WARNING!

READ CAREFULLY!

The following document has not been checked for mistakes, and although every effort possible has been made to ensure that the information in this document is correct, it might still contain some issues, either with the content or with the layout of the document. Therefore, it is given "as is", without any warranty whatsoever.

Neither the Professor nor the TAs have anything to do with such document. If you have issues, visit `https://github.com/pennal/LinAlgNotes`

Linear Algebra
Class Notes

Based on Professor Pivkin's Material

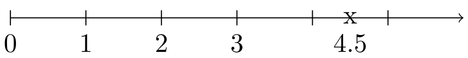
Spring 2015

Università della Svizzera Italiana
Version 0.1.3
Generated on April 23, 2015

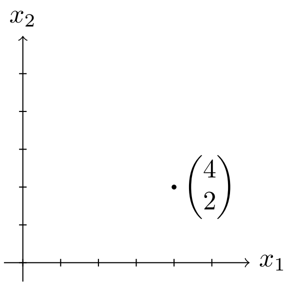
Chapter 1

Vectors

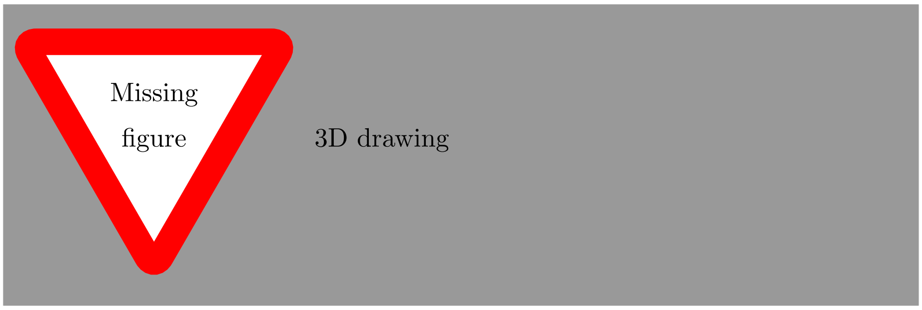
A real number can be represented by a point on a line, which is a 2-dimensional space, $\mathbb{R}$



a pair of real numbers can be represented by a point on a plane, which is a 2-dimensional space, $\mathbb{R}^2$



a triplet od real numbers can be represented by a point in 3D space, $\mathbb{R}^3$



Definition
A vector is an ordered collection of n numbers

Notation

Usually vectors are given by letters, such as u, v, w. In textbooks vectors are written with bold font. In handwriting vectors are often written with a right arrow on top, such as $\overrightarrow{u}$. We will underline vectors, like so: $\underline{u}$.

□

Definition

Let us consider vector $\underline{u} \in \mathbb{R}^n$. The $i-$th component of vector

$$\underline{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

is u_i

Example

$$\underline{u} = \begin{pmatrix} 3 \\ 7 \\ 11 \end{pmatrix} \in \mathbb{R}^3 \Rightarrow u_1 = 3, u_2 = 7, u_3 = 11$$

Definition

Let us consider vectors $\underline{u} \in \mathbb{R}^n$ and $\underline{v} \in \mathbb{R}^n$. Vector $\underline{w} \in \mathbb{R}^n$ is a sum of $\underline{u}$ and $\underline{v}$, $\underline{w} = \underline{u} + \underline{v}$, if $w_i = u_i + v_i$ for all $i = 1, \ldots, n$

Example

1.
$$\underline{u} = \begin{pmatrix} 3 \\ 5 \\ 1 \end{pmatrix}, \underline{v} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \underline{w} = \underline{u} + \underline{v} = \begin{pmatrix} 3 + (-1) \\ 5 + 0 \\ 1 + 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \\ 2 \end{pmatrix}$$

2.
$$\underline{u} = \begin{pmatrix} 3 \\ 9 \\ -2 \end{pmatrix}, \underline{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix}$$

   $\underline{u} + \underline{v}$ is not defined! Both vectors should have the same number of components.

Definition

1. Vectors $\underline{u} \in \mathbb{R}^n$ and $\underline{v} \in \mathbb{R}^n$ are equal, if $u_i = v_i$ for all $i = 1, \ldots, n$
2. A scalar is just another name for real number
3. Let us consider a scalar $\alpha \in \mathbb{R}$ and vector $\underline{u} \in \mathbb{R}^n$. A product of α and $\underline{u}$ is defined as:

$$\alpha\underline{u} = \alpha \cdot \begin{pmatrix} u_1 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \alpha \cdot u_1 \\ \vdots \\ \alpha \cdot v_n \end{pmatrix}$$

Example

$$\alpha = 3, \underline{u} = \begin{pmatrix} -1 \\ 2 \\ 5 \\ 7 \end{pmatrix} \Rightarrow \alpha \cdot \underline{u} \begin{pmatrix} 3 \cdot -1 \\ 3 \cdot 2 \\ 3 \cdot 5 \\ 3 \cdot 7 \end{pmatrix} = \begin{pmatrix} -3 \\ 6 \\ 15 \\ 21 \end{pmatrix}$$

Definition

Let us consider scalars α and β, and vectors $\underline{u} \in \mathbb{R}^n$ and $\underline{v} \in \mathbb{R}^n$. A sum of $\alpha\underline{u} + \beta \cdot \underline{v}$ is called a linear combination of vectors $\underline{u}$ and $\underline{v}$.

Example

1. $$2 \cdot \begin{pmatrix} -1 \\ 3 \\ 5 \end{pmatrix} + 3 \cdot \begin{pmatrix} 7 \\ 2 \\ 1 \end{pmatrix} + 5 \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 24 \\ 12 \\ 8 \end{pmatrix}$$

2. $$\underline{u} - \underline{v} = 1 \cdot \underline{u} + (-1) \cdot \underline{v} = \begin{pmatrix} u_1 - v_1 \\ \vdots \\ u_i - v_i \end{pmatrix}$$

3. $$\underline{u} - \underline{u} = \begin{pmatrix} u_1 - u_1 \\ \vdots \\ u_i - u_i \end{pmatrix} = \underline{0}$$

Definition

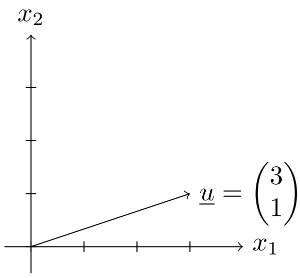
Vector $\underline{u} \in \mathbb{R}^n$ is called a zero vector if all $u_i = 0$, $i = 1, \ldots, n$. The zero vector is often written as $\underline{0} \in \mathbb{R}^n$

1.1 Graphic representation of vectors and vector operations

A vector can be represented in the following way:

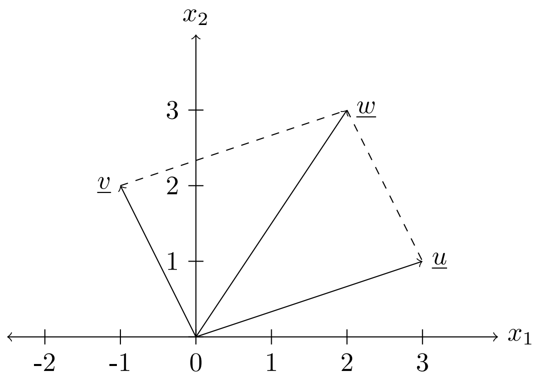
1. An ordered collection of numbers, $\underline{u} = \begin{pmatrix} 3 \\ 5 \end{pmatrix}$

2. As an arrow in space



3. A vector is a point in space, the endpoint of a vector from the origin.

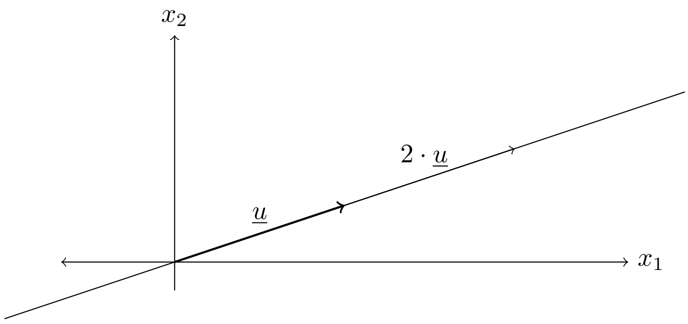
Let us consider vectors $\underline{u} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, $\underline{v} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$ and $\underline{w} = \underline{u} + \underline{v} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$



Let us consider vector $\underline{u} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$. What is $2 \cdot \underline{u}$? We can calculate as follows:

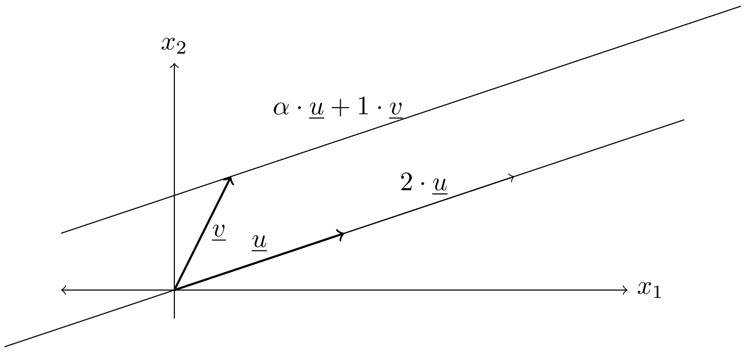
$$2 \cdot \underline{u} = 2 \cdot \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \end{pmatrix}$$

We stretch vector $\underline{u}$ 2 times along the line defined by vector $\underline{u}$. What is $-\underline{u}$? Simply reverse the direction. What will be the representation of $\alpha\underline{u}$ for all possible values of α? An endless line



Let us consider two vectors $\underline{u} \in \mathbb{R}^2$ and $\underline{v} \in \mathbb{R}^2$. What will be the representation of all linear combinations of $\underline{u}$ and $\underline{v}$, $\alpha\underline{u} + \beta\underline{v}$

1. Plane:



2. Line: $\underline{u}$ and $\underline{v}$ are on the same line.
   Note: Consider $\underline{u}, \underline{v} \in \mathbb{R}^n$. $\underline{u}$ and $\underline{v}$ are on the same line if there exists scalars α and β such that $\alpha\underline{u} + \beta\underline{v} = \underline{0}$, when α and $\beta \neq 0$

3. Point: if $\underline{u} = \underline{0}$ and $\underline{v} = \underline{0} \Rightarrow \alpha\underline{u} + \beta\underline{v} = \underline{0}$

Consider $\underline{v}, \underline{u}$. They are on the same line if $\alpha\underline{u} + \beta\underline{v} = \underline{0}$ and $\alpha, \beta \neq 0$

