The Singular Value Decomposition

Linear Algebra
MATH 2076

University of Cincinnati

Motivation—Why study quadratic functions?

Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ be a linear transformation.

Motivation—Why study quadratic functions?

Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ be a linear transformation. We know how to find the T image of a vector subspace $\mathbb{W}$.

Motivation—Why study quadratic functions?

Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ be a linear transformation. We know how to find the T image of a vector subspace $\mathbb{W}$. If $\mathbb{W} = \mathcal{S}pan\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\}$, then

Motivation—Why study quadratic functions?

Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ be a linear transformation. We know how to find the T image of a vector subspace $\mathbb{W}$. If $\mathbb{W} = \mathcal{S}pan\{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_k\}$, then $T(\mathbb{W}) = \mathcal{S}pan\{T(\vec{w}_1), T(\vec{w}_2), \dots, T(\vec{w}_k)\}$.

Motivation—Why study quadratic functions?

Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ be a linear transformation. We know how to find the T image of a vector subspace $\mathbb{W}$. If $\mathbb{W} = \mathcal{S}pan\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\}$, then $T(\mathbb{W}) = \mathcal{S}pan\{T(\vec{w}_1), T(\vec{w}_2), \ldots, T(\vec{w}_k)\}$. This allows us to find images of rectangular boxes, and so forth.

Motivation—Why study quadratic functions?

Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ be a linear transformation. We know how to find the T image of a vector subspace $\mathbb{W}$. If $\mathbb{W} = \mathcal{S}pan\{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_k\}$, then $T(\mathbb{W}) = \mathcal{S}pan\{T(\vec{w}_1), T(\vec{w}_2), \dots, T(\vec{w}_k)\}$. This allows us to find images of rectangular boxes, and so forth.

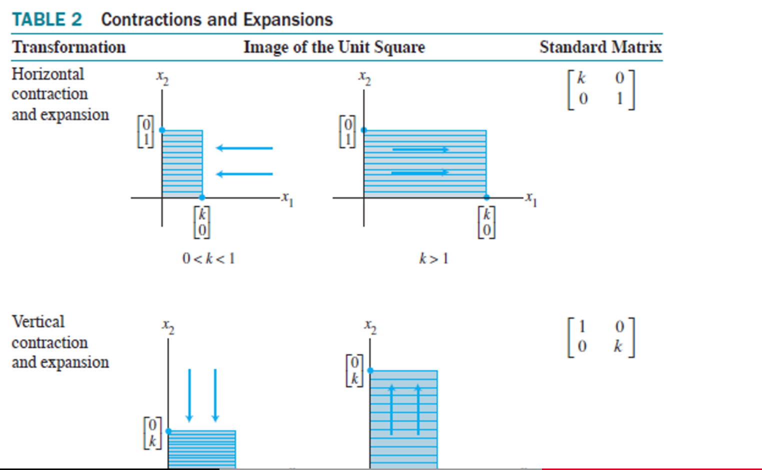
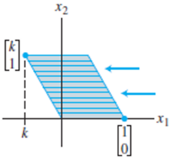
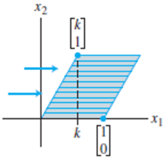
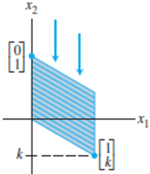
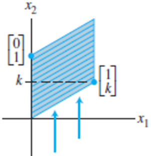


TABLE 2 Contractions and Expansions

Transformation	Image of the Unit Square		Standard Matrix
Horizontal contraction and expansion	$0 < k < 1$	$k > 1$	$\begin{bmatrix} k & 0 \\ 0 & 1 \end{bmatrix}$
Vertical contraction and expansion			$\begin{bmatrix} 1 & 0 \\ 0 & k \end{bmatrix}$

Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ be a linear transformation. We know how to find the T image of a vector subspace $\mathbb{W}$. If $\mathbb{W} = \mathcal{S}pan\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\}$, then $T(\mathbb{W}) = \mathcal{S}pan\{T(\vec{w}_1), T(\vec{w}_2), \ldots, T(\vec{w}_k)\}$. This allows us to find images of rectangular boxes, and so forth.

TABLE 3 Shears

Transformation	Image of the Unit Square		Standard Matrix
Horizontal shear	 $k < 0$	 $k > 0$	$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$
Vertical shear			$\begin{bmatrix} 1 & 0 \\ k & 1 \end{bmatrix}$

Motivation—Why study quadratic functions?

Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ be a linear transformation. We know how to find the T image of a vector subspace $\mathbb{W}$. If $\mathbb{W} = \mathcal{S}pan\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\}$, then $T(\mathbb{W}) = \mathcal{S}pan\{T(\vec{w}_1), T(\vec{w}_2), \ldots, T(\vec{w}_k)\}$. This allows us to find images of rectangular boxes, and so forth.

But what about finding the T image of the unit sphere $\mathbb{S}^{n-1} = \{\|\vec{x}\| = 1\}$?

Motivation—Why study quadratic functions?

Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ be a linear transformation. We know how to find the T image of a vector subspace $\mathbb{W}$. If $\mathbb{W} = \mathcal{S}pan\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\}$, then $T(\mathbb{W}) = \mathcal{S}pan\{T(\vec{w}_1), T(\vec{w}_2), \ldots, T(\vec{w}_k)\}$. This allows us to find images of rectangular boxes, and so forth.

But what about finding the T image of the unit sphere $\mathbb{S}^{n-1} = \{\|\vec{x}\| = 1\}$? Can we find an equation that describes $T(\mathbb{S}^{n-1})$?

