

Data Representation

• For example, We have the following data

Age	GPA	Hours Studied
20	3.4	15
21	3.6	10
19	3.2	18

- How do we represent the data?
- Every row is a vector representing one student.
 Every column is a vector representing a feature.

Data Representation

• Suppose we have the grades obtain for each student as well.

Age	GPA	Hours Studied	Grades
20	3.4	15	80
21	3.6	10	50
19	3.2	18	90

 How can we train from this data to predict the grades of any students whose features are known?

Linear Algebra

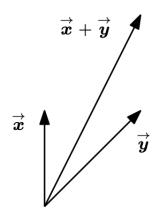
• Linear algebra is the study of vectors and certain rules associated with vectors.

- Linear algebra is the study of vectors and certain rules associated with vectors.
- The vectors many of us know from school are called "geometric vectors", which are usually denoted by a small arrow above the letter, e.g., \vec{x} and \vec{y} .
- We will discuss more general concepts of vectors and use a bold letter to represent them, e.g., x and y.
- One major idea in mathematics is the idea of "closure".
- This is the question: Does the resultant of the addition of two vectors belong from the same set or outside the set? Does the resultant of the scalar multiplication of a vectors with a scalar belong from the same set or outside the set?

• In general, vectors are special objects that can be added together and multiplied by scalars to produce another object of the same kind.

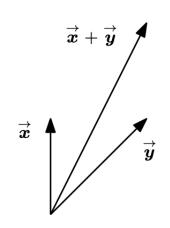
Examples:

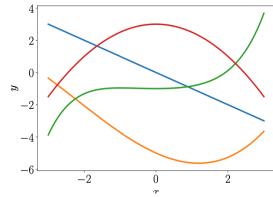
- 1. Geometric vectors. This example of a vector may be familiar from high school mathematics and physics. Two geometric vectors can be added, such that $\vec{x} + \vec{y} = \vec{z}$ is another geometric vector. Furthermore, multiplication by a scalar $\lambda \vec{x}$, $\lambda \in \mathbb{R}$, is also a geometric vector.
- Polynomials are also vectors; see Figure 2.1(b): Two polynomials can be added together, which results in another polynomial; and they can be multiplied by a scalar λ ∈ R, and the result is a polynomial as well.



Examples:

- 1. Geometric vectors. This example of a vector may be familiar from high school mathematics and physics. Two geometric vectors can be added, such that $\vec{x} + \vec{y} = \vec{z}$ is another geometric vector. Furthermore, multiplication by a scalar $\lambda \vec{x}$, $\lambda \in \mathbb{R}$, is also a geometric vector.
- Polynomials are also vectors; see Figure 2.1(b): Two polynomials can be added together, which results in another polynomial; and they can be multiplied by a scalar λ ∈ R, and the result is a polynomial as well.





Examples:

- 3. Audio signals are vectors. Audio signals are represented as a series of numbers. We can add audio signals together, and their sum is a new audio signal. If we scale an audio signal, we also obtain an audio signal. Therefore, audio signals are a type of vector, too.
- 4. Elements of \mathbb{R}^n (tuples of n real numbers) are vectors. \mathbb{R}^n is more abstract than polynomials. For instance,

$$a = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \in \mathbb{R}^n$$

is an example of a triplet of numbers. Adding two vectors $a, b \in \mathbb{R}^n$ component-wise results in another vector: $a + b = c \in \mathbb{R}^n$. Moreover, multiplying $a \in \mathbb{R}^n$ by $\lambda \in \mathbb{R}^n$ results in a scaled vector $\lambda a \in \mathbb{R}^n$.

Generalization to \mathbb{R}^n

• Column vectors with n entries, i.e.

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$
 are vectors in \mathbb{R}^n

- All the vectors with n real valued entries are form \mathbb{R}^n
- Generally, we cannot have a geometric description for \mathbb{R}^n

Linear Operations in \mathbb{R}^n

Sum and scalar multiple?

Algebraic Properties in \mathbb{R}^n

For all \mathbf{u} , \mathbf{v} , \mathbf{w} in \mathbb{R}^n and all scalars c and d:

(i)
$$\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$$

(ii)
$$(u + v) + w = u + (v + w)$$

(iii)
$$u + 0 = 0 + u = u$$

(iv)
$$\mathbf{u} + (-\mathbf{u}) = -\mathbf{u} + \mathbf{u} = \mathbf{0}$$
,
where $-\mathbf{u}$ denotes $(-1)\mathbf{u}$