1.2 Dot Product (Scalar product)

Definition

Let us consider two vectors $\underline{u} \in \mathbb{R}^n$ and $\underline{v} \in \mathbb{R}^n$. The dot (or scalar) product of vectors $\underline{u}$ and $\underline{v}$ is defined as

$$\langle \underline{u}, \underline{v} \rangle = u_1 v_1 + u_2 v_2 + \cdots + \underline{u}_n \underline{v}_n = \sum_{i=1}^{n} u_i v_i$$

Notation

We will use $\langle \underline{u}, \underline{v} \rangle$ to denote the dot product, but sometimes $\underline{u} \cdot \underline{v}$ is used

□

Example

1.

$$\underline{u} = \begin{pmatrix} 1 \\ -1 \\ 3 \end{pmatrix}, \underline{u} = \begin{pmatrix} 0 \\ \frac{1}{2} \\ -1 \end{pmatrix}$$

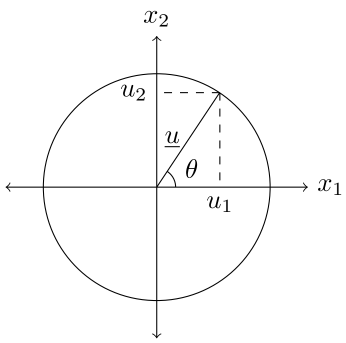
$$\langle \underline{u}, \underline{v} \rangle = 1 \cdot 0 + (-1) \cdot \frac{1}{2} + 3 \cdot (-1) = -3.5$$

2.

$$\underline{u} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \underline{u} = \begin{pmatrix} 0 \\ 1 \\ , \end{pmatrix} \langle \underline{u}, \underline{v} \rangle = 0$$

Let us consider $\mathbb{R}^2$. What is the set of all possible endpoints of unit vectors in $\mathbb{R}^2$, originating from the origin?

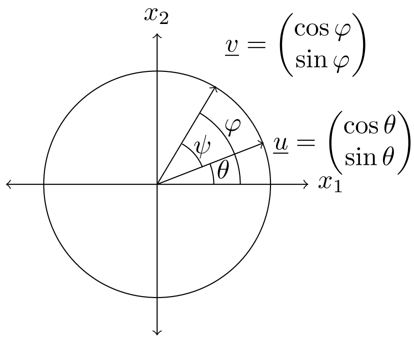
Fix positioning problem



$$\begin{aligned} \underline{u} &= \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \\ \cos(\theta) &= \frac{u_1}{\|\underline{u}\|} = u_1 \\ \sin(\theta) &= \frac{u_2}{\|\underline{u}\|} = u_2 \\ \underline{u} &= \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix} \end{aligned}$$

Now let us consider two unit vectors If $\underline{u} \neq \underline{0}$ or $\underline{v} \neq \underline{0}$ are not unit vectors we can find the angle between them as follows:

$$\begin{aligned} \langle \underline{u}, \underline{v} \rangle &= \left\langle \|\underline{u}\| \cdot \frac{1}{\|\underline{u}\|} \cdot \underline{u}, \|\underline{v}\| \cdot \frac{1}{\|\underline{v}\|} \cdot \underline{v} \right\rangle \\ &= \|\underline{u}\| \|\underline{v}\| \left\langle \underbrace{\frac{1}{\|\underline{u}\|} \cdot \underline{u}, \frac{1}{\|\underline{v}\|} \cdot \underline{v}}_{\text{Unit Vectors}} \right\rangle \\ &= \|\underline{u}\| \|\underline{v}\| \cos\left(\angle(\underline{u}, \underline{v})\right) \end{aligned}$$



$$\begin{aligned} \langle \underline{u}, \underline{v} \rangle &= \cos(\theta)\cos(\varphi) + \sin(\theta)\sin(\varphi) \\ &= \cos(\theta - \varphi) = \cos(\psi) \\ &= \cos(\angle(\underline{u}, \underline{v})) \end{aligned}$$

Lemma

If $\underline{u} \neq \underline{0}, \underline{v} \neq \underline{0}, \underline{u} \in \mathbb{R}^n, \underline{v} \in \mathbb{R}^n$, then

$$\cos\left(\angle(\underline{u}, \underline{v})\right) = \frac{\langle \underline{u}, \underline{v} \rangle}{\|\underline{u}\|\|\underline{v}\|}$$

1.3 Properties of dot product

1. $\langle \alpha \cdot \underline{u}, \underline{v} \rangle = \alpha \cdot \langle \underline{u}, \underline{v} \rangle$ for any $\alpha \in \mathbb{R}, \underline{u} \in \mathbb{R}^n, \underline{v} \in \mathbb{R}^n$. Proof:

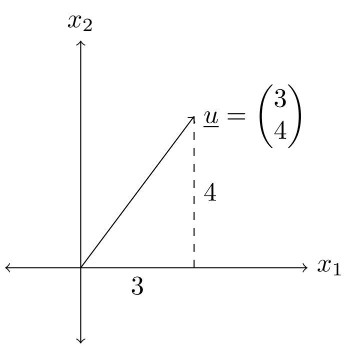
$$\begin{aligned} \langle \alpha \cdot \underline{u}, \underline{v} \rangle &= (\alpha u_1) \cdot v_1 + \cdots + (\alpha u_n) \cdot v_n \\ &= \alpha \cdot (u_1 \cdot v_1 + \cdots + u_n \cdot v_n) \\ &= \alpha \cdot \langle \underline{u}, \underline{v} \rangle \end{aligned}$$

2. $\langle \underline{u}, \alpha \underline{v} \rangle = \alpha \langle \underline{u}, \underline{v} \rangle$ for any $\alpha \in \mathbb{R}, \underline{u}, \underline{v} \in \mathbb{R}^n$

3. $\langle \alpha \underline{u} + \beta \underline{v}, \underline{w} \rangle = \alpha \cdot \langle \underline{u}, \underline{w} \rangle + \beta \langle \underline{v}, \underline{w} \rangle, \forall \alpha \in \mathbb{R}, \forall \underline{u}, \underline{vw} \in \mathbb{R}^n$

Example

Let us consider $\underline{u} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$ $. \langle \underline{u}, \underline{u} \rangle = 3 \cdot 3 + 4 \cdot 4 = 9 + 16 = 25 = 5^2$



Definition

The length of vector $\underline{u} \in \mathbb{R}^n$, $\|\underline{u}\|$, is defined as $\|\underline{u}\| = \sqrt{\langle \underline{u}, \underline{u} \rangle}$. Sometimes it is also called the Euclidian norm of $\underline{u}$.

Definition

A vector with length equal to 1 is called a unit vector

If we take vector $\underline{u} \neq \underline{0}$, how to make it a unit vector? We should multiply vector $\underline{u}$ by $\frac{1}{\|\underline{u}\|}$, we will get $\frac{\underline{u}}{\|\underline{u}\|}$ = unit vector.

In our previous example: $\underline{u} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$. Unit vector is then

$$\frac{\underline{u}}{\|\underline{u}\|} = \frac{1}{5} \cdot \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} \frac{3}{5} \\ \frac{4}{5} \end{pmatrix} = \begin{pmatrix} 0.6 \\ 0.8 \end{pmatrix}$$

We got $\langle \underline{u}, \underline{v} \rangle = \|\underline{u}\| \|\underline{v}\| \cdot \cos\left(\angle(\underline{u}, \underline{v})\right)$. Let us take the absolute value of this

$$|\langle \underline{u}, \underline{v} \rangle| = \|\underline{u}\| \|\underline{v}\| \cdot |\cos\left(\angle(\underline{u}, \underline{v})\right)|$$

Notice that $|\cos\left(\angle(\underline{u}, \underline{v})\right)| \leq 1$

Lemma

Cauchy Schwartz Inequality: for any $\underline{u} \in \mathbb{R}^n$ and $\underline{v} \in \mathbb{R}^n$

$$|\langle \underline{u}, \underline{v} \rangle| \leq \|\underline{u}\| \|\underline{v}\|$$

Remark: It is easy to see that Cauchy - Schwartz inequality is correct also for zero vectors.

Chapter 2

Matrices

Let us consider a linear combination of vectors

$$x_1 \cdot \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} + x_2 \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} + x_3 \cdot \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

This can be written using matrices in the following way:

$$\begin{pmatrix} u_1 & v_1 & w_1 \\ \vdots & \vdots & \vdots \\ u_n & v_n & w_n \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

In matrix-vector multiplication, we take dot products of rows of matrices times the vector.

Example

1.
$$\begin{pmatrix} 1 & 0 & -1 \\ 3 & 1 & 2 \\ 1 & -1 & 5 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 1 + 0 \cdot 0 + (-1) \cdot 1 \\ 3 \cdot 1 + 1 \cdot 0 + 2 \cdot 1 \\ 1 \cdot 1 + (-1) \cdot 0 + 5 \cdot 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 5 \\ 6 \end{pmatrix}$$

Notation

Matrices are usually written with capital letters, i.e. $A, B, c, \ldots$

A is an n by m matrix, $A \in \mathbb{R}^{n,m}$ if it has n rows and m columns.

The element of matrix A located in row i and column j is written as a_{ij} or $(A)_{ij}$.

□

2.

$$A = \begin{pmatrix} -1 & 2 & 3 \\ 0 & 1 & 0 \end{pmatrix}, \underline{x} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\begin{aligned} A \cdot \underline{x} &= \begin{pmatrix} -1 & 2 & 3 \\ 0 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} (-1) \cdot 1 + 2 \cdot 1 + 3 \cdot 1 \\ 0 \cdot 1 + 1 \cdot 1 + 0 \cdot 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 1 \end{pmatrix} \end{aligned}$$

For the product of matrix A with vector $\underline{x}$ to exist, matrix A should have the same number of columns as vector $\underline{x}$ components.

2.1 Matrix Operations

Definition

Let us consider matrices $A \in \mathbb{R}^{n,m}$ and $B \in \mathbb{R}^{n,m}$ where $n =$ rows, $m =$ columns. Matrix $C \in \mathbb{R}^{n,m}$ is a sum of A and $B, C = A + B$ if $C_{ij} = A_{ij} + B_{ij}$ for all $i = 1, \dots, n$, $j = 1, \dots, m$

Example

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 5 \end{pmatrix}, B = \begin{pmatrix} -1 & 0 \\ 0 & -1 \\ -1 & 0 \end{pmatrix}, C = A + B = \begin{pmatrix} 0 & 2 \\ 3 & 3 \\ 4 & 6 \end{pmatrix}$$

Definition

A product of a scalar α and a matrix $A \in \mathbb{R}^{n,m}$ is defined as $(\alpha A)_{ij} = \alpha \cdot A_{ij}$, $\forall i = 1, \dots, n; j = 1, \dots, m$.

