

# Linear Algebra and Matrix Methods

## Vector Spaces, Subspaces

# From Ordered $n$ -tuples of Real Numbers to...

So far we have dealt with vectors of sequences of real numbers (with real scalar multiplication), and how can be used in solving **systems of linear equations**.

While  $\mathbb{R}^n$  is the most useful vector space we know about right now, there are many other mathematical objects that can be described as **vectors**. The spaces they exist in are called **vector spaces**.

# Abstract Vector Spaces

If the elements of a space  $V$  have the following properties, we call  $V$  a **real vector space**.<sup>1</sup>

- ▶ **linearity** over  $\mathbb{R}$ : The operations of **scalar multiplication** by real numbers and **vector addition** are **closed** in  $V$ . This means  $V$  contains all **linear combinations** of its elements:

$$v, w \in V \text{ and } a, b \in \mathbb{R} \implies av + bw \in V.$$

---

<sup>1</sup>We can replace the scalar field  $\mathbb{R}$  with  $\mathbb{C}$  for a **complex vector space**. This happens, for example, in quantum mechanics.

# Abstract Vector Spaces

- ▶ **vector addition is commutative:** If  $v, w \in V$ , then

$$v + w = w + v.$$

- ▶ **vector addition is associative:** If  $v, w, x \in V$ , then

$$v + (w + x) = (v + w) + x.$$

# Abstract Vector Spaces

- ▶ **scalar multiplication is associative**: if  $a, b \in \mathbb{R}$  and  $v \in V$ , then

$$a(bv) = (ab)v$$

- ▶ **scalar multiplication is distributive<sup>2</sup>**: if  $a \in \mathbb{R}$  and  $v, w \in V$ , then

$$a(v + w) = av + aw.$$

---

<sup>2</sup>We don't need to mention properties of real number arithmetic because we are focusing here on how scalars operate on vectors. Also, if we replace  $\mathbb{R}$  with  $\mathbb{C}$  as the scalar field, scalar multiplication is *not* commutative. As that is not one of the properties we require, complex vector spaces are OK.

# Abstract Vector Spaces

- ▶ **scalar multiplicative identity:** The real number 1 acts as a multiplicative identity for scalars: for any  $v \in V$ ,

$$1v = v.$$

- ▶ **vector additive identity:** There exists a vector called **zero** (denoted 0) in  $V$  such that

$$\forall v \in V, \quad v + 0 = 0 + v = v.$$

# Abstract Vector Spaces

- ▶ **vector additive inverses:** For each  $v \in V$ , there exists a unique  $w \in V$ , called the **additive inverse** of  $v$ , such that

$$v + w = w + v = 0.$$

We denote  $w = -v$ , and call  $w$  “negative  $v$ ”.

Note that vector multiplication is *not* considered in the vector space properties.

# Some Examples of Real Vector Spaces

Certainly,  $\mathbb{R}^n$  is a real vector space, and the one we will focus on. However, there are other very important vector spaces that may not be as obvious at first glance.

$$\blacktriangleright F(\mathbb{R}^n, \mathbb{R}^m) = \{f \mid f : \mathbb{R}^n \rightarrow \mathbb{R}^m\}$$

The set of real-vector-valued functions of a real vector contains all linear combinations of functions.

Let  $f(x) = \sin(x)$  and  $g(x) = x^3 - 5x + 2$ . Then  $f, g \in F(\mathbb{R}, \mathbb{R})$ , and so we know that

$$h := 10f - 3.4g \in F(\mathbb{R}, \mathbb{R})$$

as well. In fact,

$$h(x) = 10 \sin(x) - 3.4x^3 + 17x - 6.8.$$



# Some Examples of Real Vector Spaces

- ▶  $M_{n \times n}(\mathbb{R}) = \{M \mid M \text{ is an } n \times n \text{ matrix with entries in } \mathbb{R}\}$

Yes, it may look strange to call matrices “vectors”, but the space of square matrices of the same size satisfies all the properties of a vector space.

$$A = \begin{pmatrix} 4 & 5 \\ -2 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 2 \\ 0 & 6 \end{pmatrix} \in M_{2 \times 2}(\mathbb{R})$$

# Vector Subspaces

A subset  $W \subseteq V$  is called a **vector subspace** of  $V$  if  $W$  *also* satisfies the linearity property of a vector space (closed under linear combinations). This means  $W$  is a vector space in its own right.