(v)
$$c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$$

(vi)
$$(c+d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$$

(vii)
$$c(d\mathbf{u}) = (cd)\mathbf{u}$$

(viii)
$$1\mathbf{u} = \mathbf{u}$$

Let $a, b \in \mathbb{R}^m$ be two m-dimensional vectors given as

$$m{a} = egin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}$$
 , $m{b} = egin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$

Dot Product

The dot product between a and b gives a scalar value and is defined as

$$\mathbf{a}.\,\mathbf{b} = \mathbf{a}^T\mathbf{b} = (a_1 \quad a_2 \quad \dots \quad a_m) \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = a_1b_1 + a_2b_2 + \dots + a_mb_m$$

Let $a, b \in \mathbb{R}^m$ be two m-dimensional vectors given as

$$m{a} = egin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}, m{b} = egin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Length

The Euclidean norm or length of a vector a is defined as

$$\|a\| = \sqrt{a^T a} = \sqrt{a_1^2 + a_2^2 + \dots + a_m^2}$$

The Euclidean norm is a special case of a general class of norms, known as ${\cal L}_p$ -norm, defined as

$$\|\boldsymbol{a}\|_{p} = (|a_{1}|^{p} + |a_{2}|^{p} + \dots + |a_{m}|^{p})^{1/p}$$

Let $a, b \in \mathbb{R}^m$ be two m-dimensional vectors given as

$$m{a} = egin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}, m{b} = egin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Distance

From Euclidean norm we can define the Euclidean distance between vectors a and b defined as

$$\|\boldsymbol{a} - \boldsymbol{b}\| = \sqrt{(\boldsymbol{a} - \boldsymbol{b})^T (\boldsymbol{a} - \boldsymbol{b})} = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \dots + (a_m - b_m)^2}$$

Similarly L_p -norm distance is defined as

$$\|\boldsymbol{a} - \boldsymbol{b}\|_p = (|a_1 - b_1|^p + |a_2 - b_2|^p + \dots + |a_m - b_m|^p)^{1/p}$$

Let $a, b \in \mathbb{R}^m$ be two m-dimensional vectors given as

$$m{a} = egin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}, m{b} = egin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Unit Vector

The unit vector in the direction of a is given as

$$u = \frac{a}{\|a\|} = \left(\frac{1}{\|a\|}\right)a$$

This in turn makes ||u|| = 1, and it is also called **normalized vector.**

Let $a, b \in \mathbb{R}^m$ be two m-dimensional vectors given as

$$m{a} = egin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}, m{b} = egin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Angle

The cosine of the smallest angle between vectors a and b is given as

$$\cos \theta = \frac{\boldsymbol{a}.\boldsymbol{b}}{\|\boldsymbol{a}\|\|\boldsymbol{b}\|} = \left(\frac{\boldsymbol{a}}{\|\boldsymbol{a}\|}\right)^T \left(\frac{\boldsymbol{b}}{\|\boldsymbol{b}\|}\right)$$

Let $a, b \in \mathbb{R}^m$ be two m-dimensional vectors given as

$$m{a} = egin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_m \end{pmatrix}, m{b} = egin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$$

Orthogonality

Two vectors \mathbf{a} and \mathbf{b} are said to be orthogonal if and only if \mathbf{a} . $\mathbf{b} = \mathbf{a}^T \mathbf{b} = 0$, which in turn implies that $\cos \theta = 0$ and 90°. In this case, we say that the vectors have no similarity.

Orthogonal Projection

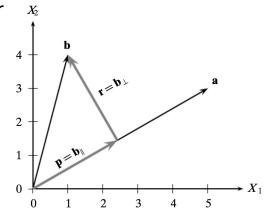
Often in data mining we need to project a point or vector onto another vector, for example, to obtain a new point after a change of the basis vectors.

Let $a, b \in \mathbb{R}^m$ be two m-dimensional vectors. An **orthogonal decomposition** of the vector b in the direction of another vector a is given as

$$\boldsymbol{b} = \boldsymbol{b}_{\parallel} + \boldsymbol{b}_{\perp} = \boldsymbol{p} + \boldsymbol{r}$$

where $p = b_{\parallel}$ is parallel to a, and $r = b_{\perp}$ is perpendicular or orthogonal to a.

The vector p is called the **orthogonal projection** or simply **projection** of p on the vector p, denoted by $proj_{q}(p)$.



Orthogonal Projection

- Note that the point $p \in \mathbb{R}^m$ is the point closest to b on the line passing through a.
- Thus, the magnitude of the vector r = b p gives the **perpendicular distance** between b and a, which is often interpreted as the residual or error between the points b and p.
- The vector r is also called the error vector.

Orthogonal Projection

We can derive an expression for $p = proj_a(b)$ by noting that p = ca for some scalar c, as p is parallel to a. Thus,

$$r = b - p = b - ca$$

Because p and r are orthogonal, we have

$$p \cdot r = p^{T} r = 0$$

$$(ca)^{T} (b - ca) = ca^{T} b - c^{2} a^{T} a = 0$$

$$c = \frac{a^{T} b}{a^{T} a} = \frac{a \cdot b}{a \cdot a} = \frac{a \cdot b}{\|a\|^{2}}$$

Therefore,

$$p = proj_a(b) = \left(\frac{a.b}{\|a\|^2}\right)a$$

An array of numbers is called a matrix.

A matrix with n number of rows and m number of columns is known as an $n \times m$ matrix.

Example:

Data can often be represented or abstracted as an $n \times d$ data matrix, with n rows and d columns, where rows correspond to entities in the dataset, and columns represent attributes or properties of interest.

$$D = \begin{pmatrix} x_1 & x_{11} & x_{12} & \dots & x_{1d} \\ x_2 & x_{21} & x_{22} & \dots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_n & x_{n1} & x_{n2} & \dots & x_{nd} \end{pmatrix}$$

Example:.

$$D = \begin{pmatrix} x_1 & x_{11} & x_{12} & \dots & x_{1d} \\ x_2 & x_{21} & x_{22} & \dots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_n & x_{n1} & x_{n2} & \dots & x_{nd} \end{pmatrix}$$

where x_i denotes the *i*th row, which is a d-tuple given as

$$\boldsymbol{x_i} = (x_{i1}, x_{i2}, \dots, x_{id})$$

And X_i denotes the jth column, which is an n-tuple given as

$$X_j = (x_{1j}, x_{2j}, \dots, x_{nj})$$

- Depending on the application domain, rows may also be referred to as entities, instances, examples, records, transactions, objects, points, feature-vectors, tuples, and so on.
- Likewise, columns may also be called attributes, properties, features, dimensions, variables, fields, and so on.
- The number of instances n is referred to as the **size** of the data, whereas the number of attributes d is called the **dimensionality** of the data.
- The analysis of a single attribute is referred to as univariate analysis.
- The simultaneous analysis of two attributes is called bivariate analysis.
- The simultaneous analysis of more than two attributes is called multivariate analysis.

The following are the operations of matrices

- **1.** Matrix addition M + N (only possible of both matrices have same size)
- **2.** Scalar multiplication λM ($\lambda \in \mathbb{R}$)
- **3. Matrix multiplication** *MN* (only possible if the number of column of the first matrix is equal to the number of rows of the second matrix).

If M is a $m \times p$ matrix and N is a $p \times n$ matrix then MN will be a $m \times n$ matrix

Example:

The multiplication of matrices

$$M = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 1 & 4 \\ 6 & 7 & 9 \end{bmatrix}, N = \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix}$$

ls

$$MN = \begin{bmatrix} 3 & 3 \\ 7 & 7 \\ 22 & 22 \end{bmatrix}$$

Example:

For a linear system of linear equations

$$2x + 3y = 2$$
$$2x + 2y = 1$$

The matrix equation is written as Ax = b

$$\begin{bmatrix} 2 & 3 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Why?

Because when we multiply

$$A\mathbf{x} = \begin{bmatrix} 2 & 3 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 2x + 3y \\ 2x + 2y \end{bmatrix}$$

Which can also be written as

$$x \begin{bmatrix} 2 \\ 2 \end{bmatrix} + y \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

Which is a **linear combination** of the vectors $v_1 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$ and $v_2 = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$. Here the vectors v_1 and v_2 are called **column vectors** of the matrix A.

We will come back to discuss the solution of system linear equation.

Linear Combinations of Vectors

- Given a set of vectors v_1, v_2, \dots, v_k , where k is an integer,
- then the vector y defined as

$$y = x_1 \boldsymbol{v_1} + x_2 \boldsymbol{v_2} + \dots + x_k \boldsymbol{v_k}$$

is a **linear combination** of vectors v_1, v_2, \dots, v_k , where $x_{1, x_2, \dots, x_k} \in \mathbb{R}$.

• Or equivalently, the vectors v_1, v_2, \dots, v_k said to **generate** the vector y.

Example

$$2\begin{bmatrix} -1\\1 \end{bmatrix} - 0.5\begin{bmatrix} 2\\4 \end{bmatrix} + 2\begin{bmatrix} 0.5\\-0.5 \end{bmatrix} = \begin{bmatrix} -2\\-1 \end{bmatrix}$$

Example

Let
$$a_1 = \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}$$
, $a_2 = \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$ and $b = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$

- Determine whether b can be generated (or written as a linear combination) of a_1 and a_2 .
- That is, determine whether weights x_1 and x_2 exist such that $x_1 a_1 + x_2 a_2 = b$.
- The above equation is known as a vector equation.
- Find if this vector equation has a solution.

Solution

Use the definitions of scalar multiplication and vector addition to rewrite the vector equation

$$x_{1}a_{1} + x_{2}a_{2} = b$$

$$x_{1}\begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + x_{2}\begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

$$\begin{bmatrix} x_{1} \\ -2x_{1} \\ -5x_{1} \end{bmatrix} + \begin{bmatrix} 2x_{2} \\ 5x_{2} \\ 6x_{2} \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

$$\begin{bmatrix} x_{1} + 2x_{2} \\ -2x_{1} + 5x_{2} \\ -5x_{1} + 6x_{2} \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

$$x_1 + 2x_2 = 7$$

$$-2x_1 + 5x_2 = 4$$

$$-5x_1 + 6x_2 = -3$$

The solution of above system is $x_1 = 3$ and $x_2 = 2$. Hence **b** is a linear combination of a_1 and a_2 , with weights $x_1 = 3$ and $x_2 = 2$. That is,

$$3\begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + 2\begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

Vector Equation

A vector equation

$$x_1 \boldsymbol{a_1} + x_2 \boldsymbol{a_2} + \dots + x_n \boldsymbol{a_n} = \boldsymbol{b}$$

has the same solution set as the linear system whose augmented matrix is

$$[a_1 \ a_2 \ \dots \ a_n \ b]$$

• In particular, b can be generated by a linear combination of a_1, a_2, \ldots, a_n if and only if there exists a solution to the linear system corresponding to the augmented matrix.

An Important Note

A Key Idea

One of the key ideas in linear algebra is to study the set of all vectors that can be generated or written as a linear combination of a fixed set $\{v_1, \dots, v_p\}$ of vectors.

Definition: Span

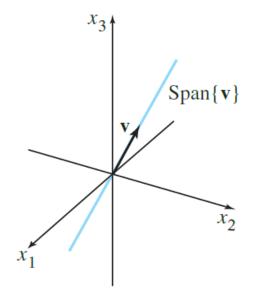
If v_1, \ldots, v_p are in \mathbb{R}^n , then the set of all linear combinations of v_1, \ldots, v_p is denoted by $Span\{v_1, \ldots, v_p\}$ and is called the subset of \mathbb{R}^n spanned (or generated) by $\{v_1, \ldots, v_p\}$.

That is, $Span\{v_1, \ldots, v_p\}$ is the collection of all vectors that can be written in the form

$$c_1 \boldsymbol{v_1} + c_2 \boldsymbol{v_2} + \ldots + c_p \boldsymbol{v_p}$$

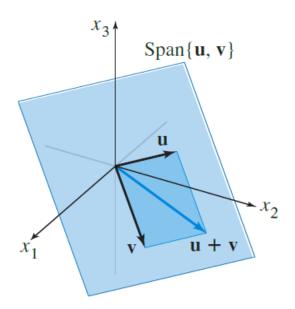
with c_1, \ldots, c_p scalars.

Geometric Description of Span



(a) $\mathsf{Span}\{v\}$ as a line through the origin

Geometric Description of Span



(b) $\mathsf{Span}\{\mathsf{u},\mathsf{v}\}$ as a plane through u,v and the origin

Example

Let
$$a_1 = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$$
, $a_2 = \begin{bmatrix} 5 \\ -13 \\ -3 \end{bmatrix}$ and $b = \begin{bmatrix} -3 \\ 8 \\ 1 \end{bmatrix}$

- Then $Span\{a_1, a_2\}$ is a plane through the origin in \mathbb{R}^3 .
- Is **b** in that plane?

Solution:

· We need to find solution of the following equation

$$x_1 \boldsymbol{a}_1 + x_2 \boldsymbol{a}_2 = \boldsymbol{b}$$

- Solve augmented matrix and check whether solution exists or not.
- If there is no solution, then b is not in $Span\{a_1, a_2\}$.

Linear Independence

We say that the vectors $v_1, v_2, ..., v_k$ are **linearly dependent** if at least one vector can be written as a linear combination of the others.

On the other hand, the vectors $v_1, v_2, ..., v_k$ are **linearly independent** if and only if $c_1v_1, c_2v_2, ..., c_kv_k = \mathbf{0}$

Only have trivial solution. Or, none of the vector can be written as a linear combination of the remaining vectors.

Linear Independence

Example:

The vectors

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{v}_k = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

Are linearly independent vector of \mathbb{R}^3 .

Vector Space

A **vector space** is a nonempty set V of objects, called *vectors*, on which are defined two operations, called *addition* and *multiplication* by *scalars* (real numbers), subject to the ten axioms (or rules) listed below. The axioms must hold for all vectors **u**, **v**, and **w** in V and for all scalars c and d.

- 1. The sum of \mathbf{u} and \mathbf{v} , denoted by $\mathbf{u}+\mathbf{v}$, is in V.
- 2. u+v=v+u.
- 3. (u+v)+w=u+(v+w).
- 4. There is a **zero** vector **0** in V such that **u+0=u**.
- 5. For each \mathbf{u} in \mathbf{V} , there is a vector $-\mathbf{u}$ in \mathbf{V} such that $\mathbf{u}+(-\mathbf{u})=\mathbf{0}$.
- 6. The scalar multiple of **u** by c, denoted by c**u**, is in V.
- 7. $c(\mathbf{u}+\mathbf{v})=c\mathbf{u}+c\mathbf{v}$.
- 8. (c+d)u=cu+du.
- 9. c(du)=(cd) u.
- 10. 1**u=u**.

Vector Space

Example:

The Euclidean space \mathbb{R}^m is an m dimensional vector space.

Subspace

Any subset W of a vector space V, which itself is a vector space, is called a **subspace** of V.

How to check for a subset W of a vector space V is a subspace?

- Any linear combination of any pair of vectors from the subset belongs to the subset as well. That is, if $v, u \in W$ then $av + bu \in W$.
- The set W should contain the zero vector.

Subspace

Examples:

- Given any set of vectors $v_1, ..., v_k$ in an m-dimensional vector space \mathbb{R}^m the span of k vectors is itself a vector space being a subspace of \mathbb{R}^m .
- The set $\left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$ is a subspace of \mathbb{R}^3 .
- Let v be a vector of a vector space V. The set $span\{v\}$ is a subspace of V which is the collection of all possible scalar multiples of v.

Definition

• If $span(v_1, ..., v_k) = \mathbb{R}^m$, then we say that $v_1, ..., v_k$ is a spanning set for \mathbb{R}^m .

Row and Column Space

Given the data matrix D

$$D = \begin{pmatrix} x_1 & x_{11} & x_{12} & \dots & x_{1d} \\ x_2 & x_{21} & x_{22} & \dots & x_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ x_n & x_{n1} & x_{n2} & \dots & x_{nd} \end{pmatrix}$$

We can also write D as

$$D = (x_1 \quad x_2 \quad ... \quad x_n)^T = (X_1 \quad X_2 \quad ... \quad X_d)$$

Row and Column Space

There are several interesting vector spaces associated with the data matrix D, two of which are the column space and row space of D.

The **column space** of D, denoted col(D), is the set of all linear combinations of the d attributes $X_i \in \mathbb{R}^n$.

That is,

$$col(D) = span(X_1, X_2, ..., X_d)$$

By definition col(D) is a subspace of \mathbb{R}^n .

The **row space** of D, denoted row(D), is the set of all linear combinations of the n points $x_i \in \mathbb{R}^d$, that is,

$$row(D) = span(\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_n)$$

By definition row(D) is a subspace of \mathbb{R}^d .

Basis and Dimension

- The **basis** of a vector space V is a set of vectors from V that are linearly independent, and they also span the entire vector space.
- In other words, basis are the minimal spanning set.
- If the vectors in the basis are pairwise orthogonal, they are called an orthogonal basis.
- An orthogonal basis with all vectors of length 1 is called an orthonormal basis.
- Any two basis of a vector space V must have the same number of vectors in it.
- The number of vectors in a basis of a vector space V is called the dimension of V.

Basis and Dimension

Example:

The standard basis of \mathbb{R}^3 is

$$\left\{ \begin{pmatrix} 1\\0\\0 \end{pmatrix}, \begin{pmatrix} 0\\1\\0 \end{pmatrix}, \begin{pmatrix} 0\\0\\1 \end{pmatrix} \right\}$$

Which is also an orthonormal basis of \mathbb{R}^3 .

Example:

Consider the matrix

$$\begin{pmatrix} 1 & 3 & 8 \\ 1 & 2 & 6 \\ 0 & 1 & 2 \end{pmatrix}$$

Find the basis for the column space and row space of the matrix.

Questions:

- What do we mean by Ax = b?
- What is the solution of Ax = b?
- Why do we care about the solution?
- How do we solve the system?
- What does it mean if there is no solution to the system?
- What is a linear combination?
- Can we interpret Ax = b in terms of a linear combination?
- In terms of linear combination, what is the meaning of Ax=b being consistent or being inconsistent?