Motivation—Why study quadratic functions?

Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ be a linear transformation. We know how to find the T image of a vector subspace $\mathbb{W}$. If $\mathbb{W} = \mathcal{S}pan\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\}$, then $T(\mathbb{W}) = \mathcal{S}pan\{T(\vec{w}_1), T(\vec{w}_2), \ldots, T(\vec{w}_k)\}$. This allows us to find images of rectangular boxes, and so forth.

But what about finding the T image of the unit sphere $\mathbb{S}^{n-1} = \{\|\vec{x}\| = 1\}$? Can we find an equation that describes $T(\mathbb{S}^{n-1})$?

Assume $m = n$ and T is invertible. Let A be the (standard) matrix for T^{-1}.

Motivation—Why study quadratic functions?

Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ be a linear transformation. We know how to find the T image of a vector subspace $\mathbb{W}$. If $\mathbb{W} = \mathcal{S}pan\{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_k\}$, then $T(\mathbb{W}) = \mathcal{S}pan\{T(\vec{w}_1), T(\vec{w}_2), \dots, T(\vec{w}_k)\}$. This allows us to find images of rectangular boxes, and so forth.

But what about finding the T image of the unit sphere $\mathbb{S}^{n-1} = \{\|\vec{x}\| = 1\}$? Can we find an equation that describes $T(\mathbb{S}^{n-1})$?

Assume $m = n$ and T is invertible. Let A be the (standard) matrix for T^{-1}. Let $\vec{y} = T(\vec{x})$, so $\vec{x} = T^{-1}(y) = A\vec{y}$.

Motivation—Why study quadratic functions?

Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ be a linear transformation. We know how to find the T image of a vector subspace $\mathbb{W}$. If $\mathbb{W} = \mathcal{S}pan\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\}$, then $T(\mathbb{W}) = \mathcal{S}pan\{T(\vec{w}_1), T(\vec{w}_2), \ldots, T(\vec{w}_k)\}$. This allows us to find images of rectangular boxes, and so forth.

But what about finding the T image of the unit sphere $\mathbb{S}^{n-1} = \{\|\vec{x}\| = 1\}$? Can we find an equation that describes $T(\mathbb{S}^{n-1})$?

Assume $m = n$ and T is invertible. Let A be the (standard) matrix for T^{-1}. Let $\vec{y} = T(\vec{x})$, so $\vec{x} = T^{-1}(y) = A\vec{y}$. Now suppose $\|\vec{x}\| = 1$. Then

Motivation—Why study quadratic functions?

Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ be a linear transformation. We know how to find the T image of a vector subspace $\mathbb{W}$. If $\mathbb{W} = \mathcal{S}pan\{\vec{w}_1, \vec{w}_2, \ldots, \vec{w}_k\}$, then $T(\mathbb{W}) = \mathcal{S}pan\{T(\vec{w}_1), T(\vec{w}_2), \ldots, T(\vec{w}_k)\}$. This allows us to find images of rectangular boxes, and so forth.

But what about finding the T image of the unit sphere $\mathbb{S}^{n-1} = \{\|\vec{x}\| = 1\}$? Can we find an equation that describes $T(\mathbb{S}^{n-1})$?

Assume $m = n$ and T is invertible. Let A be the (standard) matrix for T^{-1}. Let $\vec{y} = T(\vec{x})$, so $\vec{x} = T^{-1}(y) = A\vec{y}$. Now suppose $\|\vec{x}\| = 1$. Then

$$1 = \vec{x} \cdot \vec{x} = (\vec{x})^T \vec{x} = (A\vec{y})^T A\vec{y} = \vec{y}^T (A^T A)\, \vec{y}.$$

Motivation—Why study quadratic functions?

Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ be a linear transformation. We know how to find the T image of a vector subspace $\mathbb{W}$. If $\mathbb{W} = \mathcal{S}pan\{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_k\}$, then $T(\mathbb{W}) = \mathcal{S}pan\{T(\vec{w}_1), T(\vec{w}_2), \dots, T(\vec{w}_k)\}$. This allows us to find images of rectangular boxes, and so forth.

But what about finding the T image of the unit sphere $\mathbb{S}^{n-1} = \{\|\vec{x}\| = 1\}$? Can we find an equation that describes $T(\mathbb{S}^{n-1})$?

Assume $m = n$ and T is invertible. Let A be the (standard) matrix for T^{-1}. Let $\vec{y} = T(\vec{x})$, so $\vec{x} = T^{-1}(y) = A\vec{y}$. Now suppose $\|\vec{x}\| = 1$. Then

$$1 = \vec{x} \cdot \vec{x} = (\vec{x})^T \vec{x} = (A\vec{y})^T A\vec{y} = \vec{y}^T (A^T A)\, \vec{y}.$$

Thus we do have an equation for $T(\mathbb{S}^{n-1})$, namely

$$\vec{y}^T M \vec{y} = 1 \quad \text{where} \quad M = A^T A.$$

Motivation—Why study quadratic functions?

Let $\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m$ be a linear transformation. We know how to find the T image of a vector subspace $\mathbb{W}$. If $\mathbb{W} = \mathcal{S}pan\{\vec{w}_1, \vec{w}_2, \dots, \vec{w}_k\}$, then $T(\mathbb{W}) = \mathcal{S}pan\{T(\vec{w}_1), T(\vec{w}_2), \dots, T(\vec{w}_k)\}$. This allows us to find images of rectangular boxes, and so forth.

