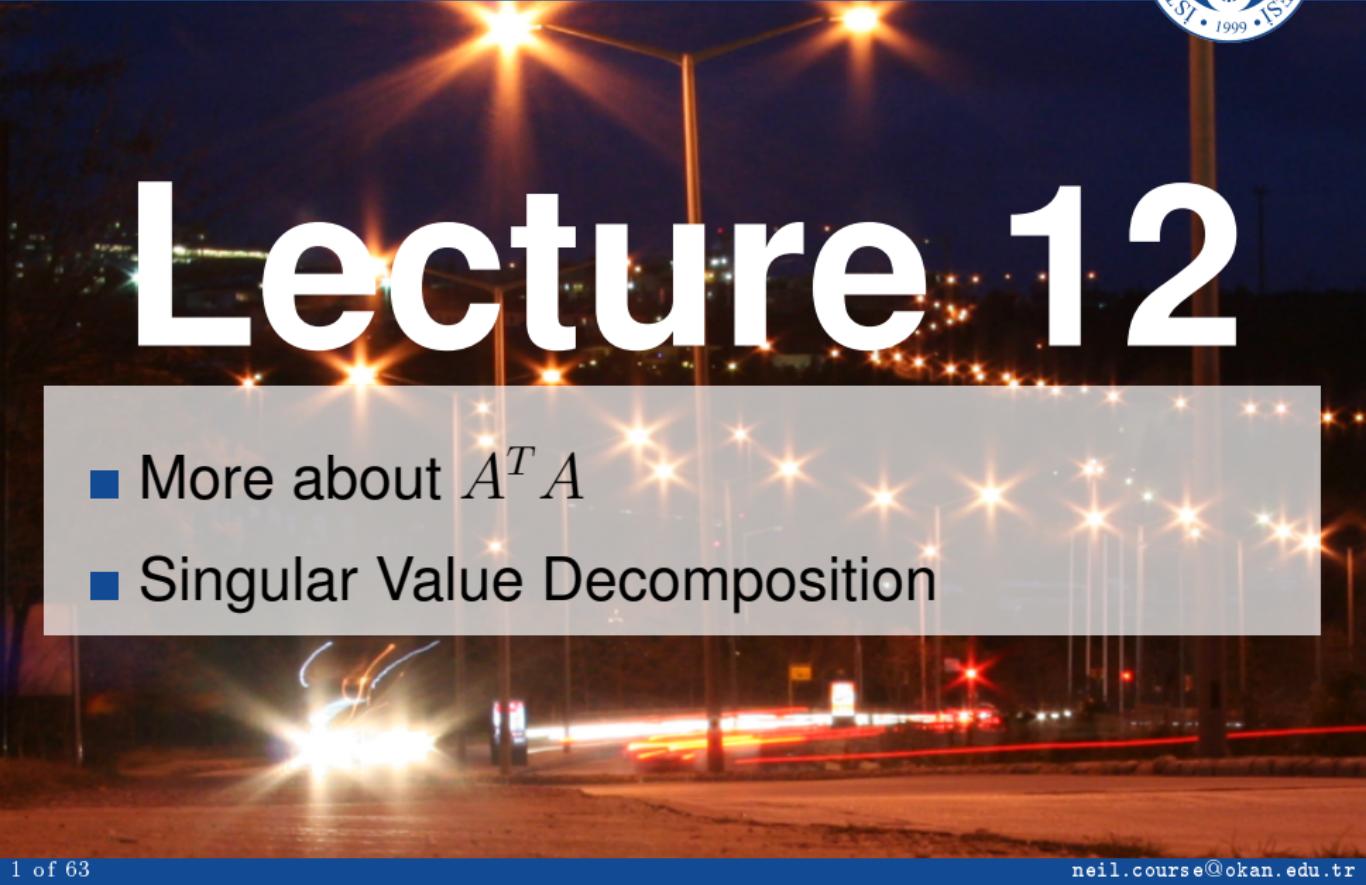


Lecture 12

- More about $A^T A$
- Singular Value Decomposition





More about $A^T A$

Revision from Lecture 3

Note that if A is an $m \times n$ matrix, then A^T is an $n \times m$ matrix, so the products AA^T and A^TA are both square matrices.

Revision from Lecture 3

Note that if A is an $m \times n$ matrix, then A^T is an $n \times m$ matrix, so the products AA^T and A^TA are both square matrices. Moreover, since $(AB)^T = B^TA^T$, we have

$$(AA^T)^T = (A^T)^T A^T$$

Revision from Lecture 3

Note that if A is an $m \times n$ matrix, then A^T is an $n \times m$ matrix, so the products AA^T and A^TA are both square matrices. Moreover, since $(AB)^T = B^TA^T$, we have

$$(AA^T)^T = (A^T)^T A^T = AA^T$$

More about $A^T A$

Revision from Lecture 3

Note that if A is an $m \times n$ matrix, then A^T is an $n \times m$ matrix, so the products AA^T and A^TA are both square matrices. Moreover, since $(AB)^T = B^T A^T$, we have

$$(A\textcolor{brown}{A}^T)^T = (\textcolor{green}{A}^T)^T \textcolor{brown}{A}^T = \textcolor{green}{A}\textcolor{brown}{A}^T$$

and

$$(A^T A)^T = A^T (A^T)^T = A^T A$$

which shows that both AA^T and A^TA are symmetric.

More about $A^T A$

Example

Let

$$A = \begin{bmatrix} 1 & -2 & 4 \\ 3 & 0 & -5 \end{bmatrix}.$$

Please check that

$$A^T A = \begin{bmatrix} 1 & 3 \\ -2 & 0 \\ 4 & -5 \end{bmatrix} \begin{bmatrix} 1 & -2 & 4 \\ 3 & 0 & -5 \end{bmatrix} = \begin{bmatrix} 10 & -2 & -11 \\ -2 & 4 & -8 \\ -11 & -8 & 41 \end{bmatrix}$$

and

$$AA^T = \begin{bmatrix} 1 & -2 & 4 \\ 3 & 0 & -5 \end{bmatrix} \begin{bmatrix} 1 & 3 \\ -2 & 0 \\ 4 & -5 \end{bmatrix} = \begin{bmatrix} 21 & -17 \\ -17 & 34 \end{bmatrix}.$$

More about $A^T A$

Theorem

If A is an invertible matrix, then AA^T and $A^T A$ are also invertible.

Proof.

A is invertible $\implies A^T$ is invertible. Recall that the product of two invertible matrices is invertible. □



Six Lemmata about $A^T A$

Let A be an $m \times n$ matrix. I want to state and prove 6 lemmata¹ about $A^T A$.

¹“lemmata” is the plural of “lemma”

More about $A^T A$

Lemma (1/6)

A and $A^T A$ have the same null space.

More about $A^T A$

Lemma (1/6)

A and $A^T A$ have the same null space.

Proof.

We must show that

$$\begin{array}{ccc} \mathbf{x}_0 \text{ is a solution of} & \iff & \mathbf{x}_0 \text{ is a solution of} \\ A\mathbf{x} = \mathbf{0} & & A^T A\mathbf{x} = \mathbf{0} \end{array}$$

More about $A^T A$

Lemma (1/6)

A and $A^T A$ have the same null space.

Proof.

We must show that

$$\begin{array}{ccc} \mathbf{x}_0 \text{ is a solution of} & \iff & \mathbf{x}_0 \text{ is a solution of} \\ A\mathbf{x} = \mathbf{0} & & A^T A\mathbf{x} = \mathbf{0} \end{array}$$

Note that

$$\begin{array}{ccc} \mathbf{x}_0 \text{ is a solution of} & \implies & A^T A\mathbf{x}_0 = A^T(A\mathbf{x}_0) = A^T \mathbf{0} = \mathbf{0}. \\ A\mathbf{x} = \mathbf{0} & & \end{array}$$

Next we need to prove “ \Leftarrow ”.

More about $A^T A$

Proof Continued.

Suppose that \mathbf{x}_0 is a solution of $A^T A \mathbf{x} = \mathbf{0}$, which is the same as supposing that $\mathbf{x}_0 \in \text{Nul}(A^T A)$. We need to show that $A\mathbf{x}_0 = \mathbf{0}$ too.

