Singular Value Decomposition
Matrix Computations — CPSC 5006 E

Julien Dompierre

Department of Mathematics and Computer Science
Laurentian University

Sudbury, November 16, 2009

Singular Value Decomposition

- Diagonalization. Orthogonal diagonalization.
- The URV decomposition – orthogonal spaces — four fundamental subspaces
- The SVD – existence – properties.
- Pseudo-inverses and the SVD
- Use of SVD for least-squares problems
- Use of SVD as a theoretical tool
- Applications of the SVD
- Sections 2.5.1 – 2.5.5 and 5.5.1 – 5.5.4 of the textbook

Similar Matrices

Definition

Let A and B be *square* matrices of the same size. A **is similar** to B if there is an invertible matrix P such that $P^{-1}AP = B$, or equivalently, $A = PBP^{-1}$. Writing Q for P^{-1}, we have $Q^{-1}BQ = A$. So B is also similar to A, and we say simply that A and B **are similar**. Changing A into $P^{-1}AP$ is called a **similarity transformation**.

Example: Let

$$A = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}, B = \begin{bmatrix} -3 & 2 \\ -10 & 6 \end{bmatrix}, P = \begin{bmatrix} -2 & 1 \\ 3 & -1 \end{bmatrix}, P^{-1} = \begin{bmatrix} 1 & 1 \\ 3 & 2 \end{bmatrix}.$$

Then

$$P^{-1}AP = \begin{bmatrix} 1 & 1 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} -2 & 1 \\ 3 & -1 \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ -10 & 6 \end{bmatrix} = B.$$

Similar Matrices and Eigenvalues

Theorem (Theorem 4)

If $n \times n$ matrices A and B are similar, then they have the same characteristic polynomial and hence the same eigenvalues (with the same multiplicities).

Proof.

Let A and B be similar matrices. Hence there exists a matrix P such that $B = P^{-1}AP$. The characteristic polynomial of B is $\det(B - \lambda I) = |B - \lambda I|$.

$$\begin{aligned} |B - \lambda I| &= |P^{-1}AP - \lambda I| = |P^{-1}AP - P^{-1}\lambda IP| \\ &= |P^{-1}(A - \lambda I)P| = |P^{-1}|\,|A - \lambda I|\,|P| \\ &= |A - \lambda I| \end{aligned}$$

The characteristic polynomials of A and B are identical. This means that their eigenvalues are the same. $\square$

Similar Matrices and Eigenvalues

Let A and B be similar matrices:

$$A = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}, B = \begin{bmatrix} -3 & 2 \\ -10 & 6 \end{bmatrix}.$$

The characteristic polynomial of A is

$$|A - \lambda I| = \begin{vmatrix} 2-\lambda & 0 \\ 1 & 1-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda) = \lambda^2 - 3\lambda + 2.$$

The characteristic polynomial of B is

$$|B - \lambda I| = \begin{vmatrix} -3-\lambda & 2 \\ -10 & 6-\lambda \end{vmatrix} = (-3-\lambda)(6-\lambda) - (2)(-10) = \lambda^2 - 3\lambda + 2.$$

Diagonalizable Matrices

Definition

A *square* matrix A is said to be **diagonalizable** if A is similar to a diagonal matrix, that is, there exists an *invertible* matrix P such that $A = PDP^{-1}$ where D is a diagonal matrix.

Theorem

Let A be an $n \times n$ matrix.

(a) *If A has n linearly independent eigenvectors, it is diagonalizable.*

(b) *If A is diagonalizable, then it has n linearly independent eigenvectors.*

In fact, $A = PDP^{-1}$, with D a diagonal matrix, if and only if the columns of P are n linearly independent eigenvectors of A. In this case, the diagonal entries of D are eigenvalues of A that correspond, respectively, to the eigenvectors in P.

In other words, A is diagonalizable if and only if there are enough eigenvectors to form a basis of $\mathbb{R}^n$. We call such a basis an **eigenvector basis**.

The Diagonalization Theorem

Proof.

