

Singular Value Decomposition

Sources:

1. `Linear Algebra : Theory, Intuition, Code` author: Mike X Cohen, publisher: sincXpress
2. `Matrix Methods for Computational Modeling and Data Analytics` author: Mark Embree, Virginia Tech

This notebook has been setup to understand how singular value decomposition works. The content is mainly an adaption of the chapter 16 of `Linear Algebra : Theory, Intuition, Code` and of chapter 5 `The singular value decomposition of Matrix Methods for Computational Modeling and Data Analytics`. I have used mostly `Matrix Methods for Computational Modeling and Data Analytics` because to me it seems more accessible than `Linear Algebra : Theory, Intuition, Code`.

```
In [1]: # some imports for numerical experiments
import numpy as np
import scipy as sc
import math
```

Derivation of Singular Value Decomposition

mainly from `Matrix Methods for Computational Modeling and Data Analytics`

All columns of matrix $\mathbf{A} \in \mathbb{R}^{m \times n} : m \geq n$ shall be linearly independent. Thus the rank of the matrix is as large as possible.

From \mathbf{A} a square matrix $\mathbf{A}^T \cdot \mathbf{A} \in \mathbb{R}^{m \times m}$ is constructed. This matrix is symmetric and it is also *positive-definite* $\mathbf{x}^T \cdot \mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{x} \geq 0$

An implication of $\mathbf{A}^T \cdot \mathbf{A}$ being *positive-definite* is that its eigenvalues are positive.

We can therefore find pairs of eigenvalues/eigenvectors $\lambda_j, \mathbf{v}_j : 1 \leq j \leq n$. Also the ordering of these pair is arbitrary we assume descending ordering of eigenvalues such as $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n > 0$.

Eigenvectors \mathbf{v}_j have unit length $||\mathbf{v}_j|| = 1$. Moreover they are mutually orthonormal $\mathbf{v}_j^T \cdot \mathbf{v}_k = 0$ for $j \neq k$.

Since we have

$$\mathbf{v}_j^T \cdot \mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{v}_j = (\mathbf{A} \cdot \mathbf{v}_j)^T \cdot (\mathbf{A} \cdot \mathbf{v}_j) = \lambda_j \cdot \mathbf{v}_j^T \cdot \mathbf{v}_j = \lambda_j = ||\mathbf{A} \cdot \mathbf{v}_j||^2$$

$$\rightarrow$$

$$\sqrt{\lambda_j} = ||\mathbf{A} \cdot \mathbf{v}_j||$$

The positive square root of eigenvalue λ_j is defined as *singular value* σ_j .

$$\sigma_j = \sqrt{\lambda_j} = ||\mathbf{A} \cdot \mathbf{v}_j||$$

and because we arranged eigenvalues in decreasing order we get $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$ (singular values of \mathbf{A}).

Now we define new vectors $\mathbf{u}_j = \mathbf{A} \cdot \frac{\mathbf{v}_j}{\sigma_j}$.

The vectors have *unit length* and they are *mutually orthonormal*. These properties are derived here:

The unit length property follows directly from

$$||\mathbf{u}_j||^2 = \left(\mathbf{A} \cdot \frac{\mathbf{v}_j}{\sigma_j} \right)^T \cdot \mathbf{A} \cdot \frac{\mathbf{v}_j}{\sigma_j} = \frac{1}{\sigma_j^2} \cdot ||\mathbf{A} \cdot \mathbf{v}_j||^2 = 1$$

Orthogonality is proved like this:

$$\mathbf{u}_j^T \cdot \mathbf{u}_k = \left(\mathbf{A} \cdot \frac{\mathbf{v}_j}{\sigma_j} \right)^T \cdot \mathbf{A} \cdot \frac{\mathbf{v}_k}{\sigma_k} = \frac{1}{\sigma_j \cdot \sigma_k} \cdot \mathbf{v}_j^T \cdot \mathbf{A}^T \mathbf{A} \cdot \mathbf{v}_k = \frac{\lambda_k}{\sigma_j \cdot \sigma_k} \cdot \underbrace{\mathbf{v}_j^T \cdot \mathbf{v}_k}_{0 : j \neq k}$$

Summary

$$\sigma_j \cdot \mathbf{u}_j = \mathbf{A} \cdot \mathbf{v}_j$$

Arranging the column vectors of both sides of the equation yields a matrix equation:

$$\begin{bmatrix} | & | & & | \\ \mathbf{A}\mathbf{v}_1 & \mathbf{A}\mathbf{v}_2 & \dots & \mathbf{A}\mathbf{v}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \sigma_1 \mathbf{u}_1 & \sigma_2 \mathbf{u}_2 & \dots & \sigma_n \mathbf{u}_n \\ | & | & & | \end{bmatrix}$$