More about $A^T A$

Proof Continued.

Suppose that \mathbf{x}_0 is a solution of $A^T A \mathbf{x} = \mathbf{0}$, which is the same as supposing that $\mathbf{x}_0 \in \text{Nul}(A^T A)$. We need to show that $A\mathbf{x}_0 = \mathbf{0}$ too.

Recall that in Lecture 7 we saw that

$$\text{Nul } B = (\text{Row } B)^\perp.$$

This means that \mathbf{x}_0 is orthogonal to every vector in the row space of $A^T A$.

More about $A^T A$

Proof Continued.

Suppose that \mathbf{x}_0 is a solution of $A^T A \mathbf{x} = \mathbf{0}$, which is the same as supposing that $\mathbf{x}_0 \in \text{Nul}(A^T A)$. We need to show that $A\mathbf{x}_0 = \mathbf{0}$ too.

Recall that in Lecture 7 we saw that

$$\text{Nul } B = (\text{Row } B)^\perp.$$

This means that \mathbf{x}_0 is orthogonal to every vector in the row space of $A^T A$.

But $A^T A$ is symmetric, so \mathbf{x}_0 is also orthogonal to every vector in the column space of $A^T A$.

More about $A^T A$

Proof Continued.

Since the vector $\mathbf{y}_0 = (A^T A)\mathbf{x}_0$ is in $\text{Col}(A^T A)$ we must have

$$0 = \mathbf{x}_0 \cdot \mathbf{y}_0 = \quad = \quad = .$$

More about $A^T A$

Proof Continued.

Since the vector $\mathbf{y}_0 = (A^T A)\mathbf{x}_0$ is in $\text{Col}(A^T A)$ we must have

$$0 = \mathbf{x}_0 \cdot \mathbf{y}_0 = \mathbf{x}_0 \cdot A^T A \mathbf{x}_0 = \quad = \quad .$$

More about $A^T A$

Proof Continued.

Since the vector $\mathbf{y}_0 = (A^T A)\mathbf{x}_0$ is in $\text{Col}(A^T A)$ we must have

$$0 = \mathbf{x}_0 \cdot \mathbf{y}_0 = \mathbf{x}_0 \cdot A^T A \mathbf{x}_0 = A\mathbf{x}_0 \cdot A\mathbf{x}_0 = .$$

More about $A^T A$

Proof Continued.

Since the vector $\mathbf{y}_0 = (A^T A)\mathbf{x}_0$ is in $\text{Col}(A^T A)$ we must have

$$0 = \mathbf{x}_0 \cdot \mathbf{y}_0 = \mathbf{x}_0 \cdot A^T A \mathbf{x}_0 = A\mathbf{x}_0 \cdot A\mathbf{x}_0 = \|A\mathbf{x}_0\|^2.$$

More about $A^T A$

Proof Continued.

Since the vector $\mathbf{y}_0 = (A^T A)\mathbf{x}_0$ is in $\text{Col}(A^T A)$ we must have

$$0 = \mathbf{x}_0 \cdot \mathbf{y}_0 = \mathbf{x}_0 \cdot A^T A \mathbf{x}_0 = A\mathbf{x}_0 \cdot A\mathbf{x}_0 = \|A\mathbf{x}_0\|^2.$$

Hence $A\mathbf{x}_0 = \mathbf{0}$ and we are finished.



More about $A^T A$

Lemma (2/6)

A and $A^T A$ have the same rank

More about $A^T A$

Lemma (2/6)

A and $A^T A$ have the same rank

Proof.

Recall that for any matrix B we have

$$\text{rank } B + \text{nullity } B = \begin{matrix} \text{number of} \\ \text{columns in } B. \end{matrix}$$

More about $A^T A$

Lemma (2/6)

A and $A^T A$ have the same rank

Proof.

Recall that for any matrix B we have

$$\text{rank } B + \text{nullity } B = \frac{\text{number of columns in } B}{}$$

Since A and $A^T A$ have the same null space, and since they have the same number of columns we must have that

$$\text{rank } A = \text{rank}(A^T A).$$



More about $A^T A$

Lemma (3/6)

A and $A^T A$ have the same row space.

More about $A^T A$

Lemma (3/6)

A and $A^T A$ have the same row space.

Proof.

Because $A^T A$ is symmetric, we have $\text{Row}(A^T A) = \text{Col}(A^T A)$.

Each column of $A^T A$ is a linear combination of the columns in A^T because of the way in which we multiply matrices together (think about it).

More about $A^T A$

Lemma (3/6)

A and $A^T A$ have the same row space.

Proof.

Because $A^T A$ is symmetric, we have $\text{Row}(A^T A) = \text{Col}(A^T A)$.

Each column of $A^T A$ is a linear combination of the columns in A^T because of the way in which we multiply matrices together (think about it).

So every column of $A^T A$ is in $\text{Col } A^T = \text{Row } A$.

More about $A^T A$

Lemma (3/6)

A and $A^T A$ have the same row space.

Proof.

Because $A^T A$ is symmetric, we have $\text{Row}(A^T A) = \text{Col}(A^T A)$.

Each column of $A^T A$ is a linear combination of the columns in A^T because of the way in which we multiply matrices together (think about it).

So every column of $A^T A$ is in $\text{Col } A^T = \text{Row } A$. So every row of $A^T A$ is in $\text{Row}(A)$.

More about $A^T A$

Lemma (3/6)

A and $A^T A$ have the same row space.

Proof.

Because $A^T A$ is symmetric, we have $\text{Row}(A^T A) = \text{Col}(A^T A)$.

Each column of $A^T A$ is a linear combination of the columns in A^T because of the way in which we multiply matrices together (think about it).

So every column of $A^T A$ is in $\text{Col } A^T = \text{Row } A$. So every row of $A^T A$ is in $\text{Row}(A)$. This proves that $\text{Row}(A^T A)$ is a subspace of $\text{Row } A$.

More about $A^T A$

Lemma (3/6)

A and $A^T A$ have the same row space.

Proof.

Because $A^T A$ is symmetric, we have $\text{Row}(A^T A) = \text{Col}(A^T A)$.

Each column of $A^T A$ is a linear combination of the columns in A^T because of the way in which we multiply matrices together (think about it).

So every column of $A^T A$ is in $\text{Col } A^T = \text{Row } A$. So every row of $A^T A$ is in $\text{Row}(A)$. This proves that $\text{Row}(A^T A)$ is a subspace of $\text{Row } A$.

But since $A^T A$ and A have the same rank, we must have that $\text{Row } A = \text{Row}(A^T A)$. □

More about $A^T A$

Lemma (4/6)

A^T and $A^T A$ have the same column space.

More about $A^T A$

Lemma (4/6)

A^T and $A^T A$ have the same column space.

Proof.

$$\text{Col } A^T = \text{Row } A = \text{Row}(A^T A) = \text{Col}(A^T A).$$



More about $A^T A$

Lemma (5/6)

$A^T A$ is orthogonally diagonalisable.

More about $A^T A$

Lemma (5/6)

$A^T A$ is orthogonally diagonalisable.

Proof.

$A^T A$ is symmetric and all symmetric matrices are orthogonally diagonalisable. □

More about $A^T A$

Lemma (6/6)

The eigenvalues of $A^T A$ are nonnegative (i.e. $\lambda \geq 0$).

More about $A^T A$

Lemma (6/6)

The eigenvalues of $A^T A$ are nonnegative (i.e. $\lambda \geq 0$).

Proof.

Since $A^T A$ is orthogonally diagonalisable, there is an orthonormal basis for \mathbb{R}^n consisting of eigenvectors of $A^T A$, say

$$\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}.$$

More about $A^T A$

Lemma (6/6)

The eigenvalues of $A^T A$ are nonnegative (i.e. $\lambda \geq 0$).

Proof.

Since $A^T A$ is orthogonally diagonalisable, there is an orthonormal basis for \mathbb{R}^n consisting of eigenvectors of $A^T A$, say

$$\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}.$$

Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the corresponding eigenvalues (these may not be all different).

More about $A^T A$

Lemma (6/6)

The eigenvalues of $A^T A$ are nonnegative (i.e. $\lambda \geq 0$).