Let A have eigenvalues $\lambda_1, ..., \lambda_n$, (which need not to be distinct), with corresponding *linearly independent* eigenvectors $v_1, ..., v_n$. Let P be the matrix having $v_1, ..., v_n$ as column vectors:

$$P = [v_1 \ v_2 \ \cdots \ v_n].$$

Since $Av_1 = \lambda_1 v_1$, ..., $Av_n = \lambda_1 v_n$, matrix multiplication in terms of columns gives

$$\begin{aligned} AP &= A[v_1 \ \cdots \ v_n] = [Av_1 \ \cdots \ Av_n] = [\lambda_1 v_1 \ \cdots \ \lambda_n v_n] \\ &= [v_1 \ \cdots \ v_n] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} = P \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \end{aligned}$$

Since the columns of P are linearly independent, P is non singular and invertible. Thus

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} = D.$$

□

Diagonalization Algorithm

To diagonalize an $n \times n$ matrix A, the diagonalization theorem can be implemented in four steps.

Step 1: Find the eigenvalues of A.
Step 2: Find n linearly independent eigenvectors of A. This is a critical step. If it fails, the diagonalization theorem says that A cannot be diagonalized.
Step 3: Construct P from the n linearly independent eigenvectors in step 2.
Step 4: Construct D from the **corresponding** eigenvalues.
Step 5 (Optional): Check that it works! A should be equal to PDP^{-1}. To avoid computing P^{-1}, simply verify that $AP = PD$.

Diagonalization Algorithm

Let

$$A = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix}$$

Step 1: Find the eigenvalues of A.

$$\begin{aligned} 0 &= \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 3 & 3 \\ -3 & -5-\lambda & -3 \\ 3 & 3 & 1-\lambda \end{vmatrix} \\ &= -\lambda^3 - 3\lambda^2 + 4 = -(\lambda - 1)(\lambda + 2)^2 \end{aligned}$$

The eigenvalues are $\lambda = 1$ and $\lambda = -2$ which is of multiplicity 2.

Diagonalization Algorithm

Step 2: Find n linearly independent eigenvectors of A.
With $\lambda = 1$, the linear system to solve is

$$A - \lambda I = \begin{bmatrix} 0 & 3 & 3 \\ -3 & -6 & -3 \\ 3 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Using Gaussian elimination, we get the row reduced echelon form

$$A - \lambda I = \begin{bmatrix} 0 & 3 & 3 \\ -3 & -6 & -3 \\ 3 & 3 & 0 \end{bmatrix} \approx \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

We deduced than $x_3 = r \in \mathbb{R}$, $x_2 = -r$ and $x_1 = r$. The eigenvector v_1 corresponding to the eigenvalue $\lambda = 1$ is

$$v_1 = r \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}.$$

Diagonalization Algorithm

With $\lambda = -2$, the linear system to solve is

$$A - \lambda I = \begin{bmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Using Gaussian elimination, we get the row reduced echelon form

$$A - \lambda I = \begin{bmatrix} 3 & 3 & 3 \\ -3 & -3 & -3 \\ 3 & 3 & 3 \end{bmatrix} \approx \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

We deduced than $x_3 = s \in \mathbb{R}$, $x_2 = r \in \mathbb{R}$ and $x_1 = -r - s$. The eigenvectors $v_{2,3}$ corresponding to the eigenvalue $\lambda = -2$ are

$$v_{2,3} = \begin{bmatrix} -r - s \\ r \\ s \end{bmatrix} = r \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

Diagonalization Algorithm

Step 3: Construct P from the 3 linearly independent eigenvectors in step 2.

$$P = [v_1 \ v_2 \ v_3] = \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}.$$

Diagonalization Algorithm

Step 4: Construct D from the **corresponding** eigenvalues.

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

Diagonalization Algorithm

Step 5 (Optional): Check that it works! A should be equal to PDP^{-1}. To avoid computing P^{-1}, simply verify that $AP = PD$.

$$AP = \begin{bmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & -1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ -1 & -2 & 0 \\ 1 & 0 & -2 \end{bmatrix}.$$

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

Power of a Matrix

For any diagonal matrix D

$$D = \begin{bmatrix} d_{11} & 0 & & 0 \\ 0 & d_{22} & & \\ & & \ddots & \\ 0 & & & d_{nn} \end{bmatrix}, \quad D^k = \begin{bmatrix} d_{11}^k & 0 & & 0 \\ 0 & d_{22}^k & & \\ & & \ddots & \\ 0 & & & d_{nn}^k \end{bmatrix}.$$

If A is similar to a diagonal matrix D under the similarity transformation $A = PDP^{-1}$, then

$$A^k = (PDP^{-1})^k = \underbrace{(PDP^{-1})\cdots(PDP^{-1})}_{k \text{ times}} = PD^kP^{-1}.$$

Symmetric Matrices

A **symmetric matrix** is a matrix such that $A^T = A$. Such a matrix is necessarily *square*. Its main diagonal entries are arbitrary, but its other entries occur in pairs — on opposite side of the diagonal.