Each side of the matrix equation can be further decomposed into a product of matrices:

$$\mathbf{A} \cdot \underbrace{\begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix}}_{\mathbf{V}} = \underbrace{\begin{bmatrix} | & | & & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \\ | & | & & | \end{bmatrix}}_{\mathbf{U}} \cdot \underbrace{\begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \end{bmatrix}}_{\mathbf{\Sigma}}$$

$$\mathbf{A} \cdot \mathbf{V} = \mathbf{U} \cdot \mathbf{\Sigma}$$

$\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{V} \in \mathbb{R}^{n \times n}$, $\mathbf{U} \in \mathbb{R}^{m \times n}$ and $\mathbf{\Sigma} \in \mathbb{R}^{n \times n}$

Since $\mathbf{V}^T \mathbf{V} = \mathbf{I}$ we have $\mathbf{V}^{-1} = \mathbf{V}^T$. Therefore $\mathbf{V} \cdot \mathbf{V}^T = \mathbf{V} \cdot \mathbf{V}^{-1} = \mathbf{I}$.

$$\begin{aligned} \mathbf{A} \cdot \mathbf{V} \cdot \mathbf{V}^T &= \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T \\ &\rightarrow \\ \mathbf{A} &= \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T \end{aligned}$$

A worked example

Given a 3×2 matrix \mathbf{A} .

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

$$\mathbf{A}^T \mathbf{A} = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$$

Eigenvalues of $\mathbf{A}^T \mathbf{A}$ are computed from the characteristic polynomial.

$$\det \begin{pmatrix} 3-\lambda & -1 \\ -1 & 3-\lambda \end{pmatrix} = (9 - 6 \cdot \lambda + \lambda^2 - 1) = \lambda^2 - 6 \cdot \lambda + 8$$

Eigenvalues are: $\lambda_1 = 4$ and $\lambda_2 = 2$. The corresponding eigenvectors \mathbf{v}_1 and \mathbf{v}_2 are computed from:

$$\begin{bmatrix} 3-4 & -1 \\ -1 & 3-4 \end{bmatrix} = \begin{bmatrix} -1 & -1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

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

$$\mathbf{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

Accordingly singular values are: $\sigma_1 = \sqrt{\lambda_1} = 2$ and $\sigma_2 = \sqrt{\lambda_2} = \sqrt{2}$.

In the next step vectors \mathbf{u}_1 and \mathbf{u}_2 are computed.

$$\mathbf{u}_1 = \mathbf{A} \cdot \frac{\mathbf{v}_1}{\sigma_1} = \frac{1}{2} \cdot \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ \sqrt{2} & -\sqrt{2} \end{bmatrix} \cdot \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

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

We are now able to write the matrix decomposition:

$$\underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 0 \\ \sqrt{2} & -\sqrt{2} \end{bmatrix}}_{\mathbf{A}} = \underbrace{\begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}}_{\mathbf{U}} \cdot \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & \sqrt{2} \end{bmatrix}}_{\mathbf{\Sigma}} \cdot \underbrace{\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}}_{\mathbf{V}^T}$$

Below is a numerical example using `Numpy`.

<https://numpy.org/doc/stable/reference/generated/numpy.linalg.svd.html>

```
In [2]: # numerical example
Amat = np.array([[1,1], [0,0], [math.sqrt(2), -math.sqrt(2)]])

Umat, SingularVals, Vtmat = np.linalg.svd(Amat, full_matrices=False)

print(f"Amat          :\n{Amat}\n")
print(f"Umat           :\n{Umat}\n")
print(f"SingularVals      :\n{SingularVals}\n")
print(f"Vtmat             :\n{Vtmat}\n")
```

```

Amat      :
[[ 1.      1.      ]
 [ 0.      0.      ]
 [ 1.41421356 -1.41421356]]

Umat      :
[[ 4.01572963e-17 -1.00000000e+00]
 [ 0.00000000e+00  0.00000000e+00]
 [-1.00000000e+00 -1.79477306e-16]]

SingularVals :
[2.      1.41421356]

Vtmat     :
[[-0.70710678  0.70710678]
 [-0.70710678 -0.70710678]]

```

Note

`Umat` is equivalent to $-\mathbf{U}$ and `Vtmat` is equivalent to $-\mathbf{V}^T$. So these signs cancel each other.

Another Way to formulate the SVD

Again a matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ with linearly independent columns is assumed.

The singular value decomposition has been found as:

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T$$

With the matrix product $\mathbf{U} \cdot \mathbf{\Sigma}$ expressed as:

$$\mathbf{U} \cdot \mathbf{\Sigma} = \begin{bmatrix} | & | & & | \\ \sigma_1 \mathbf{u}_1 & \sigma_2 \mathbf{u}_2 & \cdots & \sigma_n \mathbf{u}_n \\ | & | & & | \end{bmatrix}$$

matrix \mathbf{A} is written like this:

$$\mathbf{A} = \begin{bmatrix} | & | & & | \\ \sigma_1 \mathbf{u}_1 & \sigma_2 \mathbf{u}_2 & \cdots & \sigma_n \mathbf{u}_n \\ | & | & & | \end{bmatrix} \cdot \begin{bmatrix} - & \mathbf{v}_1^T & - \\ - & \mathbf{v}_2^T & - \\ & \cdots & \\ - & \mathbf{v}_n^T & - \end{bmatrix} = \sum_{j=1}^n \sigma_j \cdot \mathbf{u}_j \cdot \mathbf{v}_j^T$$