Proof.

Since $A^T A$ is orthogonally diagonalisable, there is an orthonormal basis for \mathbb{R}^n consisting of eigenvectors of $A^T A$, say

$$\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}.$$

Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the corresponding eigenvalues (these may not be all different). For $1 \leq i \leq n$ we have that

$$0 \leq \|A\mathbf{u}_i\|^2 = A\mathbf{u}_i \cdot A\mathbf{u}_i$$

More about $A^T A$

Lemma (6/6)

The eigenvalues of $A^T A$ are nonnegative (i.e. $\lambda \geq 0$).

Proof.

Since $A^T A$ is orthogonally diagonalisable, there is an orthonormal basis for \mathbb{R}^n consisting of eigenvectors of $A^T A$, say

$$\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}.$$

Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the corresponding eigenvalues (these may not be all different). For $1 \leq i \leq n$ we have that

$$0 \leq \|A\mathbf{u}_i\|^2 = A\mathbf{u}_i \cdot A\mathbf{u}_i = \mathbf{u}_i \cdot A^T A \mathbf{u}_i$$

More about $A^T A$

Lemma (6/6)

The eigenvalues of $A^T A$ are nonnegative (i.e. $\lambda \geq 0$).

Proof.

Since $A^T A$ is orthogonally diagonalisable, there is an orthonormal basis for \mathbb{R}^n consisting of eigenvectors of $A^T A$, say

$$\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}.$$

Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the corresponding eigenvalues (these may not be all different). For $1 \leq i \leq n$ we have that

$$0 \leq \|A\mathbf{u}_i\|^2 = A\mathbf{u}_i \cdot A\mathbf{u}_i = \mathbf{u}_i \cdot A^T A \mathbf{u}_i = \mathbf{u}_i \cdot \lambda_i \mathbf{u}_i$$

More about $A^T A$

Lemma (6/6)

The eigenvalues of $A^T A$ are nonnegative (i.e. $\lambda \geq 0$).

Proof.

Since $A^T A$ is orthogonally diagonalisable, there is an orthonormal basis for \mathbb{R}^n consisting of eigenvectors of $A^T A$, say

$$\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}.$$

Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the corresponding eigenvalues (these may not be all different). For $1 \leq i \leq n$ we have that

$$0 \leq \|A\mathbf{u}_i\|^2 = A\mathbf{u}_i \cdot A\mathbf{u}_i = \mathbf{u}_i \cdot A^T A \mathbf{u}_i = \mathbf{u}_i \cdot \lambda_i \mathbf{u}_i = \lambda_i (\mathbf{u}_i \cdot \mathbf{u}_i)$$

More about $A^T A$

Lemma (6/6)

The eigenvalues of $A^T A$ are nonnegative (i.e. $\lambda \geq 0$).

Proof.

Since $A^T A$ is orthogonally diagonalisable, there is an orthonormal basis for \mathbb{R}^n consisting of eigenvectors of $A^T A$, say

$$\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}.$$

Let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the corresponding eigenvalues (these may not be all different). For $1 \leq i \leq n$ we have that

$$0 \leq \|A\mathbf{u}_i\|^2 = A\mathbf{u}_i \cdot A\mathbf{u}_i = \mathbf{u}_i \cdot A^T A \mathbf{u}_i = \mathbf{u}_i \cdot \lambda_i \mathbf{u}_i = \lambda_i (\mathbf{u}_i \cdot \mathbf{u}_i) = \lambda_i$$

□

Singular Values

Definition

If A is an $m \times n$ matrix, and if $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of $A^T A$, then the numbers

$$\sigma_1 = \sqrt{\lambda_1}, \quad \sigma_2 = \sqrt{\lambda_2}, \quad \dots, \quad \sigma_n = \sqrt{\lambda_n}$$

are called the *singular values* of A .

σ is the lowercase of the Greek letter sigma, and Σ is capital sigma.

Remark

From now on, we are going to assume that the eigenvalues of $A^T A$ are named so that

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$$

and hence that

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

More about $A^T A$

Example

Find the singular values of the matrix $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$.

More about $A^T A$

Example

Find the singular values of the matrix $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$.

First we need to find the eigenvalues of the matrix

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

More about $A^T A$

Example

Find the singular values of the matrix $A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$.

First we need to find the eigenvalues of the matrix

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

The characteristic equation is

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

More about $A^T A$



So the eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = 1$. (Note that we put the biggest eigenvalue first.)

More about $A^T A$

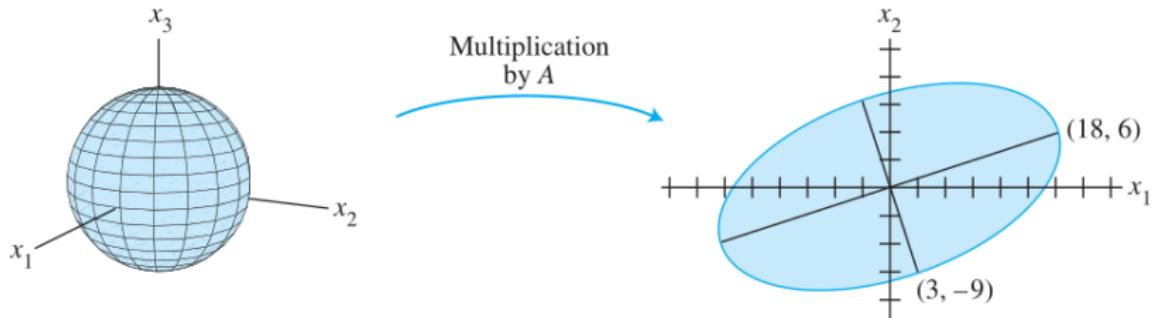


So the eigenvalues are $\lambda_1 = 3$ and $\lambda_2 = 1$. (Note that we put the biggest eigenvalue first.)

Therefore the singular values of A are

$$\sigma_1 = \sqrt{\lambda_1} = \sqrt{3}, \quad \sigma_2 = \sqrt{\lambda_2} = 1.$$

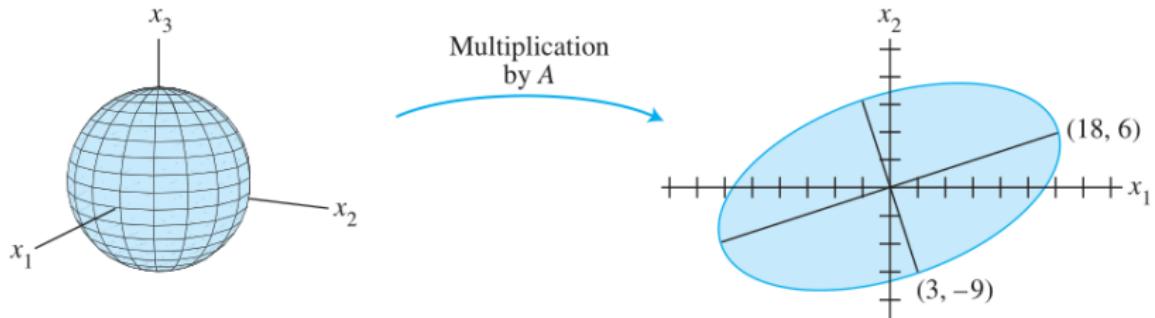
More about $A^T A$



Example

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

More about $A^T A$



Example

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

We need to find the eigenvalues of the following matrix, then take their square roots:

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

More about $A^T A$



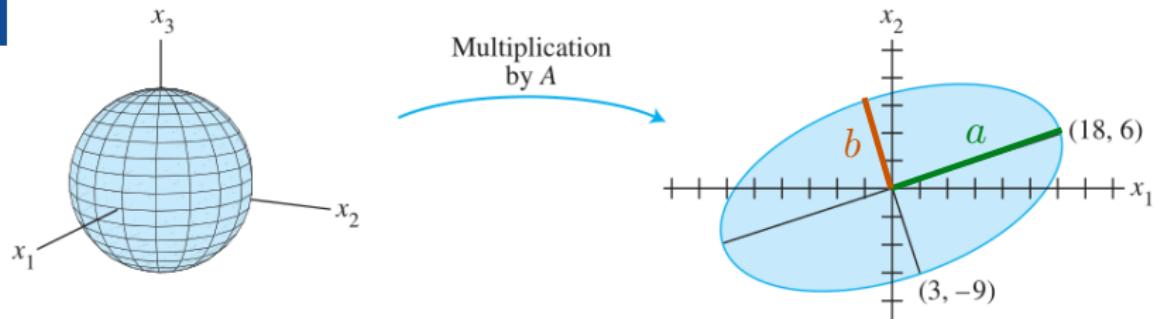
The eigenvalues of $A^T A$ are $\lambda_1 = 360$, $\lambda_2 = 90$ and $\lambda_3 = 0$ (please check!).