Symmetric Matrices and Eigenvectors

Theorem

If A is a symmetric matrix, then any two eigenvectors from different eigenspaces are orthogonal.

Proof.

Let v_1 and v_2 be eigenvectors that correspond to distinct eigenvalues, say λ_1 and λ_2. To show that $v_1 \cdot v_2 = 0$, compute

$$\begin{array}{rcll} \lambda_1 v_1 \cdot v_2 & = & (\lambda_1 v_1)^T v_2 = (Av_1)^T v_2 & \text{Since } v_1 \text{ is an eigenvector} \\ & = & (v_1^T A^T) v_2 = v_1^T (Av_2) & \text{Since } A^T = A \\ & = & v_1^T (\lambda_2 v_2) & \text{Since } v_2 \text{ is an eigenvector} \\ & = & \lambda_2 v_1^T v_2 = \lambda_2 v_1 \cdot v_2. & \end{array}$$

Hence $(\lambda_1 - \lambda_2)(v_1 \cdot v_2) = 0$. But $\lambda_1 - \lambda_2 \neq 0$, so $v_1 \cdot v_2 = 0$. $\square$

Orthogonal Diagonalization

Recall that an $n \times n$ matrix P is **orthogonal** if $P^{-1} = P^T$. The columns of P are pairwise orthogonal and are of length one.

Definition

A square matrix A is said to be **orthogonally diagonalizable** if there are an orthogonal matrix P and a diagonal matrix D such that

$$A = PDP^{-1} = PDP^T.$$

Orthogonal Diagonalization

Theorem

An $n \times n$ matrix A is orthogonally diagonalizable if and only if A is a symmetric matrix.

Orthogonally diagonalizable $\Rightarrow$ symmetric matrix.

Assume that A is orthogonally diagonalizable. Thus there exists an orthogonal matrix P and a diagonal matrix D such that $A = PDP^T$. Use the properties of transpose to get

$$A^T = (PDP^T)^T = (P^T)^T D^T P^T = PDP^T = A.$$

Thus A is symmetric. $\square$

Symmetric matrix $\Rightarrow$ orthogonally diagonalizable.

This one is difficult. $\square$

Orthogonal Diagonalization Algorithm

To diagonalize an $n \times n$ symmetric matrix A, the orthogonal diagonalization theorem can be implemented in four steps.

Step 1: Find the eigenvalues of A.
Step 2: For each eigenvalue, find the corresponding eigenspace. Find an orthonormal basis for this eigenspace. (Use the Gram-Schmidt process if necessary.)
Step 3: Construct P from the n linearly independent eigenvectors in step 2.
Step 4: Construct D from the **corresponding** eigenvalues.
Step 5 (Optional): Check that it works! A should be equal to PDP^T.

Orthogonal Diagonalization Algorithm

$$A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$$

Step 1: Find the eigenvalues of A.

$$\begin{aligned} 0 &= \det(A - \lambda I) = \begin{vmatrix} 3-\lambda & -2 & 4 \\ -2 & 6-\lambda & 2 \\ 4 & 2 & 3-\lambda \end{vmatrix} \\ &= -\lambda^3 + 12\lambda^2 - 21\lambda - 98 = -(\lambda+2)(\lambda-7)^2 \end{aligned}$$

The eigenvalues are $\lambda = -2$ and $\lambda = 7$ which is of multiplicity 2.

Orthogonal Diagonalization Algorithm

Step 2: Find n linearly independent eigenvectors of A.
With $\lambda = 7$, the linear system to solve is

$$A - \lambda I = \begin{bmatrix} -4 & -2 & 4 \\ -2 & -1 & 2 \\ 4 & 2 & -4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Using Gaussian elimination, we get the row reduced echelon form

$$A - \lambda I = \begin{bmatrix} -4 & -2 & 4 \\ -2 & -1 & 2 \\ 4 & 2 & -4 \end{bmatrix} \approx \begin{bmatrix} 1 & 1/2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

We deduced than $x_3 = s \in \mathbb{R}$, $x_2 = r \in \mathbb{R}$ and $x_1 = -r/2 + s$.
The eigenvectors $v_{1,2}$ corresponding to the eigenvalue $\lambda = 7$ are

$$v_{1,2} = \begin{bmatrix} -r/2 + s \\ r \\ s \end{bmatrix} = r \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} + s \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Orthogonal Diagonalization Algorithm