Therefore the SVD of \mathbf{A} may be expressed as the sum of n submatrices $\sigma_j \cdot \mathbf{u}_j \cdot \mathbf{v}_j^T$.

$$\mathbf{A} = \sum_{j=1}^n \sigma_j \cdot \mathbf{u}_j \cdot \mathbf{v}_j^T$$

To illustrate this equation we use again

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

and decompose it into

$$\begin{aligned} \mathbf{A} &= \sum_{j=1}^2 \sigma_j \cdot \mathbf{u}_j \cdot \mathbf{v}_j^T = \sigma_1 \cdot \mathbf{u}_1 \cdot \mathbf{v}_1^T + \sigma_2 \cdot \mathbf{u}_2 \cdot \mathbf{v}_2^T \\ \mathbf{A} &= \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ \sqrt{2} & -\sqrt{2} \end{bmatrix} = 2 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} + \sqrt{2} \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \\ \mathbf{A} &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \sqrt{2} & -\sqrt{2} \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \\ \sqrt{2} & -\sqrt{2} \end{bmatrix} \end{aligned}$$

A simple application

The **SVD** shall be used to solve the matrix equation $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$. Assumption for matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$: \mathbf{A} is invertible. So its inverse \mathbf{A}^{-1} exists.

$$\begin{aligned} \mathbf{A} &= \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T = \sum_{j=1}^n \sigma_j \cdot \mathbf{u}_j \cdot \mathbf{v}_j^T \\ \mathbf{A} \cdot \mathbf{x} &= \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T \cdot \mathbf{x} = \sum_{j=1}^n \sigma_j \cdot \mathbf{u}_j \cdot (\mathbf{v}_j^T \cdot \mathbf{x}) = \mathbf{b} \end{aligned}$$

Using $\mathbf{U}^T \cdot \mathbf{U} = \mathbf{I}$ we obtain:

$$\begin{aligned} \mathbf{U}^T \cdot \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T \cdot \mathbf{x} &= \mathbf{U}^T \cdot \mathbf{b} \\ \mathbf{\Sigma} \cdot \mathbf{V}^T \cdot \mathbf{x} &= \mathbf{U}^T \cdot \mathbf{b} \end{aligned}$$

With the inverse matrix

$$\Sigma^{-1} = \begin{bmatrix} 1/\sigma_1 & & & \\ & 1/\sigma_2 & & \\ & & \ddots & \\ & & & 1/\sigma_n \end{bmatrix}$$

we have

$$\begin{aligned}\Sigma^{-1} \cdot \Sigma \cdot V^T \cdot x &= \Sigma^{-1} \cdot U^T \cdot b \\ V^T \cdot x &= \Sigma^{-1} \cdot U^T \cdot b \\ V \cdot V^T \cdot x &= V \cdot \Sigma^{-1} \cdot U^T \cdot b \\ &\rightarrow \\ x &= V \cdot \Sigma^{-1} \cdot U^T \cdot b\end{aligned}$$

We conclude that

$$\begin{aligned}A^{-1} &= V \cdot \Sigma^{-1} \cdot U^T = \sum_{j=1}^n \frac{1}{\sigma_j} v_j \cdot u_j^T \\ x &= A^{-1} \cdot b = \sum_{j=1}^n \frac{1}{\sigma_j} v_j \cdot (u_j^T \cdot b) \\ x &= \sum_{j=1}^n \left(\frac{u_j^T \cdot b}{\sigma_j} \right) \cdot v_j\end{aligned}$$

We see that x is the weighted addition of eigenvectors v_1, v_2, \dots, v_n of symmetric matrix $A^T A$.

The next section is about the **Full SVD**.

Full SVD

Assumption are the same as for the reduced version of the **SVD**.

1. $A \in \mathbb{R}^{m \times n}$ with $m \geq n$.
2. Again the **n** columns of the matrix shall be linearly independent.

Contrary to the reduced version of the **SVD** where matrix $U \in \mathbb{R}^{m \times n}$ has **n** orthonormal columns, we generate an **augmented** matrix $\tilde{U} \in \mathbb{R}^{m \times m}$ by adding $m - n$ extra column vectors to the reduced matrix U . Finally all column vectors of \tilde{U} are mutually orthonormal.

The exact procedure to achieve these extra column vectors should be made more explicit. Here we just postulate that such augmentation is feasible.

The orthonormal column vectors of the augmented matrix $\tilde{\mathbf{U}}$ are denoted $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n, \mathbf{u}_{n+1}, \dots, \mathbf{u}_m$.

$$\tilde{\mathbf{U}} = \begin{bmatrix} | & & | & & | \\ \mathbf{u}_1 & \cdots & \mathbf{u}_n & \mathbf{u}_{n+1} & \cdots & \mathbf{u}_m \\ | & & | & & | \end{bmatrix} \in \mathbb{R}^{m \times m}$$

Since $\tilde{\mathbf{U}}$ is square with m orthonormal columns we have:

$$\tilde{\mathbf{U}}^T \tilde{\mathbf{U}} = \mathbf{I}$$

Moreover $\tilde{\mathbf{U}}^{-1} = \tilde{\mathbf{U}}^T$.