Example

$$\alpha = 3, A = \begin{pmatrix} 0 & 0 & 1 \\ 2 & 3 & 5 \end{pmatrix} \Rightarrow \alpha \cdot A = \begin{pmatrix} 0 & 0 & 3 \\ 6 & 9 & 15 \end{pmatrix}$$

Properties

- $A \in \mathbb{R}^{n,m}$ and $B \in \mathbb{R}^{n,m}$: $A + B = B + A$
  Proof:
  $$\begin{cases} (A + B)_{ij} = A_{ij} + B_{ij} \\ (B + A)_{ij} = B_{ij} + A_{ij} \end{cases}$$
- $A, B, C \in \mathbb{R}^{n,m}$: $(A + B) + C = A + (B + C)$
- $\alpha \cdot (A + B) = \alpha A + \alpha B$ for $\forall \alpha \in \mathbb{R}, A, B \in \mathbb{R}^{n,m}$

2.2 Matrix - Matrix multiplication

Definition

Let us consider matrix $A \in \mathbb{R}^{n,m}$ and $A \in \mathbb{R}^{m,l}$. Then $C = A \cdot B$ is an n by l matrix, $C \in \mathbb{R}^{n,l}$ such that

$$C_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Example

1.

$$\begin{aligned} A &= \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 4 \end{pmatrix} \in \mathbb{R}^{3,2}, B = \begin{pmatrix} 1 & 2 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{1,4} \\ C &= A \cdot B \in \mathbb{R}^{3,4} \\ &= \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 0 & 1 \\ -1 & 1 & 1 & 0 \end{pmatrix} = \begin{pmatrix} -1 & 4 & 2 & 1 \\ -1 & 1 & 1 & 0 \\ -1 & 10 & 4 & 3 \end{pmatrix} \end{aligned}$$

2.

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \\ 3 & 4 \end{pmatrix}, B = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 2 & 2 & 2 \end{pmatrix}; AB = \text{ Not defined}$$

Properties

1. AB is not always equal to BA. (most often, is the case).

$$\begin{aligned} A &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ AB &= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, BA = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

2. $C(A + B) = CA + CB$

3. $(A + B)C = AC + BC$

4. $\alpha(AB) = A(\alpha B)$, $A \in \mathbb{R}^{n,m}$, $B \in \mathbb{R}^{m,l}$. Proof:

$$(\alpha(AB))_{ij} = \alpha \sum_{k=1}^{m} a_{ik} b_{kj} = \sum_{k=1}^{m} a_{ik} (\alpha b_{kj}) = A(\alpha B)$$

5. $(AB)C = A(BC)$

Theorem

Let us consider matrices $A \in \mathbb{R}^{n,n}$ and $B \in \mathbb{R}^{n,n}$, such that A^{-1} and B^{-1} exist. Then,

$$(AB)^{-1} = B^{-1} \cdot A^{-1}$$

Proof

$$\left.\begin{aligned}(AB)(B^{-1}A^{-1}) = I\\(B^{-1}A^{-1})(AB) = I\end{aligned}\right\} \text{ Prove this}$$

$$(AB)(B^{-1}A^{-1}) = A\underbrace{BB^{-1}}_{I}A^{-1} = A \cdot I \cdot A^{-1} = I$$

$$(B^{-1}A^{-1})(AB) = B^{-1}\underbrace{A^{-1}A}_{I}B = \cdot B^{-1} \cdot I \cdot B = I$$

$\Rightarrow$ According to the definition $B^{-1}A^{-1}$ is the inverse of AB

$\square$

Lemma

$A, B, C \in \mathbb{R}^{n,n}, \exists A^{-1}, \exists B^{-1}, \exists C^{-1}$

$$(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$$

Theorem

Let us consider $A \in \mathbb{R}^{n,n}$. Let us consider that $B \in \mathbb{R}^{n,n}$ and $C \in \mathbb{R}^{n,n}$ are both inverses of A.Then $B = C$. (The inverse is unique)

Proof

$$AB = BA = I \qquad\qquad AC = CA = I$$

$$\underbrace{BA \times C} = I \times C \qquad\qquad \underbrace{B \times AC} = B \times I$$

$$\underline{\underline{C = B}}$$

$\square$

2.3 Linear system of equations

Let us consider the following system of equations

$$\begin{cases} 2x_1 + 2x_2 + 4x_3 = 2 \\ x_2 + 2x_3 = 3 \\ 4x_3 = -1 \end{cases}$$

Find x_1, x_2, x_3. We can write this system in matrix form.

$$A = \begin{pmatrix} 2 & 2 & 4 \\ 0 & 1 & 2 \\ 0 & 0 & 4 \end{pmatrix} \in \mathbb{R}^{3,3}, \underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, \underline{b} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} \Rightarrow A\underline{x} = \underline{b}$$

A is an upper triangular matrix. We can use backward substitution to find the solution:

1. $x_3 = -\frac{1}{4} = \frac{b_3}{a_{33}}$
2. $x_2 = \frac{3-2x_3}{1} = \frac{3-2\cdot\left(-\frac{1}{4}\right)}{1} = 3.5 = \frac{b_2 - a_{33}\cdot x_3}{a_{22}}$
3. $x_1 = \frac{2-4x_3-2x_2}{2} = -2 = \frac{b_1 - a_{13}x_3 - a_{12}x_2}{a_{11}}$

In general, if $A \in \mathbb{R}^{n,n}$ is an upper triangular with $a_{ii} \neq 0, i = 1, \ldots, n$ then the backward substitution works as:

1. $x_n = \frac{b_n}{a_{nn}}$
2. $x_{n-1} = \frac{b_{n-1} - a_{n-1\ n}x_n}{a_{n-1\ n-1}} \qquad \ldots \qquad x_i = \frac{b_i - a_{in}x_n - \cdots - a_{i\ i+1} + x_{i+1}}{a_{ii}}, \qquad i = 1, \ldots, n$

2.4 Inverse Matrix

Definition

Let us consider a matrix $A \in \mathbb{R}^{n,n}$ (square matrix). Matrix $B \in \mathbb{R}^{n,n}$ is called an inverse of A, if

$$A \cdot B = I \text{ AND } B \cdot A = I$$

(Both conditions are vital)

Notation

Usually, the inverse of A is written as A^{-1}

□

Note

Not all matrices have an inverse! In most cases, it is quite difficult to find an inverse matrix. But in some cases, the inverse is easy to find.

Example

$$A = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & a_{nn} \end{pmatrix}, a_{ii} \neq 0, \forall i = 1, \dots, n$$

Then

$$A = \begin{pmatrix} a_{11}^{-1} & 0 & \dots & 0 \\ 0 & a_{22}^{-1} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & a_{nn}^{-1} \end{pmatrix}$$

$$A \cdot A^{-1} = \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} a_{11}^{-1} & 0 & \dots & 0 \\ 0 & a_{22}^{-1} & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & a_{nn}^{-1} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

$$A \cdot A^{-1} = I$$

2.5 Special Matrices

- Let us consider $A \in \mathbb{R}^{n,m}$ matrix. A is called the zero matrix if all $a_{ij} = 0$, $\forall i = 1, \dots, n; j = 1, \dots, n$
- $D \in \mathbb{R}^{n,n}$ - square matrix is called diagonal matrix, if $d_{ij} = 0$ and if $i \neq j$
- Identity matrix:

$$I \in \mathbb{R}^{n,n}, I = \begin{pmatrix} 1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1 \end{pmatrix}$$

- $L \in \mathbb{R}^{n,n}$ - lower triangular matrix, if

$$l_{ij} = 0, \forall i < j, L = \begin{pmatrix} * & \dots & 0 \\ \vdots & \ddots & \vdots \\ * & \dots & * \end{pmatrix}$$

- $U \in \mathbb{R}^{n,n}$ - upper triangular matrix, if

$$u_{ij} = 0, \forall i > j, L = \begin{pmatrix} * & \dots & * \\ \vdots & \ddots & \vdots \\ 0 & \dots & * \end{pmatrix}$$

Remark:

If $A, B \in \mathbb{R}^{n,n}$ are both upper (lower) triangular matrices, then $C = A \cdot B$ is an upper triangular (lower).

If A is lower triangular, $A \in \mathbb{R}^{n,n}, a_{ii} \neq 0, i = 1, \dots, n$ then we can use forward substitution, i.e.:

$$\begin{aligned} x_1 &= \frac{b_1}{a_{11}} \\ &\vdots \\ x_i &= \frac{b_i - a_{i1}x_1 - \dots - a_{ii-1}x_{i-1}}{a_{ii}} \qquad \forall i = 2, \dots, n \end{aligned}$$

2.6 Elementary Transition Matrices

Let us consider matrix

$$A = \begin{pmatrix} 1 & 0 & -1 & 2 \\ 3 & 4 & 5 & 7 \\ 2 & -1 & 0 & 0 \\ -1 & 3 & 5 & -1 \end{pmatrix}$$

and matrix

$$I_{21} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Then

$$I_{21} \cdot A = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -1 & 2 \\ 3 & 4 & 5 & 7 \\ 2 & -1 & 0 & 0 \\ -1 & 3 & 5 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & -1 & 2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

also

$$A \cdot I_{21} = \begin{pmatrix} 1 & 0 & -1 & 2 \\ 3 & 4 & 5 & 7 \\ 2 & -1 & 0 & 0 \\ -1 & 3 & 5 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 4 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 \end{pmatrix}$$

Definition

We can define the elementary transition matrix $I_{pq} \in \mathbb{R}^{n,n}$

$$(I_{pq}) = \begin{cases} 1 & i = p, q = j \\ 0 & \text{otherwise} \end{cases}$$

If we take a matrix $A \in \mathbb{R}^{n,n}$ then when calculating I_{pq} we take row q of A, put it into row p, replace everything else with 0.

We can also define:

$$E_{pq}(l) = I + l \cdot I_{pq}, l \in \mathbb{R} \text{ - scalar}$$
$$E_{pq}(l) \cdot A = (I + lI_{pq}) \cdot A = A + l \cdot I_{pq}A$$

We take row q of A, multiply it by l, add it to row p of A

$$E_{pq}^{-1}(l) = E_{pq}(-l)$$

If we have vector $\underline{b} \in \mathbb{R}^n$, then $I_{pq}\underline{b}$ - we take component q of $\underline{b}$, put it into component p, replace everything else with zeros.

$$E_{pq}(l)\underline{b} \text{ - same as for matrices}$$

Chapter 3

Gaussian Elimination

Example

$$A = \begin{pmatrix} 2 & 4 & -2 \\ 4 & 9 & -3 \\ -2 & -3 & 7 \end{pmatrix} .\underline{b} = \begin{pmatrix} 2 \\ 8 \\ 10 \end{pmatrix}, A\underline{x} = b$$

We can write this as a system of equations:

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ 4x_1 + 9x_2 - 3x_3 = 8 \\ -2x_1 - 3x_2 7x_3 = 10 \end{cases}$$

We can multiply equation 1 by $-\frac{a_{21}}{a_{11}} = -\frac{4}{2} = -2$, and add to equation 2. This is equivalent to multiplying $A\underline{x} = \underline{b}$ by $E_{21}\left(-\frac{a_{21}}{a_{11}}\right)$ on the left.