But what about finding the T image of the unit sphere $\mathbb{S}^{n-1} = \{\|\vec{x}\| = 1\}$? Can we find an equation that describes $T(\mathbb{S}^{n-1})$?

Assume $m = n$ and T is invertible. Let A be the (standard) matrix for T^{-1}. Let $\vec{y} = T(\vec{x})$, so $\vec{x} = T^{-1}(y) = A\vec{y}$. Now suppose $\|\vec{x}\| = 1$. Then

$$1 = \vec{x} \cdot \vec{x} = (\vec{x})^T \vec{x} = (A\vec{y})^T A\vec{y} = \vec{y}^T (A^T A)\, \vec{y}.$$

Thus we do have an equation for $T(\mathbb{S}^{n-1})$, namely

$$\vec{y}^T M \vec{y} = 1 \quad \text{where} \quad M = A^T A.$$

In particular we see that $T(\mathbb{S}^{n-1})$ is *the level set of a quadratic function*!

The Matrix $M = A^T A$

What can we say about $M = A^T A$ when A is an $m \times n$ matrix?

The Matrix $M = A^T A$

What can we say about $M = A^T A$ when A is an $m \times n$ matrix?

Evidently, M is an $n \times n$ matrix, and

The Matrix $M = A^T A$

What can we say about $M = A^T A$ when A is an $m \times n$ matrix?

Evidently, M is an $n \times n$ matrix, and $M^T = (A^T A)^T = A^T A = M$, so M is *symmetric*. Thus:

The Matrix $M = A^T A$

What can we say about $M = A^T A$ when A is an $m \times n$ matrix?

Evidently, M is an $n \times n$ matrix, and $M^T = (A^T A)^T = A^T A = M$, so M is *symmetric*. Thus:

1. M has n real eigenvalues, counting according to multiplicity.

The Matrix $M = A^T A$

What can we say about $M = A^T A$ when A is an $m \times n$ matrix?

Evidently, M is an $n \times n$ matrix, and $M^T = (A^T A)^T = A^T A = M$, so M is *symmetric*. Thus:

1. M has n real eigenvalues, counting according to multiplicity.
2. M is *orthogonally diagonalizable*.

The Matrix $M = A^T A$

What can we say about $M = A^T A$ when A is an $m \times n$ matrix?

Evidently, M is an $n \times n$ matrix, and $M^T = (A^T A)^T = A^T A = M$, so M is *symmetric*. Thus:

1. M has n real eigenvalues, counting according to multiplicity.
2. M is *orthogonally diagonalizable*.

Item (1) says that M has (real) eigenvalues $\lambda_1, \lambda_2 \dots, \lambda_n$.

The Matrix $M = A^T A$

What can we say about $M = A^T A$ when A is an $m \times n$ matrix?

Evidently, M is an $n \times n$ matrix, and $M^T = (A^T A)^T = A^T A = M$, so M is *symmetric*. Thus:

1. M has n real eigenvalues, counting according to multiplicity.
2. M is *orthogonally diagonalizable*.

Item (1) says that M has (real) eigenvalues $\lambda_1, \lambda_2 \ldots, \lambda_n$. Item (2) says that M has an associated *orthonormal* eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$.
Here $M\vec{v}_i = \lambda_i \vec{v}_i$.

The Matrix $M = A^T A$

What can we say about $M = A^T A$ when A is an $m \times n$ matrix?

Evidently, M is an $n \times n$ matrix, and $M^T = (A^T A)^T = A^T A = M$, so M is *symmetric*. Thus:

1. M has n real eigenvalues, counting according to multiplicity.
2. M is *orthogonally diagonalizable*.

Item (1) says that M has (real) eigenvalues $\lambda_1, \lambda_2 \ldots, \lambda_n$. Item (2) says that M has an associated *orthonormal* eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$.
Here $M\vec{v}_i = \lambda_i \vec{v}_i$. Now look at

$$\begin{aligned} \lambda_i &= \lambda_i \|\vec{v}_i\|^2 = \lambda_i \vec{v}_i \cdot \vec{v}_i = (M\vec{v}_i) \cdot \vec{v}_i = (M\vec{v}_i)^T \vec{v}_i = \vec{v}_i^T M^T \vec{v}_i \\ &= \vec{v}_i^T M \vec{v}_i = \vec{v}_i^T A^T A \vec{v}_i = (A\vec{v}_i)^T (A\vec{v}_i) = \|A\vec{v}_i\|^2. \end{aligned}$$

The Matrix $M = A^T A$

What can we say about $M = A^T A$ when A is an $m \times n$ matrix?

Evidently, M is an $n \times n$ matrix, and $M^T = (A^T A)^T = A^T A = M$, so M is *symmetric*. Thus:

1. M has n real eigenvalues, counting according to multiplicity.
2. M is *orthogonally diagonalizable*.

Item (1) says that M has (real) eigenvalues $\lambda_1, \lambda_2 \ldots, \lambda_n$. Item (2) says that M has an associated *orthonormal* eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$.
Here $M\vec{v}_i = \lambda_i \vec{v}_i$. Now look at

$$\begin{aligned}\lambda_i &= \lambda_i \|\vec{v}_i\|^2 = \lambda_i \vec{v}_i \cdot \vec{v}_i = (M\vec{v}_i) \cdot \vec{v}_i = (M\vec{v}_i)^T \vec{v}_i = \vec{v}_i^T M^T \vec{v}_i \\ &= \vec{v}_i^T M \vec{v}_i = \vec{v}_i^T A^T A \vec{v}_i = (A\vec{v}_i)^T (A\vec{v}_i) = \|A\vec{v}_i\|^2.\end{aligned}$$

Thus each $\lambda_i \geq 0$.

