Singular Value Decomposition

Linear Algebra

This work was initially funded by an Institutional OER Grant from the Colorado Department of Higher Education (CDHE). For similar OER materials in other courses funded by this project in the Department of Mathematical and Statistical Sciences at the University of Colorado Denver, visit
https://github.com/CU-Denver-MathStats-OER

Summary of Diagonalization

- Not every matrix can be diagonalized:

  only square matrices with n linearly independent eigenvectors

- Not every square matrix can be orthogonally diagonalized:

  only square matrices A with n orthogonal eigenvectors $\Leftrightarrow$ A is symmetric

Summary of Diagonalization

- Not every matrix can be diagonalized:

  only square matrices with n linearly independent eigenvectors

- Not every square matrix can be orthogonally diagonalized:

  only square matrices A with n orthogonal eigenvectors $\Leftrightarrow$ A is symmetric

Can we do something like diagonalization for all matrices (including non-square matrices)?

Generalizing the Diagonalization Process

$A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$ is not diagonalizable (not even square). Instead let's look at $A^T A$.

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

Generalizing the Diagonalization Process

$A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$ is not diagonalizable (not even square). Instead let's look at $A^T A$.

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

Theorem

Let A denote an $m \times n$ matrix. Then $A^T A$ can be orthogonally diagonalized.

Generalizing the Diagonalization Process

$A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$ is not diagonalizable (not even square). Instead let's look at A^TA.

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

Theorem

Let A denote an $m \times n$ matrix. Then A^TA can be orthogonally diagonalized.

Proof.

The matrix A^TA is symmetric since $(A^TA)^T = A^T(A^T)^T = A^TA$.

All symmetric matrices are orthogonally diagonalizable! $\square$

The Eigenvalues of $A^T A$

Theorem

Let A be an $m \times n$ matrix. Then the eigenvalues of $A^T A$ are nonnegative real numbers.

The Eigenvalues of $A^T A$

Theorem

Let A be an $m \times n$ matrix. Then the eigenvalues of $A^T A$ are nonnegative real numbers.

Proof.

$A^T A$ is a symmetric $n \times n$ square matrix, and is orthogonally diagonalizable.

Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$ corresponding to eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$.

The Eigenvalues of $A^T A$

Theorem

Let A be an $m \times n$ matrix. Then the eigenvalues of $A^T A$ are nonnegative real numbers.

Proof.

$A^T A$ is a symmetric $n \times n$ square matrix, and is orthogonally diagonalizable.

Let $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n\}$ be an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$ corresponding to eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Then we have

$$\|A\mathbf{v}_i\|^2 = (A\mathbf{v}_i)^T (A\mathbf{v}_i) = \mathbf{v}_i^T A^T A\mathbf{v}_i =$$

The Eigenvalues of $A^T A$

Theorem

Let A be an $m \times n$ matrix. Then the eigenvalues of $A^T A$ are nonnegative real numbers.

Proof.

$A^T A$ is a symmetric $n \times n$ square matrix, and is orthogonally diagonalizable.

Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$ corresponding to eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_n$. Then we have

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

Since the length of a vector is a nonnegative real number, λ_i is a also nonnegative real num.

Hence, all the eigenvalues of $A^T A$ are nonnegative real numbers. $\square$

Singular Values of an $m \times n$ Matrix

Let A be an $m \times n$ matrix, and let

$$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n \geq 0$$

denote the nonnegative eigenvalues of $A^T A$, arranged in descending order (λ_1 is the largest eigenvalue and λ_n is the smallest).

Definition

The singular values of A are the square roots of the eigenvalues of $A^T A$. We typically denote the singular values as $\sigma_i = \sqrt{\lambda_i}$. The singular values of A can be listed in descending order as

$$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0.$$

On the previous slide we showed that $\|A\mathbf{v}_i\|^2 = \lambda_i$.

Thus the singular value σ_i is the length of $A\mathbf{v}_i$: $\sigma_i = \|A\mathbf{v}_i\|$.

Example

Find the singular values of $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$.

Example

Find the singular values of $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$.

1. Compute the product $A^T A$.

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

2. Find the eigenvalues of $A^T A$. $\lambda_1 = 360$, $\lambda_2 = 90$, and $\lambda_3 = 0$.

3. The singular values of A are the square roots of the eigenvalues of $A^T A$. We have $\sigma_1 = \sqrt{360}$, $\sigma_2 = \sqrt{90}$, and $\sigma_3 = \sqrt{0} = 0$.

Singular Value Decomposition

Theorem

Let A be an $m \times n$ matrix with rank r. Then there exists an $m \times n$ matrix

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} \quad \text{where } D \text{ is an } r \times r \text{ diagonal submatrix}$$

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

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}V^T_{n\times n}.$$

- The factorization $A = U\Sigma V^T$ is called a singular value decomposition (or SVD for short).
- As with the diagonalization process, the matrices U and V are not uniquely determined.
- The matrix Σ is unique, if the diagonal entries are listed in descending order.

Steps for Finding Singular Value Decomposition

Example

Find a singular value decomposition of $A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}$.