This basis may be converted via orthogonal projection to an orthogonal basis for the eigenspace.

$$z_1 = v_1 = \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}$$

$$\begin{aligned} z_2 &= v_2 - proj_{z_1} v_2 = v_2 - \frac{v_2 \cdot z_1}{z_1 \cdot z_1} z_1 \\ &= \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} - \frac{-1}{5} \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 4/5 \\ 2/5 \\ 1 \end{bmatrix} \end{aligned}$$

The vectors z_1 and z_2 can be normalized to get

$$u_1 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \\ 0 \end{bmatrix} \quad \text{and} \quad u_2 = \begin{bmatrix} 4/\sqrt{45} \\ 2/\sqrt{45} \\ 5/\sqrt{45} \end{bmatrix}$$

Orthogonal Diagonalization Algorithm

With $\lambda = -2$, the linear system to solve is

$$A - \lambda I = \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Using Gaussian elimination, we get the row reduced echelon form

$$A - \lambda I = \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix} \approx \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0.5 \\ 0 & 0 & 0 \end{bmatrix}.$$

We deduced than $x_3 = r \in \mathbb{R}$, $x_2 = -r/2$ and $x_1 = -r$. The eigenvector u_3 corresponding to the eigenvalue $\lambda = -2$ is

$$u_3 = r \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix} = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}.$$

Orthogonal Diagonalization Algorithm

Step 3: Construct P from the 3 linearly independent eigenvectors in step 2.

$$P = [u_1 \;\; u_2 \;\; u_3] = \begin{bmatrix} -1/\sqrt{5} & 4/\sqrt{45} & -2/3 \\ 2/\sqrt{5} & 2/\sqrt{45} & -1/3 \\ 0 & 5/\sqrt{45} & 2/3 \end{bmatrix}.$$

Orthogonal Diagonalization Algorithm

Step 4: Construct D from the **corresponding** eigenvalues.

$$D = \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix}.$$

Orthogonal Diagonalization Algorithm

Step 5 (Optional): Check that it works! A should be equal to PDP^T.

$$PDP^T =$$

$$\begin{bmatrix} -1/\sqrt{5} & 4/\sqrt{45} & -2/3 \\ 2/\sqrt{5} & 2/\sqrt{45} & -1/3 \\ 0 & 5/\sqrt{45} & 2/3 \end{bmatrix} \begin{bmatrix} 7 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} -1/\sqrt{5} & 2/\sqrt{5} & 0 \\ 4/\sqrt{45} & 2/\sqrt{45} & 5/\sqrt{45} \\ -2/3 & -1/3 & 2/3 \end{bmatrix}$$

$$= \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix} = A$$

The Spectral Theorem for Symmetric Matrices

Theorem

An $n \times n$ symmetric matrix A has the following properties:

a. *A has n real eigenvalues, counting multiplicities.*
b. *The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation.*
c. *The eigenspaces are mutually orthogonal, in the sense that eigenvectors corresponding to different eigenvalues are orthogonal.*
d. *A is orthogonally diagonalizable.*

The set of eigenvalues of a matrix A is sometimes called the **spectrum** of A.

Spectral Decomposition

Let A be an $n \times n$ orthogonally diagonalizable matrix. A can be written as $A = PDP^{-1}$, where the columns of P are orthonormal eigenvectors $u_1, \ldots, u_n$ of A and the corresponding eigenvalues $\lambda_1, \ldots, \lambda_n$ are in the diagonal matrix D. Then, we can write A as follow

$$A = \lambda_1 u_1 u_1^T + \lambda_2 u_2 u_2^T + \cdots + \lambda_n u_n u_n^T.$$

Each term in this equation is an $n \times n$ matrix of rank 1. This representation of A is called a **spectral decomposition** of A because it breaks up A into pieces determined by the spectrum (eigenvalues) of A.

The Singular Values of an $m \times n$ Matrix

Let A be an $m \times n$ matrix. Then $A^T A$ is symmetric and can be orthogonally diagonalized. Let $\{v_1, ..., v_n\}$ be an orthonormal basis for $\mathbb{R}^n$ consisting of normalized eigenvectors of $A^T A$, and let $\lambda_1, ..., \lambda_n$ be associated eigenvalues of $A^T A$.