The Matrix $M = A^T A$

What can we say about $M = A^T A$ when A is an $m \times n$ matrix?

Evidently, M is an $n \times n$ matrix, and $M^T = (A^T A)^T = A^T A = M$, so M is *symmetric*. Thus:

1. M has n real eigenvalues, counting according to multiplicity.
2. M is *orthogonally diagonalizable*.

Item (1) says that M has (real) eigenvalues $\lambda_1, \lambda_2 \ldots, \lambda_n$. Item (2) says that M has an associated *orthonormal* eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$.
Here $M\vec{v}_i = \lambda_i \vec{v}_i$. Now look at

$$\begin{aligned}\lambda_i &= \lambda_i \|\vec{v}_i\|^2 = \lambda_i \vec{v}_i \cdot \vec{v}_i = (M\vec{v}_i) \cdot \vec{v}_i = (M\vec{v}_i)^T \vec{v}_i = \vec{v}_i^T M^T \vec{v}_i \\ &= \vec{v}_i^T M \vec{v}_i = \vec{v}_i^T A^T A \vec{v}_i = (A\vec{v}_i)^T (A\vec{v}_i) = \|A\vec{v}_i\|^2.\end{aligned}$$

Thus each $\lambda_i \geq 0$. By relabeling, we can assume $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$.

The Matrix $M = A^T A$

When A is an $m \times n$ matrix, $M = A^T A$ is an $n \times n$ *symmetric* matrix with non-negative real eigenvalues, say $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$, and it has an associated *orthonormal* eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$. Here $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$.

The Matrix $M = A^T A$

When A is an $m \times n$ matrix, $M = A^T A$ is an $n \times n$ *symmetric* matrix with non-negative real eigenvalues, say $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$, and it has an associated *orthonormal* eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$. Here $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$.

The numbers $\boxed{\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|}$ are called the *singular values of A*.

The Matrix $M = A^T A$

When A is an $m \times n$ matrix, $M = A^T A$ is an $n \times n$ *symmetric* matrix with non-negative real eigenvalues, say $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$, and it has an associated *orthonormal* eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$. Here $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$.

The numbers $\boxed{\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|}$ are called the *singular values of A*. We really only care about the *non-zero* singular values.

The Matrix $M = A^T A$

When A is an $m \times n$ matrix, $M = A^T A$ is an $n \times n$ *symmetric* matrix with non-negative real eigenvalues, say $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$, and it has an associated *orthonormal* eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$. Here $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$.

The numbers $\boxed{\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|}$ are called the *singular values of A*. We really only care about the *non-zero* singular values. Suppose

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$

The Matrix $M = A^T A$

When A is an $m \times n$ matrix, $M = A^T A$ is an $n \times n$ *symmetric* matrix with non-negative real eigenvalues, say $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$, and it has an associated *orthonormal* eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$. Here $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$.

The numbers $\boxed{\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|}$ are called the *singular values of A*. We really only care about the *non-zero* singular values. Suppose

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$

Here $1 \leq r \leq n$ and when $r < n$, $A\vec{v}_{r+1} = \cdots = A\vec{v}_n = \vec{0}$.

The Matrix $M = A^T A$

When A is an $m \times n$ matrix, $M = A^T A$ is an $n \times n$ *symmetric* matrix with non-negative real eigenvalues, say $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$, and it has an associated *orthonormal* eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$. Here $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$.

The numbers $\boxed{\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|}$ are called the *singular values of A*. We really only care about the *non-zero* singular values. Suppose

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$

Here $1 \leq r \leq n$ and when $r < n$, $A\vec{v}_{r+1} = \cdots = A\vec{v}_n = \vec{0}$. So, if $\vec{x} = \sum_{i=1}^{n} c_i \vec{v}_i$, then

The Matrix $M = A^T A$

When A is an $m \times n$ matrix, $M = A^T A$ is an $n \times n$ *symmetric* matrix with non-negative real eigenvalues, say $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$, and it has an associated *orthonormal* eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$. Here $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$.

The numbers $\boxed{\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|}$ are called the *singular values of A*. We really only care about the *non-zero* singular values. Suppose

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$

Here $1 \leq r \leq n$ and when $r < n$, $A\vec{v}_{r+1} = \cdots = A\vec{v}_n = \vec{0}$. So, if $\vec{x} = \sum_{i=1}^{n} c_i \vec{v}_i$, then

$$A\vec{x} = \sum_{i=1}^{n} c_i A\vec{v}_i = \sum_{i=1}^{r} c_i A\vec{v}_i.$$

The Matrix $M = A^TA$

When A is an $m \times n$ matrix, $M = A^TA$ is an $n \times n$ *symmetric* matrix with non-negative real eigenvalues, say $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$, and it has an associated *orthonormal* eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$. Here $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$.

The numbers $\boxed{\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|}$ are called the *singular values of A*. We really only care about the *non-zero* singular values. Suppose

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$

Here $1 \leq r \leq n$ and when $r < n$, $A\vec{v}_{r+1} = \cdots = A\vec{v}_n = \vec{0}$. So, if $\vec{x} = \sum_{i=1}^n c_i \vec{v}_i$, then

$$A\vec{x} = \sum_{i=1}^{n} c_i A\vec{v}_i = \sum_{i=1}^{r} c_i A\vec{v}_i.$$

Thus, $\mathcal{CS}(A) = \mathcal{S}pan\{A\vec{v}_1, \ldots, A\vec{v}_r\}$.