More about $A^T A$

The eigenvalues of $A^T A$ are $\lambda_1 = 360$, $\lambda_2 = 90$ and $\lambda_3 = 0$ (please check!). Therefore the singular values of A are

$$\sigma_1 = \sqrt{360} = 6\sqrt{10}, \quad \sigma_2 = \sqrt{90} = 3\sqrt{10}, \quad \sigma_3 = \sqrt{0} = 0.$$

More about $A^T A$



Remark

The matrix transformation $T_A : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, $\mathbf{x} \mapsto A\mathbf{x}$ maps the unit sphere in \mathbb{R}^3 to an ellipse in \mathbb{R}^2 as shown above. This ellipse has semimajor axis²

$$a = \sqrt{18^2 + 6^2} = \sqrt{360} = \sigma_1$$

and semiminor axis

$$b = \sqrt{3^2 + (-9)^2} = \sqrt{90} = \sigma_2.$$

²see Thomas' Calculus

More about $A^T A$

Theorem

Suppose $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ 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

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0,$$

and suppose A has r nonzero singular values.

More about $A^T A$

Theorem

Suppose $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ 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

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0,$$

and suppose A has r nonzero singular values.

Then $\{A\mathbf{u}_1, A\mathbf{u}_2, \dots, A\mathbf{u}_r\}$ is an orthogonal basis for $\text{Col } A$, and

$$\text{rank } A = r.$$

More about $A^T A$

Proof.

Suppose that $i \neq j$. Since $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ is an orthonormal basis, we know that

$$\mathbf{u}_i \cdot \mathbf{u}_j = 0.$$

More about $A^T A$

Proof.

Suppose that $i \neq j$. Since $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ is an orthonormal basis, we know that

$$\mathbf{u}_i \cdot \mathbf{u}_j = 0.$$

Therefore

$$A\mathbf{u}_i \cdot A\mathbf{u}_j = \mathbf{u}_i \cdot A^T A\mathbf{u}_j =$$

More about $A^T A$

Proof.

Suppose that $i \neq j$. Since $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ is an orthonormal basis, we know that

$$\mathbf{u}_i \cdot \mathbf{u}_j = 0.$$

Therefore

$$A\mathbf{u}_i \cdot A\mathbf{u}_j = \mathbf{u}_i \cdot A^T A\mathbf{u}_j = \mathbf{u}_i \cdot \lambda_j \mathbf{u}_j =$$

More about $A^T A$

Proof.

Suppose that $i \neq j$. Since $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ is an orthonormal basis, we know that

$$\mathbf{u}_i \cdot \mathbf{u}_j = 0.$$

Therefore

$$A\mathbf{u}_i \cdot A\mathbf{u}_j = \mathbf{u}_i \cdot A^T A\mathbf{u}_j = \mathbf{u}_i \cdot \lambda_j \mathbf{u}_j = \lambda_j (\mathbf{u}_i \cdot \mathbf{u}_j) = 0.$$

More about $A^T A$

Proof.

Suppose that $i \neq j$. Since $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ is an orthonormal basis, we know that

$$\mathbf{u}_i \cdot \mathbf{u}_j = 0.$$

Therefore

$$A\mathbf{u}_i \cdot A\mathbf{u}_j = \mathbf{u}_i \cdot A^T A\mathbf{u}_j = \mathbf{u}_i \cdot \lambda_j \mathbf{u}_j = \lambda_j (\mathbf{u}_i \cdot \mathbf{u}_j) = 0.$$

Hence $\{A\mathbf{u}_1, A\mathbf{u}_2, \dots, A\mathbf{u}_n\}$ is an orthogonal set.

More about $A^T A$

Proof Continued.

Moreover since the theorem said that A has r nonzero singular values, and since

$$\|A\mathbf{u}_i\| = \sqrt{A\mathbf{u}_i \cdot A\mathbf{u}_i} = \sqrt{\lambda_i} \sqrt{\mathbf{u}_i \cdot \mathbf{u}_i} = \sigma_i,$$

we have that

$$A\mathbf{u}_i \neq \mathbf{0} \iff 1 \leq i \leq r.$$

More about $A^T A$

Proof Continued.

Moreover since the theorem said that A has r nonzero singular values, and since

$$\|A\mathbf{u}_i\| = \sqrt{A\mathbf{u}_i \cdot A\mathbf{u}_i} = \sqrt{\lambda_i} \sqrt{\mathbf{u}_i \cdot \mathbf{u}_i} = \sigma_i,$$

we have that

$$A\mathbf{u}_i \neq \mathbf{0} \iff 1 \leq i \leq r.$$

Thus

- $A\mathbf{u}_1, A\mathbf{u}_2, \dots, A\mathbf{u}_r$ are linearly independent vectors; and
- $A\mathbf{u}_1, A\mathbf{u}_2, \dots, A\mathbf{u}_r$ are in $\text{Col } A$.

More about $A^T A$

Proof Continued.

Finally, let \mathbf{y} be any vector in $\text{Col } A$. Let $\mathbf{x} \in \mathbb{R}^n$ satisfy $A\mathbf{x} = \mathbf{y}$. Since $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ is an orthonormal basis for \mathbb{R}^n , we can write

$$\mathbf{x} = c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2 + \dots + c_n \mathbf{u}_n.$$

More about $A^T A$

Proof Continued.

Finally, let \mathbf{y} be any vector in $\text{Col } A$. Let $\mathbf{x} \in \mathbb{R}^n$ satisfy $A\mathbf{x} = \mathbf{y}$. Since $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ is an orthonormal basis for \mathbb{R}^n , we can write

$$\mathbf{x} = c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2 + \dots + c_n \mathbf{u}_n.$$

But then

$$\begin{aligned}\mathbf{y} &= A\mathbf{x} \\ &= c_1 A\mathbf{u}_1 + c_2 A\mathbf{u}_2 + \dots + c_r A\mathbf{u}_r + c_{r+1} A\mathbf{u}_{r+1} + \dots + c_n A\mathbf{u}_n \\ &= c_1 A\mathbf{u}_1 + c_2 A\mathbf{u}_2 + \dots + c_r A\mathbf{u}_r + 0 + \dots + 0.\end{aligned}$$

More about $A^T A$

Proof Continued.

Finally, let \mathbf{y} be any vector in $\text{Col } A$. Let $\mathbf{x} \in \mathbb{R}^n$ satisfy $A\mathbf{x} = \mathbf{y}$. Since $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ is an orthonormal basis for \mathbb{R}^n , we can write

$$\mathbf{x} = c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2 + \dots + c_n \mathbf{u}_n.$$

But then

$$\begin{aligned}\mathbf{y} &= A\mathbf{x} \\ &= c_1 A\mathbf{u}_1 + c_2 A\mathbf{u}_2 + \dots + c_r A\mathbf{u}_r + c_{r+1} A\mathbf{u}_{r+1} + \dots + c_n A\mathbf{u}_n \\ &= c_1 A\mathbf{u}_1 + c_2 A\mathbf{u}_2 + \dots + c_r A\mathbf{u}_r + 0 + \dots + 0.\end{aligned}$$

This proves that $\mathbf{y} \in \text{span}\{A\mathbf{u}_1, A\mathbf{u}_2, \dots, A\mathbf{u}_r\}$ and thus that $\{A\mathbf{u}_1, A\mathbf{u}_2, \dots, A\mathbf{u}_r\}$ is an orthogonal basis for $\text{Col } A$.

More about $A^T A$

Proof Continued.

Finally, let \mathbf{y} be any vector in $\text{Col } A$. Let $\mathbf{x} \in \mathbb{R}^n$ satisfy $A\mathbf{x} = \mathbf{y}$. Since $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ is an orthonormal basis for \mathbb{R}^n , we can write

$$\mathbf{x} = c_1 \mathbf{u}_1 + c_2 \mathbf{u}_2 + \dots + c_n \mathbf{u}_n.$$

But then

$$\begin{aligned}\mathbf{y} &= A\mathbf{x} \\ &= c_1 A\mathbf{u}_1 + c_2 A\mathbf{u}_2 + \dots + c_r A\mathbf{u}_r + c_{r+1} A\mathbf{u}_{r+1} + \dots + c_n A\mathbf{u}_n \\ &= c_1 A\mathbf{u}_1 + c_2 A\mathbf{u}_2 + \dots + c_r A\mathbf{u}_r + 0 + \dots + 0.\end{aligned}$$

This proves that $\mathbf{y} \in \text{span}\{A\mathbf{u}_1, A\mathbf{u}_2, \dots, A\mathbf{u}_r\}$ and thus that $\{A\mathbf{u}_1, A\mathbf{u}_2, \dots, A\mathbf{u}_r\}$ is an orthogonal basis for $\text{Col } A$. We conclude that

$$\text{rank } A = \dim \text{Col } A = r.$$





Singular Value Decomposition

Singular Value Decomposition



In Lecture 9 we studied diagonalisation

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

and in Lecture 11 we studied orthogonal diagonalisation

$$A = PDP^T.$$

Unfortunately, as we have learned, not all matrices can be factorised like this.

Singular Value Decomposition



In Lecture 9 we studied diagonalisation

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

and in Lecture 11 we studied orthogonal diagonalisation

$$A = PDP^T.$$

Unfortunately, as we have learned, not all matrices can be factorised like this.

There is however, another type of factorisation that is possible for any $m \times n$ matrix A . This is called *Singular Value Decomposition* (SVD) and takes the form

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

Singular Value Decomposition



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

Remark

Note here that we have two different matrices U and V .

Singular Value Decomposition



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

Remark

Note here that we have two different matrices U and V .

Remark

Σ will be the same size as A ($m \times n$).

Singular Value Decomposition



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

Remark

Note here that we have two different matrices U and V .

Remark

Σ will be the same size as A ($m \times n$).

Remark

U will be an $m \times m$ orthogonal matrix.

V will be an $n \times n$ orthogonal matrix.

Singular Value Decomposition



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

The matrix Σ will be an $m \times n$ matrix of the form

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

where D is a diagonal matrix and each 0 is a zero matrix. (If $r = m$ or $r = n$ then some of the zero matrices do not appear.)

Singular Value Decomposition



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

The matrix Σ will be an $m \times n$ matrix of the form

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

$r \quad n - r$

Diagram illustrating the structure of the matrix Σ . The matrix is partitioned into four quadrants. The top-right quadrant contains a zero matrix of size $r \times (n-r)$. The bottom-left quadrant contains a zero matrix of size $(m-r) \times r$. The bottom-right quadrant contains a zero matrix of size $(m-r) \times (n-r)$. The top-left quadrant is labeled D , which is a diagonal matrix of size $r \times r$. Braces on the right side group the top-right and bottom-left zero matrices, both labeled r . Braces at the bottom group the bottom-left and bottom-right zero matrices, both labeled $n-r$.

where D is a diagonal matrix and each 0 is a zero matrix. (If $r = m$ or $r = n$ then some of the zero matrices do not appear.)

Singular Value Decomposition



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

How do we find V ?

Singular Value Decomposition



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

How do we find V ?

$$\textcolor{brown}{A^T} A = (\textcolor{brown}{U\Sigma V^T})^T U\Sigma V^T$$

=

=

Singular Value Decomposition



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

How do we find V ?

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

Singular Value Decomposition



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

How do we find V ?

$$\begin{aligned} \textcolor{brown}{A^T} A &= (\textcolor{brown}{U}\Sigma V^T)^T \textcolor{brown}{U}\Sigma V^T \\ &= V\Sigma^T \textcolor{red}{U^T} \textcolor{red}{U}\Sigma V^T \\ &= \end{aligned}$$

Singular Value Decomposition



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

How do we find V ?

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

Singular Value Decomposition



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

How do we find V ?

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

Note that $\Sigma^T \Sigma$ will be a diagonal matrix (think about it).
Hence V will be a matrix which orthogonally diagonalises $A^T A$.

Singular Value Decomposition



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

And how do we find U ?

$$A\textcolor{brown}{A}^T = U\Sigma V^T (\textcolor{brown}{U}\Sigma V^T)^T$$

=

=

Singular Value Decomposition



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

And how do we find U ?

$$\begin{aligned} A\textcolor{brown}{A^T} &= U\Sigma V^T (\textcolor{brown}{U\Sigma V^T})^T \\ &= U\Sigma \textcolor{red}{V^T} \textcolor{red}{V} \Sigma^T U^T \\ &= \end{aligned}$$

Singular Value Decomposition



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

And how do we find U ?

$$\begin{aligned} A\cancel{A^T} &= U\Sigma V^T (\cancel{U\Sigma V^T})^T \\ &= U\Sigma \cancel{V^T} \cancel{V} \Sigma^T U^T \\ &= U \underbrace{\Sigma \Sigma^T}_{\text{diagonal}} U^T. \end{aligned}$$

Hence U will be a matrix which orthogonally diagonalises AA^T .

Singular Value Decomposition



Theorem (The Singular Value Decomposition Theorem)

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

Singular Value Decomposition



Theorem (The Singular Value Decomposition Theorem)

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

- an $m \times n$ matrix $\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$ for which the diagonal entries in D are the first r singular values of A ,

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0;$$

- an $m \times m$ orthogonal matrix U ; and
- an $n \times n$ orthogonal matrix V ,

Singular Value Decomposition



Theorem (The Singular Value Decomposition Theorem)

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

- an $m \times n$ matrix $\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$ for which the diagonal entries in D are the first r singular values of A ,

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0;$$

- an $m \times m$ orthogonal matrix U ; and
- an $n \times n$ orthogonal matrix V ,

such that

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

Singular Value Decomposition



Theorem (The Singular Value Decomposition Theorem)

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

- an $m \times n$ matrix $\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$ for which the diagonal entries in D are the first r singular values of A ,

$$\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0;$$

- an $m \times m$ orthogonal matrix U ; and
- an $n \times n$ orthogonal matrix V ,

such that

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

Singular Value Decomposition



Any factorisation

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

where

- U and V are orthogonal matrices;

- $\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$; and

- $D = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_r \end{bmatrix}$

is called a *singular value decomposition* (or SVD) of A .

Singular Value Decomposition



$$A = U\Sigma V^T = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^T$$

Remark

The matrices U and V are not necessarily unique.

Singular Value Decomposition



$$A = U\Sigma V^T = U \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix} V^T$$

Remark

The matrices U and V are not necessarily unique.

Remark

D (and hence Σ) is unique, because the diagonal entries of D must be the first r singular values of A .

Singular Value Decomposition



Proof of the The Singular Value Decomposition Theorem.

Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be 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

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0.$$

Singular Value Decomposition



Proof of the The Singular Value Decomposition Theorem.

Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be 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

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0.$$

Let

$$V = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \end{bmatrix}.$$

Singular Value Decomposition



Proof of the The Singular Value Decomposition Theorem.

Let $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ be 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

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0.$$

Let

$$V = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \end{bmatrix}.$$

Note that

- V is an orthogonal $n \times n$ matrix; and
- this is almost exactly the same as the matrix P that we use for orthogonal diagonalisation. The only difference here is that we require that $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0$.

Singular Value Decomposition

Proof Continued.

By the previous theorem, we know that $\{A\mathbf{v}_1, A\mathbf{v}_2, \dots, A\mathbf{v}_r\}$ is an orthogonal basis for $\text{Col } A$. We need to normalise these vectors.

³We talked briefly about this in the “Dimension” section of Lecture 6

Singular Value Decomposition

Proof Continued.

By the previous theorem, we know that $\{A\mathbf{v}_1, A\mathbf{v}_2, \dots, A\mathbf{v}_r\}$ is an orthogonal basis for $\text{Col } A$. We need to normalise these vectors. For each $i \in \{1, 2, \dots, r\}$ define

$$\mathbf{u}_i = \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|} = \frac{A\mathbf{v}_i}{\sigma_i}.$$

Then $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\}$ is an orthonormal basis for $\text{Col } A$.

³We talked briefly about this in the “Dimension” section of Lecture 6

Singular Value Decomposition



Proof Continued.

By the previous theorem, we know that $\{A\mathbf{v}_1, A\mathbf{v}_2, \dots, A\mathbf{v}_r\}$ is an orthogonal basis for $\text{Col } A$. We need to normalise these vectors. For each $i \in \{1, 2, \dots, r\}$ define

$$\mathbf{u}_i = \frac{A\mathbf{v}_i}{\|A\mathbf{v}_i\|} = \frac{A\mathbf{v}_i}{\sigma_i}.$$

Then $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r\}$ is an orthonormal basis for $\text{Col } A$.

By adding in an extra $m - r$ vectors, we can extend³ this to be an orthonormal basis for \mathbb{R}^m :

$$\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_r, \mathbf{u}_{r+1}, \dots, \mathbf{u}_m\}.$$

³We talked briefly about this in the “Dimension” section of Lecture 6

Singular Value Decomposition



Proof Continued.

Let

$$U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_m \end{bmatrix}.$$

Singular Value Decomposition



Proof Continued.

Let

$$U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_m \end{bmatrix}.$$

Note that

- U is an orthogonal $m \times m$ matrix.

Singular Value Decomposition

$$\mathbf{u}_i = \frac{A\mathbf{v}_i}{\sigma_i}$$



Proof Continued.

Moreover, note that

$$AV = A \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \end{bmatrix}$$

=

=

Singular Value Decomposition

$$\mathbf{u}_i = \frac{A\mathbf{v}_i}{\sigma_i}$$



Proof Continued.

Moreover, note that

$$\begin{aligned} AV &= A \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \end{bmatrix} \\ &= \begin{bmatrix} A\mathbf{v}_1 & A\mathbf{v}_2 & \dots & A\mathbf{v}_r & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix} \\ &= \end{aligned}$$

Singular Value Decomposition

$$\mathbf{u}_i = \frac{A\mathbf{v}_i}{\sigma_i}$$



Proof Continued.

Moreover, note that

$$\begin{aligned} AV &= A \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \end{bmatrix} \\ &= \begin{bmatrix} A\mathbf{v}_1 & A\mathbf{v}_2 & \dots & A\mathbf{v}_r & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1 \mathbf{u}_1 & \sigma_2 \mathbf{u}_2 & \dots & \sigma_r \mathbf{u}_r & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix} \end{aligned}$$

Singular Value Decomposition

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$



Proof Continued.

and that

$$U\Sigma = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_m \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \sigma_r & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix}$$

=

=

Singular Value Decomposition



Proof Continued.

and that

$$\begin{aligned} U\Sigma &= \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_m \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \sigma_r & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1 \mathbf{u}_1 & \sigma_2 \mathbf{u}_2 & \dots & \sigma_r \mathbf{u}_r & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix} \\ &= \end{aligned}$$

Singular Value Decomposition

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$



Proof Continued.

and that

$$\begin{aligned} U\Sigma &= \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_m \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & \sigma_2 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & \sigma_r & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1 \mathbf{u}_1 & \sigma_2 \mathbf{u}_2 & \dots & \sigma_r \mathbf{u}_r & \mathbf{0} & \dots & \mathbf{0} \end{bmatrix} \\ &= AV. \end{aligned}$$

Singular Value Decomposition

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$



Proof Continued.

So we have

$$U\Sigma = AV.$$

It follows that

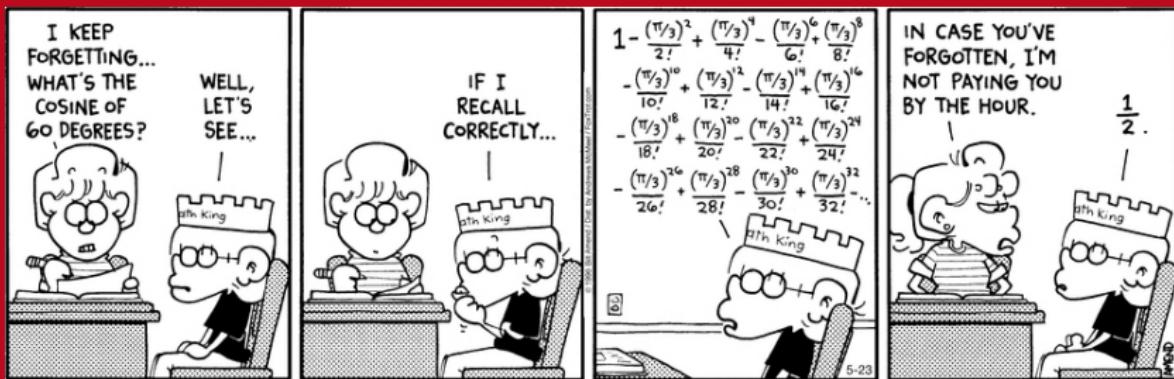
$$U\Sigma V^T = A V V^T = A$$

since V is an orthogonal matrix. □

(This was the last proof in this course.)

Break

We will continue at 3pm



How to do a Singular Value Decomposition

Let A be an $m \times n$ matrix of rank r .

- 1 Find⁴ an orthonormal basis for \mathbb{R}^n consisting of eigenvectors of $A^T A$. After rearranging as necessary, call this

$$\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$$

where

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0.$$

⁴Gram-Schmidt process

How to do a Singular Value Decomposition

Let A be an $m \times n$ matrix of rank r .

- 1 Find⁴ an orthonormal basis for \mathbb{R}^n consisting of eigenvectors of $A^T A$. After rearranging as necessary, call this

$$\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$$

where

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \geq 0.$$

- 2 Let $V = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \dots & \mathbf{v}_n \end{bmatrix}$. This will be an $n \times n$ orthogonal matrix.

⁴Gram-Schmidt process

Singular Value Decomposition



$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

3 Define

$$\mathbf{u}_1 = \frac{A\mathbf{v}_1}{\sigma_1}, \quad \mathbf{u}_2 = \frac{A\mathbf{v}_2}{\sigma_2}, \quad \dots, \quad \mathbf{u}_r = \frac{A\mathbf{v}_r}{\sigma_r}$$

where $\sigma_i = \sqrt{\lambda_i}$ are the singular values of A .

Singular Value Decomposition



$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

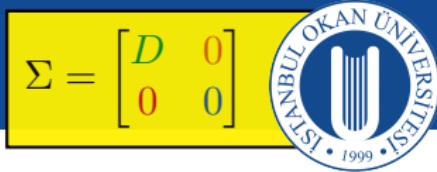
- 3 Define

$$\mathbf{u}_1 = \frac{A\mathbf{v}_1}{\sigma_1}, \quad \mathbf{u}_2 = \frac{A\mathbf{v}_2}{\sigma_2}, \quad \dots, \quad \mathbf{u}_r = \frac{A\mathbf{v}_r}{\sigma_r}$$

where $\sigma_i = \sqrt{\lambda_i}$ are the singular values of A .

- 4 Add in $m - r$ extra vectors so that $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$ is an orthonormal basis for \mathbb{R}^m .

Singular Value Decomposition



$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$

- 3 Define

$$\mathbf{u}_1 = \frac{A\mathbf{v}_1}{\sigma_1}, \quad \mathbf{u}_2 = \frac{A\mathbf{v}_2}{\sigma_2}, \quad \dots, \quad \mathbf{u}_r = \frac{A\mathbf{v}_r}{\sigma_r}$$

where $\sigma_i = \sqrt{\lambda_i}$ are the singular values of A .

- 4 Add in $m - r$ extra vectors so that $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m\}$ is an orthonormal basis for \mathbb{R}^m .

- 5 Let $U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_m \end{bmatrix}$. This will be an $m \times m$ orthogonal matrix.

Singular Value Decomposition

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix}$$



6 Write

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

where Σ is the $m \times n$ matrix of the form

$$\Sigma = \left[\begin{array}{cccc|c} \sigma_1 & 0 & \cdots & 0 & 0_{r \times (n-r)} \\ 0 & \sigma_2 & \cdots & 0 & \vdots \\ \vdots & \vdots & & \vdots & \\ 0 & 0 & \cdots & \sigma_r & \\ \hline & & & & \\ & 0_{(m-r) \times n} & & 0_{(m-r) \times (n-r)} & \end{array} \right]$$

Singular Value Decomposition



Example

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

Singular Value Decomposition



Example

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

We saw earlier that the eigenvalues of $A^T A$ are $\lambda_1 = 3$ and $\lambda_2 = 1$, and that the singular values of A are $\sigma_1 = \sqrt{3}$ and $\sigma_2 = 1$.

Singular Value Decomposition



Example

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

We saw earlier that the eigenvalues of $A^T A$ are $\lambda_1 = 3$ and $\lambda_2 = 1$, and that the singular values of A are $\sigma_1 = \sqrt{3}$ and $\sigma_2 = 1$. I leave it to you to check that

$$\mathbf{v}_1 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \begin{bmatrix} \frac{\sqrt{2}}{2} \\ -\frac{\sqrt{2}}{2} \end{bmatrix}$$

are eigenvectors corresponding to λ_1 and λ_2 , respectively, and that $\{\mathbf{v}_1, \mathbf{v}_2\}$ is an orthonormal basis for \mathbb{R}^2 .

Singular Value Decomposition



So we have

$$V = \begin{bmatrix} \textcolor{green}{v}_1 & \textcolor{orange}{v}_2 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} \end{bmatrix}.$$

(Note that V orthogonally diagonalises $A^T A$.)

Singular Value Decomposition

Next define

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A \mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} = \frac{\sqrt{3}}{3} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{6}}{3} \\ \frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} \end{bmatrix}$$

and

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

Singular Value Decomposition



Next define

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A \mathbf{v}_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} = \frac{\sqrt{3}}{3} \begin{bmatrix} \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \\ \frac{\sqrt{2}}{2} \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{6}}{3} \\ \frac{\sqrt{6}}{6} \\ \frac{\sqrt{6}}{6} \end{bmatrix}$$

and

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

Recall that U will be a 3×3 matrix. We have found the first two columns of U , now we need to find the third column.

Singular Value Decomposition



We need to find a unit vector \mathbf{u}_3 such that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal basis for \mathbb{R}^3 .

Singular Value Decomposition

We need to find a unit vector \mathbf{u}_3 such that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal basis for \mathbb{R}^3 . So we need to find a unit vector \mathbf{u}_3 that is orthogonal to both \mathbf{u}_1 and \mathbf{u}_2 .

Singular Value Decomposition

We need to find a unit vector \mathbf{u}_3 such that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal basis for \mathbb{R}^3 . So we need to find a unit vector \mathbf{u}_3 that is orthogonal to both \mathbf{u}_1 and \mathbf{u}_2 .

So as to make our calculations a little easier, let us instead look for a unit vector \mathbf{u}_3 that is orthogonal to both

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

Singular Value Decomposition



We need to find a unit vector \mathbf{u}_3 such that $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ is an orthonormal basis for \mathbb{R}^3 . So we need to find a unit vector \mathbf{u}_3 that is orthogonal to both \mathbf{u}_1 and \mathbf{u}_2 .

So as to make our calculations a little easier, let us instead look for a unit vector \mathbf{u}_3 that is orthogonal to both

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

So we want to find a unit vector \mathbf{u}_3 that is a solution to the homogeneous linear system

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

Singular Value Decomposition

I leave it to you to check that the general solution to

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

is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = t \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}.$$

Singular Value Decomposition

I leave it to you to check that the general solution to

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

is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = t \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}.$$

Therefore the unit vector that we want is

$$\mathbf{u}_3 = \frac{\begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix}}{\left\| \begin{bmatrix} -1 \\ 1 \\ 1 \end{bmatrix} \right\|} = \begin{bmatrix} -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}.$$

Singular Value Decomposition



So we have

$$U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{6}}{3} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} & \frac{1}{\sqrt{3}} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

Singular Value Decomposition



So we have

$$U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} \frac{\sqrt{6}}{3} & 0 & -\frac{1}{\sqrt{3}} \\ \frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{2} & \frac{1}{\sqrt{3}} \\ \frac{\sqrt{6}}{6} & \frac{\sqrt{2}}{2} & \frac{1}{\sqrt{3}} \end{bmatrix}.$$

Therefore a singular value decomposition of A is

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

Singular Value Decomposition



Example

Find a singular value decomposition of the matrix

$$A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}.$$

We saw earlier that the eigenvalues of $A^T A$ are $\lambda_1 = 360$, $\lambda_2 = 90$ and $\lambda_3 = 0$.

Singular Value Decomposition



Example

Find a singular value decomposition of the matrix

$$A = \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix}.$$

We saw earlier that the eigenvalues of $A^T A$ are $\lambda_1 = 360$, $\lambda_2 = 90$ and $\lambda_3 = 0$. I leave it to you to check that

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

are corresponding orthonormal eigenvectors of $A^T A$.

Singular Value Decomposition



So we have our first matrix

$$V = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & -\frac{2}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{1}{3} & -\frac{2}{3} \\ \frac{2}{3} & \frac{2}{3} & \frac{1}{3} \end{bmatrix}.$$

Singular Value Decomposition



The singular values of A are

$$\sigma_1 = 6\sqrt{10}, \quad \sigma_2 = 3\sqrt{10}, \quad \sigma_3 = 0.$$

Therefore

$$D = \begin{bmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \end{bmatrix}.$$

Singular Value Decomposition



The singular values of A are

$$\sigma_1 = 6\sqrt{10}, \quad \sigma_2 = 3\sqrt{10}, \quad \sigma_3 = 0.$$

Therefore

$$D = \begin{bmatrix} 6\sqrt{10} & 0 \\ 0 & 3\sqrt{10} \end{bmatrix}.$$

Recall that the matrix Σ is the same size as A (i.e. 2×3), with D in its upper left and zeros everywhere else:

$$\Sigma = [D \mid 0] = \left[\begin{array}{cc|c} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{array} \right].$$

Finally we need to construct U .

Singular Value Decomposition



$$\sigma_1 = 6\sqrt{10}, \quad \sigma_2 = 3\sqrt{10}, \quad \sigma_3 = 0$$

Since A has two nonzero singular values, we have $\text{rank } A = 2$.

Singular Value Decomposition



$$\sigma_1 = 6\sqrt{10}, \quad \sigma_2 = 3\sqrt{10}, \quad \sigma_3 = 0$$

Since A has two nonzero singular values, we have $\text{rank } A = 2$.

We calculate that

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A \mathbf{v}_1 = \frac{1}{6\sqrt{10}} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{3}{3} \\ \frac{2}{3} \\ \frac{3}{3} \end{bmatrix} = \frac{1}{6\sqrt{10}} \begin{bmatrix} 18 \\ 6 \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} \end{bmatrix}$$

$$\mathbf{u}_2 = \frac{1}{\sigma_2} A \mathbf{v}_2 = \frac{1}{3\sqrt{10}} \begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} \begin{bmatrix} -\frac{2}{3} \\ -\frac{1}{3} \\ \frac{2}{3} \end{bmatrix} = \frac{1}{3\sqrt{10}} \begin{bmatrix} 3 \\ -9 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{10}} \\ -\frac{3}{\sqrt{10}} \end{bmatrix}$$

Since $\{\mathbf{u}_1, \mathbf{u}_2\}$ is already a basis for \mathbb{R}^2 , we do not need to add any extra vectors in.

Singular Value Decomposition



So our third and final matrix is

$$U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \end{bmatrix}.$$

Singular Value Decomposition



So our third and final matrix is

$$U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \end{bmatrix}.$$

Therefore an SVD of A is

$$\begin{bmatrix} 4 & 11 & 14 \\ 8 & 7 & -2 \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{10}} & \frac{1}{\sqrt{10}} \\ \frac{1}{\sqrt{10}} & -\frac{3}{\sqrt{10}} \end{bmatrix} \begin{bmatrix} 6\sqrt{10} & 0 & 0 \\ 0 & 3\sqrt{10} & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{3} & \frac{2}{3} & \frac{2}{3} \\ -\frac{2}{3} & -\frac{1}{3} & \frac{2}{3} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

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

Example (Final Example of This Course)

Find a singular value decomposition of the matrix $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$.

Example (Final Example of This Course)

Find a singular value decomposition of the matrix $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$.

First we calculate that

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

Example (Final Example of This Course)

Find a singular value decomposition of the matrix $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$.

First we calculate that

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

The eigenvalues of $A^T A$ are $\lambda_1 = 18$ and $\lambda_2 = 0$ with corresponding eigenvectors

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{and} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Example (Final Example of This Course)

Find a singular value decomposition of the matrix $A = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix}$.

First we calculate that

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

The eigenvalues of $A^T A$ are $\lambda_1 = 18$ and $\lambda_2 = 0$ with corresponding eigenvectors

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{and} \quad \mathbf{x}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

These are orthogonal, but are not orthonormal so we need to normalise them:

$$\mathbf{v}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} \quad \text{and} \quad \mathbf{v}_2 = \frac{\mathbf{x}_2}{\|\mathbf{x}_2\|} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Singular Value Decomposition

Thus

$$V = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

Singular Value Decomposition



Thus

$$V = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

The singular values are $\sigma_1 = \sqrt{\lambda_1} = \sqrt{18} = 3\sqrt{2}$ and $\sigma_2 = \sqrt{\lambda_2} = 0$. Since there is only one nonzero singular value, D will be a 1×1 matrix. The matrix Σ is the same size as A , with D in its upper left corner.

Singular Value Decomposition



Thus

$$V = \begin{bmatrix} \mathbf{v}_1 & \mathbf{v}_2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

The singular values are $\sigma_1 = \sqrt{\lambda_1} = \sqrt{18} = 3\sqrt{2}$ and $\sigma_2 = \sqrt{\lambda_2} = 0$. Since there is only one nonzero singular value, D will be a 1×1 matrix. The matrix Σ is the same size as A , with D in its upper left corner. Hence

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

Singular Value Decomposition



To construct U , we will need $A\mathbf{v}_1$ and $A\mathbf{v}_2$:

Singular Value Decomposition

To construct U , we will need $A\mathbf{v}_1$ and $A\mathbf{v}_2$:

$$A\mathbf{v}_1 = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{2}} \\ -\frac{4}{\sqrt{2}} \\ \frac{4}{\sqrt{2}} \end{bmatrix}$$

$$A\mathbf{v}_2 = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Singular Value Decomposition



To construct U , we will need $A\mathbf{v}_1$ and $A\mathbf{v}_2$:

$$A\mathbf{v}_1 = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{2}} \\ -\frac{4}{\sqrt{2}} \\ \frac{4}{\sqrt{2}} \end{bmatrix}$$

$$A\mathbf{v}_2 = \begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Then we normalise the first one of these

$$\mathbf{u}_1 = \frac{1}{\sigma_1} A\mathbf{v}_1 = \frac{1}{3\sqrt{2}} \begin{bmatrix} \frac{2}{\sqrt{2}} \\ -\frac{4}{\sqrt{2}} \\ \frac{4}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix}.$$

Singular Value Decomposition

$$\mathbf{u}_1 = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ \frac{2}{3} \end{bmatrix}.$$



Since the matrix U needs to be a 3×3 matrix, we need another two orthogonal unit vectors \mathbf{u}_2 and \mathbf{u}_3 .

Singular Value Decomposition

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



Since the matrix U needs to be a 3×3 matrix, we need another two orthogonal unit vectors \mathbf{u}_2 and \mathbf{u}_3 .

Each of \mathbf{u}_2 and \mathbf{u}_3 must be orthogonal to \mathbf{u}_1 , but the general solution of

$$0 = 3\mathbf{u}_1 \cdot \mathbf{x} = x_1 - 2x_2 + 2x_3$$

is

$$\mathbf{x} = s\mathbf{w}_2 + t\mathbf{w}_3 = s \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}.$$

I leave it to you to check that each of \mathbf{w}_2 and \mathbf{w}_3 is orthogonal to \mathbf{u}_1 .

Singular Value Decomposition

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



Since the matrix U needs to be a 3×3 matrix, we need another two orthogonal unit vectors \mathbf{u}_2 and \mathbf{u}_3 .

Each of \mathbf{u}_2 and \mathbf{u}_3 must be orthogonal to \mathbf{u}_1 , but the general solution of

$$0 = 3\mathbf{u}_1 \cdot \mathbf{x} = x_1 - 2x_2 + 2x_3$$

is

$$\mathbf{x} = s\mathbf{w}_2 + t\mathbf{w}_3 = s \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}.$$

I leave it to you to check that each of \mathbf{w}_2 and \mathbf{w}_3 is orthogonal to \mathbf{u}_1 .

However \mathbf{w}_2 and \mathbf{w}_3 are not orthogonal to each other, so we need to apply the Gram-Schmidt process.

Singular Value Decomposition

$$\mathbf{u}_1 = \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{w}_3 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$



We define

$$\mathbf{z}_2 = \mathbf{w}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

$$\mathbf{z}_3 = \mathbf{w}_3 - \frac{\mathbf{w}_3 \cdot \mathbf{z}_2}{\|\mathbf{z}_2\|^2} \mathbf{z}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} - \frac{(-4)}{5} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{2}{5} \\ \frac{4}{5} \\ 1 \end{bmatrix}$$

Singular Value Decomposition

$$\mathbf{u}_1 = \begin{bmatrix} \frac{1}{3} \\ -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix}, \quad \mathbf{w}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \quad \mathbf{w}_3 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}$$



We define

$$\mathbf{z}_2 = \mathbf{w}_2 = \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}$$

$$\mathbf{z}_3 = \mathbf{w}_3 - \frac{\mathbf{w}_3 \cdot \mathbf{z}_2}{\|\mathbf{z}_2\|^2} \mathbf{z}_2 = \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} - \frac{(-4)}{5} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{2}{5} \\ \frac{4}{5} \\ 1 \end{bmatrix}$$

and then

$$\mathbf{u}_2 = \frac{\mathbf{z}_2}{\|\mathbf{z}_2\|} = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{bmatrix}, \quad \mathbf{u}_3 = \frac{\mathbf{z}_3}{\|\mathbf{z}_3\|} = \begin{bmatrix} -\frac{2}{\sqrt{45}} \\ \frac{4}{\sqrt{45}} \\ \frac{5}{\sqrt{45}} \end{bmatrix}.$$

Singular Value

$$\mathbf{u}_1 = \begin{bmatrix} \frac{1}{3} \\ \frac{2}{3} \\ -\frac{2}{3} \end{bmatrix}, \quad \mathbf{u}_2 = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \\ 0 \end{bmatrix}, \quad \mathbf{u}_3 = \begin{bmatrix} -\frac{2}{\sqrt{45}} \\ \frac{4}{\sqrt{45}} \\ \frac{5}{\sqrt{45}} \end{bmatrix}.$$



Thus

$$U = \begin{bmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \mathbf{u}_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{2}{\sqrt{5}} & -\frac{2}{\sqrt{45}} \\ -\frac{2}{3} & \frac{1}{\sqrt{5}} & \frac{4}{\sqrt{45}} \\ \frac{2}{3} & 0 & \frac{5}{\sqrt{45}} \end{bmatrix}.$$

Now we have everything that we need.

Singular Value Decomposition



Therefore an SVD for A is

$$\begin{bmatrix} 1 & -1 \\ -2 & 2 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} \frac{1}{3} & \frac{2}{\sqrt{5}} & -\frac{2}{\sqrt{45}} \\ -\frac{2}{3} & \frac{1}{\sqrt{5}} & \frac{4}{\sqrt{45}} \\ \frac{2}{3} & 0 & \frac{5}{\sqrt{45}} \end{bmatrix} \begin{bmatrix} 3\sqrt{2} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}.$$

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



The End

