the vectors.

Starting with $\{3\vec{u}_1\}$ instead of $\{\vec{u}_1\}$ makes the arithmetic a bit easier. It is easy to verify that

$$\left\{ \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \right\}$$

is a basis of \mathbb{R}^3 . Set

$$\vec{f}_1 = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix}, \vec{x}_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \vec{x}_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix},$$

and apply the Gram-Schmidt orthogonalization algorithm to $\{\vec{f}_1, \vec{x}_2, \vec{x}_3\}$.

Solution (continued)

This gives us

$$\vec{f}_2 = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \vec{f}_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

Therefore,

$$\vec{u}_2 = \frac{1}{\sqrt{18}} \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{u}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix},$$

and

$$U = \begin{bmatrix} -\frac{1}{3} & \frac{4}{\sqrt{18}} & 0 \\ \frac{2}{3} & \frac{1}{\sqrt{18}} & \frac{1}{\sqrt{2}} \\ \frac{2}{3} & \frac{1}{\sqrt{18}} & -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Solution (continued)

This gives us

$$\vec{f}_2 = \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix} \quad \text{and} \quad \vec{f}_3 = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}.$$

Therefore,

$$\vec{u}_2 = \frac{1}{\sqrt{18}} \begin{bmatrix} 4 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{u}_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix},$$

and

$$U = \begin{bmatrix} -\frac{1}{3} & \frac{4}{\sqrt{18}} & 0 \\ \frac{2}{3} & \frac{1}{\sqrt{18}} & \frac{1}{\sqrt{2}} \\ \frac{2}{3} & \frac{1}{\sqrt{18}} & -\frac{1}{\sqrt{2}} \end{bmatrix}.$$

Finally,

$$A = \begin{bmatrix} -1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} & \frac{4}{\sqrt{18}} & 0 \\ \frac{2}{3} & \frac{1}{\sqrt{18}} & \frac{1}{\sqrt{2}} \\ \frac{2}{3} & \frac{1}{\sqrt{18}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} [1].$$



Problem

Find a singular value decomposition of $A = \begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix}$.

Problem

Find a singular value decomposition of $A = \begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix}$.

Solution

$$\begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix} = \left(\frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} \right) \begin{bmatrix} \sqrt{85} & 0 \\ 0 & 0 \end{bmatrix} \left(\frac{1}{\sqrt{17}} \begin{bmatrix} 1 & -4 \\ 4 & 1 \end{bmatrix} \right).$$



Problem

Find a singular value decomposition of $A = \begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix}$.

Solution

$$\begin{bmatrix} 1 & 4 \\ 2 & 8 \end{bmatrix} = \left(\frac{1}{\sqrt{5}} \begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix} \right) \begin{bmatrix} \sqrt{85} & 0 \\ 0 & 0 \end{bmatrix} \left(\frac{1}{\sqrt{17}} \begin{bmatrix} 1 & -4 \\ 4 & 1 \end{bmatrix} \right).$$



Remark

Since there is only one non-zero eigenvalue, \vec{u}_2 (the second column of U) can not be found using the formula $\vec{u}_2 = \frac{1}{\sigma_2} A\vec{v}_2$. However, \vec{u}_2 can be chosen to be any unit vector orthogonal to \vec{u}_1 ; in this case, $\vec{u}_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

Problem

Find a singular value decomposition of $A = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$.

Problem

Find a singular value decomposition of $A = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix}$.

Solution

$$\begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \\ \parallel \\ \left(\frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} \right) \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \left(\frac{1}{\sqrt{6}} \begin{bmatrix} 1 & -2 & 1 \\ -\sqrt{3} & 0 & \sqrt{3} \\ \sqrt{2} & \sqrt{2} & \sqrt{2} \end{bmatrix} \right)$$



Singular Value Decomposition

Examples

Fundamental Subspaces

Applications

Fundamental Subspaces

Full Singular Value Decomposition

$$\begin{matrix} m & & n \\ & & A \end{matrix} \approx \begin{matrix} & & k & & \Sigma & & V^T \\ & & \begin{matrix} | \\ | \\ | \\ | \end{matrix} & & \begin{matrix} \diagdown \\ \diagup \end{matrix} & & \begin{matrix} \equiv \\ \equiv \\ \equiv \end{matrix} \end{matrix}$$

$$= \begin{matrix} & & k & & \Sigma & & V^T \\ & & \begin{matrix} | \\ | \\ | \\ | \end{matrix} & \begin{matrix} | \\ | \\ | \\ | \end{matrix} & \begin{matrix} \diagdown \\ \diagup \\ \bigcirc \end{matrix} & & \begin{matrix} \square \end{matrix} \end{matrix}$$

The diagram illustrates the Full Singular Value Decomposition (SVD) of a matrix A of size $m \times n$. The matrix A is decomposed into three matrices: U (size $m \times k$), Σ (size $m \times n$), and V^T (size $k \times n$). The matrix U is represented by a vertical rectangle with four vertical lines. The matrix Σ is represented by a square with a diagonal line from the top-left to the bottom-right. The matrix V^T is represented by a square with three horizontal lines. The decomposition is shown as $A \approx U \Sigma V^T$. The second part of the diagram shows a more detailed view of the matrices, where U is a vertical rectangle with four vertical lines, Σ is a square with a diagonal line and a circle in the bottom-right corner, and V^T is a square. The matrix U is labeled with k above it, and the matrix Σ is labeled with Σ above it. The matrix V^T is labeled with V^T above it.