The eigenvalues of $A^T A$ are all non negative and by renumbering, we may assume they are arranged in decreasing order $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$.

The **singular values** of an $m \times n$ matrix A are the square roots of the eigenvalues of $A^T A$, denoted by $\sigma_1 = \sqrt{\lambda_1}, \cdots, \sigma_n = \sqrt{\lambda_n}$ and arranged in decreasing order.

The singular values of A are the lengths of vectors $Av_1, \cdots, Av_n$, where $\{v_1, \cdots, v_n\}$ forms an orthonormal basis for $\mathbb{R}^n$ consisting of eigenvectors of $A^T A$.

The Singular Values and Col Space

Theorem

Suppose $\{v_1, \cdots, v_n\}$ *is an orthonormal basis of* $\mathbb{R}^n$ *consisting of eigenvectors of* $A^T A$*, arranged so that the corresponding eigenvalues of* $A^T A$ *satisfy* $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$*, and suppose* A *has* r *nonzero singular values. Then* $\{Av_1, \cdots, Av_n\}$ *is an orthogonal basis for Col* A*, and rank* $A = r$*.*

The singular Value Decomposition

The SVD decomposition of A involves an $m \times n$ "diagonal" matrix Σ of the form

$$\Sigma_{m\times n} = \begin{bmatrix} D_{r\times r} & 0_{r\times n-r} \\ 0_{m-r\times r} & 0_{m-r\times n-r} \end{bmatrix}$$

where D is an $r \times r$ diagonal matrix for some r not exceeding the smaller of m and n.

Theorem

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

Definition

Any factorization $A = U\Sigma V^T$, with U and V orthogonal, Σ as before, and positive diagonal entries in D, is called a **singular value decomposition** (or **SVD**) of A. The matrices U and V are not uniquely determined by A, but the diagonal entries of Σ are necessarily the singular values 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.

Method for the Singular Value Decomposition

To construct a singular value decomposition of a matrix A:

Step 1. **Find an orthogonal diagonalization of** $A^T A$; i.e. find the eigenvalues of $A^T A$ and a corresponding orthonormal set of eigenvectors.

Step 2. **Set up** V **and** Σ; Arrange the eigenvalues of $A^T A$ in decreasing order. The corresponding unit eigenvectors are the right singular vectors of A and form the columns of V. The square roots of the eigenvalues are the singular values. The nonzero singular values $\sigma_1, \cdots, \sigma_r$ are the diagonal entries of D. The matrix Σ is the same size of A, with D in its upper-left corner and with 0's elsewhere.

Step 3. **Construct U**; The first r columns of U are the normalized vectors obtained from $Av_1, \cdots, Av_r$. Add $n - r$ columns in U to form an orthonormal basis.

Step 4. (optional) **Check that it works!** A should be equal to $U \Sigma V^T$.

Find the SVD of the matrix $A = \begin{bmatrix} 2 & -1 \\ 2 & 2 \end{bmatrix}$.

Step 1. Find an orthogonal diagonalization of $A^T A$; i.e. find the eigenvalues of $A^T A$ and a corresponding orthonormal set of eigenvectors.

Step 1 is itself composed of 5 steps...
Step 1.1: Find the eigenvalues of $A^T A$.
Step 1.2: For each eigenvalue, find the corresponding eigenspace. Find an orthonormal basis for this eigenspace. (Use the Gram-Schmidt process if necessary.)
Step 1.3: Construct P from the n linearly independent eigenvectors in step 2.
Step 1.4: Construct D from the **corresponding** eigenvalues.
Step 1.5 (Optional): Check that it works! $A^T A$ should be equal to PDP^T.

Step 1.1: Find the eigenvalues of $A^T A$.

$$A = \begin{bmatrix} 2 & -1 \\ 2 & 2 \end{bmatrix}, \quad A^T A = \begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix}$$

$$A^T A - \lambda I = \begin{bmatrix} 8-\lambda & 2 \\ 2 & 5-\lambda \end{bmatrix}$$

$$\det(A^T A - \lambda I) = \lambda^2 - 13\lambda + 36 = (\lambda - 9)(\lambda - 4)$$

$$\lambda_1 = 9, \quad \lambda_2 = 4.$$

Step 1.2: For each eigenvalue, find the corresponding eigenspace.
For $\lambda_1 = 9$,

$$A^T A - 9I = \begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -1 & 2 \\ 2 & -4 \end{bmatrix} \approx \begin{bmatrix} -1 & 2 \\ 0 & 0 \end{bmatrix}$$