1. Find an orthogonal diagonalization of $A^T A$.

1. Find an orthogonal diagonalization of A^TA.

- We have $A^TA = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}$.
- Find the eigenvalues of A^TA. We have $\lambda_1 = 360$, $\lambda_2 = 90$, and $\lambda_3 = 0$.

1. Find an orthogonal diagonalization of $A^T A$.

- We have $A^T A = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}$.
- Find the eigenvalues of $A^T A$. We have $\lambda_1 = 360$, $\lambda_2 = 90$, and $\lambda_3 = 0$.
- For each eigenvalue, find an orthogonal basis for the eigenspace (use Gram–Schmidt if needed). Corresponding to $\lambda_1 = 360$, $\lambda_2 = 90$, and $\lambda_3 = 0$, respectively, we have

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

1. Find an orthogonal diagonalization of A^TA.

- We have $A^TA = \begin{bmatrix} 4 & 8 \\ 11 & 7 \\ 14 & -2 \end{bmatrix} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} 80 & 100 & 40 \\ 100 & 170 & 140 \\ 40 & 140 & 200 \end{bmatrix}$.
- Find the eigenvalues of A^TA. We have $\lambda_1 = 360$, $\lambda_2 = 90$, and $\lambda_3 = 0$.
- For each eigenvalue, find an orthogonal basis for the eigenspace (use Gram–Schmidt if needed). Corresponding to $\lambda_1 = 360$, $\lambda_2 = 90$, and $\lambda_3 = 0$, respectively, we have

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

- Normalize to find the orthonormal columns of P.

$$A^TA = PDP^T \quad \text{where} \quad P = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix} \text{ and } D = \begin{bmatrix} 360 & 0 & 0 \\ 0 & 90 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$A^TA = PDP^T \quad \text{where} \quad P = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix} \text{ and } D = \begin{bmatrix} 360 & 0 & 0 \\ 0 & 90 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

We need to find the decomposition $A = U\Sigma V^T$.

$$A^T A = PDP^T \quad \text{where} \quad P = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix} \text{ and } D = \begin{bmatrix} 360 & 0 & 0 \\ 0 & 90 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

We need to find the decomposition $A = U\Sigma V^T$.

2. Find the matrices V and Σ.

- We have $V = P = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}$.

  The column vectors of V are called right singular vectors of A.

- We have $D = \begin{bmatrix} \sqrt{360} & 0 \\ 0 & \sqrt{90} \end{bmatrix}$ so $\Sigma = \begin{bmatrix} \sqrt{360} & 0 & 0 \\ 0 & \sqrt{90} & 0 \end{bmatrix}$.

So far we have the decomposition $A = U\Sigma V^T = U \begin{bmatrix} \sqrt{360} & 0 & 0 \\ 0 & \sqrt{90} & 0 \end{bmatrix} \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}^T$.

So far we have the decomposition $A = U\Sigma V^T = U \begin{bmatrix} \sqrt{360} & 0 & 0 \\ 0 & \sqrt{90} & 0 \end{bmatrix} \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}^T$.

3. Construct the matrix U. The i^{th} column of U is $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$ for each $\sigma_i \neq 0$.

- $\mathbf{u}_1 = \frac{1}{\sigma_1} A\mathbf{v}_1 = \frac{1}{\sqrt{360}} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}$
- $\mathbf{u}_2 = \frac{1}{\sigma_2} A\mathbf{v}_2 = \frac{1}{\sqrt{90}} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$

So far we have the decomposition $A = U\Sigma V^T = U \begin{bmatrix} \sqrt{360} & 0 & 0 \\ 0 & \sqrt{90} & 0 \end{bmatrix} \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3 & -1/3 & -2/3 \\ 2/3 & 2/3 & 1/3 \end{bmatrix}^T.$

3. Construct the matrix U. The i^{th} column of U is $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$ for each $\sigma_i \neq 0$.
   - $\mathbf{u}_1 = \frac{1}{\sigma_1} A\mathbf{v}_1 = \frac{1}{\sqrt{360}} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} 1/3 \\ 2/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{10} \\ 1/\sqrt{10} \end{bmatrix}$
   - $\mathbf{u}_2 = \frac{1}{\sigma_2} A\mathbf{v}_2 = \frac{1}{\sqrt{90}} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{10} \\ -3/\sqrt{10} \end{bmatrix}$

4. Express the complete decomposition in the form $A = U\Sigma V^T$.

$$\begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{10} & 1/\sqrt{10} \\ 1/\sqrt{10} & -3\sqrt{10} \end{bmatrix} \begin{bmatrix} \sqrt{360} & 0 & 0 \\ 0 & \sqrt{90} & 0 \end{bmatrix} \begin{bmatrix} 1/3 & 2/3 & 2/3 \\ -2/3 & -1/3 & 2/3 \\ 2/3 & -2/3 & 1/3 \end{bmatrix}$$

Example 2

Compute the SVD of $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$.

Example 2

Compute the SVD of $A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$.