Fundamental Subspaces


Full Singular Value Decomposition

$$\begin{aligned} {}^m \boxed{A} &= \boxed{U} \boxed{\Sigma} \boxed{V^T} \\ &= \boxed{U} \boxed{\Sigma} \boxed{V^T} \end{aligned}$$

The diagram illustrates the Full Singular Value Decomposition (SVD) of a matrix A . The matrix A is shown as a rectangle with a diagonal line from the top-left to the bottom-right, representing its fundamental subspaces. The decomposition is shown as $A = U \Sigma V^T$. The matrix U is a square matrix with three vertical bars, representing its columns. The matrix Σ is a rectangle with a diagonal line from the top-left to the bottom-right and a circle in the bottom-right corner, representing its singular values and the zero block. The matrix V^T is a rectangle with three horizontal bars, representing its rows.

Fundamental Subspaces

$$A = U \Sigma V'$$

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & & \\ & \sigma_r & & & & \\ & & 0 & & & \\ & & & \ddots & & \\ & & & & 0 & \\ & & & & & 0 \end{bmatrix}$$


$$U = [u_1, \dots, u_m]$$

$$V = [v_1, \dots, v_n]$$

base B:

$$\{u_1, \dots, u_r\}$$

$$\dim \begin{cases} \text{col}(A^T) \\ \parallel \\ \text{row}(A) \\ \parallel \\ \text{im}(A^T) \end{cases}$$

+

$$\dim \sim \text{null}(A)$$

$$\parallel \\ m$$

$$\text{base B: } \{u_{r+1}, \dots, u_m\}$$

$$\text{base B: } \{v_1, \dots, v_r\}$$

$$\dim(A) = \dim(\text{im}(A)) \sim \dim(\text{row}(A^T))$$

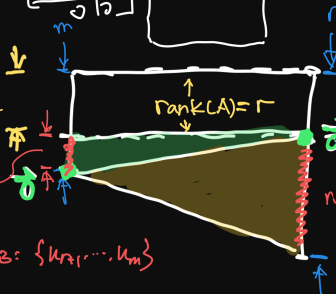
$$\dim(\text{null}(A^T)) \sim \dim(\text{row}(A))$$

+

||

n

$$\text{base B: } \{v_{r+1}, \dots, v_n\}$$



Singular Value Decomposition

Examples

Fundamental Subspaces

Applications

Applications

Example (Polar Decomposition)

$$a + bi = \underbrace{\sqrt{a^2 + b^2}}_{\text{radius}} \underbrace{e^{i\theta}}_{\text{rotation}}.$$

Similarly, any square matrix

$$A = U\Sigma V^T = \underbrace{U\Sigma U^T}_{\text{nonneg. def.}} \underbrace{UV^T}_{\text{rotation}}$$

Definition

A real $n \times n$ matrix G is **nonnegative definite** (or **positive** in the book) if it is symmetric and for all $\vec{x} \in \mathbb{R}^n$

$$\vec{x}^T G \vec{x} \geq 0.$$

Example (Generalized inverse)

$$\boxed{A} \begin{array}{|c} \hline \\ \hline \end{array} = \begin{array}{|c} \hline \\ \hline \end{array}$$

↓

$$\begin{array}{|c} \hline \\ \hline \end{array} = \boxed{A^{-1}} \begin{array}{|c} \hline \\ \hline \end{array}$$

Example (Generalized inverse)

$$\boxed{A} \begin{array}{|} \hline \\ \hline \end{array} = \begin{array}{|} \hline \\ \hline \end{array}$$

\Downarrow

$$\begin{array}{|} \hline \\ \hline \end{array} \cong \begin{array}{|} \hline A^+ \\ \hline \end{array} \begin{array}{|} \hline \\ \hline \end{array}$$

\parallel
 $A^T(AA^T)^{-1}$

Example (Generalized inverse)

$$\boxed{A} \quad \downarrow \quad = \quad |$$

$$| \quad = \quad \boxed{A^+} \quad | \quad = \quad (A^T A)^{-1} A^T$$

The diagram illustrates the relationship between a matrix A and its generalized inverse A^+ . A red vertical line is drawn between the two equations, with a blue double-headed arrow pointing downwards, indicating a correspondence or mapping between the two states.