From row 2, we have $x_2 = r \in \mathbb{R}$. From row one, we have $-1x_1 + 2x_2 = 0$, then $x_1 = 2r$. The eigenspace is then

$$v_1 = r \begin{bmatrix} 2 \\ 1 \end{bmatrix}, r \in \mathbb{R}.$$

An orthonormal basis of this eigenspace is

$$v_1 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}.$$

Step 1.2: For each eigenvalue, find the corresponding eigenspace.
For $\lambda_2 = 4$,

$$A^T A - 4I = \begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 4 & 2 \\ 2 & 1 \end{bmatrix} \approx \begin{bmatrix} 4 & 2 \\ 0 & 0 \end{bmatrix}$$

From row 2, we have $x_2 = r \in \mathbb{R}$. From row one, we have $4x_1 + 2x_2 = 0$, then $x_1 = -r/2$. The eigenspace is then

$$v_2 = r \begin{bmatrix} -1/2 \\ 1 \end{bmatrix}, r \in \mathbb{R}.$$

An orthonormal basis of this eigenspace is

$$v_2 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}.$$

Step 1.3: Construct P from the n linearly independent eigenvectors in step 2.

$$P = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

Step 1.4: Construct D from the **corresponding** eigenvalues.

$$D = \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix}$$

Step 1.5 (Optional): Check that it works! $A^T A$ should be equal to PDP^T.

$$A^T A = \begin{bmatrix} 8 & 2 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 9 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}.$$

Step 2: Set up V and Σ; Arrange the eigenvalues of $A^T A$ in decreasing order. The corresponding unit eigenvectors are the right singular vectors of A and form the columns of V. The square roots of the eigenvalues are the singular values. The nonzero singular values $\sigma_1, \cdots, \sigma_r$ are the diagonal entries of D. The matrix Σ is the same size of A, with D in its upper-left corner and with 0's elsewhere.

$\sigma_1 = \sqrt{\lambda_1} = \sqrt{9} = 3$ and $\sigma_2 = \sqrt{\lambda_2} = \sqrt{4} = 2$.

$$\Sigma = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$$

$$V = P = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

Step 3; Construct U; The first r columns of U are the normalized vectors obtained from $Av_1, \cdots, Av_r$. Add $n-r$ columns in U to form an orthonormal basis.

$$u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{3} \begin{bmatrix} 2 & -1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}$$

$$u_2 = \frac{1}{\sigma_2} A v_2 = \frac{1}{2} \begin{bmatrix} 2 & -1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$$

Since $\{u_1, u_2\}$ is a basis for $\mathbb{R}^2$, let

$$U = \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

Step 4: (optional) Check that it works! A should be equal to $U\Sigma V^T$.

$$\begin{aligned} A &= U\Sigma V^T \\ &= \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \\ &= \begin{bmatrix} 2 & -1 \\ 2 & 2 \end{bmatrix} \end{aligned}$$

Singular Value Decomposition — Step 1

Find the SVD of the matrix $A = \begin{bmatrix} 7 & 1 \\ 0 & 0 \\ 5 & 5 \end{bmatrix}$.

Step 1. Find an orthogonal diagonalization of $A^T A$; i.e. find the eigenvalues of $A^T A$ and a corresponding orthonormal set of eigenvectors.

Step 1.1: Find the eigenvalues of $A^T A$.

$$A^T A = \begin{bmatrix} 74 & 32 \\ 32 & 26 \end{bmatrix}$$

$$A^T A - \lambda I = \begin{bmatrix} 74 - \lambda & 32 \\ 32 & 26 - \lambda \end{bmatrix}$$

$$\det(A^T A - \lambda I) = \lambda^2 - 100\lambda + 900 = (\lambda - 90)(\lambda - 10)$$

$$\lambda_1 = 90, \quad \lambda_2 = 10.$$

Step 1.2: For each eigenvalue, find the corresponding eigenspace. For $\lambda_1 = 90$,

$$A^T A - 90I = \begin{bmatrix} -16 & 32 \\ 32 & -64 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} -16 & 32 \\ 32 & -64 \end{bmatrix} \approx \begin{bmatrix} -1 & 2 \\ 0 & 0 \end{bmatrix}$$