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

1. Find an orthogonal diagonalization of $A^T A = PDP^T$.

$$p(\lambda) = (2-\lambda)(2-\lambda) - 1 = \lambda^2 - 4\lambda + 3 = (\lambda - 3)(\lambda - 1) = 0 \quad \lambda_1 = 3 \text{ and } \lambda_2 = 1$$

$$\begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \qquad \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \qquad \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

Finding Matrices Σ and V

$$A^T A = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad \text{has} \quad \lambda_1 = 3 \;\; \mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \text{ and } \lambda_2 = 1 \;\; \mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

2. Find the matrices V and Σ.

- The eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are already orthogonal. $V = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$
- We have $\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$.

Finding the Matrix U

3. Find the matrix U. So far we have $V = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$ and $\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$.

For each nonzero σ_i, we have $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$.

Finding the Matrix U

3. Find the matrix U. So far we have $V = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$ and $\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$.

For each nonzero σ_i, we have $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$.

$$\mathbf{u}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} \sqrt{2/3} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \frac{1}{1} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

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

Finding the Matrix U

3. Find the matrix U. So far we have $V = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$ and $\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$.

For each nonzero σ_i, we have $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$.

$$\mathbf{u}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} \sqrt{2/3} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} \quad \text{and} \quad \mathbf{u}_2 = \frac{1}{1} \begin{bmatrix} 1 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

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

But U needs to be 3×3! We need another column.

Finding the Remaining Columns of U

We have found $\mathbf{u}_1 = \begin{bmatrix} \sqrt{2/3} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$.

We need one more column vector (which is orthogonal to the first two vectors!).

Finding the Remaining Columns of U

We have found $\mathbf{u}_1 = \begin{bmatrix} \sqrt{2/3} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$ and $\mathbf{u}_2 = \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$.

We need one more column vector (which is orthogonal to the first two vectors!).

$$\begin{aligned} \mathbf{z}^T\mathbf{u}_1 &= \sqrt{2/3}z_1 + 1/\sqrt{6}z_2 + 1/\sqrt{6}z_3 = 0 \\ \mathbf{z}^T\mathbf{u}_2 &= 0z_1 - 1/\sqrt{2}z_2 + 1/\sqrt{2}z_3 = 0 \end{aligned} \qquad \begin{bmatrix} \sqrt{2/3} & 1/\sqrt{6} & 1/\sqrt{6} \\ 0 & -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\mathbf{z} = \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \quad \mathbf{u}_3 = \begin{bmatrix} -1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \quad U = \begin{bmatrix} \sqrt{2/3} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix}$$

Success!

We want to find $A = U\Sigma V^T$.

- A is a 3×2 matrix.
- Σ is a 3×2 matrix (size matches A).
- U is a 3×3 orthogonal matrix.
- V is a 2×2 orthogonal matrix.

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

Summary of Steps

Given an $m \times n$ matrix A, to find a singular value decomposition $A = U\Sigma V^T$:

1. Find an orthogonal diagonalization of $A^TA = PDP^T$.
   - Find eigenvalues and eigenvectors of A^TA.
   - Use Gram–Schmidt (if needed) to make orthogonal basis for $\mathbb{R}^n$. Then normalize.
   - Give the orthogonal matrix P.

2. Find the matrices V and Σ.
   - $V = P$ from the previous step.
   - $\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$, D is $r \times r$ diagonal matrix with nonzero singular values on the diagonal.

3. Find the matrix U.
   - For each nonzero σ_i, we have $\mathbf{u}_i = \frac{1}{\sigma_i} A\mathbf{v}_i$.
   - If there are r nonzero σ_i, then the remaining columns of U are an orthogonal basis for the orthogonal complement of $\text{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_r\}$.

An Application to Image Compression

We can expand the singular value decomposition as follows:

$$A = U\Sigma V^T = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \ldots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T.$$

An Application to Image Compression

We can expand the singular value decomposition as follows:

$$A = U\Sigma V^T = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \ldots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T.$$

Note that $\mathbf{u}_i \mathbf{v}_i^T$ are $m \times n$ rank-1 matrices.

An Application to Image Compression

We can expand the singular value decomposition as follows:

$$A = U\Sigma V^T = \sigma_1 \mathbf{u}_1 \mathbf{v}_1^T + \sigma_2 \mathbf{u}_2 \mathbf{v}_2^T + \ldots + \sigma_r \mathbf{u}_r \mathbf{v}_r^T.$$

Note that $\mathbf{u}_i \mathbf{v}_i^T$ are $m \times n$ rank-1 matrices.

If we want to approximate A with rank-1 matrices, then we can drop the terms with small singular values. See Python notebook.

Linear Transformations

Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and let $A_{m\times n}$ be the associated matrix.

Let S denote the unit sphere in $\mathbb{R}^n$ (ie, all points at distance 1 from the origin).

What does the image of S under T look like? See Python notebook.

Linear Transformations

Let $T\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and let $A_{m\times n}$ be the associated matrix.

Let S denote the unit sphere in $\mathbb{R}^n$ (ie, all points at distance 1 from the origin).

What does the image of S under T look like? See Python notebook.

The image of S under T is an ellipsoid in $\mathbb{R}^m$. We can find the axes and size of the ellipsoid using the SVD of A.