The matrix $\mathbf{\Sigma} \in \mathbb{R}^{n \times n}$ (already defined in the reduced SVD) is augmented to another matrix $\tilde{\mathbf{\Sigma}} \in \mathbb{R}^{m \times n}$ by inserting $m - n$ zero values row vectors below the n 'th row of $\mathbf{\Sigma}$.

$$\tilde{\mathbf{\Sigma}} = \begin{bmatrix} \mathbf{\Sigma} \\ \mathbf{0} \\ \text{---} \\ (m-n) \times n \end{bmatrix}$$

Having augmented matrix \mathbf{U} by $m - n$ additional orthonormal columns to a matrix $\tilde{\mathbf{U}}$ and extended matrix $\mathbf{\Sigma}$ by inserting $m - n$ zero rows to a matrix $\tilde{\mathbf{\Sigma}}$ we can now write the Full SVD.

$$\mathbf{A} = \tilde{\mathbf{U}} \cdot \tilde{\mathbf{\Sigma}} \cdot \mathbf{V}^T$$

As for the reduced SVD a numerical example shows some more details.

```
In [3]: # numerical example
Amat = np.array([[1,1], [0,0], [math.sqrt(2), -math.sqrt(2)]])

Umat, SingularVals, Vtmat = np.linalg.svd(Amat, full_matrices=True)

print(f"Amat                : \n{Amat}\n")
print(f"Umat (augmented)      : \n{Umat}\n")
print(f"SingularVals (augmented) : \n{SingularVals}\n")
print(f"Vtmat                     : \n{Vtmat}\n")
```



```

Amat          :
[[ 1.          1.          ]
 [ 0.          0.          ]
 [ 1.41421356 -1.41421356]]

Umat (augmented) :
[[ 4.01572963e-17 -1.00000000e+00 -1.81298661e-16]
 [ 0.00000000e+00  0.00000000e+00  1.00000000e+00]
 [-1.00000000e+00 -1.79477306e-16  1.28197512e-16]]

SingularVals (augmented) :
[2.          1.41421356]

Vtmat          :
[[-0.70710678  0.70710678]
 [-0.70710678 -0.70710678]]

```

Review (so far)

We have so far dealt with two representations of the SVD :

1. the reduced SVD
2. the full SVD

But for both cases we have made the same assumption regarding matrix \mathbf{A} : The $m \times n: m \geq n$ matrix has full columns rank. $\text{rank}(\mathbf{A}) = n$.

Reduced SVD : General case $m \geq n$

Now we consider a more general case:

1. $m \geq n$
2. the columns of \mathbf{A} need to be all linearly independent.

If there are linearly dependent columns of \mathbf{A} there exist some vector $\mathbf{v} \in \mathbb{R}^n$ for which we have:

$$\mathbf{A} \cdot \mathbf{v} = \mathbf{0} : \mathbf{v} \neq \mathbf{0}$$

Multiplying this equation from the left by \mathbf{A}^T yields:

$$\begin{aligned} \mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{v} &= \mathbf{0} \\ \mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{v} &= \mathbf{0} \cdot \mathbf{v} \end{aligned}$$

Vector \mathbf{v} is therefore an eigenvector of $\mathbf{A}^T \cdot \mathbf{A}$ with eigenvalue $\lambda = 0$.

All vectors \mathbf{v} which satisfy $\mathbf{A} \cdot \mathbf{v} = \mathbf{0}$ are then eigenvectors with eigenvalue 0 .

There a r : $0 \leq r \leq n$ non-zero eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_r > 0$. And there are $n - r$ zero eigenvalues $\lambda_{r+1} = \dots = \lambda_n = 0$.

To these eigenvalues belong eigenvectors $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$.

Without Proof

1. Eigenvectors $\mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ form a basis of the null space $R(\mathbf{A})$. The basis vectors are orthonormal (proof required)
2. Moreover $\mathbf{v}_{r+1}, \dots, \mathbf{v}_n$ are orthonormal to $\mathbf{v}_1, \dots, \mathbf{v}_r$ (proof required)

Building the expression for the SVD proceeds in similar steps as before ..

Define **singular values** σ_j :

$$\sigma_j = \sqrt{(\mathbf{A} \cdot \mathbf{v})^T \cdot (\mathbf{A} \cdot \mathbf{v})} = ||\mathbf{A} \cdot \mathbf{v}|| = \sqrt{\lambda_j} \quad j = 1, \dots, n$$

Clearly $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ and $\sigma_{r+1} = \dots = \sigma_n = 0$

Define orthonormal vectors \mathbf{u}_j

$$\mathbf{u}_j = \frac{1}{\sigma_j} \mathbf{A} \cdot \mathbf{v} : j = 1, \dots, r$$

Augment the set of vectors $\mathbf{u}_1, \dots, \mathbf{u}_r$ by vectors $\mathbf{u}_{r+1}, \dots, \mathbf{u}_n$. (provide steps to do this)

Due to its definition we also have:

$$\mathbf{A} \cdot \mathbf{v}_j = \sigma_j \cdot \mathbf{u}_j : j = 1, \dots, n$$

And since $\mathbf{A} \cdot \mathbf{v}_j = \mathbf{0} : j = r + 1, \dots, n$ also $\mathbf{u}_j = \mathbf{u}_j : j = r + 1, \dots, n$.