When A is an $m \times n$ matrix, $M = A^T A$ is an $n \times n$ *symmetric* matrix with non-negative real eigenvalues, say $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$, and it has an associated *orthonormal* eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$. Here $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$.

The numbers $\boxed{\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|}$ are called the *singular values of A*. We really only care about the *non-zero* singular values. Suppose

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$

Here $1 \leq r \leq n$ and when $r < n$, $A\vec{v}_{r+1} = \cdots = A\vec{v}_n = \vec{0}$. So, if $\vec{x} = \sum_{i=1}^{n} c_i \vec{v}_i$, then

$$A\vec{x} = \sum_{i=1}^{n} c_i A\vec{v}_i = \sum_{i=1}^{r} c_i A\vec{v}_i.$$

Thus, $\mathcal{CS}(A) = \mathcal{S}pan\{A\vec{v}_1, \ldots, A\vec{v}_r\}$. Is $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ LI?

$M = A^T A$ where A is $m \times n$

$M = A^T A$ where A is $m \times n$

- $M = A^T A$ is an $n \times n$ symmetric matrix

$M = A^T A$ where A is $m \times n$

- $M = A^T A$ is an $n \times n$ symmetric matrix
- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M

$M = A^T A$ where A is $m \times n$

- $M = A^T A$ is an $n \times n$ symmetric matrix
- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- orthonormal eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ assoc'd with M

$M = A^T A$ where A is $m \times n$

- $M = A^T A$ is an $n \times n$ symmetric matrix
- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- orthonormal eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ assoc'd with M
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$

$M = A^T A$ where A is $m \times n$

- $M = A^T A$ is an $n \times n$ symmetric matrix
- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- orthonormal eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ assoc'd with M
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$

$M = A^T A$ where A is $m \times n$

- $M = A^T A$ is an $n \times n$ symmetric matrix
- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- orthonormal eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ assoc'd with M
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$

$M = A^T A$ where A is $m \times n$

- $M = A^T A$ is an $n \times n$ symmetric matrix
- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- orthonormal eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ assoc'd with M
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$

Look at

$$A\vec{v}_i \cdot A\vec{v}_j = (A\vec{v}_i)^T A\vec{v}_j = \vec{v}_i^T (A^T A)\vec{v}_j = \vec{v}_i^T (M\vec{v}_j) = \lambda_j \vec{v}_i^T \vec{v}_j = \lambda_j \vec{v}_i \cdot \vec{v}_j = \lambda_j \delta_{ij}.$$

$M = A^T A$ where A is $m \times n$

- $M = A^T A$ is an $n \times n$ symmetric matrix
- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- orthonormal eigenbasis $\{\vec{v}_1, \dots, \vec{v}_n\}$ assoc'd with M
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$

Look at

$$A\vec{v}_i \cdot A\vec{v}_j = (A\vec{v}_i)^T A\vec{v}_j = \vec{v}_i^T (A^T A)\vec{v}_j = \vec{v}_i^T (M\vec{v}_j) = \lambda_j \vec{v}_i^T \vec{v}_j = \lambda_j \vec{v}_i \cdot \vec{v}_j = \lambda_j \delta_{ij}.$$

Thus $\{A\vec{v}_1, \dots, A\vec{v}_r\}$ is an orthogonal set, so it is LI.

$M = A^T A$ where A is $m \times n$

- $M = A^T A$ is an $n \times n$ symmetric matrix
- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- orthonormal eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ assoc'd with M
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$

Look at

$$A\vec{v}_i \cdot A\vec{v}_j = (A\vec{v}_i)^T A\vec{v}_j = \vec{v}_i^T (A^T A)\vec{v}_j = \vec{v}_i^T (M\vec{v}_j) = \lambda_j \vec{v}_i^T \vec{v}_j = \lambda_j \vec{v}_i \cdot \vec{v}_j = \lambda_j \delta_{ij}.$$

Thus $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an orthogonal set, so it is LI.
$\therefore$ $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an *orthogonal basis* for $\mathcal{CS}(A)$.

$M = A^T A$ where A is $m \times n$

- $M = A^T A$ is an $n \times n$ symmetric matrix
- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- orthonormal eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ assoc'd with M
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$

$M = A^T A$ where A is $m \times n$

- $M = A^T A$ is an $n \times n$ symmetric matrix
- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- orthonormal eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ assoc'd with M
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$
- $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an *orthogonal basis* for $\mathcal{CS}(A)$

$M = A^T A$ where A is $m \times n$

- $M = A^T A$ is an $n \times n$ symmetric matrix
- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- orthonormal eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ assoc'd with M
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$
- $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an *orthogonal basis* for $\mathcal{CS}(A)$

It follows that $\operatorname{rank}(A) = \dim \mathcal{CS}(A) = r$.

$M = A^T A$ where A is $m \times n$

- $M = A^T A$ is an $n \times n$ symmetric matrix
- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- orthonormal eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ assoc'd with M
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$
- $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an *orthogonal basis* for $\mathcal{CS}(A)$

It follows that $\text{rank}(A) = \dim \mathcal{CS}(A) = r$. Recall that $\mathcal{CS}(A)$ is vector subspace of $\mathbb{R}^m$.