There are many subsets of  $V$  that are *not* subspaces of  $V$ .

# Vector Subspaces

Some observations:

- ▶ If  $W$  is a subspace of  $V$ , then  $0 \in W$ .
- ▶ The simplest subspaces of  $V$  to describe are  $V$  itself and the **trivial subspace**  $\{0\}$ . The rest are “in between”, subset-wise.

Note that the trivial subspace  $\{0\}$  is *not* the empty set  $\emptyset$ .

The empty set is *not* a vector space (why)?

# Vector Subspaces: Examples

Some examples of vector subspaces are:

- ▶  $C(\mathbb{R}) \subseteq F(\mathbb{R}, \mathbb{R})$ , the subspace of continuous real-valued functions, is a vector subspace of all real-valued functions.
- ▶ The space of  $n \times n$  *diagonal* matrices is a subspace of  $M_{n \times n}(\mathbb{R})$ .
- ▶ Any line through the origin is a vector subspace of  $\mathbb{R}^n$ .
- ▶ Any plane through the origin is a vector subspace of  $\mathbb{R}^n$ .

# Vector Subspaces of $\mathbb{R}^n$

Consider a plane through the origin of  $\mathbb{R}^3$ . This plane can be described as a two-dimensional subspace, since the plane can be represented by one equation, with fixed  $a, b, c \in \mathbb{R}$ ,  $a \neq 0$ :

$$ax_1 + bx_2 + cx_3 = 0.$$

This one equation only needs two variables to describe its set of vectors. Call this plane  $P$ . Then,

$$\begin{aligned} P &= \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid ax_1 + bx_2 + cx_3 = 0\} \\ &= \left\{ (x_1, x_2, x_3) \in \mathbb{R}^3 \mid x_1 = \frac{1}{a}(-bx_2 - cx_3) \right\} \\ &= \left\{ \left( \frac{1}{a}(-bx_2 - cx_3), x_2, x_3 \right) \in \mathbb{R}^3 \right\}. \end{aligned}$$

We will call  $x_1$  a **pivot variable**, and  $x_2$  and  $x_3$  **free variables**.

# Linear span of a set of vectors

Let  $S = \{v_1, v_2, \dots, v_m\} \subseteq V$  be a set of vectors.

The **(linear) span** of the set  $S$  is the set of all linear combinations of vectors in  $S$ :

$$\text{span}(S) := \{c_1 v_1 + c_2 v_2 + \cdots + c_m v_m \mid c_i \in \mathbb{R}, i = 1, 2, \dots, m\}.$$

Note that, since setting all  $c_i = 0$  is a possibility,  $0 \in \text{span}(S)$ .

$\text{span}(S)$  is a vector subspace of  $V$ .

This type of subspace is crucial to our analysis.

# A vector subspace is a linear span

Back to our example plane:

$$\begin{aligned} P &= \{(x_1, x_2, x_3) \in \mathbb{R}^3 \mid ax_1 + bx_2 + cx_3 = 0\} \\ &= \left\{ \left( \frac{1}{a}(-bx_2 - cx_3), x_2, x_3 \right) \in \mathbb{R}^3 \right\}. \end{aligned}$$

We can decompose this representation of vectors in  $P$  by rewriting per variable, getting  $P$  as the span of two vectors in  $\mathbb{R}^3$ . We call these two vectors **special solutions**.