Now with $\mathbf{A} \cdot \mathbf{v}_j = \sigma_j \cdot \mathbf{u}_j : j = 1, \dots, n$ everything is ready to formulate the **SVD**.

$$\begin{bmatrix} | & & | & & | & & | \\ \mathbf{A} \cdot \mathbf{v}_1 & \cdots & \mathbf{A} \cdot \mathbf{v}_r & \mathbf{A} \cdot \mathbf{v}_{r+1} & \cdots & \mathbf{A} \cdot \mathbf{v}_n \\ | & & | & & | & & | \end{bmatrix} = \begin{bmatrix} | & & | & & | & & | \\ \sigma_1 \cdot \mathbf{u}_1 & \cdots & \sigma_r \cdot \mathbf{u}_r & \sigma_{r+1} \cdot \mathbf{u}_{r+1} & \cdots & \sigma_n \cdot \mathbf{u}_n \\ | & & | & & | & & | \end{bmatrix} :$$

As before the matrices on either side of the equation can be expressed as products of matrices:

$$\mathbf{A} \cdot \begin{bmatrix} | & & | & & | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_r & \mathbf{v}_{r+1} & \cdots & \mathbf{v}_n \\ | & & | & & | & & | \end{bmatrix} = \begin{bmatrix} | & & | \\ \mathbf{u}_1 & \cdots & \mathbf{u}_r \\ | & & | \end{bmatrix} \cdot \begin{bmatrix} \sigma_1 & & | & 0 & \cdots & 0 \\ & \ddots & & \vdots & \ddots & \vdots \\ & & \sigma_r & 0 & \cdots & 0 \end{bmatrix}$$

In the next step we define two matrices \mathbf{V} and \mathbf{V}_\perp :

$$\mathbf{V} = \begin{bmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_r \\ | & & | \end{bmatrix} : \in \mathbb{R}^{n \times r} \quad \mathbf{V}_\perp = \begin{bmatrix} | & & | \\ \mathbf{v}_{r+1} & \cdots & \mathbf{v}_n \\ | & & | \end{bmatrix} : \in \mathbb{R}^{n \times (n-r)}$$

These matrices are concatenated into a single matrix $\tilde{\mathbf{V}} \in \mathbb{R}^{n \times n}$.

$$\tilde{\mathbf{V}} = \begin{bmatrix} \mathbf{V} & \mathbf{V}_\perp \end{bmatrix}$$

All columns of $\tilde{\mathbf{V}}$ are mutually orthonormal. Therefore

$$\begin{aligned} \tilde{\mathbf{V}}^T \cdot \tilde{\mathbf{V}} &= \tilde{\mathbf{V}} \cdot \tilde{\mathbf{V}}^T = \mathbf{I} \\ \tilde{\mathbf{V}}^{-1} &= \tilde{\mathbf{V}}^T \end{aligned}$$

Similarly we define matrix \mathbf{U} as the concatenation of r column vectors $\mathbf{u}_1, \dots, \mathbf{u}_r$.

$$\mathbf{U} = \begin{bmatrix} | & & | \\ \mathbf{u}_1 & \cdots & \mathbf{u}_r \\ | & & | \end{bmatrix}$$

And the r singular values are put into a diagonal matrix $\Sigma \in \mathbb{R}^{n \times n}$.

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

With these definitions we get

$$\mathbf{A} \cdot \tilde{\mathbf{V}} = \mathbf{U} \cdot \begin{bmatrix} \Sigma & \mathbf{0} \end{bmatrix}$$

Multiplying both sides of the equation from the right by $\tilde{\mathbf{V}}^T$ we obtain:

$$\begin{aligned} \mathbf{A} \cdot \tilde{\mathbf{V}} \tilde{\mathbf{V}}^T &= \mathbf{U} \cdot \begin{bmatrix} \Sigma & \mathbf{0} \end{bmatrix} \cdot \tilde{\mathbf{V}}^T \\ \mathbf{A} &= \mathbf{U} \cdot \begin{bmatrix} \Sigma & \mathbf{0} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{V} & \mathbf{V}_\perp \end{bmatrix}^T \\ \mathbf{A} &= \mathbf{U} \cdot \begin{bmatrix} \Sigma & \mathbf{0} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{V}^T \\ \mathbf{V}_\perp^T \end{bmatrix} \\ \mathbf{A} &= \mathbf{U} \cdot \Sigma \cdot \mathbf{V}^T \end{aligned}$$

The last equation looks like the **reduced SVD** for the full rank case. However the dimension of the matrices involved have been changed.

$\mathbf{U} \in \mathbb{R}^{m \times r}$, $\mathbf{\Sigma} \in \mathbb{R}^{r \times r}$ and $\mathbf{V} \in \mathbb{R}^{n \times r}$.