From row 2, we have $x_2 = r \in \mathbb{R}$. From row one, we have $-1x_1 + 2x_2 = 0$, then $x_1 = 2r$. The eigenspace is then

$$v_1 = r \begin{bmatrix} 2 \\ 1 \end{bmatrix}, r \in \mathbb{R}.$$

An orthonormal basis of this eigenspace is

$$v_1 = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}.$$

Step 1.2: For each eigenvalue, find the corresponding eigenspace. For $\lambda_2 = 10$,

$$A^T A - 10I = \begin{bmatrix} 64 & 32 \\ 32 & 16 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 64 & 32 \\ 32 & 16 \end{bmatrix} \approx \begin{bmatrix} 2 & 1 \\ 0 & 0 \end{bmatrix}$$

From row 2, we have $x_2 = r \in \mathbb{R}$. From row one, we have $2x_1 + 1x_2 = 0$, then $x_1 = -r/2$. The eigenspace is then

$$v_2 = r \begin{bmatrix} -1/2 \\ 1 \end{bmatrix}, r \in \mathbb{R}.$$

An orthonormal basis of this eigenspace is

$$v_2 = \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix}.$$

Step 1.3: Construct P from the n linearly independent eigenvectors in step 2.

$$P = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

Step 1.4: Construct D from the **corresponding** eigenvalues.

$$D = \begin{bmatrix} 90 & 0 \\ 0 & 10 \end{bmatrix}$$

Step 1.5 (Optional): Check that it works! A^TA should be equal to PDP^T.

$$A^TA = \begin{bmatrix} 74 & 32 \\ 32 & 26 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \begin{bmatrix} 90 & 0 \\ 0 & 10 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

Singular Value Decomposition — Step 2

Step 2: Set up V and Σ; Arrange the eigenvalues of A^TA in decreasing order. The corresponding unit eigenvectors are the right singular vectors of A and form the columns of V. The square roots of the eigenvalues are the singular values. The nonzero singular values $\sigma_1, \cdots, \sigma_r$ are the diagonal entries of D. The matrix Σ is the same size of A, with D in its upper-left corner and with 0's elsewhere.

$\sigma_1 = \sqrt{\lambda_1} = \sqrt{90} = 3\sqrt{10}$ and $\sigma_2 = \sqrt{\lambda_2} = \sqrt{10}$.

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

$$V = P = \begin{bmatrix} 2/\sqrt{5} & -1/\sqrt{5} \\ 1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix}$$

Step 3; Construct U; The first r columns of U are the normalized vectors obtained from $Av_1, \cdots, Av_r$. Add $n-r$ columns in U to form an orthonormal basis.

$$u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{3\sqrt{10}} \begin{bmatrix} 7 & 1 \\ 0 & 0 \\ 5 & 5 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$$

$$u_2 = \frac{1}{\sigma_2} A v_2 = \frac{1}{\sqrt{10}} \begin{bmatrix} 7 & 1 \\ 0 & 0 \\ 5 & 5 \end{bmatrix} \begin{bmatrix} -1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} = \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$$

Singular Value Decomposition — Step 3

Since $\{u_1, u_2\}$ is not a basis for $\mathbb{R}^3$, we need a unit vector u_3 that is orthogonal to both u_1 and u_2. The vector $u_3 = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^T$ must satisfy the set of equations $u_1 \cdot u_3 = u_1^T u_3 = 0$ and $u_2 \cdot u_3 = u_2^T u_3 = 0$. These are equivalent to the linear equations

$$\begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ -1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} \approx \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

From row 2, we have that $x_3 = 0$, From row 1, we have that $x_1 = 0$, and no condition on x_2. So $x_2 = r \in \mathbb{R}$. The eigenspace orthogonal to u_1 and u_2 is $\begin{bmatrix} 0 & r & 0 \end{bmatrix}^T$ and a normal basis is $\begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$.

Singular Value Decomposition — Step 4

Therefore let

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

Step 4: (optional) Check that it works! A should be equal to $U\Sigma V^T$.

$$\begin{aligned} A &= U\Sigma V^T \\ &= \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} 3\sqrt{10} & 0 \\ 0 & \sqrt{10} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2/\sqrt{5} & 1/\sqrt{5} \\ -1/\sqrt{5} & 2/\sqrt{5} \end{bmatrix} \\ &= \begin{bmatrix} 7 & 1 \\ 0 & 0 \\ 5 & 5 \end{bmatrix}. \end{aligned}$$