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ 4x_1 + 9x_2 - 3x_3 = 8 \\ -2x_1 - 3x_2 7x_3 = 10 \end{cases} \Leftrightarrow E_{21}\left(-\frac{a_{21}}{a_{11}}\right) \times A\underline{x} = E_{21}\left(-\frac{a_{21}}{a_{11}}\right)\underline{b}$$

$$E_{21}\left(-\frac{a_{21}}{a_{11}}\right) = \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ x_2 + x_3 = 4 \\ x_2 + 5x_3 = 12 \end{cases}$$

$$\Leftrightarrow E_{31}\left(-\frac{a_{31}}{a_{11}}\right) E_{21}\left(-\frac{a_{21}}{a_{11}}\right) \times A\underline{x} = E_{31}\left(-\frac{a_{31}}{a_{11}}\right) E_{21}\left(-\frac{a_{21}}{a_{11}}\right)\underline{b}$$

$$E_{31}\left(-\frac{a_{31}}{a_{11}}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

We are done with the first column. Let us denote the resulting matrix by $A^{(1)}$

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ x_2 + x_3 = 4 \\ 4x_3 = 8 \end{cases}$$

$$\Leftrightarrow E_{32}\left(-\frac{a_{32}^{(1)}}{a_{22}^{(1)}}\right) E_{31}\left(-\frac{a_{31}}{a_{11}}\right) E_{21}\left(-\frac{a_{21}}{a_{11}}\right) \times \underbrace{A\underline{x}}_{\underline{b}}$$

$$E_{32}\left(-\frac{a_{32}^{(1)}}{a_{22}^{(1)}}\right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

We are done with the second column, so we can denote the resulting matrix by $A^{(2)}$.

In fact, we got an upper triangular matrix. We can solve it using backward compatibility. Let us denote

$$E_{32}\left(-\frac{a_{32}^{(1)}}{a_{22}^{(1)}}\right) E_{31}\left(-\frac{a_{31}}{a_{11}}\right) E_{21}\left(-\frac{a_{21}}{a_{11}}\right) = U$$

where U is the upper triangular matrix. Then the inverse of it is

$$\begin{aligned} &\left[E_{32}\left(-\frac{a_{32}^{(1)}}{a_{22}^{(1)}}\right) E_{31}\left(-\frac{a_{31}}{a_{11}}\right) E_{21}\left(-\frac{a_{21}}{a_{11}}\right)\right]^{-1} \\ &= E_{21}\left(-\frac{a_{21}}{a_{11}}\right) E_{31}\left(-\frac{a_{31}}{a_{11}}\right) E_{32}\left(-\frac{a_{32}^{(1)}}{a_{22}^{(1)}}\right) \\ A &= \underbrace{E_{21}\left(-\frac{a_{21}}{a_{11}}\right) E_{31}\left(-\frac{a_{31}}{a_{11}}\right) E_{32}\left(-\frac{a_{32}^{(1)}}{a_{22}^{(1)}}\right)}_{L} \cdot U \end{aligned}$$

All matrices $E_{xx}(x)$ are lower triangular $\rightarrow$ the product is also lower triangular $(A = l \cdot U)$. So using Gaussian elimination, we represented A as a product of lower and upper triangular matrices

$$A\underline{x} = \underline{b} \Rightarrow LU\underline{x} = \underline{b}$$

Let us denote $U\underline{x}$ by $\underline{y}$, then we get

$$\begin{cases} l\underline{y} = \underline{b} & \text{Solve by forward substitution, find } \underline{y} \\ U\underline{x} = \underline{y} & \text{Solve by backward substitution} \end{cases}$$

<u>**Remark:**</u>

Gaussian elimination works if all elements $a_{11}, a_{22}^{(1)}, a_{33}^{(2)}, \ldots, , a_{ii}^{(i-1)}$ are non-zero! These elements are called <u>PIVOT</u> elements.

Example

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ 4x_1 + 8x_2 - 3x_3 = 6 \\ -2x_1 - 3x_2 + 7x_3 = 10 \end{cases} \Leftrightarrow A\underline{x} = \underline{b}$$

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ x_3 = 2 \\ -2x_1 - 3x_2 + 7x_3 = 10 \end{cases} \Leftrightarrow E_{21}\left(-\frac{a_{21}}{a_{11}}\right) A\underline{x} = E_{21}\left(-\frac{a_{21}}{a_{11}}\right)$$

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ x_3 = 2 \\ x_2 + 5x_3 = 12 \end{cases} \Leftrightarrow E_{31}\left(-\frac{a_{31}}{a_{11}}\right) A\underline{x} = E_{31}\left(-\frac{a_{31}}{a_{11}}\right) E_{21}\left(-\frac{a_{21}}{a_{11}}\right)\underline{b}$$

We denote the resulting matrix by $A^{(1)}$. In order to proceed we need $a_{22}^{(1)} \neq 0$. Let us consider matrix P_{pq}-matrix, which you get from identity matrix by exchanging rows p and q. It is easy to show that $P_{pq} \cdot A$ is equal to matrix A with rows p and q exchanged.

Definition

Permutation matrix P is an identity matrix with rows in any order.

Remark:

$P^{-1} = P$. The product of permutation on matrices is a permutation matrix.

We want to exchange rows 2 and 3. We need to multiply by the permutation matrix P_{23}

$$\begin{cases} 2x_1 + 4x_2 - 2x_3 = 2 \\ x_2 + 5x_3 = 12 \\ x_3 = 2 \end{cases}$$

$$\begin{aligned} &\Leftrightarrow P_{23} \cdot E_{31}\left(-\frac{a_{31}}{a_{11}}\right) E_{21}\left(-\frac{a_{21}}{a_{11}}\right) A\underline{x} \\ &= P_{23} \cdot E_{31}\left(-\frac{a_{31}}{a_{11}}\right) E_{21}\left(-\frac{a_{21}}{a_{11}}\right) \underline{b} \end{aligned}$$

In general, the Gaussian elimination proceeds like this:

$$E_{xx} \dots E_{xx} P_{xx} E_{xx} \dots E_{xx} A\underline{x} = E_{xx} \dots E_{xx} P_{xx} E_{xx} \dots E_{xx} \underline{b}$$

Turns out, that we can exchange the rows, or in other words multiply A by

$(P_{xx}\dots P_{xx}$ before doing the Gaussian elimination

$$\underbrace{\underbrace{(E_{xx}\dots E_{xx})}_{E}\underbrace{(P_{xx}\dots P_{xx})}_{P}A}_{U}\,\underline{x} = (E_{xx}\dots E_{xx})(P_{xx}\dots P_{xx})\underline{b}$$

$$EPA = U$$

$$PA = E^{-1}U = LU \leftarrow \text{Lower triangular}$$

Theorem

There exists permutation matrix P, such that $PA = LU$. The only necessary condition for that is that A^{-1} exists.

3.1 Matrix Transposition

Definition

Let us consider matrix $A \in \mathbb{R}^{m,n}$. Matrix $B \in \mathbb{R}^{n,m}$ is called the transpose of A if $(B)_{ij} = (A)_{ji}, i = 1\dots n, j = 1\dots n$

Notation

Usually the transpose of A is written as A^T

□

Example

$$A = \begin{pmatrix} 2 & 3 \\ 4 & 5 \\ 6 & 7 \\ 9 & 10 \end{pmatrix} \in \mathbb{R}^{4,2} \Rightarrow A = \begin{pmatrix} 2 & 4 & 6 & 9 \\ 3 & 5 & 7 & 10 \end{pmatrix} \in \mathbb{R}^{2,4}$$

Properties

1. $\left(A^T\right)^T = A$
2. $(A+B)^T = A^T + B^T$
3. $(AB)^T = B^T \cdot A^T$
4. $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$

Proof

3.

$$
\begin{aligned}
A \in \mathbb{R}^{m,n} &= \begin{pmatrix} -\text{row } 1 \rightarrow \\ \vdots \\ -\text{row } n \rightarrow \end{pmatrix}, B \in \mathbb{R}^{n,l} = \begin{pmatrix} \text{-col 1}\downarrow & \cdots & \text{-col n}\downarrow \end{pmatrix} \\
(AB)_{ij} &= \langle \text{row } i \text{ of } A, \text{ column } j \text{ of } B \rangle \\
\left((AB)^T\right)_{pq} &= (AB)_{qp} = \langle \text{row } q \text{ of } A, \text{ column } p \text{ of } B \rangle \\
B^T &= \begin{pmatrix} -\text{col } 1 \rightarrow \\ \vdots \\ -\text{col } n \rightarrow \end{pmatrix}, A^T = \begin{pmatrix} \text{-col 1}\downarrow & \cdots & \text{-col n}\downarrow \end{pmatrix} \\
\left(B^T A^T\right)_{pq} &= \langle \text{column } p \text{ of } B, \text{ row } q \text{ of } A \rangle \\
\Rightarrow \left((AB)^T\right)_{pq} &= \left(B^T A^T\right)_{pq}; p = 1, \dots, l; q = 1, \dots, m. \\
\Rightarrow (AB)^T &= B^T A^T
\end{aligned}
$$

4. Assume that $A \in \mathbb{R}^{n,n}, \exists A^{-1}$

$$
\begin{aligned}
AA^{-1} &= I \rightarrow (AA^{-1})^T = \left(A^{-1}\right)^T \cdot A^T = I^T = I \\
A^{-1}A &= I \rightarrow (A^{-1}A)^T = A^T \cdot \left(A^{-1}\right)^T = I^T = I \\
\left(A^T\right)^{-1} &= \left(A^{-1}\right)^T
\end{aligned}
$$

□

Let us consider vector $\underline{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \in \mathbb{R}^{n,1}$ - column vector. Then $\underline{u}^T \in \mathbb{R}^{1,n} = (u_1 \dots u_n)$ - row vector. Let us also consider $\underline{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^{n,1}$. Then

$$
\underline{u}^T \cdot \underline{v} = (u_1 \dots u_n) \cdot \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n = \langle \underline{u}, \underline{v} \rangle
$$

$$
\underline{v} \cdot \underline{u}^T = n \times n \text{ matrix}
$$

Definition

Matrix A is called symmetric if $A^t = A$. Matrix A should be a square matrix, $A \in \mathbb{R}^{n,n}$

$$
\text{e.g. } A = \begin{pmatrix} 0 & 3 \\ 3 & 4 \end{pmatrix} \rightarrow A^T = \begin{pmatrix} 0 & 3 \\ 3 & 4 \end{pmatrix} \Rightarrow A^T = A
$$

$$
\text{e.g. } A = I \in \mathbb{R}^{n,n} \rightarrow I^T = I
$$

Chapter 4

Vector Spaces

Definition

A vector space V is a set of objects, such that any two objects can be added together, any object can be multiplied by a scalar.

If two objects belong to the vector space, then their sum also belongs to the vector space.

If an object belongs to V, then the product of any scalar with this object belongs to V and the following properties are satisfied:

1. $\forall \underline{u}, \underline{v}, \underline{w} \in V$; $(\underline{u}+\underline{v})+\underline{w} = \underline{u}+(\underline{v}+\underline{w})$
2. $\forall \underline{u}, \underline{v} \in V$; $\underline{u}+\underline{v} = \underline{v}+\underline{u}$
3. There exists unique elements $\underline{0} \in V$, such that $\forall \underline{u} \in V$; $\underline{u}+\underline{0} = \underline{0}+\underline{u} = \underline{u}$
4. For any $\underline{u} \in V, \exists!(-\underline{u}) \in V$, such that $\underline{u}+(-\underline{u}) = \underline{0}$
5. $\forall \underline{u}, \underline{v} \in V$; $\forall \alpha \in \mathbb{R}$; $\alpha(\underline{u}+\underline{v}) = \alpha\underline{u}+\alpha\underline{v}$
6. $\forall \underline{u} \in V$; $\forall \alpha, \beta \in \mathbb{R}$; $(\alpha+\beta)\underline{u} = \alpha\underline{u}+\beta\underline{u}$
7. $\forall \underline{u} \in V$; $\forall \alpha, \beta \in \mathbb{R}$; $(\alpha\beta)\underline{u} = \alpha(\beta\underline{u})$
8. $\forall \underline{u} \in V$; $1 \cdot \underline{u} = \underline{u}$ (1 is a scalar here)

Remark:

The "vectors" in the vector space, are not necessarily vectors $(\in \mathbb{R}^n)$, but can be other objects, as long as the definition is satisfied.

Example

Let us consider a set of all 2×2 matrices. It is a vector space. Proof:
If $A, B \in \mathbb{R}^{2,2}$ $\quad (A+B) \in \mathbb{R}^{2,2}$
$\alpha \in \mathbb{R}, A \in \mathbb{R}^{2,2}$ $\quad \alpha A \in \mathbb{R}^{2,2}$

1. $A, B, C \in \mathbb{R}^{2,2}$; $(A+B)+C = A+(B+C)$
2. ...
3. $$\underline{0} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{2,2}, \forall A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \Rightarrow A + \underline{0} = A$$
4. $$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \Rightarrow (-A) = \begin{pmatrix} -a_{11} & -a_{12} \\ -a_{21} & -a_{22} \end{pmatrix}$$

Example

Let us consider a set consisting of a single object, $\underline{0}$. It is a vector space. Note: There is no vector space, which does not contain $\underline{0}$

4.1 Subspace of the vector space

Definition

A subspace W of the vector space V, is a set of vectors in V, such that:

1. If $\underline{u}, \underline{v} \in W$ then $\underline{u} + \underline{v} \in W$
2. If $\alpha \in \mathbb{R}$, $\underline{u} \in W$ then $\alpha \underline{u} \in W$

Definition

Let us consider a set of vectors $\{\underline{u}_1, \dots, \underline{u}_n\}$. The span of vectors $\{\underline{u}_1, \dots, \underline{u}_n\}$ is defined as

$$\mathcal{S} = \operatorname{span}\{\underline{u}_1, \dots, \underline{u}_n\} = \{\alpha_1 \underline{u}_1 + \dots + \alpha_n \underline{u}_n \mid \forall \alpha_1 \dots \alpha_n \in \mathbb{R}\}$$

Example

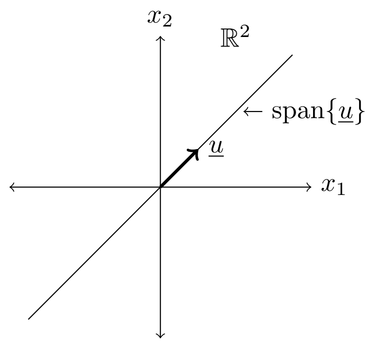
Is $\operatorname{span}\{\underline{u}\}$ a subspace in $\mathbb{R}^2$? Proof:

$$\underline{v} = \alpha \underline{u} \in \operatorname{span}\{\underline{u}\}$$
$$\underline{w} = \beta \underline{u} \in \operatorname{span}\{\underline{u}\}$$

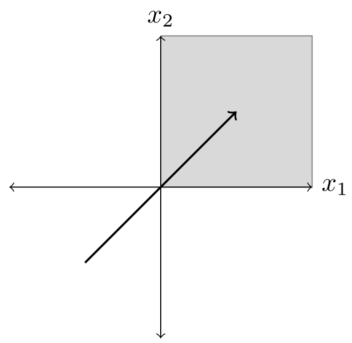
1. $\underline{v} + \underline{w} = \alpha \underline{u} + \beta \underline{u} = (\alpha + \beta)\underline{u} \in \operatorname{span}\{\underline{u}\}$
2. $\gamma \in \mathbb{R}, \gamma \underline{v} = \gamma \cdot (\alpha \underline{u}) = (\gamma \cdot \alpha)\underline{u} \in \operatorname{span}\{\underline{u}\}$

Example

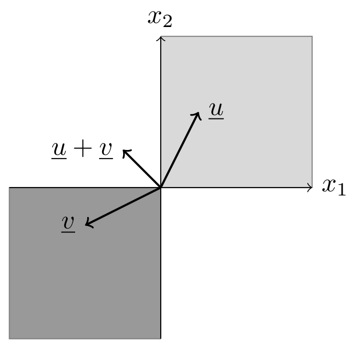
1.



2.



3.



4.2 Linear Independence

Definition

Let us consider vector space V and $\underline{v}_1, \ldots, \underline{v}_n \in V$. $\underline{v}_1, \ldots, \underline{v}_n$ are linearly dependent if there exists scalars $\alpha_1, \ldots, \alpha_n$ not all equal to zero, such that $\alpha_1 \underline{v}_1 + \cdots + \alpha_n \underline{v}_n = \underline{0}$

If no such scalars exist, the vectors $\underline{v}_1, \ldots, \underline{v}_n$ are linearly independent.

Definition

Vectors $\underline{v}_1, \ldots, \underline{v}_n \in V$ are linearly independent if the following is true:

$$\alpha_1 \underline{v}_1 + \cdots + \alpha_n \underline{v}_n = 0 \Rightarrow \text{ all } \alpha_i = 0, i = 1, \ldots, n$$

Example

1. Let us consider $\mathbb{R}^n$ and vectors
$$\underline{E}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \underline{E}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \underline{E}_i = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \underline{E}_n = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$
$\underline{E}_1, \ldots, \underline{E}_n$ are linearly independent.

2. Let us consider $\mathbb{R}^2$, $\underline{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\underline{u}_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$. Are they linearly independent? See proof 2.

Proof

1. Assume that
$$\alpha_1 \underline{E}_1 + \cdots + \alpha_n \underline{E}_n = 0 \Rightarrow \alpha_1 \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \cdots + \alpha_n \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$
$\Rightarrow$ then all $\alpha_i = 0$ for $i = 1, \ldots, n$,then based on the definition $\underline{E}_1, \ldots, \underline{E}_n$ are linearly independent.

2. Let us consider $\alpha_1 \underline{u}_1 + \alpha_2 \underline{u}_2 = \underline{0} \Rightarrow \alpha_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$
$$\begin{cases} \alpha_1 + 3\alpha_2 = 0 \\ \alpha_1 + \alpha_2 = 0 \end{cases} \rightarrow \begin{cases} 2\alpha_2 = 0 \\ \alpha_1 + \alpha_2 = 0 \end{cases} \rightarrow \begin{cases} \alpha_2 = 0 \\ \alpha_1 = 0 \end{cases}$$
If we assume $\alpha_1 \underline{u}_1 + \alpha_2 \underline{u}_2 = \underline{0}$, we have to show that all α_i are zeroes $\Rightarrow$ vectors are linearly independent.

$\square$

Example

Let us consider $\mathbb{R}^2$, $\underline{u}_1 = \begin{pmatrix}1\\1\end{pmatrix}$, $\underline{u}_2 = \begin{pmatrix}2\\2\end{pmatrix}$. Let us assume that

$$\alpha_1\underline{u}_1 + \alpha_2\underline{u}_2 = \underline{0} \Rightarrow \alpha_1\begin{pmatrix}1\\1\end{pmatrix} + \alpha_2\begin{pmatrix}2\\2\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

$$\begin{cases}\alpha_1 + 2\alpha_2 = 0\\ \alpha_1 + 2\alpha_2 = 0\end{cases} \rightarrow \begin{cases}\alpha_1 + 2\alpha_2 = 0\\ 0 = 0\end{cases}$$

One possible solution:

$$\begin{cases}\alpha_1 = -2\\ \alpha_2 = 1\end{cases}$$

Linearly dependent.

Recap

If we consider vectors $\underline{v}_1, \dots, \underline{v}_n \in V$, then

$$\mathbf{span}\{\underline{v}_1, \dots, \underline{v}_n\} = \{\alpha_1\underline{v}_1, \dots, \alpha_n\underline{v}_n \mid \text{ for all possible } \alpha_1, \dots, \alpha_n \in \mathbb{R}\}$$

Definition

If vector space v is generated by $\{\underline{v}_1, \dots, \underline{v}_n\}$ (in other words, $V = \text{span}\{\underline{v}_1, \dots, \underline{v}_n\}$) and $\underline{v}_1, \dots, \underline{v}_n$ are linearly independent, then $\{\underline{v}_1, \dots, \underline{v}_n\}$ is called basis of V

Example

Let us consider $\mathbb{R}^n$ and $\underline{E}_1, \dots, \underline{E}_n$. They form basis of $\mathbb{R}^n$.

Proof

1. "V is generated by $\underline{v}_1, \dots, \underline{v}_n$". Let us consider any vector $\underline{u} \in \mathbb{R}^n$

$$\underline{u} = \begin{pmatrix}\underline{u}_1\\ \vdots\\ \underline{u}_n\end{pmatrix}, \text{ we have}$$

$$\underline{u} = \begin{pmatrix}\underline{u}_1\\ \vdots\\ \underline{u}_n\end{pmatrix} = \underline{u}_1\underline{E}_1 + \dots + \underline{u}_n\underline{E}_n \Rightarrow \mathbb{R}^n = \text{span}\{\underline{E}_1, \dots, \underline{E}_n\}$$

2. "Linear independence" already proven before.

$\square$

Example

Let us consider $\mathbb{R}^2$ and $\underline{u}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\underline{u}_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}$, is it a basis?

1. Is $\mathbb{R}^2 = \text{span}\left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 3 \\ 1 \end{pmatrix} \right\}$? Let us consider an arbitrary vector $\underline{v} = \begin{pmatrix} \underline{v}_1 \\ \underline{v}_2 \end{pmatrix} \in \mathbb{R}^2$. We should check that there exists scalars α_1, α_2 such that

$$\underline{v} = \alpha_1 \underline{u}_1 + \alpha_2 \underline{u}_2 \rightarrow \underline{v} = \alpha_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 3 \\ 1 \end{pmatrix} = \begin{pmatrix} \underline{v}_1 \\ \underline{v}_2 \end{pmatrix}$$

$$\begin{cases} \alpha_1 + 3\alpha_2 = \underline{v}_1 \\ \alpha_1 + \alpha_2 = \underline{v}_2 \end{cases} \rightarrow \begin{cases} 2\alpha_2 = \underline{v}_1 - \underline{v}_2 \\ \alpha_1 + \alpha_2 = \underline{v}_2 \end{cases} \rightarrow \begin{cases} \alpha_2 = \frac{\underline{v}_1 - \underline{v}_2}{2} \\ \alpha_1 = \underline{v}_2 - \frac{\underline{v}_1 - \underline{v}_2}{2} = \frac{3\underline{v}_2 - \underline{v}_1}{2} \end{cases}$$

2. $\underline{u}_1, \underline{u}_2$ = linearly independent (We showed it before).

Definition

Let us consider vector space V and vectors $\underline{v}_1, \dots, \underline{v}_n$ that form a basis of V. If vector $\underline{x} \in V$ can be written as $\underline{x} = x_1 \underline{v}_1 + \dots + x_n \underline{v}_n$ then $(x_1, \dots, x_n)$ are called the coordinates of $\underline{x}$ with respect to basis $\{v_1, \dots, v_n\}$

Theorem

Let us consider vector space V and $v_1, \dots, v_n$ that are linearly independent. Let us assume that $\underline{x} = \alpha_1 v_1 + \dots + \alpha_n v_n$ and $\underline{x} = \beta_1 \underline{v}_1 + \dots + \beta_n \underline{v}_n$, then

$$\alpha_i = \beta_i \quad \forall i = 1, \dots, n$$

Proof

We have

$$x = \alpha_1 \underline{v}_1 + \dots + \alpha_n \underline{v}_n = \beta_1 \underline{v}_1 + \dots + \beta_n \underline{v}_n \rightarrow (\alpha_1 - \beta_1)\underline{v}_1 + \dots + (\alpha_n - \beta_n)\underline{v}_n = \underline{0}$$

Since $v_1, \dots, v_n$ are linearly independent $\Rightarrow \alpha_i = \beta_i$, $\forall i = 1, \dots, n$

$\square$

<u>Remark:</u>

The coordinates of any vector $\underline{x}$ with respect to given basis $\{\underline{v}_1, \dots, \underline{v}_n\}$ are unique.

Theorem

Let us consider vector space V. The number of vectors in any basis of V is always the same.

Remark:

The number of vectors in the basis of vector space V is called the dimension of vector space V.

4.3 Rank of matrix

Definition

The row rank of matrix A is a maximum number of linearly independent rows of matrix A.

Definition

The column rank of matrix A is a maximum number of linearly independent columns of matrix A.

Remark:

For any matrix $A \in \mathbb{R}^{m,n}$, the row rank is equal to the column rank. Therefore the row rank and column rank are sometimes called rank of matrix A, $\operatorname{rank}(A)$.

Example

1.
$$A = \begin{pmatrix} 1 & 3 \\ 1 & 1 \end{pmatrix}$$
We have shown before that $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\begin{pmatrix} 3 \\ 1 \end{pmatrix}$ are linearly independent, therefore $\operatorname{rank}(A) = 2$.

2.
$$A = \begin{pmatrix} 1 & 0 \\ 7 & 0 \\ 3 & 0 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{4,2}$$
The column vectors $\begin{pmatrix} 1 \\ 7 \\ 3 \\ -1 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$ are linearly dependent, thus $\operatorname{rank}(A) = 1$ (i.e. the maximum number of linearly independent columns is 1).

Remark:

Two vectors are orthogonal if $\langle \underline{u}, \underline{v} \rangle = \underline{u}^T \underline{v} = 0$ (they basically must be perpendicular, i.e. the angle between $\underline{u}$ and $\underline{v}$ is 90 degrees).

Definition

Two subspaces U and W of vector space V are orthogonal, if $\forall \underline{u} \in U$ and $\forall \underline{w} \in W$, we have $\langle \underline{u}, \underline{w} \rangle = 0$

Definition

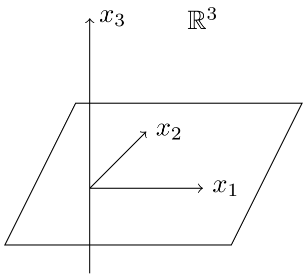
Orthogonal complement of subspace M of vector space V contains every vector orthogonal to M. This subspace is usually denoted by $M^{\perp}$

Remark:

$\dim M + \dim M^{\perp} = \dim V$

Example

Consider $\mathbb{R}^3$



line α plane - orthogonal subspace. Orthogonal complement of each other

Example

Not orthogonal subspace!
$\underline{u} \neq 0$ ADD FIGURE
$\underline{u} \in I$ & $\underline{u} \in W$
$\langle \underline{u} \in U, \underline{u} \in W \rangle = 0$

Add missing figure to minipage

Note

If vector $\underline{u}$ belongs to 2 orthogonal subspaces, this vector is necessarily a zero vector, $\underline{u} = 0$ because we should have

$$\langle \underline{u}, \underline{u} \rangle = \underline{u}^T \underline{u} = 0 \Rightarrow \underline{u} = \underline{0}$$

Chapter 5

Linear Mapping

Definition

Let us consider 2 vector spaces V and W. A function $\mathcal{L} : V \to W$ is called a linear mapping, if:

1. For any $\underline{v} \in V$ and $\underline{v}' \in V$, $\mathcal{L}(\underline{v} + \underline{v}') = \mathcal{L}(\underline{v}) + \mathcal{L}(\underline{v}')$
2. For any $\underline{v} \in V$ and any scalar α, $\mathcal{L}(\alpha \underline{v}) = \alpha \cdot \mathcal{L}(\underline{v})$

Example

Let us consider matrix $A \in \mathbb{R}^{n,m}$. We can define linear mapping $\mathcal{L}_A$ as follows:

$$\mathcal{L}_A(\underline{v}) = A\underline{v} \qquad \mathcal{L}_A : \mathbb{R}^m \to \mathbb{R}^n$$

Is $\mathcal{L}_A$ a linear mapping? Yes!

Proof

1. $\forall \underline{v}, \underline{v}' \in \mathbb{R}^m$, we have:

$$\mathcal{L}_A(\underline{v} + \underline{v}') = A(\underline{v} + \underline{v}') = A\underline{v} + A\underline{v}' = \mathcal{L}_A(\underline{v}) + \mathcal{L}_A(\underline{v}')$$

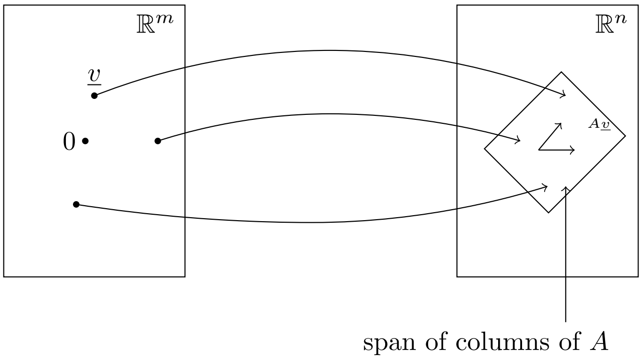
2. $\forall \underline{v} \in \mathbb{R}^m, \forall \alpha$ (α is scalar), we have:

$$\mathcal{L}_A(\alpha \underline{v}) = A(\alpha \underline{v}) = \alpha \cdot A\underline{v} = \alpha \mathcal{L}_A(\underline{v})$$

$\square$

Let us consider matrix $A \in \mathbb{R}^{n,m}$, $A : \mathbb{R}^m \to \mathbb{R}^n$. Let us consider vector $\underline{v} \in \mathbb{R}^m$
.

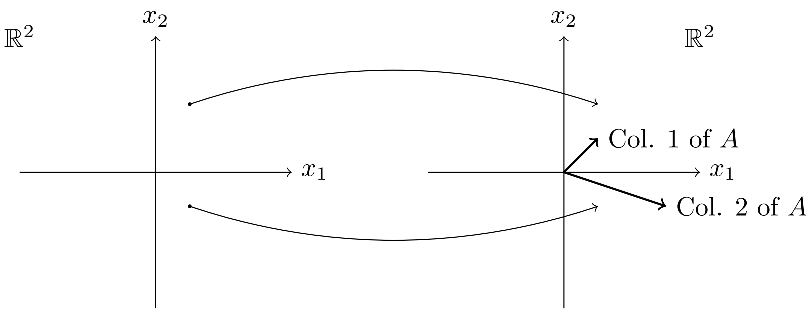
$$A\underline{v} = \underbrace{v_1 \cdot \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + v_2 \cdot \begin{pmatrix} a_{12} \\ \vdots \\ a_{n2} \end{pmatrix} + \ldots + v_m \cdot \begin{pmatrix} a_{1m} \\ \vdots \\ a_{nm} \end{pmatrix}}_{\text{Linear combination of columns of } A}$$



Example

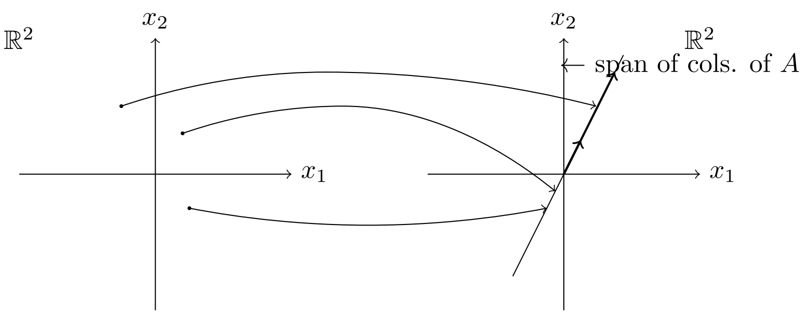
1.

$$A \in \mathbb{R}^{2,2} = \begin{pmatrix} 1 & 3 \\ 1 & -1 \end{pmatrix}$$



2.

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$$



Note

In order for solution of $A\underline{x} = \underline{b}$ to exist, $\underline{b}$ should belong to a span of columns of matrix A.

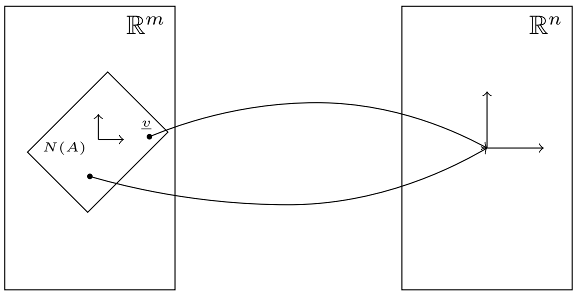
Definition

The span of columns of matrix $A \in \mathbb{R}^{n,m}$ is called a column space of A, denoted by $C(A)$, where $C(A) \subset R^n$.

Definition

Let us consider matrix $A \in \mathbb{R}^{n,m}, A : \mathbb{R}^m \to \mathbb{R}^n$. The null space of A is defined as

$$N(A) = \{\underline{v} \in \mathbb{R}^m \mid A\underline{v} = \underline{0}\}, N(A) \subset \mathbb{R}^m$$



Example

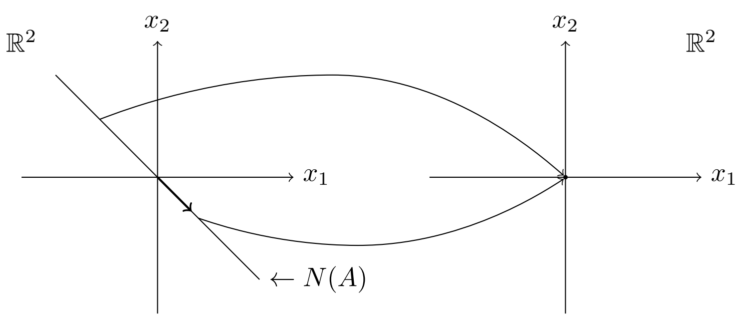
$$A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix}$$

What is $N(A)$? We should find all solutions of $A\underline{x} = \underline{0}$, this will give us $N(A)$.

$$\begin{cases} x_1 + 3x_2 = 0 \\ 2x_1 + 6x_2 = 0 \end{cases} \to \begin{cases} x_1 + 3x_2 = 0 \\ 0 = 0 \end{cases} \to \begin{cases} x_1 = -3x_2 \\ 0 = 0 \end{cases}$$

The nullspace of this matrix will be a line formed by a linear combination of the vector $\begin{pmatrix} -3 \\ 1 \end{pmatrix}$ ($\alpha \cdot \begin{pmatrix} -3 \\ 1 \end{pmatrix}$, for all possible α), or in other words it will be the $span(\begin{pmatrix} -3 \\ 1 \end{pmatrix})$.

$$x_1 = -3x_2 = -3\alpha, x_2 = \alpha \to \alpha \begin{pmatrix} -3 \\ 1 \end{pmatrix}, \alpha \begin{pmatrix} -6 \\ 2 \end{pmatrix}$$



Theorem

The nullspace, $N(A)$, of $A \in \mathbb{R}^{n,m}$ is a subspace of $\mathbb{R}^m$.

Proof

Let us assume that $\underline{x}, \underline{x}' \in N(A)$ and α is arbitrarily scalar.

1. $A(\underline{x} + \underline{x}') = A\underline{x} + A\underline{x}' = \underline{0} + \underline{0} = \underline{0} \Rightarrow (\underline{x} + \underline{x}') \in N(A)$

2. $A(\alpha\underline{x}) = \alpha(A\underline{x}) = \alpha \cdot \underline{0} = \underline{0} \Rightarrow \alpha\underline{x} \in N(A)$

□

Theorem

The column space, $C(A)$, of $A \in \mathbb{R}^{n,m}$ is a subspace of $\mathbb{R}^n$.

Definition

The row space of matrix $A \in R^{n,m}$ is a span of rows of A. Clearly, $R(A) = C(A^T)$ and $R(A) \subset \mathbb{R}^m$.

Definition

The left nullspace of A is defined as $N(A^T)$. $N(A^T) \subset \mathbb{R}^n$.

Theorem

$R(A)$ is a subspace of $\mathbb{R}^m$

Proof

Same as for the proof that $C(A)$ is a subspace of $\mathbb{R}^n$, but for A^T

□

Theorem

$N(A^T)$ is a subspace of $\mathbb{R}^n$

Proof

Same as for $N(A)$ but replace A with A^T

□

Theorem

$R(A)$ and $N(A)$ are orthogonal subspaces in $\mathbb{R}^m$ for $A \in \mathbb{R}^{n,m}$

Proof

Let us consider $\forall \underline{x} \in N(A), A\underline{x} = \underline{0}$

$$A\underline{x} = \begin{pmatrix} - \text{ row 1 of } A \rightarrow \\ \vdots \\ - \text{ row } n \text{ of } A \rightarrow \end{pmatrix} \cdot \begin{pmatrix} | \\ x \\ \downarrow \end{pmatrix} = \begin{pmatrix} < \text{ row 1 of } A, \underline{x} > \\ \vdots \\ < \text{ row } n \text{ of } A, \underline{x} > \end{pmatrix} \overset{\underline{x} \in N(A)}{=} \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

$\underline{x}$ is orthogonal to every row of A. $\underline{x}$ is orthogonal to every linear combination of rows of A. $\underline{x}$ is orthogonal to $R(A)$. In fact, what we just showed is that $N(A)$ & $R(A)$ are orthogonal complements.

$\square$

Theorem

$N(A^T)$ & $C(A) = R(A^T)$ are orthogonal complements in $\mathbb{R}^n$

$A \in \mathbb{R}^{n,m} : \mathbb{R}^m \rightarrow \mathbb{R}^n$. Row rank of $A = \text{rank}(A) = \dim(R(A)) = \dim(C(A))$

$$\begin{aligned} N(A) &: A\underline{x} = \underline{0} \qquad \forall x \in \mathbb{R}^m \\ C(A) &: A\underline{v} = \text{ Linear combinations of columns of } A \\ &\qquad = v_1 \cdot \text{col 1 of } A + \cdots + v_n \cdot \text{col } n \text{ of } A \in \mathbb{R}^n \end{aligned}$$

Theorem

$N(A)$ is an orthogonal complement of $R(A)$ in $\mathbb{R}^m$,

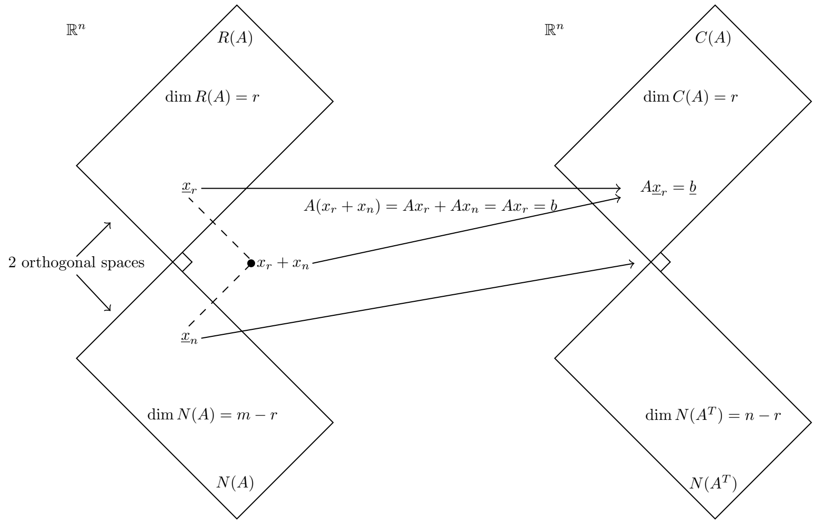
$$\dim N(A) + \underbrace{\dim R(A)}_{=\text{rank}(A)} = m$$

Theorem

$N(A^T)$ is an orthogonal complement of $R(A^T) = C(A)$ in $\mathbb{R}^n$,

$$\dim N(A^T) + \underbrace{\dim C(A)}_{=\text{rank}(A)} = n$$

Let us consider $A \in \mathbb{R}^{n,m}, A : \mathbb{R}^m \rightarrow \mathbb{R}^n, \text{rank}(A) = r$



Lemma

For any vector $\underline{b}$ in $C(A)$, there exists one and only one vector $\underline{x}_r \in R(A)$ such that $A\underline{x}_r = b$

Proof

Let us assume that $\underline{x}_r$ and $\underline{x}'_r$ are in the row space, $R(A)$. Let us assume that $A\underline{x}_r = A\underline{x}'_r$. We have

$$\underline{x}_r \in R(A) - \underline{x}'_r \in R(A) \in R(A)$$

But we also have

$$A\underline{x}_r - A\underline{x}'_r = \underbrace{A(\underline{x}_r - \underline{x}'_r}_{\in N(A)} = \underline{0}$$

It means that $(\underline{x}_r - \underline{x}'_r)$ is in $R(A)$ and $N(A)$, but they are orthogonal subspaces, therefore

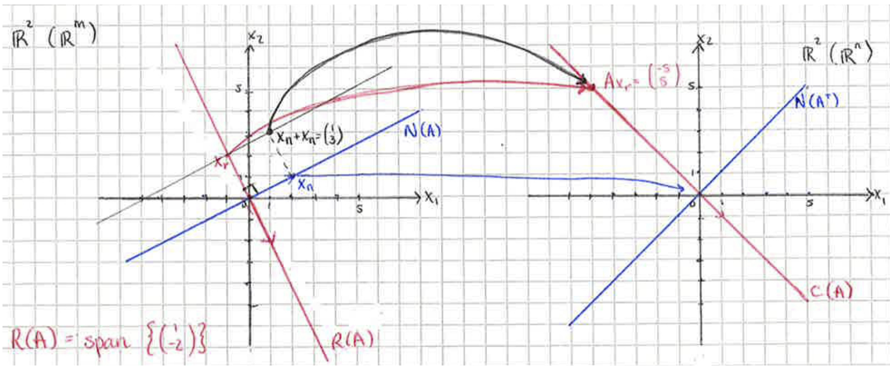
$$\underline{x}_r - \underline{x}'_r = \underline{0} \Rightarrow \underline{x}_r = \underline{x}'_r$$

$\square$

Example

Let us consider

$$A = \begin{pmatrix} 1 & -2 \\ -1 & 2 \end{pmatrix} \in \mathbb{R}^{2,2}$$



Row space: rank $A = 1 \Rightarrow \dim R(A) = 1$

$$R(A) = \text{span}\left\{\begin{pmatrix}1\\-2\end{pmatrix}, \begin{pmatrix}-1\\2\end{pmatrix}\right\} = \text{span}\left\{\begin{pmatrix}1\\-2\end{pmatrix}\right\}$$

Null space: $\dim N(A) = 2 - 1 = 1$

$$A\underline{x} = 0 \Rightarrow \begin{cases} x_1 - 2x_2 = 0 \\ -x_1 + 2x_2 = 0 \end{cases} \Rightarrow \begin{cases} x_1 - 2x_2 = 0 \\ 0 = 0 \end{cases} \Rightarrow x_1 = 2x_2 \text{ (Line)}$$

Column space: $\dim C(A) = \dim R(A) =\perp$

$$C(A) = \text{span}\left\{\begin{pmatrix}1\\-1\end{pmatrix}, \begin{pmatrix}-2\\2\end{pmatrix}\right\} = \text{span}\left\{\begin{pmatrix}1\\-1\end{pmatrix}\right\}$$

Left Null space: $\dim N(A^T) = 2 - 1 = 1$. Consider

$$\underline{x}_r = \begin{pmatrix}-1\\2\end{pmatrix} \Rightarrow A\underline{x}_r = \begin{pmatrix}1 & -2\\-1 & 2\end{pmatrix}\begin{pmatrix}-1\\2\end{pmatrix} = \begin{pmatrix}-5\\5\end{pmatrix}$$

$$\underline{x}_n = \begin{pmatrix}2\\1\end{pmatrix} \Rightarrow A\underline{x}_n = \begin{pmatrix}1 & -2\\-1 & 2\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

5.1 Orthogonal Basis and Gram-Schmidt process

Definition

Vectors $\underline{q}_1, \ldots, \underline{q}_m$ are orthogonal if:

$$\langle \underline{q}_i, \underline{q}_j \rangle = \underline{q}_i^T \underline{q}_j = 0 \quad \text{if } i \neq j$$

Definition

Vectors $\underline{q}_1, \ldots, \underline{q}_m$ are orthonormal if:

$$\langle \underline{q}_i, \underline{q}_j \rangle = \underline{q}_i^T \underline{q}_j = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$$

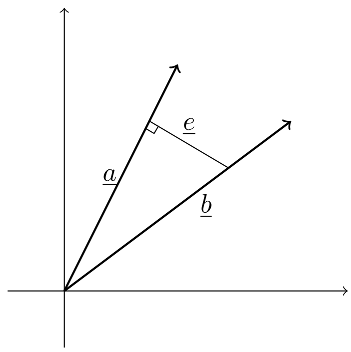
If the columns of the matrix are orthonormal vectors, then this matrix is usually denoted by Q, In this case, we have $Q^TQ = I$. If Q is not a square matrix then QQ^T is not necessarily I.

Definition

A square matrix is called orthogonal (if its columns are orthonormal vectors) if $Q^TQ = I$. In this case, since it is a square matrix, $QQ^T = I$

5.1.1 Projection on the line

Let us assume that we have a line given by vector $\underline{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^n$ and vector $\underline{b} \in \mathbb{R}^n$. We want to find vector $\underline{p}$ belonging to the line, closest to vector $\underline{b}$. In other words, we are looking for $\underline{p}$ which is orthogonal projection of $\underline{b}$ onto the line given by $\underline{a}$



$\underline{p}$ is proportional to $\underline{a}$, $\underline{p} = \hat{x}\underline{a}$, where $\hat{x}$ is some scalar. Let us define vector $\underline{e} = \underline{b} - \underline{p} = \underline{b} - \hat{x}\underline{a}$ (error vector). $\underline{e}$ is orthogonal to the line, therefore

$$
\begin{aligned}
&\langle \underline{a}, \underline{e} \rangle = 0 \\
&\langle \underline{a}, \underline{e} \rangle = \underline{a}^T(\underline{(}b) - \hat{x}\underline{a}) = \underline{a}^T\underline{b} - \hat{x}\underline{a}^T\underline{a} = 0 \\
&\Rightarrow \hat{x} = \frac{\underline{a}^T\underline{b}}{\underline{a}^T\underline{a}} \\
&\Rightarrow \underline{p} = \hat{x}\underline{a} = \underline{a}\hat{x} = \underline{a}\frac{\underline{a}^T\underline{b}}{\underline{a}^T\underline{a}} = \underbrace{\frac{\underline{a}\underline{a}^T}{\underline{a}^T\underline{a}}}_{P \in \mathbb{R}^{n,n} \text{ (projection matrix)}} \cdot \underline{b}
\end{aligned}
$$

Example

Let us consider $\underline{a} = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix} \in \mathbb{R}^3$

$$
P = \frac{\underline{a}^T\underline{b}}{\underline{a}^T\underline{a}} = \langle \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 2 \end{pmatrix} \rangle \cdot \frac{1}{9} = \frac{1}{9} \cdot \begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{pmatrix}
$$

Let us take

$$\underline{b} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \underline{p} = P\underline{b} = \frac{1}{9}\begin{pmatrix} 1 & 2 & 2 \\ 2 & 4 & 4 \\ 2 & 4 & 4 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{9}\begin{pmatrix} 5 \\ 10 \\ 10 \end{pmatrix}$$

Note

$\underline{p}^2 = \underline{p}$

Note

$(I - P)-$ projection onto subspace orthogonal to the line given by $\underline{a}$

5.2 Gram-Schmidt process

Given linear independent vectors $\underline{a}, \underline{b}, \underline{c}, \ldots$ we first find orthogonal vectors $\underline{a}', \underline{b}', \underline{c}', \ldots$ which span the same subspace as $\underline{a}, \underline{b}, \underline{c}, \ldots$ and then we normalise them,

$$\underline{q}_1 = \frac{\underline{a}'}{\|\underline{a}'\|}, \underline{q}_2 = \frac{\underline{b}'}{\|\underline{b}'\|}, \underline{q}_3 = \frac{\underline{c}'}{\|\underline{c}'\|}, \ldots$$

So, Gram-Schmidt process allows us to construct an orthogonal basis of span$\underline{a}, \underline{b}, \underline{c}, \ldots \in \mathbb{R}^n$

1. Choose $\underline{a}' = \underline{a}$
2. It is likely that $\underline{b}$ is not orthogonal to $\underline{a}'$, so we need to subtract its projection on the line defined by $\underline{a}'$

$$\underline{b}' = \underline{b} - \frac{\underline{a}^T\underline{b}}{\underline{a}^T\underline{a}}\underline{a}'$$

3. $\underline{c}'$ is likely not orthogonal to $\underline{a}'$ and $\underline{b}'$. Again, subtract its projections

$$\underline{c}' = \underline{c} - \frac{\underline{a}^T\underline{c}}{\underline{a}^T\underline{a}}\underline{a}' - \frac{\underline{b}'^T\underline{c}}{\underline{b}'^T\underline{b}'}\underline{b}'$$

and so on. Finally, normalise $\underline{q}_1, \underline{q}_2, \underline{q}_3, \ldots$

Example

With

$$\underline{a} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \underline{b} = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}, \underline{c} = \begin{pmatrix} 3 \\ -3 \\ 3 \end{pmatrix}$$

find $\underline{a}', \underline{b}', \underline{c}', \underline{q}_1, \underline{q}_2, \underline{q}_3$

1.

$$\underline{a}' = \underline{a} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$$

2.

$$\underline{b'} = \underline{b} - \frac{\underline{a'}^T \underline{b}}{\underline{a'}^T \underline{a'}} \underline{a'} = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} - \frac{\langle \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} \rangle}{\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

3.

$$\underline{c'} = \underline{c} - \frac{\underline{a}^T \underline{c}}{\underline{a}^T \underline{a}} \underline{a'} - \frac{\underline{b'}^T \underline{c}}{\underline{b'}^T \underline{b'}} \underline{b'} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

$$\langle \underline{a'}, \underline{c'} \rangle = 0, \langle \underline{b'}, \underline{c'} \rangle = 0$$

Finally normalise:

$$\underline{q}_1 = \frac{\underline{a'}}{\|\underline{a'}\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \underline{q}_2 = \frac{\underline{b'}}{\|\underline{b'}\|} = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \underline{q}_3 = \frac{\underline{c'}}{\|\underline{c'}\|} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

5.3 Projection onto subspace

Assume we have linearly independent vectors $a_1, \dots, a_m \in \mathbb{R}^n$. We want to project vector $\underline{b} \in \mathbb{R}^n$ onto subspace spanned by $a_1, \dots, a_m$. Subspace consists of all linear combinations

$$x_1 a_1 + \cdots + x_m a_m = \underbrace{\begin{pmatrix} | & & | \\ a_1 & \dots & a_m \\ \downarrow & & \downarrow \end{pmatrix}}_{A \in \mathbb{R}^{n,m}} \cdot \underbrace{\hat{x}}_{\in \mathbb{R}^m}$$

We are looking for the projection $\underline{p}$ of $\underline{b}$ onto his subspace. We can define $\underline{e} = \underline{b} - \underline{p}$, $\underline{e}$ should be orthogonal to all $a_1, \dots, a_m$

$$\left.\begin{array}{c} \langle a_1, \underline{e} \rangle = \underline{a_1}^T \cdot (\underline{b} - A\hat{x}) = 0 \\ \vdots \\ \langle a_m, \underline{e} \rangle = \underline{a_m}^T \cdot (\underline{b} - A\hat{x}) = 0 \end{array}\right\} \Rightarrow \underbrace{\begin{pmatrix} -\underline{a}_1^T \rightarrow \\ \vdots \\ -\underline{a}_m^T \rightarrow \end{pmatrix}}_{A^T} (\underline{b} - A\hat{x}) = 0$$

$$A^T(\underline{b} - A\hat{x} = 0$$
$$A^T\underline{b} - A^T A\hat{x} = 0$$

Theorem

A has linearly independent columns. Then $A^T A$ is:

- Square
- Symmetric

- Invertible

$$\underline{\hat{x}} = A(A^TA)^{-1}A^T\underline{b}$$
$$\underline{p} = A\hat{x} = \underbrace{A(A^TA)^{-1}A^T}_{P\text{ - Proj. matrix}} \cdot \underline{b} \text{ - Projection vector}$$