Summary

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T = \underbrace{\begin{bmatrix} | & & | \\ \mathbf{u}_1 & \cdots & \mathbf{u}_r \\ | & & | \end{bmatrix}}_{m \times r} \cdot \underbrace{\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}}_{r \times r} \cdot \underbrace{\begin{bmatrix} - & \mathbf{v}_1^T & - \\ & \vdots & \\ - & \mathbf{v}_r^T & - \end{bmatrix}}_{r \times n} = \sum_{j=1}^r \sigma_j \cdot \mathbf{u}_j \cdot \mathbf{v}_j^T$$

The full SVD : General case $m \geq n$

As in the reduced SVD case we use

$$\mathbf{V} = \begin{bmatrix} | & & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_r \\ | & & | \end{bmatrix} : \in \mathbb{R}^{n \times r} \quad \mathbf{V}_{\perp} = \begin{bmatrix} | & & | \\ \mathbf{v}_{r+1} & \cdots & \mathbf{v}_n \\ | & & | \end{bmatrix} : \in \mathbb{R}^{n \times (n-r)} \quad \tilde{\mathbf{V}} = [\mathbf{V} \quad \mathbf{V}_{\perp}]$$

In a similar fashion we define a matrix $\tilde{\mathbf{U}} \in \mathbb{R}^{m \times m}$ which is constructed from the concatenation of matrices $\mathbf{U} \in \mathbb{R}^{m \times r}$ and $\mathbf{U}_{\perp} \in \mathbb{R}^{m \times (m-r)}$.

$$\mathbf{U} = \begin{bmatrix} | & & | \\ \mathbf{u}_1 & \cdots & \mathbf{u}_r \\ | & & | \end{bmatrix} : \in \mathbb{R}^{m \times r} \quad \mathbf{U}_{\perp} = \begin{bmatrix} | & & | \\ \mathbf{u}_{r+1} & \cdots & \mathbf{u}_m \\ | & & | \end{bmatrix} : \in \mathbb{R}^{m \times (m-r)}$$

$$\tilde{\mathbf{U}} = [\mathbf{U} \quad \mathbf{U}_{\perp}]$$

The diagonal matrix $\mathbf{\Sigma}$

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

is extended to a matrix $\tilde{\mathbf{\Sigma}} \in \mathbb{R}^{m \times n}$

$$\tilde{\mathbf{\Sigma}} = \begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

$$\mathbf{A} = \tilde{\mathbf{U}} \cdot \tilde{\mathbf{\Sigma}} \cdot \tilde{\mathbf{V}}^T = [\mathbf{U} \quad \mathbf{U}_{\perp}] \cdot \begin{bmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \cdot \begin{bmatrix} \mathbf{V}^T \\ \mathbf{V}_{\perp}^T \end{bmatrix}$$

Example

To clarify these procedures an example is taken from **Matrix Methods for Computational Modeling and Data Analytics** . (Example 5-4)

A 3×2 matrix \mathbf{A} is defined:

$$\mathbf{A} = \begin{bmatrix} 1 & \sqrt{2} \\ 1 & \sqrt{2} \\ 1 & \sqrt{2} \end{bmatrix}$$

The two columns are linearly dependent. Thus the column rank is **1** . (The dimension of the column space is **1**).

Eigenvalues are obtained from $\mathbf{A}^T \mathbf{A}$.

$$\mathbf{A}^T \mathbf{A} = \begin{bmatrix} 3 & 3\sqrt{2} \\ 3\sqrt{2} & 6 \end{bmatrix}$$

(it is easy to realise that the matrix has linearly dependent rows / columns.)

Eigenvalues are computed from the characteristic polynomial.

$$\det \begin{pmatrix} 3 - \lambda & 3\sqrt{2} \\ 3\sqrt{2} & 6 - \lambda \end{pmatrix} = (3 - \lambda) \cdot (6 - \lambda) - 18 = \lambda^2 - 9\lambda = 0$$

$\lambda_1 = 9$ and $\lambda_2 = 0$.

Corresponding eigenvectors \mathbf{v}_1 and \mathbf{v}_2 are found as:

$$\mathbf{v}_1 = \begin{bmatrix} 1/\sqrt{3} \\ \sqrt{2}/\sqrt{3} \end{bmatrix}$$

$$\mathbf{v}_2 = \begin{bmatrix} \sqrt{2}/\sqrt{3} \\ -1/\sqrt{3} \end{bmatrix}$$

Singular values are:

$$\sigma_1 = \sqrt{\lambda_1} = 3 \text{ and } \sigma_2 = \sqrt{\lambda_2} = 0$$

Since $r = 1$ the singular vector \mathbf{u}_1 is computed:

$$\mathbf{u}_1 = \frac{1}{\sigma_1} \mathbf{A} \mathbf{v}_1 = \frac{1}{3} \begin{bmatrix} 3/\sqrt{3} \\ 3/\sqrt{3} \\ 3/\sqrt{3} \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}$$

The remaining left singular vectors $\mathbf{u}_2, \mathbf{u}_3$ are computed using the orthonormality condition.