$M = A^T A$ where A is $m \times n$

- $M = A^T A$ is an $n \times n$ symmetric matrix
- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- orthonormal eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ assoc'd with M
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$
- $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an *orthogonal basis* for $\mathcal{CS}(A)$

It follows that $\text{rank}(A) = \dim \mathcal{CS}(A) = r$. Recall that $\mathcal{CS}(A)$ is vector subspace of $\mathbb{R}^m$.

For $1 \leq i \leq r$, let $\vec{u}_i = \dfrac{A\vec{v}_i}{\|A\vec{v}_i\|}$.

$M = A^T A$ where A is $m \times n$

- $M = A^T A$ is an $n \times n$ symmetric matrix
- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- orthonormal eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ assoc'd with M
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$
- $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an *orthogonal basis* for $\mathcal{CS}(A)$

It follows that $\operatorname{rank}(A) = \dim \mathcal{CS}(A) = r$. Recall that $\mathcal{CS}(A)$ is vector subspace of $\mathbb{R}^m$.

For $1 \leq i \leq r$, let $\vec{u}_i = \dfrac{A\vec{v}_i}{\|A\vec{v}_i\|}$. Then $\boxed{A\vec{v}_i = \sigma_i \vec{u}_i}$.

$M = A^T A$ where A is $m \times n$

- $M = A^T A$ is an $n \times n$ symmetric matrix
- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- orthonormal eigenbasis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ assoc'd with M
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$
- $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$
- $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an *orthogonal basis* for $\mathcal{CS}(A)$

It follows that $\text{rank}(A) = \dim \mathcal{CS}(A) = r$. Recall that $\mathcal{CS}(A)$ is vector subspace of $\mathbb{R}^m$.

For $1 \leq i \leq r$, let $\vec{u}_i = \dfrac{A\vec{v}_i}{\|A\vec{v}_i\|}$. Then $\boxed{A\vec{v}_i = \sigma_i \vec{u}_i}$. Add unit orthogonal vectors to get orthonormal basis $\{\vec{u}_1, \ldots, \vec{u}_r, \vec{u}_{r+1}, \ldots, \vec{u}_m\}$ for $\mathbb{R}^m$.

Matrices: A ($m \times n$), $M = A^T A$ ($n \times n$), and U, Σ, V

- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- $\{\vec{v}_1, \ldots, \vec{v}_n\}$ orthon eigenbasis assoc'd with M (so basis for $\mathbb{R}^n$)
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$ ($\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$)
- $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an *orthogonal basis* for $\mathcal{CS}(A)$
- $\{\vec{u}_1, \ldots, \vec{u}_r, \ldots, \vec{u}_m\}$ orthon basis for $\mathbb{R}^m$

Matrices: A $(m \times n)$, $M = A^T A$ $(n \times n)$, and U, Σ, V

- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- $\{\vec{v}_1, \ldots, \vec{v}_n\}$ orthon eigenbasis assoc'd with M (so basis for $\mathbb{R}^n$)
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$ $(\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0)$
- $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an *orthogonal basis* for $\mathcal{CS}(A)$
- $\{\vec{u}_1, \ldots, \vec{u}_r, \ldots, \vec{u}_m\}$ orthon basis for $\mathbb{R}^m$

Let $U = \left[\vec{u}_1\ \vec{u}_2 \ldots \vec{u}_m\right]$ (U is $m \times m$)

Matrices: A $(m \times n)$, $M = A^T A$ $(n \times n)$, and U, Σ, V

- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- $\{\vec{v}_1, \ldots, \vec{v}_n\}$ orthon eigenbasis assoc'd with M (so basis for $\mathbb{R}^n$)
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$ $(\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0)$
- $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an *orthogonal basis* for $\mathcal{CS}(A)$
- $\{\vec{u}_1, \ldots, \vec{u}_r, \ldots, \vec{u}_m\}$ orthon basis for $\mathbb{R}^m$

Let $U = \left[\vec{u}_1 \ \vec{u}_2 \ldots \vec{u}_m\right]$ (U is $m \times m$) and $V = \left[\vec{v}_1 \ \vec{v}_2 \ldots \vec{v}_n\right]$ (V is $n \times n$).

Matrices: A $(m \times n)$, $M = A^T A$ $(n \times n)$, and U, Σ, V

- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- $\{\vec{v}_1, \ldots, \vec{v}_n\}$ orthon eigenbasis assoc'd with M (so basis for $\mathbb{R}^n$)
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$ $(\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0)$
- $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an *orthogonal basis* for $\mathcal{CS}(A)$
- $\{\vec{u}_1, \ldots, \vec{u}_r, \ldots, \vec{u}_m\}$ orthon basis for $\mathbb{R}^m$

Let $U = \left[\vec{u}_1 \ \vec{u}_2 \ldots \vec{u}_m\right]$ (U is $m \times m$) and $V = \left[\vec{v}_1 \ \vec{v}_2 \ldots \vec{v}_n\right]$ (V is $n \times n$).
Now define an $r \times r$ diagonal matrix D and an $m \times n$ matrix Σ by

$$\Sigma = \begin{bmatrix} D & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad \text{where} \quad D = \begin{bmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ 0 & \vdots & \vdots & 0 \\ 0 & 0 & \ldots & \sigma_r \end{bmatrix}.$$

Matrices: A $(m \times n)$, $M = A^T A$ $(n \times n)$, and U, Σ, V

- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- $\{\vec{v}_1, \ldots, \vec{v}_n\}$ orthon eigenbasis assoc'd with M (so basis for $\mathbb{R}^n$)
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$ $(\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0)$
- $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an *orthogonal basis* for $\mathcal{CS}(A)$
- $\{\vec{u}_1, \ldots, \vec{u}_r, \ldots, \vec{u}_m\}$ orthon basis for $\mathbb{R}^m$

Let $U = \begin{bmatrix}\vec{u}_1 \; \vec{u}_2 \ldots \vec{u}_m\end{bmatrix}$ (U is $m \times m$) and $V = \begin{bmatrix}\vec{v}_1 \; \vec{v}_2 \ldots \vec{v}_n\end{bmatrix}$ (V is $n \times n$).
Now define an $r \times r$ diagonal matrix D and an $m \times n$ matrix Σ by

$$\Sigma = \begin{bmatrix} D & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad \text{where} \quad D = \begin{bmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ 0 & \vdots & \vdots & 0 \\ 0 & 0 & \ldots & \sigma_r \end{bmatrix}.$$

Then, $\boxed{A = U\Sigma V^T}$.

Matrices: A $(m \times n)$, $M = A^T A$ $(n \times n)$, and U, Σ, V

- eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ for M
- $\{\vec{v}_1, \ldots, \vec{v}_n\}$ orthon eigenbasis assoc'd with M (so basis for $\mathbb{R}^n$)
- $M\vec{v}_i = \lambda_i \vec{v}_i$ and $\lambda_i = \lambda_i \|\vec{v}_i\|^2 = \|A\vec{v}_i\|^2$
- *singular values* (of A) $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$ $(\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0)$
- $\{A\vec{v}_1, \ldots, A\vec{v}_r\}$ is an *orthogonal basis* for $\mathcal{CS}(A)$
- $\{\vec{u}_1, \ldots, \vec{u}_r, \ldots, \vec{u}_m\}$ orthon basis for $\mathbb{R}^m$

Let $U = \begin{bmatrix}\vec{u}_1 \ \vec{u}_2 \ldots \vec{u}_m\end{bmatrix}$ (U is $m \times m$) and $V = \begin{bmatrix}\vec{v}_1 \ \vec{v}_2 \ldots \vec{v}_n\end{bmatrix}$ (V is $n \times n$).
Now define an $r \times r$ diagonal matrix D and an $m \times n$ matrix Σ by

$$\Sigma = \begin{bmatrix} D & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} \quad \text{where} \quad D = \begin{bmatrix} \sigma_1 & 0 & \ldots & 0 \\ 0 & \sigma_2 & \ldots & 0 \\ 0 & \vdots & \vdots & 0 \\ 0 & 0 & \ldots & \sigma_r \end{bmatrix}.$$

Then, $\boxed{A = U\Sigma V^T}$. This is a *singular value decomposition* for A.

Finding matrices U, Σ, V so that $A = U\Sigma V^T$

Start with *any* $m \times n$ matrix A.

Finding matrices U, Σ, V so that $A = U\Sigma V^T$

Start with *any* $m \times n$ matrix A.

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

Find the eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ of M and $\{\vec{v}_1, \ldots, \vec{v}_n\}$ (an orthon eigenbasis assoc'd with M, so a basis for $\mathbb{R}^n$).

Finding matrices U, Σ, V so that $A = U\Sigma V^T$

Start with *any* $m \times n$ matrix A.

1. Find an orthogonal diagonalization of $M = A^T A$.
2. Write down V and Σ.

Find the eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ of M and $\{\vec{v}_1, \ldots, \vec{v}_n\}$ (an orthon eigenbasis assoc'd with M, so a basis for $\mathbb{R}^n$).

Get $n \times n$ matrix $V = \left[\vec{v}_1 \ \vec{v}_2 \ldots \vec{v}_n\right]$.

Finding matrices U, Σ, V so that $A = U\Sigma V^T$

Start with *any* $m \times n$ matrix A.

1. Find an orthogonal diagonalization of $M = A^T A$.
2. Write down V and Σ.

Find the eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ of M and $\{\vec{v}_1, \ldots, \vec{v}_n\}$ (an orthon eigenbasis assoc'd with M, so a basis for $\mathbb{R}^n$).

Get $n \times n$ matrix $V = \left[\vec{v}_1 \ \vec{v}_2 \ldots \vec{v}_n\right]$. Also, $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$, which gives D and then Σ.

Finding matrices U, Σ, V so that $A = U\Sigma V^T$

Start with *any* $m \times n$ matrix A.

1. Find an orthogonal diagonalization of $M = A^T A$.
2. Write down V and Σ.
3. Find $\vec{u}_i = \dfrac{A\vec{v}_i}{\|A\vec{v}_i\|}$ and then U.

Find the eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ of M and $\{\vec{v}_1, \ldots, \vec{v}_n\}$ (an orthon eigenbasis assoc'd with M, so a basis for $\mathbb{R}^n$).

Get $n \times n$ matrix $V = \left[\vec{v}_1 \ \vec{v}_2 \ldots \vec{v}_n\right]$. Also, $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$, which gives D and then Σ.

Need $\{\vec{u}_1, \ldots, \vec{u}_r, \ldots, \vec{u}_m\}$, an orthon basis for $\mathbb{R}^m$.

Finding matrices U, Σ, V so that $A = U\Sigma V^T$

Start with *any* $m \times n$ matrix A.

1. Find an orthogonal diagonalization of $M = A^T A$.
2. Write down V and Σ.
3. Find $\vec{u}_i = \dfrac{A\vec{v}_i}{\|A\vec{v}_i\|}$ and then U.

Find the eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n \geq 0$ of M and $\{\vec{v}_1, \ldots, \vec{v}_n\}$ (an orthon eigenbasis assoc'd with M, so a basis for $\mathbb{R}^n$).

Get $n \times n$ matrix $V = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \ldots \vec{v}_n \end{bmatrix}$. Also, $\sigma_i = \sqrt{\lambda_i} = \|A\vec{v}_i\|$, which gives D and then Σ.

Need $\{\vec{u}_1, \ldots, \vec{u}_r, \ldots, \vec{u}_m\}$, an orthon basis for $\mathbb{R}^m$. Then have $m \times m$ matrix $U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 \ldots \vec{u}_m \end{bmatrix}$.

For $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$, $A^T A = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$ which has evs $\lambda_1 = 25, \lambda_2 = 9$.

For $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$, $A^T A = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$ which has evs $\lambda_1 = 25, \lambda_2 = 9$.

Thus $\sigma_1 = 5, \sigma_2 = 3$ and $\Sigma = \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$.

For $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$, $A^T A = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$ which has evs $\lambda_1 = 25, \lambda_2 = 9$.

Thus $\sigma_1 = 5, \sigma_2 = 3$ and $\Sigma = \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$. Get assoc'd e$\vec{v}$s $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

For $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$, $A^T A = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$ which has evs $\lambda_1 = 25, \lambda_2 = 9$.

Thus $\sigma_1 = 5, \sigma_2 = 3$ and $\Sigma = \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$. Get assoc'd e$\vec{v}$s $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ so

take $\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and then

For $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$, $A^T A = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$ which has evs $\lambda_1 = 25, \lambda_2 = 9$.

Thus $\sigma_1 = 5, \sigma_2 = 3$ and $\Sigma = \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$. Get assoc'd e$\vec{v}$s $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ so

take $\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and then $V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

For $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$, $A^T A = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$ which has evs $\lambda_1 = 25, \lambda_2 = 9$.

Thus $\sigma_1 = 5, \sigma_2 = 3$ and $\Sigma = \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$. Get assoc'd e$\vec{v}$s $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ so

take $\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and then $V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

Get $A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 0 \end{bmatrix}, A \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$

For $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$, $A^T A = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$ which has evs $\lambda_1 = 25, \lambda_2 = 9$.

Thus $\sigma_1 = 5, \sigma_2 = 3$ and $\Sigma = \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$. Get assoc'd e$\vec{v}$s $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ so

take $\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and then $V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

Get $A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 0 \end{bmatrix}, A \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$ so $\vec{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \vec{u}_2 = \frac{1}{3\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$.

For $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$, $A^T A = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$ which has evs $\lambda_1 = 25, \lambda_2 = 9$.

Thus $\sigma_1 = 5, \sigma_2 = 3$ and $\Sigma = \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$. Get assoc'd e$\vec{v}$s $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ so

take $\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and then $V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

Get $A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 0 \end{bmatrix}, A \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$ so $\vec{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \vec{u}_2 = \frac{1}{3\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$.

Need $\vec{u}_3 \cdot \vec{u}_1 = 0 = \vec{u}_3 \cdot \vec{u}_2$.

For $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$, $A^T A = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$ which has evs $\lambda_1 = 25, \lambda_2 = 9$.

Thus $\sigma_1 = 5, \sigma_2 = 3$ and $\Sigma = \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$. Get assoc'd e$\vec{v}$s $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ so

take $\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and then $V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

Get $A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 0 \end{bmatrix}$, $A \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$ so $\vec{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\vec{u}_2 = \frac{1}{3\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$.

Need $\vec{u}_3 \cdot \vec{u}_1 = 0 = \vec{u}_3 \cdot \vec{u}_2$. Use $\vec{u}_3 = \frac{1}{3} \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$.

For $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$, $A^T A = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$ which has evs $\lambda_1 = 25, \lambda_2 = 9$.

Thus $\sigma_1 = 5, \sigma_2 = 3$ and $\Sigma = \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$. Get assoc'd e$\vec{v}$s $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$ so

take $\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and then $V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

Get $A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 0 \end{bmatrix}$, $A \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$ so $\vec{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$, $\vec{u}_2 = \frac{1}{3\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$.

Need $\vec{u}_3 \cdot \vec{u}_1 = 0 = \vec{u}_3 \cdot \vec{u}_2$. Use $\vec{u}_3 = \frac{1}{3} \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$. Then

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

For $A = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$, $A^T A = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$ which has evs $\lambda_1 = 25, \lambda_2 = 9$.

Thus $\sigma_1 = 5, \sigma_2 = 3$ and $\Sigma = \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$. Get assoc'd e$\vec{v}$s $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ so

take $\vec{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ and then $V = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$.

Get $A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 5 \\ 0 \end{bmatrix}, A \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$ so $\vec{u}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \vec{u}_2 = \frac{1}{3\sqrt{2}} \begin{bmatrix} 1 \\ -1 \\ 4 \end{bmatrix}$.

Need $\vec{u}_3 \cdot \vec{u}_1 = 0 = \vec{u}_3 \cdot \vec{u}_2$. Use $\vec{u}_3 = \frac{1}{3} \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix}$. Then

$U = \frac{1}{3\sqrt{2}} \begin{bmatrix} 3 & 1 & -2\sqrt{2} \\ 3 & -1 & 2\sqrt{2} \\ 0 & 4 & \sqrt{2} \end{bmatrix}$. $\frac{1}{6} \begin{bmatrix} 3 & 1 & -2\sqrt{2} \\ 3 & -1 & 2\sqrt{2} \\ 0 & 4 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$