$$\begin{aligned} P &= \left\{ x_2 \begin{pmatrix} -\frac{b}{a} \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -\frac{c}{a} \\ 0 \\ 1 \end{pmatrix} \mid x_2, x_3 \in \mathbb{R} \right\} \\ &= \text{span} \left( \left\{ \begin{pmatrix} -\frac{b}{a} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{c}{a} \\ 0 \\ 1 \end{pmatrix} \right\} \right). \end{aligned}$$

# Any vector (sub)space is a linear span

Seeing the original space  $\mathbb{R}^3$  in the same fashion, using the **standard basis vectors**, we have

$$\mathbb{R}^3 = \left\{ x_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \mid x_1, x_2, x_3 \in \mathbb{R} \right\}$$

$$= \text{span} \left( \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \right).$$

$$P = \left\{ x_2 \begin{pmatrix} -\frac{b}{a} \\ 1 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -\frac{c}{a} \\ 0 \\ 1 \end{pmatrix} \mid x_2, x_3 \in \mathbb{R} \right\}$$

$$= \text{span} \left( \left\{ \begin{pmatrix} -\frac{b}{a} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{c}{a} \\ 0 \\ 1 \end{pmatrix} \right\} \right).$$



# Column space of a matrix $A$

Consider a real-entry  $m \times n$  matrix  $A$ .

We can use the “column view” of  $A$  to write  $A$  as

$$A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix},$$

an ordered collection of  $n$  column vectors,  $a_1, a_2, \dots, a_n \in \mathbb{R}^m$ .

# Column space of a matrix $A$

The column view of  $A$  turns the matrix equation

$$A\vec{x} = b$$

with vector  $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$ , vector  $b \in \mathbb{R}^m$ , into

$$x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = b,$$

a linear combination equation.

Note that  $x_1, x_2, \dots, x_n \in \mathbb{R}$  are the scalars here.

## Column space of a matrix $A$ : a subspace of $\mathbb{R}^m$

When does  $A\vec{x} = b$  have a solution? Precisely when

$$x_1 a_1 + x_2 a_2 + \cdots + x_n a_n = b.$$

In other words, if  $b$  is a linear combination of the columns of  $A$ , then  $\vec{x}$  is a solution to  $A\vec{x} = b$ .

# Column space of a matrix $A$ : a subspace of $\mathbb{R}^m$

Define the **column space**  $C(A)$  of the matrix  $A$  as the set of all linear combinations of the columns of  $A$ ; that is,

$$C(A) := \text{span}(\{a_1, a_2, \dots, a_n\}).$$

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

# Whither solutions $\vec{x}$ of $A\vec{x} = b$ ?

If  $b \in C(A)$ , then  $A\vec{x} = b$  has a solution  $\vec{x}$ .

We'll restate the structure of  $C(A)$ :

$$C(A) = \{b \in \mathbb{R}^m \mid \exists \vec{x} \in \mathbb{R}^n : A\vec{x} = b\}.$$

Certainly, the zero vector<sup>3</sup> with  $m$  zeroes  $0 = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in C(A)$ .

---

<sup>3</sup>By now you should be comfortable with the fact that the symbol  $0$  will be used to represent any size zero vector, and  $I$  any size identity matrix, that context demands.

First, check  $A\vec{x} = 0$ : Null space of  $A$ : a subspace of  $\mathbb{R}^n$

The special case  $b = 0 \in \mathbb{R}^m$  will be especially helpful, as it helps define a complementary subspace.

First, note that the equation  $A\vec{x} = 0$  always has a solution:  $\vec{x} = 0$ .

Is this the *only* solution, or are there *infinitely many more*?

First, check  $A\vec{x} = 0$ : Null space of  $A$ : a subspace of  $\mathbb{R}^n$

Define the **null space**  $N(A)$  of the matrix  $A$  as the set of all solutions  $\vec{x}$  to the *null equation*  $A\vec{x} = 0$ :

$$N(A) := \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = 0\}.$$

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

## A solution of $A\vec{x} = b$ is a sum

If  $\vec{x}_n \in N(A)$ , then  $A\vec{x}_n = 0$ .