without justification

Remaining vectors are computed from $\mathbf{A} \cdot \mathbf{u} = \mathbf{0}$

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

We must solve for $x_1 + x_2 + x_3 = 0$ and assume that x_1, x_2 are so called **free** variables.

$$x_3 = -x_1 - x_2$$

Then $\mathbf{u}_2, \mathbf{u}_3$ must be of the form

$$\mathbf{u} = \begin{bmatrix} x_1 \\ x_2 \\ -x_1 - x_2 \end{bmatrix}$$

To determine \mathbf{u}_2 we may freely choose x_1, x_2 and normalise the vector.

$$\mathbf{u}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ -1/\sqrt{2} \end{bmatrix}$$

To determine \mathbf{u}_3 we exploit the condition $\mathbf{u}_2^T \cdot \mathbf{u} = 0$.

$$\mathbf{u}_3 = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ -x_1 - x_2 \end{bmatrix} = 2x_1/\sqrt{2} + x_2/\sqrt{2} = 0$$

$$x_2 = -2x_1$$

$$\mathbf{u}_3 = \begin{bmatrix} x_1 \\ -2x_1 \\ x_1 \end{bmatrix}$$

$$\mathbf{u}_3 = \frac{1}{\sqrt{6}} \cdot \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{6} \\ -2/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix}$$

At this point we can express matrices \mathbf{U} , $\mathbf{\Sigma}$ and \mathbf{V} :

$$\mathbf{U} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \quad \mathbf{\Sigma} = [3] \quad \mathbf{V} = \begin{bmatrix} 1/\sqrt{3} \\ \sqrt{2}/\sqrt{3} \end{bmatrix}$$

$$\mathbf{A} = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \cdot [3] \cdot \begin{bmatrix} 1/\sqrt{3} & \sqrt{2}/\sqrt{3} \end{bmatrix}$$

$$= \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \cdot \begin{bmatrix} \sqrt{3} & \sqrt{2}\sqrt{3} \end{bmatrix} = \begin{bmatrix} 1 & \sqrt{2} \\ 1 & \sqrt{2} \\ 1 & \sqrt{2} \end{bmatrix}$$

Older stuff /Definitions

The singular value decomposition is defined for the general case of rectangular $m \times n$ matrix \mathbf{B} . It includes these cases:

1. square matrix $m = n$
2. tall matrix $m > n$
3. wide matrix $m < n$

The singular value decomposition is defined as a matrix product of three matrices:

1. a square $m \times m$ matrix \mathbf{U} . The matrix has orthonormal column vectors.
2. a rectangular $n \times m$ matrix $\mathbf{\Sigma}$. The diagonal elements are the *singular values*
3. a square $n \times n$ matrix \mathbf{V}^T . The matrix \mathbf{V} has orthonormal column vectors.

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T$$

The following steps are needed to get the matrices of the singular value decomposition.

We consider the $n \times n$ matrix $\mathbf{A}^T \cdot \mathbf{A}$. The matrix is square and symmetric. Moreover its eigendecomposition exists and its eigenvalues are positive $\lambda \geq 0$. Its eigenvectors are orthonormal.

$$\begin{aligned}\mathbf{A}^T \cdot \mathbf{A} &= (\mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T)^T \cdot (\mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T) \\ \mathbf{A}^T \cdot \mathbf{A} &= \mathbf{V} \cdot \mathbf{\Sigma}^T \cdot \underbrace{\mathbf{U}^T \cdot \mathbf{U}}_{\mathbf{I}} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T \\ \mathbf{A}^T \cdot \mathbf{A} &= \mathbf{V} \cdot (\mathbf{\Sigma}^T \cdot \mathbf{\Sigma}) \cdot \mathbf{V}^T\end{aligned}$$

Matrix $(\mathbf{\Sigma}^T \cdot \mathbf{\Sigma})$ is again $n \times n$ diagonal matrix. Its diagonal entries are the eigenvalues of the eigendecomposition of $\mathbf{A}^T \cdot \mathbf{A}$.

The expression

$$\mathbf{A}^T \cdot \mathbf{A} = \mathbf{V} \cdot (\mathbf{\Sigma}^T \cdot \mathbf{\Sigma}) \cdot \mathbf{V}^T$$

serves to get matrices \mathbf{V} , $(\mathbf{\Sigma}^T \cdot \mathbf{\Sigma})$.

Using a similar procedure for the $m \times m$ matrix $\mathbf{A} \cdot \mathbf{A}^T$

$$\begin{aligned}\mathbf{A} \cdot \mathbf{A}^T &= (\mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T) \cdot (\mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T)^T \\ \mathbf{A} \cdot \mathbf{A}^T &= \mathbf{U} \cdot (\mathbf{\Sigma} \cdot \mathbf{\Sigma}^T) \cdot \mathbf{U}^T\end{aligned}$$

From this equation which again a eigendecomposition the matrix \mathbf{U} is derived.

Todo

Show how get matrix $\mathbf{\Sigma}$ from either squares matrix $(\mathbf{\Sigma}^T \cdot \mathbf{\Sigma})$ or $(\mathbf{\Sigma} \cdot \mathbf{\Sigma}^T)$.