Example (Image of unit ball under linear transform A)

Let $A = U\Sigma V^T$ be the full SVD for an $m \times n$ matrix A . We will see how the unit ball will be mapped:

$$\{A\vec{x} \mid \|\vec{x}\| \leq 1\}$$

Example (Image of unit ball under linear transform A)

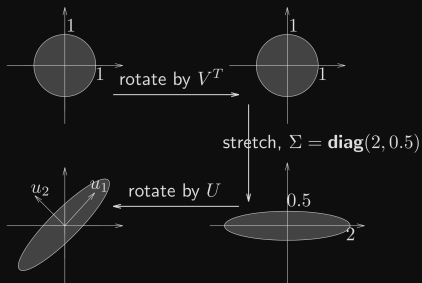
Let $A = U\Sigma V^T$ be the full SVD for an $m \times n$ matrix A. We will see how the unit ball will be mapped:

$$\{A\vec{x} \mid \|\vec{x}\| \leq 1\}$$

The linear map $\vec{y} = A\vec{x}$ is trying to do the following things:

1. Rotate the n-vector \vec{x} by V^T
2. Stretch along axes by σ_i with $\sigma_i = 0$ for $i > \text{rank}(A)$
3. Zero-pad for tall matrix (i.e., $m > n$) or truncate for fat matrix (i.e., $m < n$) to get m-vector
4. Rotate the m-vector by U^T

Example (Image of unit ball under linear transform A – continued)



Example (Image Compression)

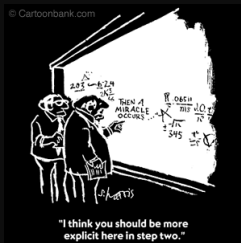
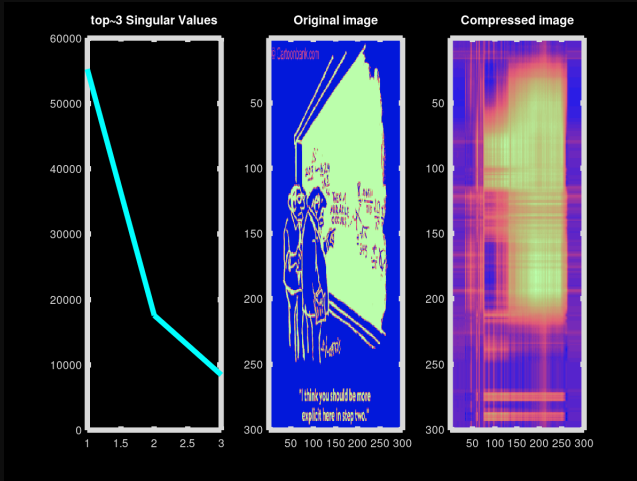


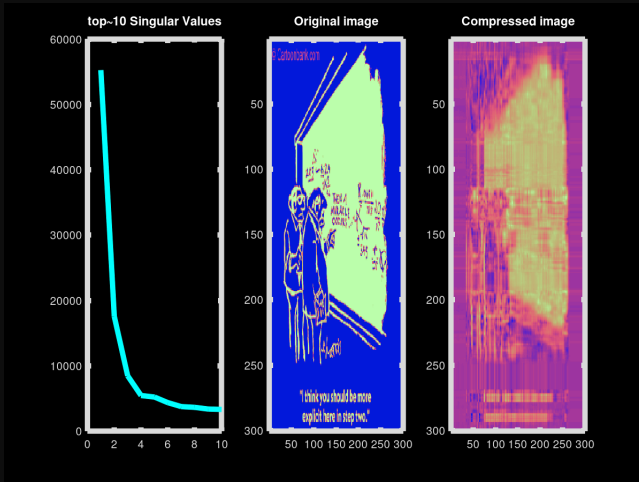
Image is a A is a 300×300 matrix.

$$A \approx \sum_{i=1}^n \sigma_i \vec{u}_i \vec{v}_i^T$$

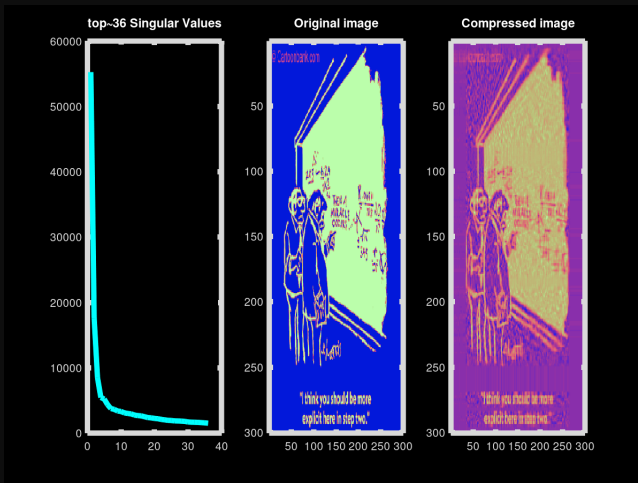
Example (Image Compression)



Example (Image Compression)



Example (Image Compression)



Example (Image Compression)