If  $\vec{x}_p$  solves  $A\vec{x} = b$ , we will call  $\vec{x}_p$  a **particular solution** of the equation.

By linearity, we have

$$A(\vec{x}_n + \vec{x}_p) = A\vec{x}_n + A\vec{x}_p = 0 + b = b.$$

Hence,  $\vec{x}_n + \vec{x}_p$  is *also* a solution to  $A\vec{x} = b$ .

To find *all* solutions to  $A\vec{x} = b$ , we need to solve  $A\vec{x} = 0$ .



To solve  $A\vec{x} = 0$ :

What are the solutions to  $A\vec{x} = 0$ ?

We will solve using the same augmented matrix technique we used to solve square systems of equations, generalizing to a matrix  $A$  of any rectangular shape. Start with

$$[A \mid 0].$$

Use downward, forward elimination, starting in the upper-left corner. Pivot on the diagonal. If you ever zero out a diagonal entry, pivot to the right of that diagonal and continue.

If you zero out any of the bottom rows, that's okay.

## To solve $A\vec{x} = 0$ : Reduced row echelon form

Then, scale the pivots to 1, and back substitute to get as close as possible to the identity matrix.

The bottom left portion of the matrix should only have pivot 1s, and 0s below. The upper right portion ideally has only 0s above pivots; some nonzero entries above pivots may be unavoidable.<sup>4</sup>

The goal is a form called **reduced row echelon form (rref)**, which has as-diagonal-as-possible 1 entries (still) called **pivots**.

The variable on any column with a pivot is called a **pivot variable**; any other column has a **free variable**.

---

<sup>4</sup>We will see several numerical examples in class.

We will call the resulting matrix  $R = rref(A)$ . The number of pivots in  $R$  is called the **rank** of the matrix  $A$ , and denoted by  $r$ .

Note that  $\vec{x}$  solves  $A\vec{x} = 0$  if and only if  $\vec{x}$  also solves  $R\vec{x} = 0$ .

$$A\vec{x} = 0 \iff R\vec{x} = 0. \therefore N(R) = N(A).$$

To construct solutions  $\vec{x}$  to  $R\vec{x} = 0$ , solve each pivot variable as a function of the free variables.

Looking back at our  $1 \times 3$  system  $ax_1 + bx_2 + cx_3 = 0$ :

$$P = \left\{ \begin{pmatrix} -\frac{b}{a}x_2 + -\frac{c}{a}x_3 \\ x_2 \\ x_3 \end{pmatrix} \mid x_2, x_3 \in \mathbb{R} \right\}.$$

There is  $r = 1$  pivot variable and  $n - r = 2$  free variables.

# Nullity, Special Solutions

For each free variable in  $\vec{x}$  solving  $R\vec{x} = 0$ , set that free variable to 1 and the other free variables to 0. The resulting vector (which has only numbers in it) is called a **special solution** of  $A\vec{x} = 0$ .

Looking back at our example of a plane in space:

$$P = \text{span} \left( \left\{ \begin{pmatrix} -\frac{b}{a} \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -\frac{c}{a} \\ 0 \\ 1 \end{pmatrix} \right\} \right).$$

Let  $S$  be the set of special solutions of  $A$ . Then  $|S| = n - r$  and

$$N(A) = \text{span}(S).$$

If  $S = \emptyset$ , then  $N(A) = \{0\}$ . We call  $n - r$  the **nullity** of  $A$ .

# Cases on the dimensions of $A$

We relate the rank  $r$  to the dimensions  $m \times n$  of  $A$ :

- ▶  $1 \leq r \leq \min(m, n)$
- ▶  $C(A)$  is spanned by  $r$  vectors;  $C(A) = \mathbb{R}^m$  if  $r = m$
- ▶  $R$  has  $n - r$  zero rows if  $r < n$
- ▶  $N(A)$  is the span of  $n - r$  special solutions;  
 $N(A) = \{0\}$  if  $r = n$

# Cases on the dimensions of $A$

- ▶  $A\vec{x} = b$ ,  $b \neq 0$ , has
  - ▶ no solutions if  $r < m$  and “ $0 = 1$ ” occurs ( $b \notin C(A)$ )  
(**overdetermined**: a contradiction exists in the equations)
  - ▶ a unique solution if  $r = m = n$  (and so  $A^{-1}$  exists)  
(one equation per variable in the solution:  $x_i = c_i$ ,  $i = 1, \dots, r$ )
  - ▶ infinitely many solutions if  $r < n$  and only “ $0 = 0$ ” occurs  
(**underdetermined**: not enough pivots to fill all variables)

If  $r = m$ , we say  $A$  has **full row rank**, since all rows have pivots.

Similarly, if  $r = n$ , we say  $A$  has **full column rank**.

# Linear Independence

A set of vectors are called **linearly dependent** if there is some nonzero linear combination of them that makes the zero vector. This means the vectors all share a subspace.

$$\{v_1, \dots, v_k\} \text{ dependent} \iff \exists c_1, \dots, c_k \in \mathbb{R}, \text{ not all } 0 : \\ c_1 v_1 + \dots + c_k v_k = 0.$$

- ▶ Two vectors on the same line are linearly dependent.
- ▶ Three vectors in the same plane are linearly dependent.

If the only linear combination of a set of vectors that makes 0 is 0 of each of them, we call the set **linearly independent**.



# Dimension of a Vector Space

The **dimension** of a vector space  $V$  is the maximum number of linearly independent vectors from  $V$  that span all of  $V$ .

We will denote the dimension of the vector space  $V$  by  $\dim(V)$ .

If  $S$  is a set of vectors and  $A$  is the matrix using the elements of  $S$  as columns, then

$$S \text{ is a linearly independent set} \iff N(A) = \{0\}.$$

If  $S \subseteq \mathbb{R}^m$ , then  $S$  cannot be a linearly independent set if  $|S| > m$ .

## Dimension of $\mathbb{R}^m$ is $\dim(\mathbb{R}^m) = m$

$S$  is called a **maximally linearly independent set**, or a **basis**, of  $V$  if  $S$  is a linearly independent set and  $\text{span}(S) = V$ .

In this case,  $\dim(V) = |S|$ .

The **standard basis** of  $\mathbb{R}^m$  is the set of  $m$  orthonormal vectors with 1 in one coordinate and 0 elsewhere:

$$\left\{ e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_m = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix} \right\}.$$

Thus,  $\dim(\mathbb{R}^m) = m$ , and any *other* basis will also have  $m$  vectors.

# A matrix of basis columns is invertible

If  $S$  is a basis of  $\mathbb{R}^m$ , and  $A$  is a matrix whose columns are the elements of  $S$ , then  $A$  is square and invertible.

The order of the columns does not matter; this is always true. (Although, our rref technique may require some initial transformations to get pivots in the “correct” places.)

There are infinitely many bases of  $\mathbb{R}^m$ , and all contain exactly  $m$  vectors. The vectors need only be linearly independent, not necessarily orthogonal, or unit length.

# The Four Fundamental Subspaces of a Matrix $A$

Recall that the **transpose** of an  $m \times n$  matrix  $A$ , denoted  $A^t$ , is the  $n \times m$  matrix where the rows of  $A$  are the columns of  $A^t$ .

We have seen two vector spaces related to the matrix  $A$ :

- ▶ the **column space**  $C(A)$  (also called the **image** of  $A$ ), consisting of all  $b \in \mathbb{R}^m$  that allow  $A\vec{x} = b$  to be solved, and
- ▶ the **null space**  $N(A)$  (also called the **kernel** of  $A$ ), consisting of all  $\vec{x} \in \mathbb{R}^n$  solving  $A\vec{x} = 0$ .