Without proof:

Since $(\mathbf{\Sigma}^T \cdot \mathbf{\Sigma})$ or $(\mathbf{\Sigma} \cdot \mathbf{\Sigma}^T)$ are diagonal matrices the matrix $\mathbf{\Sigma}$ has only non-zero entries for its diagonal elements.

Example (solved manually)

Matrix \mathbf{A} is defined as a 2×3 matrix.

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

The matrix product $\mathbf{A}^T \cdot \mathbf{A}$ which is a symmetric 3×3 matrix the eigendecomposition exists.

So lets compute (manually):

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

The eigen decomposition of $\mathbf{A}^T \cdot \mathbf{A}$ can be expressed like this:

$$\mathbf{B} = \mathbf{V} \cdot \mathbf{\Lambda} \cdot \mathbf{V}^T$$

computing the eigenvalues

The eigenvalues are computed from the roots of the characteristic polynomial $p(\lambda)$.

$$p(\lambda) = \det \begin{bmatrix} 1-\lambda & 1 & 0 \\ 1 & 2-\lambda & 1 \\ 0 & 1 & 1-\lambda \end{bmatrix} = \lambda \cdot (1-\lambda) \cdot (3-\lambda)$$

There are three real-valued distinct eigenvalues $\lambda_1 = 3$, $\lambda_2 = 1$, $\lambda_3 = 0$. For each the corresponding eigenvector is computed:

computing eigenvectors

For $\lambda_1 = 3$ the elements of eigenvector \mathbf{v}_1 are found solving this equation:

$$\begin{bmatrix} 1-3 & 1 & 0 \\ 1 & 2-3 & 1 \\ 0 & 1 & 1-3 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}$$

$$\mathbf{v}_1 = t \cdot \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = t \cdot \sqrt{6} \cdot \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{2}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \end{bmatrix}$$

For $\lambda_2 = 1$ the elements of eigenvector \mathbf{v}_2 are found solving this equation:

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

$$\mathbf{v}_2 = t \cdot \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} = t \cdot \sqrt{2} \cdot \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

For $\lambda_3 = 0$ the elements of eigenvector \mathbf{v}_3 are found solving this equation:

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}$$

$$\mathbf{v}_3 = t \cdot \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} = t \cdot \sqrt{3} \cdot \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

The eigenvectors \mathbf{v}_1 , \mathbf{v}_2 , \mathbf{v}_3 have been computed as vectors each with length 1 and a unknown scaling factor.

Finally the orthogonal matrix \mathbf{V} can be setup with the orthonormal eigenvectors as column vectors.

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

The column vectors of \mathbf{V} are mutually orthonormal. The inverse matrix \mathbf{V}^{-1} is just the transpose of \mathbf{V} .

Then the eigendecomposition of matrix $\mathbf{A}^T \cdot \mathbf{A}$ may be written like this:

$$\mathbf{A}^T \cdot \mathbf{A} = \mathbf{V} \cdot \mathbf{\Lambda} \cdot \mathbf{V}^T$$

$$\mathbf{A}^T \cdot \mathbf{A} = \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{-1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \\ \frac{2}{\sqrt{6}} & 0 & \frac{-1}{\sqrt{3}} \\ \frac{1}{\sqrt{6}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{3}} \end{bmatrix} \cdot \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

summary

We have obtained the matrix \mathbf{V} as:

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

and the diagonal matrix

$$\begin{aligned} & \left(\mathbf{\Sigma}^T \cdot \mathbf{\Sigma} \right) \\ & \left(\mathbf{\Sigma}^T \cdot \mathbf{\Sigma} \right) = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \end{aligned}$$

Then matrix $\mathbf{\Sigma}$ is a rectangular 2×3 matrix.

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

To get orthonormal matrix \mathbf{U} we compute (manually) the eigendecomposition of $\mathbf{A} \cdot \mathbf{A}^T$.

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

$$p(\lambda) = \det \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix} = (2-\lambda)^2 - 1 = 0$$

Again eigenvalues $\lambda_1 = 3$ and $\lambda_2 = 1$ are found as solutions of the characteristic polynomial.

computing eigenvectors

For $\lambda_1 = 3$ the elements of eigenvector \mathbf{u}_1 is found as:

$$\mathbf{u}_1 = t \cdot \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

For $\lambda_2 = 1$ the elements of eigenvector \mathbf{u}_2 is found as:

$$\mathbf{u}_2 = t \cdot \begin{bmatrix} \frac{-1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$$

This gives us matrix \mathbf{U} as:

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

Finally all *ingredients* of the **SVD** are in place. Thus the matrix decomposition is written like this:

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T = \frac{1}{\sqrt{2}} \cdot \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} \sqrt{3} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sqrt{6}} & \frac{2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \\ \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{3}} \end{bmatrix}$$

$$\mathbf{A} = \mathbf{U} \cdot \mathbf{\Sigma} \cdot \mathbf{V}^T = \frac{1}{\sqrt{2}} \cdot \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{2}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{-1}{2} & 0 & \frac{1}{2} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

The example demonstrates that the **SVD** has indeed reproduced rectangular matrix **A**.

In []: