Machine Learning for Data Mining

Linear Algebra Review

Andres Mendez-Vazquez

May 14, 2015

Outline

Outline

What is a Vector?

A ordered tuple of numbers

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

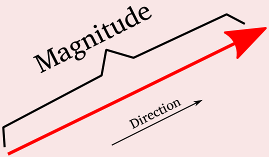
Expressing a magnitude and a direction

What is a Vector?

A ordered tuple of numbers

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Expressing a magnitude and a direction



Outline

Cinvestav

Vector Spaces

Definition

A vector is an element of a vector space

Vector Space V

It is a set that contains all linear combinations of its elements:

Vector Spaces

Definition

A vector is an element of a vector space

Vector Space V

It is a set that contains all linear combinations of its elements:

1. If $\boldsymbol{x}, \boldsymbol{y} \in V$ then $\boldsymbol{x} + \boldsymbol{y} \in V$.
2. If $\boldsymbol{x} \in V$ then $\alpha \boldsymbol{x} \in V$ for any scalar α.
3. There exists $0 \in V$ then $\boldsymbol{x} + 0 = \boldsymbol{x}$ for any $\boldsymbol{x} \in V$.

Vector Spaces

Definition

A vector is an element of a vector space

Vector Space V

It is a set that contains all linear combinations of its elements:

1. If $\boldsymbol{x}, \boldsymbol{y} \in V$ then $\boldsymbol{x} + \boldsymbol{y} \in V$.
2. If $\boldsymbol{x} \in V$ then $\alpha \boldsymbol{x} \in V$ for any scalar α.
3. There exists $0 \in V$ then $x + 0 = x$ for any $x \in V$.

A subspace

It is a subset of a vector space that is also a vector space

Vector Spaces

Definition

A vector is an element of a vector space

Vector Space V

It is a set that contains all linear combinations of its elements:

1. If $\boldsymbol{x}, \boldsymbol{y} \in V$ then $\boldsymbol{x} + \boldsymbol{y} \in V$.
2. If $\boldsymbol{x} \in V$ then $\alpha \boldsymbol{x} \in V$ for any scalar α.
3. There exists $0 \in V$ then $\boldsymbol{x} + \boldsymbol{0} = \boldsymbol{x}$ for any $\boldsymbol{x} \in V$.

A subspace

It is a subset of a vector space that is also a vector space

Cinvestav

Vector Spaces

Definition

A vector is an element of a vector space

Vector Space V

It is a set that contains all linear combinations of its elements:

1. If $\boldsymbol{x}, \boldsymbol{y} \in V$ then $\boldsymbol{x} + \boldsymbol{y} \in V$.
2. If $\boldsymbol{x} \in V$ then $\alpha \boldsymbol{x} \in V$ for any scalar α.
3. There exists $0 \in V$ then $\boldsymbol{x} + \mathbf{0} = \boldsymbol{x}$ for any $\boldsymbol{x} \in V$.

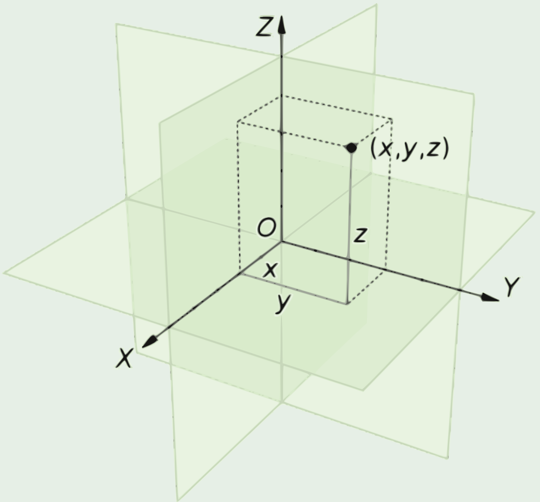
A subspace

It is a subset of a vector space that is also a vector space

Cinvestav

Classic Example

Euclidean Space $\mathbb{R}^3$



Span

Definition

The span of any set of vectors $\{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n\}$ is defined as:

$$\text{span}\left(\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n\right) = \alpha_1\boldsymbol{x}_1 + \alpha_2\boldsymbol{x}_2 + ... + \alpha_n\boldsymbol{x}_n$$

What Examples can you Imagine?

Give it a shot!!!

Span

Definition

The span of any set of vectors $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\}$ is defined as:

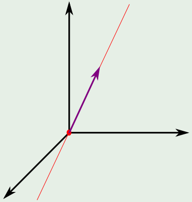
$$\text{span}\left(\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\right) = \alpha_1 \boldsymbol{x}_1 + \alpha_2 \boldsymbol{x}_2 + \ldots + \alpha_n \boldsymbol{x}_n$$

What Examples can you Imagine?

Give it a shot!!!

Subspaces of $\mathbb{R}^n$

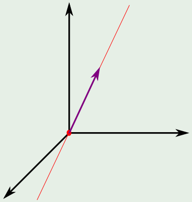
A line through the origin in $\mathbb{R}^n$



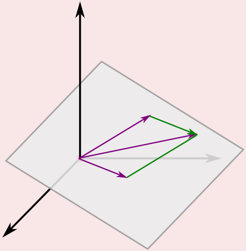
A plane in $\mathbb{R}^n$

Subspaces of $\mathbb{R}^n$

A line through the origin in $\mathbb{R}^n$



A plane in $\mathbb{R}^n$



Outline

Cinvestav

Linear Independence and Basis of Vector Spaces

Fact 1

A vector $\boldsymbol{x}$ is a linearly independent of a set of vectors $\{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n\}$ if it does not lie in their span.

Fact 2

A set of vectors is linearly independent if every vector is linearly independent of the rest.

Linear Independence and Basis of Vector Spaces

Fact 1

A vector $\boldsymbol{x}$ is a linearly independent of a set of vectors $\{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n\}$ if it does not lie in their span.

Fact 2

A set of vectors is linearly independent if every vector is linearly independent of the rest.

The Rest

1. A basis of a vector space V is a linearly independent set of vectors whose span is equal to V
2. If the basis has d vectors then the vector space V has dimensionality d

Cinvestav

Linear Independence and Basis of Vector Spaces

Fact 1

A vector $\boldsymbol{x}$ is a linearly independent of a set of vectors $\{\boldsymbol{x}_1, \boldsymbol{x}_2, ..., \boldsymbol{x}_n\}$ if it does not lie in their span.

Fact 2

A set of vectors is linearly independent if every vector is linearly independent of the rest.

The Rest

1. A basis of a vector space V is a linearly independent set of vectors whose span is equal to V
2. If the basis has *d* vectors then the vector space V has dimensionality *d*.

Outline

Norm of a Vector

Definition

A norm $\|u\|$ measures the magnitud of the vector.

Norm of a Vector

Definition

A norm $\|u\|$ measures the magnitud of the vector.

Properties

1. Homogeneity: $\|\alpha \boldsymbol{x}\| = \alpha \|x\|$.
2. Triangle inequality: $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.
3. Point Separation $\|x\| = 0$ if and only if $x = 0$.

4. Euclidean or ℓ_2-norm : $\|x\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$

Norm of a Vector

Definition

A norm $\|u\|$ measures the magnitud of the vector.

Properties

1. Homogeneity: $\|\alpha \boldsymbol{x}\| = \alpha \|x\|$.
2. Triangle inequality: $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.
3. Point Separation $\|x\| = 0$ if and only if $x = 0$.

Examples

1. Manhattan or ℓ_1-norm : $\|x\|_1 = \sum_{i=1}^{d} |x_i|$.
2. Euclidean or ℓ_2-norm : $\|x\|_2 = \sqrt{\sum_{i=1}^{d} x_i^2}$.

Norm of a Vector

Definition

A norm $\|u\|$ measures the magnitud of the vector.

Properties

1. Homogeneity: $\|\alpha \boldsymbol{x}\| = \alpha \|x\|$.
2. Triangle inequality: $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.
3. Point Separation $\|\boldsymbol{x}\| = 0$ if and only if $\boldsymbol{x} = 0$.

Examples

1. Manhattan or ℓ_1-norm : $\|\boldsymbol{x}\|_1 = \sum_{i=1}^{d} |x_i|$.
2. Euclidean or ℓ_2-norm : $\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i=1}^{d} x_i^2}$.

Cinvestav

Norm of a Vector

Definition

A norm $\|u\|$ measures the magnitud of the vector.

Properties

1. Homogeneity: $\|\alpha \boldsymbol{x}\| = \alpha \|x\|$.
2. Triangle inequality: $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.
3. Point Separation $\|\boldsymbol{x}\| = 0$ if and only if $\boldsymbol{x} = 0$.

Examples

1. Manhattan or ℓ_1-norm : $\|\boldsymbol{x}\|_1 = \sum_{i=1}^{d} |x_i|$.
2. Euclidean or ℓ_2-norm : $\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i=1}^{d} x_i^2}$.

Cinvestav

Norm of a Vector

Definition

A norm $\|u\|$ measures the magnitud of the vector.

Properties

1. Homogeneity: $\|\alpha \boldsymbol{x}\| = \alpha \|x\|$.
2. Triangle inequality: $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$.
3. Point Separation $\|\boldsymbol{x}\| = 0$ if and only if $\boldsymbol{x} = 0$.

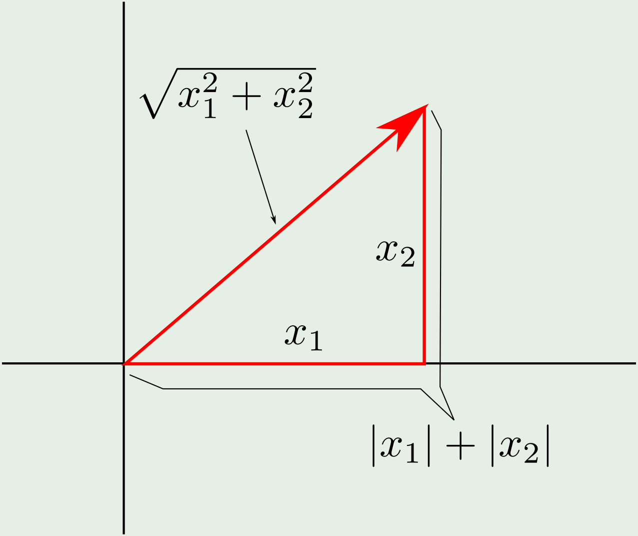
Examples

1. Manhattan or ℓ_1-norm : $\|\boldsymbol{x}\|_1 = \sum_{i=1}^{d} |x_i|$.
2. Euclidean or ℓ_2-norm : $\|\boldsymbol{x}\|_2 = \sqrt{\sum_{i=1}^{d} x_i^2}$.

Cinvestav

Examples

Example ℓ_1-norm and ℓ_2-norm



Outline

Cinvestav

Inner Product

Definition

The inner product between u and v

$$\langle u, v \rangle = \sum_{i=1}^{n} u_i v_i.$$

It is the projection of one vector onto the other one

Remark: It is related to the Euclidean norm: $\langle u, u \rangle = \|u\|_2^2$.

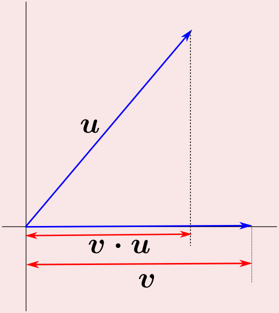
Inner Product

Definition

The inner product between u and v

$$\langle u, v\rangle = \sum_{i=1}^{n} u_i v_i.$$

It is the projection of one vector onto the other one



Properties

Meaning

The inner product is a measure of correlation between two vectors, scaled by the norms of the vectors

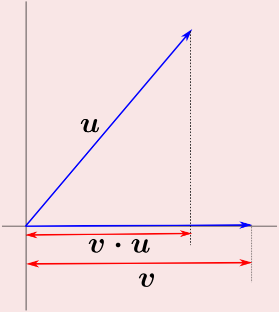
if $u \cdot v > 0$, u and v are aligned

Properties

Meaning

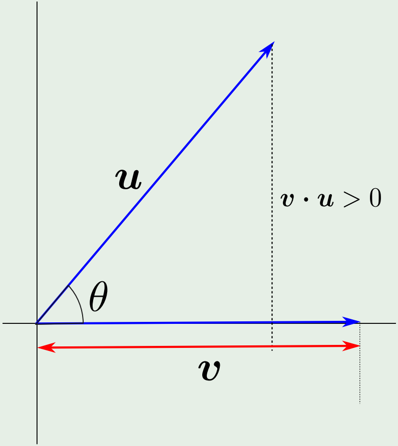
The inner product is a measure of correlation between two vectors, scaled by the norms of the vectors

if $\boldsymbol{u} \cdot \boldsymbol{v} > 0$, $\boldsymbol{u}$ and $\boldsymbol{v}$ are aligned



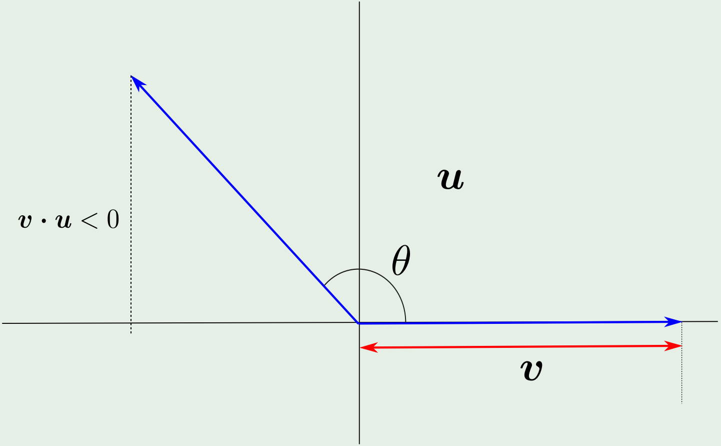
Properties

The inner product is a measure of correlation between two vectors, scaled by the norms of the vectors



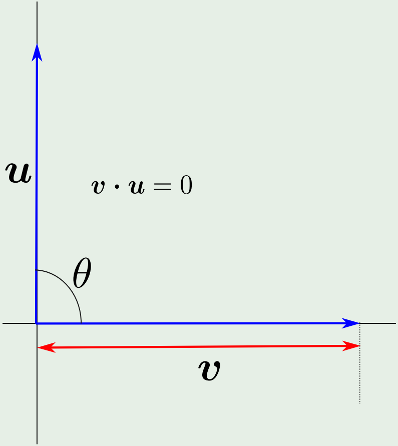
Properties

The inner product is a measure of correlation between two vectors, scaled by the norms of the vectors



Properties

The inner product is a measure of correlation between two vectors, scaled by the norms of the vectors



Definitions involving the norm

Orthonormal

The vectors in orthonormal basis have unit Euclidean norm and are orthonorgonal.

To express a vector x in an orthonormal basis

For example, given $x = \alpha_1 b_1 + \alpha_2 b_2$

Definitions involving the norm

Orthonormal

The vectors in orthonormal basis have unit Euclidean norm and are orthonorgonal.

To express a vector x in an orthonormal basis

For example, given $\boldsymbol{x} = \alpha_1 \boldsymbol{b}_1 + \alpha_2 \boldsymbol{b}_2$

$$
\begin{aligned}
\langle \boldsymbol{x}, \boldsymbol{b}_1 \rangle &= \langle \alpha_1 \boldsymbol{b}_1 + \alpha_2 \boldsymbol{b}_2, \boldsymbol{b}_1 \rangle \\
&= \alpha_1 \langle \boldsymbol{b}_1, \boldsymbol{b}_1 \rangle + \alpha_2 \langle \boldsymbol{b}_2, \boldsymbol{b}_1 \rangle \\
&= \alpha_1 + 0
\end{aligned}
$$

Likewise, $\langle \boldsymbol{x}, \boldsymbol{b}_2 \rangle = \alpha_2$

Definitions involving the norm

Orthonormal

The vectors in orthonormal basis have unit Euclidean norm and are orthonorgonal.

To express a vector x in an orthonormal basis

For example, given $\boldsymbol{x} = \alpha_1 \boldsymbol{b}_1 + \alpha_2 \boldsymbol{b}_2$

$$\begin{aligned} \langle \boldsymbol{x}, \boldsymbol{b}_1 \rangle &= \langle \alpha_1 \boldsymbol{b}_1 + \alpha_2 \boldsymbol{b}_2, \boldsymbol{b}_1 \rangle \\ &= \alpha_1 \langle \boldsymbol{b}_1, \boldsymbol{b}_1 \rangle + \alpha_2 \langle \boldsymbol{b}_2, \boldsymbol{b}_1 \rangle \\ &= \alpha_1 + 0 \end{aligned}$$

Likewise, $\langle \boldsymbol{x}, \boldsymbol{b}_2 \rangle = \alpha_2$

Outline

Cinvestav

Linear Operator

Definition

A linear operator $\mathcal{L} : U \to V$ is a map from a vector space U to another vector space V satisfies:

- $\mathcal{L}(\boldsymbol{u}_1 + \boldsymbol{u}_2) = \mathcal{L}(\boldsymbol{u}_1) + \mathcal{L}(\boldsymbol{u}_2)$

Something Notable

If the dimension n of U and m of V are finite, $\mathcal{L}$ can be represented by $m \times n$ matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Cinvestav

Linear Operator

Definition

A linear operator $\mathcal{L} : U \to V$ is a map from a vector space U to another vector space V satisfies:

- $\mathcal{L}(\boldsymbol{u}_1 + \boldsymbol{u}_2) = \mathcal{L}(\boldsymbol{u}_1) + \mathcal{L}(\boldsymbol{u}_2)$

Something Notable

If the dimension n of U and m of V are finite, $\mathcal{L}$ can be represented by $m \times n$ matrix:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

Thus, product of

The product of two linear operator can be seen as the multiplication of two matrices

$$AB = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ & & \cdots & \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{i=1}^{n} a_{1i} b_{i1} & \sum_{i=1}^{n} a_{1i} b_{i2} & \cdots & \sum_{i=1}^{n} a_{1i} b_{ip} \\ \sum_{i=1}^{n} a_{2i} b_{i1} & \sum_{i=1}^{n} a_{2i} b_{i2} & \cdots & \sum_{i=1}^{n} a_{2i} b_{ip} \\ & & \cdots & \\ \sum_{i=1}^{n} a_{mi} b_{i1} & \sum_{i=1}^{n} a_{mi} b_{i2} & \cdots & \sum_{i=1}^{n} a_{mi} b_{ip} \end{pmatrix}$$

Note: if A is $m \times n$ and B is $n \times p$, then AB is $m \times p$.

Thus, product of

The product of two linear operator can be seen as the multiplication of two matrices

$$AB = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & & \cdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ & & \cdots & \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{i=1}^{n} a_{1i} b_{i1} & \sum_{i=1}^{n} a_{1i} b_{i2} & \cdots & \sum_{i=1}^{n} a_{1i} b_{ip} \\ \sum_{i=1}^{n} a_{2i} b_{i1} & \sum_{i=1}^{n} a_{2i} b_{i2} & \cdots & \sum_{i=1}^{n} a_{2i} b_{ip} \\ & & \cdots & \\ \sum_{i=1}^{n} a_{mi} b_{i1} & \sum_{i=1}^{n} a_{mi} b_{i2} & \cdots & \sum_{i=1}^{n} a_{mi} b_{ip} \end{pmatrix}$$

Note: if A is $m \times n$ and B is $n \times p$, then AB is $m \times p$.

Cinvestav

Transpose of a Matrix

The transpose of a matrix is obtained by flipping the rows and columns

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ & & \cdots & \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix}$$

Transpose of a Matrix

The transpose of a matrix is obtained by flipping the rows and columns

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ & & \cdots & \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix}$$

Which the following properties

- $\left(A^T\right)^T = A$
- $(A+B)^T = A^T + B^T$
- $(AB)^T = B^T A^T$

Transpose of a Matrix

The transpose of a matrix is obtained by flipping the rows and columns

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ & & \cdots & \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix}$$

Which the following properties

- $\left(A^T\right)^T = A$
- $(A+B)^T = A^T + B^T$
- $(AB)^T = B^T A^T$

Not only that, we have the inner product

$$\langle u, v \rangle = u^T v$$

Transpose of a Matrix

The transpose of a matrix is obtained by flipping the rows and columns

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ & & \cdots & \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix}$$

Which the following properties

- $\left(A^T\right)^T = A$
- $(A+B)^T = A^T + B^T$
- $(AB)^T = B^T A^T$

Not only that, we have the inner product

$$\langle u, v \rangle = u^T v$$

Transpose of a Matrix

The transpose of a matrix is obtained by flipping the rows and columns

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{n1} \\ a_{12} & a_{22} & \cdots & a_{n2} \\ & & \cdots & \\ a_{1m} & a_{2m} & \cdots & a_{nm} \end{pmatrix}$$

Which the following properties

- $\left(A^T\right)^T = A$
- $(A+B)^T = A^T + B^T$
- $(AB)^T = B^T A^T$

Not only that, we have the inner product

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \boldsymbol{u}^T \boldsymbol{v}$$

As always, we have the identity operator

The identity operator in matrix multiplication is defined as

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

As always, we have the identity operator

The identity operator in matrix multiplication is defined as

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

With properties

- For any matrix A, $AI = A$.
- I is the **identity operator** for the matrix product.

As always, we have the identity operator

The identity operator in matrix multiplication is defined as

$$I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

With properties

- For any matrix A, $AI = A$.
- I is the **identity operator** for the matrix product.

Column Space, Row Space and Rank

Let A be an $m \times n$ matrix

We have the following spaces...

Column space

- Span of the columns of A.
- Linear subspace of $\mathbb{R}^m$.

Column Space, Row Space and Rank

Let A be an $m \times n$ matrix

We have the following spaces...

Column space

- Span of the columns of A.
- Linear subspace of $\mathbb{R}^m$.

Column Space, Row Space and Rank

Let A be an $m \times n$ matrix

We have the following spaces...

Column space

- Span of the columns of A.
- Linear subspace of $\mathbb{R}^m$.

Row space

- Span of the rows of A.
- Linear subspace of $\mathbb{R}^n$.

Column Space, Row Space and Rank

Let A be an $m \times n$ matrix

We have the following spaces...

Column space

- Span of the columns of A.
- Linear subspace of $\mathbb{R}^m$.

Row space

- Span of the rows of A.
- Linear subspace of $\mathbb{R}^n$.

Important facts

Something Notable

The column and row space of any matrix have the same dimension.

The rank

The dimension is the rank of the matrix.

Important facts

Something Notable

The column and row space of any matrix have the same dimension.

The rank

The dimension is the **rank** of the matrix.

Range and Null Space

Range

- Set of vectors equal to $A\boldsymbol{u}$ for some $\boldsymbol{u} \in \mathbb{R}^n$.

$$\text{Range}\,(A) = \{\boldsymbol{x} | \boldsymbol{x} = A\boldsymbol{u} \text{ for some } \boldsymbol{u} \in \mathbb{R}^n\}$$

- It is a linear subspace of $\mathbb{R}^m$ and also called the column space of A.

Range and Null Space

Range

- Set of vectors equal to $A\boldsymbol{u}$ for some $\boldsymbol{u} \in \mathbb{R}^n$.

$$\text{Range}\,(A) = \{\boldsymbol{x} | \boldsymbol{x} = A\boldsymbol{u} \text{ for some } \boldsymbol{u} \in \mathbb{R}^n\}$$

- It is a linear subspace of $\mathbb{R}^m$ and also called the column space of A.

Null Space

- We have the following definition

$$\text{Null Space}\,(A) = \{\boldsymbol{u} | A\boldsymbol{u} = \boldsymbol{0}\}$$

- It is a linear subspace of $\mathbb{R}^m$.

Range and Null Space

Range

- Set of vectors equal to $A\boldsymbol{u}$ for some $\boldsymbol{u} \in \mathbb{R}^n$.

$$\text{Range}\,(A) = \{\boldsymbol{x} | \boldsymbol{x} = A\boldsymbol{u} \text{ for some } \boldsymbol{u} \in \mathbb{R}^n\}$$

- It is a linear subspace of $\mathbb{R}^m$ and also called the column space of A.

Null Space

- We have the following definition

$$\text{Null Space}\,(A) = \{\boldsymbol{u} | A\boldsymbol{u} = \boldsymbol{0}\}$$

- It is a linear subspace of $\mathbb{R}^m$.

Range and Null Space

Range

- Set of vectors equal to $A\boldsymbol{u}$ for some $\boldsymbol{u} \in \mathbb{R}^n$.

$$\text{Range}\,(A) = \{\boldsymbol{x} | \boldsymbol{x} = A\boldsymbol{u} \text{ for some } \boldsymbol{u} \in \mathbb{R}^n\}$$

- It is a linear subspace of $\mathbb{R}^m$ and also called the column space of A.

Null Space

- We have the following definition

$$\text{Null Space}\,(A) = \{\boldsymbol{u} | A\boldsymbol{u} = \boldsymbol{0}\}$$

- It is a linear subspace of $\mathbb{R}^m$.

Important fact

Something Notable

- Every vector in the null space is orthogonal to the rows of A.
- The null space and row space of a matrix are orthogonal.

Important fact

Something Notable

- Every vector in the null space is orthogonal to the rows of A.
- The null space and row space of a matrix are orthogonal.

Cinvestav

Range and Column Space

We have another interpretation of the matrix-vector product

$$
\begin{aligned}
A\boldsymbol{u} = &\left(\boldsymbol{A}_1 \;\; \boldsymbol{A}_2 \;\; \cdots \;\; \boldsymbol{A}_n\right)\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \\
= &u_1\boldsymbol{A}_1 + u_2\boldsymbol{A}_2 + \cdots + u_n\boldsymbol{A}_n
\end{aligned}
$$

Range and Column Space

We have another interpretation of the matrix-vector product

$$
\begin{aligned}
A\boldsymbol{u} &= \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{A}_2 & \cdots & \boldsymbol{A}_n \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \\
&= u_1 \boldsymbol{A}_1 + u_2 \boldsymbol{A}_2 + \cdots + u_n \boldsymbol{A}_n
\end{aligned}
$$

Thus

- The result is a linear combination of the columns of A.
- Actually, the **range** is the **column space**.

Range and Column Space

We have another interpretation of the matrix-vector product

$$
\begin{aligned}
A\boldsymbol{u} = &\left(\boldsymbol{A}_1 \;\; \boldsymbol{A}_2 \;\; \cdots \;\; \boldsymbol{A}_n\right)\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \\
= &u_1\boldsymbol{A}_1 + u_2\boldsymbol{A}_2 + \cdots + u_n\boldsymbol{A}_n
\end{aligned}
$$

Thus

- The result is a linear combination of the columns of A.
- Actually, the **range** is the **column space**.

Matrix Inverse

Something Notable

- For an $n \times n$ matrix A: rank + dim(null space) $= n$.
- if dim(null space)$= 0$ then A is full rank.
- In this case, the action of the matrix is invertible.
- The inversion is also linear and consequently can be represented by another matrix A^{-1}.
- A^{-1} is the only matrix such that $A^{-1}A = AA^{-1} = I$.

Cinvestav

Matrix Inverse

Something Notable

- For an $n \times n$ matrix A: rank + dim(null space) $= n$.
- if dim(null space)$= 0$ then A is full rank.
- In this case, the action of the matrix is invertible.
- The inversion is also linear and consequently can be represented by another matrix A^{-1}.
- A^{-1} is the only matrix such that $A^{-1}A = AA^{-1} = I$.

Matrix Inverse

Something Notable

- For an $n \times n$ matrix A: rank $+$ dim(null space) $= n$.
- if dim(null space)$= 0$ then A is full rank.
- In this case, the action of the matrix is invertible.
- The inversion is also linear and consequently can be represented by another matrix A^{-1}.
- A^{-1} is the only matrix such that $A^{-1}A = AA^{-1} = I$.

Matrix Inverse

Something Notable

- For an $n \times n$ matrix A: rank $+$ dim(null space) $= n$.
- if dim(null space)$= 0$ then A is full rank.
- In this case, the action of the matrix is invertible.
- The inversion is also linear and consequently can be represented by another matrix A^{-1}.
- A^{-1} is the only matrix such that $A^{-1}A = AA^{-1} = I$.

Matrix Inverse

Something Notable

- For an $n \times n$ matrix A: rank $+$ dim(null space) $= n$.
- if dim(null space)$= 0$ then A is full rank.
- In this case, the action of the matrix is invertible.
- The inversion is also linear and consequently can be represented by another matrix A^{-1}.
- A^{-1} is the only matrix such that $A^{-1}A = AA^{-1} = I$.

Cinvestav

Orthogonal Matrices

Definition

An orthogonal matrix U satisfies $U^T U = I$.

Properties

U has orthonormal columns.

In addition

Applying an orthogonal matrix to two vectors does not change their inner product:

$$
\begin{aligned}
\langle U\boldsymbol{u}, U\boldsymbol{v}\rangle &= (U\boldsymbol{u})^T U\boldsymbol{v} \\
&= \boldsymbol{u}^T U^T U\boldsymbol{v} \\
&= \boldsymbol{u}^T \boldsymbol{v} \\
&= \langle \boldsymbol{u}, \boldsymbol{v}\rangle
\end{aligned}
$$

Orthogonal Matrices

Definition

An orthogonal matrix U satisfies $U^T U = I$.

Properties

U has orthonormal columns.

In addition

Applying an orthogonal matrix to two vectors does not change their inner product:

$$
\begin{aligned}
\langle Uu, Uv \rangle &= (Uu)^T Uv \\
&= u^T U^T Uv \\
&= u^T v \\
&= \langle u, v \rangle
\end{aligned}
$$

Orthogonal Matrices

Definition

An orthogonal matrix U satisfies $U^T U = I$.

Properties

U has orthonormal columns.

In addition

Applying an orthogonal matrix to two vectors does not change their inner product:

$$\begin{aligned} \langle U\boldsymbol{u}, U\boldsymbol{v}\rangle &= (U\boldsymbol{u})^T U\boldsymbol{v} \\ &= \boldsymbol{u}^T U^T U\boldsymbol{v} \\ &= \boldsymbol{u}^T \boldsymbol{v} \\ &= \langle \boldsymbol{u}, \boldsymbol{v}\rangle \end{aligned}$$

Example

A classic one

Matrices representing rotations are orthogonal.

Outline

Trace and Determinant

Definition (Trace)

The trace is the sum of the diagonal elements of a square matrix.

Definition (Determinant)

The determinant of a square matrix A, denoted by $|A|$, is defined as

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij}$$

where M_{ij} is determinant of matrix A without the row i and column j.

Trace and Determinant

Definition (Trace)

The trace is the sum of the diagonal elements of a square matrix.

Definition (Determinant)

The determinant of a square matrix A, denoted by $|A|$, is defined as

$$\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} M_{ij}$$

where M_{ij} is determinant of matrix A without the row i and column j.

Cinvestav

Special Case

For a 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

$$|A| = ad - bc$$

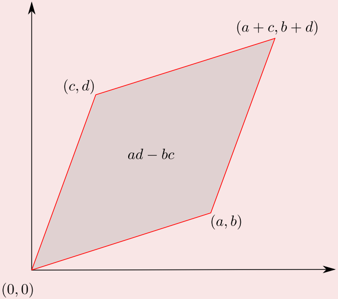
The absolute value of $|A|$ is the area of the parallelogram given by the rows of A

Special Case

For a 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

$$|A| = ad - bc$$

The absolute value of $|A|$ is the area of the parallelogram given by the rows of A



Properties of the Determinant

Basic Properties

- $|A| = \left|A^T\right|$
- $|AB| = |A|\,|B|$
- $|A| = 0$ if and only if A is not invertible
- If A is invertible, then $\left|A^{-1}\right| = \frac{1}{|A|}$.

Properties of the Determinant

Basic Properties

- $|A| = \left|A^T\right|$
- $|AB| = |A|\,|B|$
- $|A| = 0$ if and only if A is not invertible
- If A is invertible, then $\left|A^{-1}\right| = \frac{1}{|A|}$.

Properties of the Determinant

Basic Properties

- $|A| = \left|A^T\right|$
- $|AB| = |A|\,|B|$
- $|A| = 0$ if and only if A is not invertible
- If A is invertible, then $\left|A^{-1}\right| = \frac{1}{|A|}$.

Cinvestav

Properties of the Determinant

Basic Properties

- $|A| = \left|A^T\right|$
- $|AB| = |A|\,|B|$
- $|A| = 0$ if and only if A is not invertible
- If A is invertible, then $\left|A^{-1}\right| = \frac{1}{|A|}$.

Outline

Cinvestav

Eigenvalues and Eigenvectors

Eigenvalues

An eigenvalue λ of a square matrix A satisfies:

$$A\boldsymbol{u} = \lambda \boldsymbol{u}$$

for some vector , which we call an eigenvector.

Properties

Geometrically the operator A expands when $(\lambda > 1)$ or contracts $(\lambda < 1)$ eigenvectors, but does not rotate them.

Null Space relation

If u is an eigenvector of A, it is in the null space of $A - \lambda I$, which is consequently not invertible.

Eigenvalues and Eigenvectors

Eigenvalues

An eigenvalue λ of a square matrix A satisfies:

$$A\boldsymbol{u} = \lambda\boldsymbol{u}$$

for some vector , which we call an eigenvector.

Properties

Geometrically the operator A expands when $(\lambda > 1)$ or contracts $(\lambda < 1)$ eigenvectors, but does not rotate them.

Null Space relation

If $\boldsymbol{u}$ is an eigenvector of A, it is in the null space of $A - \lambda I$, which is consequently not invertible.

Eigenvalues and Eigenvectors

Eigenvalues

An eigenvalue λ of a square matrix A satisfies:

$$A\boldsymbol{u} = \lambda\boldsymbol{u}$$

for some vector , which we call an eigenvector.

Properties

Geometrically the operator A expands when $(\lambda > 1)$ or contracts $(\lambda < 1)$ eigenvectors, but does not rotate them.

Null Space relation

If $\boldsymbol{u}$ is an eigenvector of A, it is in the null space of $A - \lambda I$, which is consequently not invertible.

More properties

Given the previous relation

The eigenvalues of A are the roots of the equation $|A - \lambda I| = 0$

Remark: We do not calculate the eigenvalues this way

Something Notable

Eigenvalues and eigenvectors can be complex valued, even if all the entries of A are real.

More properties

Given the previous relation

The eigenvalues of A are the roots of the equation $|A - \lambda I| = 0$

Remark: We do not calculate the eigenvalues this way

Something Notable

Eigenvalues and eigenvectors can be complex valued, even if all the entries of A are real.

Eigendecomposition of a Matrix

Given

Let A be an $n \times n$ square matrix with n linearly independent eigenvectors $\boldsymbol{p}_1, \boldsymbol{p}_2, ..., \boldsymbol{p}_n$ and eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$

We define the matrices

$$\boldsymbol{P} = \left(\boldsymbol{p}_1 \ \boldsymbol{p}_2 \cdots \boldsymbol{p}_n\right)$$

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Eigendecomposition of a Matrix

Given

Let A be an $n \times n$ square matrix with n linearly independent eigenvectors $\boldsymbol{p}_1, \boldsymbol{p}_2, ..., \boldsymbol{p}_n$ and eigenvalues $\lambda_1, \lambda_2, ..., \lambda_n$

We define the matrices

$$\boldsymbol{P} = (\boldsymbol{p}_1 \ \boldsymbol{p}_2 \cdots \boldsymbol{p}_n)$$

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ & & \cdots & \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}$$

Properties

We have that A satisfies

$$A\boldsymbol{P} = \boldsymbol{P}\Lambda$$

In addition

$\boldsymbol{P}$ is full rank.

Thus, inverting it yields the eigendecomposition

$$A = \boldsymbol{P}\Lambda\boldsymbol{P}^{-1}$$

Properties

We have that A satisfies

$$\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}\Lambda$$

In addition

$\boldsymbol{P}$ is full rank.

Thus, inverting it yields the eigendecomposition

$$A = \boldsymbol{P}\Lambda\boldsymbol{P}^{-1}$$

Properties

We have that A satisfies

$$A\boldsymbol{P} = \boldsymbol{P}\Lambda$$

In addition

$\boldsymbol{P}$ is full rank.

Thus, inverting it yields the eigendecomposition

$$A = \boldsymbol{P}\Lambda\boldsymbol{P}^{-\mathbf{1}}$$

Cinvestav

Properties of the Eigendecomposition

We have that

- Not all matrices are diagonalizable/eigendecomposition. Example $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$
- $Trace\left(A\right) = Trace\left(\Lambda\right) = \sum_{i=1}^{n} \lambda_i$
- $|A| = |\Lambda| = \Pi_{i=1}^{n} \lambda_i$
- The rank of A is equal to the number of nonzero eigenvalues.
- If λ is anonzero eigenvalue of A, $\frac{1}{\lambda}$ is an eigenvalue of A^{-1} with the same eigenvector.
- The eigendecompositon allows to compute matrix powers efficiently:

$$A^m = \left(P\Lambda P^{-1}\right)^m = P\Lambda P^{-1} P\Lambda P^{-1} P\Lambda P^{-1} \ldots P\Lambda P^{-1} = P\Lambda^m P^{-1}$$

Properties of the Eigendecomposition

We have that

- Not all matrices are diagonalizable/eigendecomposition. Example $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$
- $Trace\,(A) = Trace\,(\Lambda) = \sum_{i=1}^{n} \lambda_i$
- $|A| = |\Lambda| = \Pi_{i=1}^{n} \lambda_i$
- The rank of A is equal to the number of nonzero eigenvalues.
- If λ is anonzero eigenvalue of A, $\frac{1}{\lambda}$ is an eigenvalue of A^{-1} with the same eigenvector.
- The eigendecompositon allows to compute matrix powers efficiently:

$$A^m = \left(P\Lambda P^{-1}\right)^m = P\Lambda P^{-1} P\Lambda P^{-1} P\Lambda P^{-1} \ldots P\Lambda P^{-1} = P\Lambda^m P^{-1}$$

Properties of the Eigendecomposition

We have that

- Not all matrices are diagonalizable/eigendecomposition. Example $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$
- $Trace\,(A) = Trace\,(\Lambda) = \sum_{i=1}^{n} \lambda_i$
- $|A| = |\Lambda| = \Pi_{i=1}^{n} \lambda_i$
- The rank of A is equal to the number of nonzero eigenvalues.
- If λ is anonzero eigenvalue of A, $\frac{1}{\lambda}$ is an eigenvalue of A^{-1} with the same eigenvector.
- The eigendecompositon allows to compute matrix powers efficiently:

$$A^m = \left(P \Lambda P^{-1}\right)^m = P \Lambda P^{-1} P \Lambda P^{-1} P \Lambda P^{-1} \dots P \Lambda P^{-1} = P \Lambda^m P^{-1}$$

Properties of the Eigendecomposition

We have that

- Not all matrices are diagonalizable/eigendecomposition. Example
$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$
- $Trace\left(A\right) = Trace\left(\Lambda\right) = \sum_{i=1}^{n} \lambda_i$
- $|A| = |\Lambda| = \Pi_{i=1}^{n} \lambda_i$
- The rank of A is equal to the number of nonzero eigenvalues.
- If λ is anonzero eigenvalue of A, $\frac{1}{\lambda}$ is an eigenvalue of A^{-1} with the same eigenvector.
- The eigendecompositon allows to compute matrix powers efficiently:
$$A^m = \left(P\Lambda P^{-1}\right)^m = P\Lambda P^{-1} P\Lambda P^{-1} P\Lambda P^{-1} \ldots P\Lambda P^{-1} = P\Lambda^m P^{-1}$$

Properties of the Eigendecomposition

We have that

- Not all matrices are diagonalizable/eigendecomposition. Example $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$
- $Trace\left(A\right) = Trace\left(\Lambda\right) = \sum_{i=1}^{n} \lambda_i$
- $|A| = |\Lambda| = \Pi_{i=1}^{n} \lambda_i$
- The rank of A is equal to the number of nonzero eigenvalues.
- If λ is anonzero eigenvalue of A, $\frac{1}{\lambda}$ is an eigenvalue of A^{-1} with the same eigenvector.
- The eigendecompositon allows to compute matrix powers efficiently:

$$A^m = \left(P\Lambda P^{-1}\right)^m = P\Lambda P^{-1} P\Lambda P^{-1} P\Lambda P^{-1} \ldots P\Lambda P^{-1} = P\Lambda^m P^{-1}$$

Properties of the Eigendecomposition

We have that

- Not all matrices are diagonalizable/eigendecomposition. Example $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$
- $Trace\,(A) = Trace\,(\Lambda) = \sum_{i=1}^{n} \lambda_i$
- $|A| = |\Lambda| = \Pi_{i=1}^{n} \lambda_i$
- The rank of A is equal to the number of nonzero eigenvalues.
- If λ is anonzero eigenvalue of A, $\frac{1}{\lambda}$ is an eigenvalue of A^{-1} with the same eigenvector.
- The eigendecompositon allows to compute matrix powers efficiently:

$$A^m = \left(\boldsymbol{P}\Lambda\boldsymbol{P}^{-1}\right)^m = \boldsymbol{P}\Lambda\boldsymbol{P}^{-1}\boldsymbol{P}\Lambda\boldsymbol{P}^{-1}\boldsymbol{P}\Lambda\boldsymbol{P}^{-1}\ldots\boldsymbol{P}\Lambda\boldsymbol{P}^{-1} = \boldsymbol{P}\Lambda^m\boldsymbol{P}^{-1}$$

Cinvestav

Eigendecomposition of a Symmetric Matrix

When A symmetric, we have

- If $A = A^T$ then A is symmetric.
- The eigenvalues of symmetric matrices are real.
- The eigenvectors of symmetric matrices are orthonormal.
- Consequently, the eigendecomposition becomes $A = U \Lambda U^T$ for Λ real and U orthogonal.
- The eigenvectors of A are an orthonormal basis for the column space and row space.

Eigendecomposition of a Symmetric Matrix

When A symmetric, we have

- If $A = A^T$ then A is symmetric.
- The eigenvalues of symmetric matrices are real.
- The eigenvectors of symmetric matrices are orthonormal.
- Consequently, the eigendecomposition becomes $A = U \Lambda U^T$ for Λ real and U orthogonal.
- The eigenvectors of A are an orthonormal basis for the column space and row space.

Eigendecomposition of a Symmetric Matrix

When A symmetric, we have

- If $A = A^T$ then A is symmetric.
- The eigenvalues of symmetric matrices are real.
- The eigenvectors of symmetric matrices are orthonormal.
- Consequently, the eigendecomposition becomes $A = U \Lambda U^T$ for Λ real and U orthogonal.
- The eigenvectors of A are an orthonormal basis for the column space and row space.

Eigendecomposition of a Symmetric Matrix

When A symmetric, we have

- If $A = A^T$ then A is symmetric.
- The eigenvalues of symmetric matrices are real.
- The eigenvectors of symmetric matrices are orthonormal.
- Consequently, the eigendecomposition becomes $A = U\Lambda U^T$ for Λ real and U orthogonal.
- The eigenvectors of A are an orthonormal basis for the column space and row space.

Eigendecomposition of a Symmetric Matrix

When A symmetric, we have

- If $A = A^T$ then A is symmetric.
- The eigenvalues of symmetric matrices are real.
- The eigenvectors of symmetric matrices are orthonormal.
- Consequently, the eigendecomposition becomes $A = U\Lambda U^T$ for Λ real and U orthogonal.
- The eigenvectors of A are an orthonormal basis for the column space and row space.

Cinvestav

We can see the action of a symmetric matrix on a vector $\boldsymbol{u}$ as...

We can decompose the action $A\boldsymbol{u} = U\Lambda U^T$u as

- Projection of $\boldsymbol{u}$ onto the column space of A (Multiplication by U^T).
- Scaling of each coefficient $\langle U_i, \boldsymbol{u}\rangle$ by the corresponding eigenvalue (Multiplication by Λ).
- Linear combination of the eigenvectors scaled by the resulting coefficient (Multiplication by U).

We can see the action of a symmetric matrix on a vector $\boldsymbol{u}$ as...

We can decompose the action $A\boldsymbol{u} = U\Lambda U^T$u as

- Projection of $\boldsymbol{u}$ onto the column space of A (Multiplication by U^T).
- Scaling of each coefficient $\langle \boldsymbol{U}_i, \boldsymbol{u} \rangle$ by the corresponding eigenvalue (Multiplication by Λ).
- Linear combination of the eigenvectors scaled by the resulting coefficient (Multiplication by U).

Final equation

$$A\boldsymbol{u} = \sum_{i=1}^{n} \lambda_i \langle \boldsymbol{U}_i, \boldsymbol{u} \rangle \boldsymbol{U}_i$$

It would be great to generalize this to all matrices!!!

We can see the action of a symmetric matrix on a vector $\boldsymbol{u}$ as...

We can decompose the action $A\boldsymbol{u} = U\Lambda U^T$u as

- Projection of $\boldsymbol{u}$ onto the column space of A (Multiplication by U^T).
- Scaling of each coefficient $\langle \boldsymbol{U}_i, \boldsymbol{u} \rangle$ by the corresponding eigenvalue (Multiplication by Λ).
- Linear combination of the eigenvectors scaled by the resulting coefficient (Multiplication by U).

Final equation

$$A\boldsymbol{u} = \sum_{i=1}^{n} \lambda_i \langle \boldsymbol{U}_i, \boldsymbol{u} \rangle \boldsymbol{U}_i$$

It would be great to generalize this to all matrices!!!

We can see the action of a symmetric matrix on a vector $\boldsymbol{u}$ as...

We can decompose the action $A\boldsymbol{u} = U\Lambda U^T$u as

- Projection of $\boldsymbol{u}$ onto the column space of A (Multiplication by U^T).
- Scaling of each coefficient $\langle \boldsymbol{U}_i, \boldsymbol{u} \rangle$ by the corresponding eigenvalue (Multiplication by Λ).
- Linear combination of the eigenvectors scaled by the resulting coefficient (Multiplication by U).

Final equation

$$A\boldsymbol{u} = \sum_{i=1}^{n} \lambda_i \langle \boldsymbol{U}_i, \boldsymbol{u} \rangle \boldsymbol{U}_i$$

It would be great to generalize this to all matrices!!!

We can see the action of a symmetric matrix on a vector $\boldsymbol{u}$ as...

We can decompose the action $A\boldsymbol{u} = U\Lambda U^T$u as

- Projection of $\boldsymbol{u}$ onto the column space of A (Multiplication by U^T).
- Scaling of each coefficient $\langle \boldsymbol{U}_i, \boldsymbol{u} \rangle$ by the corresponding eigenvalue (Multiplication by Λ).
- Linear combination of the eigenvectors scaled by the resulting coefficient (Multiplication by U).

Final equation

$$A\boldsymbol{u} = \sum_{i=1}^{n} \lambda_i \langle \boldsymbol{U}_i, \boldsymbol{u} \rangle \boldsymbol{U}_i$$

It would be great to generalize this to all matrices!!!

Outline

Cinvestav

Singular Value Decomposition

Every Matrix has a singular value decomposition

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

Where

- The columns of U are an orthonormal basis for the column space.

Singular Value Decomposition

Every Matrix has a singular value decomposition

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

Where

- The columns of U are an orthonormal basis for the column space.
- The columns of V are an orthonormal basis for the row space.
- The Σ is diagonal and the entries on its diagonal $\sigma_i = \Sigma_{ii}$ are positive real numbers, called the singular values of A.

Cinvestav

Singular Value Decomposition

Every Matrix has a singular value decomposition

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

Where

- The columns of U are an orthonormal basis for the column space.
- The columns of V are an orthonormal basis for the row space.
- The Σ is diagonal and the entries on its diagonal $\sigma_i = \Sigma_{ii}$ are positive real numbers, called the singular values of A.

The action of A on a vector u can be decomposed into

$$Au = \sum_{i=1}^{n} \sigma_i \langle V_i, u\rangle U_i$$

Singular Value Decomposition

Every Matrix has a singular value decomposition

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

Where

- The columns of U are an orthonormal basis for the column space.
- The columns of V are an orthonormal basis for the row space.
- The Σ is diagonal and the entries on its diagonal $\sigma_i = \Sigma_{ii}$ are positive real numbers, called the singular values of A.

The action of Aon a vector $\boldsymbol{u}$ can be decomposed into

$$A\boldsymbol{u} = \sum_{i=1}^{n} \sigma_i \left\langle \boldsymbol{V}_i, \boldsymbol{u} \right\rangle \boldsymbol{U}_i$$

Cinvestav

Properties of the Singular Value Decomposition

First

The eigenvalues of the symmetric matrix $A^T A$ are equal to the square of the singular values of A:

$$A^T A = V \Sigma U^T U^T \Sigma V^T = V \Sigma^2 V^T$$

Second

The rank of a matrix is equal to the number of nonzero singular values.

Third

The largest singular value σ_1 is the solution to the optimization problem:

$$\sigma_1 = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$$

Properties of the Singular Value Decomposition

First

The eigenvalues of the symmetric matrix $A^T A$ are equal to the square of the singular values of A:

$$A^T A = V \Sigma U^T U^T \Sigma V^T = V \Sigma^2 V^T$$

Second

The rank of a matrix is equal to the number of nonzero singular values.

Third

The largest singular value σ_1 is the solution to the optimization problem:

$$\sigma_1 = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2}$$

Properties of the Singular Value Decomposition

First

The eigenvalues of the symmetric matrix $A^T A$ are equal to the square of the singular values of A:

$$A^T A = V \Sigma U^T U^T \Sigma V^T = V \Sigma^2 V^T$$

Second

The rank of a matrix is equal to the number of nonzero singular values.

Third

The largest singular value σ_1 is the solution to the optimization problem:

$$\sigma_1 = \max_{\boldsymbol{x} \neq 0} \frac{\|A\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2}$$

Properties of the Singular Value Decomposition

Remark

It can be verified that the largest singular value satisfies the properties of a norm, it is called the spectral norm of the matrix.

Finally

In statistics analyzing data with the singular value decomposition is called **Principal Component Analysis**.

Properties of the Singular Value Decomposition

Remark

It can be verified that the largest singular value satisfies the properties of a norm, it is called the spectral norm of the matrix.

Finally

In statistics analyzing data with the singular value decomposition is called **Principal Component Analysis**.