# The Four Fundamental Subspaces of a Matrix $A$

The transpose  $A^t$  also has a column space and a null space; in reference to the original matrix  $A$ , we call them

- ▶ the **row space**  $C(A^t)$  (also called the **coimage** of  $A$ ), consisting of all  $c \in \mathbb{R}^n$  that allow  $A^t \vec{y} = c$  to be solved, and
- ▶ the **left null space**  $N(A^t)$  (also called the **cokernel** of  $A$ ), consisting of all  $\vec{y} \in \mathbb{R}^m$  solving  $A^t \vec{y} = 0$ .

These can both be stated in terms of the matrix  $A$  with left-multiplication by the vector  $\vec{y}^t$  by applying the transpose:

$$A^t \vec{y} = c \iff \vec{y}^t A = c^t.$$

# The Four Fundamental Subspaces of a Matrix $A$ : Why?

Why mention  $A^t$  in describing how to solve systems using  $A$ ?  
Because the two systems

$$A\vec{x} = b \text{ and } A^t\vec{y} = c$$

are deeply related; in particular,

$$\vec{x}^t c = b^t \vec{y}, \text{ or, equivalently, } c^t \vec{x} = \vec{y}^t b.$$

Notice that  $\vec{x}, c \in \mathbb{R}^n$  but  $b, \vec{y} \in \mathbb{R}^m$ . In dot product notation,

$$\vec{x} \cdot c = b \cdot \vec{y}.$$

# Sum of Two Vector Subspaces

Let  $V$  and  $W$  be two vector spaces.

Define the **sum** of  $V$  and  $W$  to be the vector space of all the sums of one element of  $V$  and one element of  $W$ :

$$V + W = \{v + w \mid v \in V, w \in W\}.$$

The sum makes linear combinations of elements of  $V$  and  $W$  to form a (larger) vector space (if neither contains the other).

If  $V$  and  $W$  have nontrivial overlap (the zero vector is considered trivial), then there are multiple ways to represent vectors in the intersection as a sum of vectors from each space.

# Sum of Two Vector Subspaces

## Example

Let  $V$  and  $W$  be two plane subspaces of  $\mathbb{R}^3$  defined by

$$V = \text{span}(\{e_1, e_2\}), \quad W = \text{span}(\{e_1, e_3\}),$$

where  $\{e_1, e_2, e_3\}$  is the standard basis of  $\mathbb{R}^3$ .

Then any vector in  $V + W$  that has an element of the line intersection

$$V \cap W = \text{span}(\{e_1\})$$

can be written in two ways inside  $V + W$ :



# Sum of Two Vector Subspaces: non-unique representation

$v_1 = ae_1 + be_2 \in V$ ,  $w_1 = ce_3 \in W$ :

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = ae_1 + be_2 + ce_3 = v_1 + w_1 = \begin{pmatrix} a \\ b \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ c \end{pmatrix} ;$$

$v_2 = be_2 \in V$ ,  $w_2 = ae_1 + ce_3 \in W$ :

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = ae_1 + be_2 + ce_3 = v_2 + w_2 = \begin{pmatrix} 0 \\ b \\ 0 \end{pmatrix} + \begin{pmatrix} a \\ 0 \\ c \end{pmatrix} .$$

# Direct Sum of Two Independent Vector Subspaces

Let  $V$  and  $W$  be two vector subspaces of the same space  $Z$  such that  $V \cap W = \{0\}$ .

Then any vector  $z \in Z$  has a *unique* representation as a sum  $z = v + w$  with  $v \in V$  and  $w \in W$ . When this is the case, we call  $V + W$  the **direct sum** of  $V$  and  $W$ , and denote it  $V \oplus W$ .

# Direct Sum: unique representation

## Example

Let  $V$  and  $W$  be two line subspaces of  $\mathbb{R}^4$  defined by

$$V = \text{span}(\{e_1 + e_3\}), \quad W = \text{span}(\{e_3 + e_4\}).$$

Then any vector in  $z \in V \oplus W$  has only one representation:

$$z = \begin{pmatrix} a \\ 0 \\ a + b \\ b \end{pmatrix} = v + w : v = a(e_1 + e_3), \quad w = b(e_3 + e_4).$$

# Orthogonal Vector Subspaces

Let  $V$  and  $W$  be two vector subspaces of the same space  $Z$ .

We call  $V$  and  $W$  **orthogonal** if every pair of vectors  $(v, w) \in V \times W$  are orthogonal:

$$V \perp W \iff \forall v \in V, w \in W : v \perp w.$$

If  $V \perp W$ , then necessarily we have  $V \cap W = \{0\}$ .

Note that two vector subspaces may have a direct sum  $V \oplus W$  whether or not they are orthogonal.

# Orthogonal Direct Sum: unique orthogonal representation

## Example

Let  $V$  and  $W$  be two line subspaces of  $\mathbb{R}^4$  defined by

$$V = \text{span}(\{e_1 + e_3\}), \quad W = \text{span}(\{e_2 + e_4\}).$$

Then any vector in  $z \in V \oplus W$  has only one representation:

$$z = \begin{pmatrix} a \\ b \\ a \\ b \end{pmatrix} = v + w : v = a(e_1 + e_3), \quad w = b(e_2 + e_4).$$

Furthermore,  $v \perp w$ .

# The Fundamental Theorem of Linear Algebra, Part I

If the function  $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$  is an  $m \times n$  matrix with rank  $r$ , then:

- ▶ the function  $A^t : \mathbb{R}^m \rightarrow \mathbb{R}^n$  is an  $n \times m$  matrix (obviously),
- ▶  $\dim(C(A)) = \dim(C(A^t)) = r$ ,
- ▶  $\dim(N(A)) = n - r$  and  $\dim(N(A^t)) = m - r$ ,
- ▶  $C(A) \oplus N(A^t) = \mathbb{R}^m$ ,
- ▶  $C(A^t) \oplus N(A) = \mathbb{R}^n$ ,
- ▶  $C(A) \perp N(A^t)$  and  $C(A^t) \perp N(A)$ .

The two direct sums above are orthogonal.

# FTLA: Consequences

Thus, every vector  $v \in \mathbb{R}^n$  has a unique representation as a sum

$$v = \vec{x}_n + c,$$

where  $\vec{x}_n \in N(A)$ ,  $c \in C(A^t)$ , and  $\vec{x}_n \perp c$ .

Likewise, every vector  $z \in \mathbb{R}^m$  has a unique representation as a sum

$$z = \vec{y}_n + b,$$

where  $\vec{y}_n \in N(A^t)$ ,  $b \in C(A)$ , and  $\vec{y}_n \perp b$ .

Note that, if  $R = rref(A)$ , then

$$N(A) = N(R), \text{ and } C(A^t) = C(R^t), \text{ but } C(A) \neq C(R).$$

# Multiplying by a matrix might mean information loss

If we apply  $A$  to  $v = \vec{x}_n + c$ , we get

$$Av = A(\vec{x}_n + c) = A\vec{x}_n + Ac = Ac,$$

since  $\vec{x}_n \in N(A)$ , so  $A\vec{x}_n = 0$ .

If  $\vec{x}_n \neq 0$ , then  $\dim(N(A)) > 0$ ,  $A$  is not invertible, and there is a kind of “information loss” when applying  $A$ : we move from a point in an  $m$ -dimensional space,  $v \in \mathbb{R}^m$ , to a point in an  $r$ -dimensional space,  $Av \in C(A)$ , with  $r < m$ .

We will explore this notion in the next chapter, under the concept of **projection**.