Applications

- Digital signal processing
  - filtering
  - moving averages
  - auralization

- Communication Systems
  - Error detection and correction in encoded messages using check matrices & linear codes

Topics

- Vector spaces (ALA 2.1)
  - Defn, examples, properties

- Subspaces (ALA 2.2)
  - Defn, examples, properties

- Span & Linear Independence (ALA 2.3)
  - linear combinations & span
  - linear dependence & independence
  - checking via solution to a linear system

- Basis & Dimension (ALA 2.4)
  - coordinate systems
  - Change of basis
  - dimension of a vector space
  - all finite dimensional vector spaces "look like" $\mathbb{R}^n$

Abstraction in Mathematics

A theme in mathematics is recognizing that seemingly unrelated settings, objects, or models all share common properties. By viewing them at the right level of abstraction, they can all be reasoned about together in the same way. This is a very powerful tool!

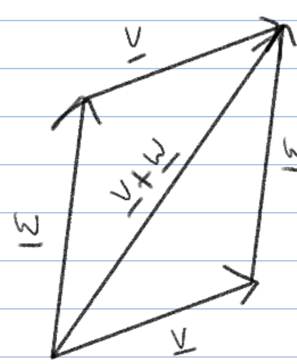
This lecture and the next will introduce the abstract notion of a vector space which unifies the seemingly disparate spaces of ordinary vectors, spaces of functions (e.g., polynomials, exponentials, trigonometric), spaces of matrices, and even (infinite dimensional) linear operators (we will see these later in the course), and more, under a common conceptual framework.

For many of you, this will be your first foray into "abstraction," and it will take some time and effort to get used to these ideas. We suggest making sure that you understand new concepts in the context of ordinary Euclidean Space, and then work through how they apply to more abstract spaces, such as the space of polynomials, vector valued sampled signals over an interval, or symmetric matrices (yes, we will see that these are all examples of vector spaces!)

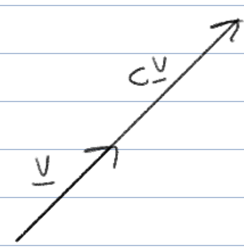
Real Vector Spaces

Our previous developments rested on certain simple & intuitive algebraic properties of how matrices & vectors can be added together and scaled.

Let us consider now the space of all $n \times 1$ real-valued vectors, denoted by $\mathbb{R}^n$. You have seen in previous classes that adding two vectors $\underline{v}, \underline{w} \in \mathbb{R}^n$ can be viewed geometrically through a parallelogram, and that scalar multiplication of $\underline{v} \in \mathbb{R}^n$ by a scalar $c \in \mathbb{R}$ is a stretch/shrinking of $\underline{v}$ by factor c:



$\underline{v}+\underline{w}$



$c\underline{v}$ $(c>1)$

Our goal is to extract out, or abstract, these properties, so that generic "vectors" living in a "vector space" behave in the same way when we add them or multiply them by a scalar.

Definition 2.1. A *vector space* is a set V equipped with two operations:

(*i*) *Addition*: adding any pair of vectors $\mathbf{v}, \mathbf{w} \in V$ produces another vector $\mathbf{v} + \mathbf{w} \in V$;

(*ii*) *Scalar Multiplication*: multiplying a vector $\mathbf{v} \in V$ by a scalar $c \in \mathbb{R}$ produces a vector $c\mathbf{v} \in V$.

These are subject to the following axioms, valid for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and all scalars $c, d \in \mathbb{R}$:

(a) *Commutativity of Addition*: $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$.

(b) *Associativity of Addition*: $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$.

(c) *Additive Identity*: There is a zero element $\mathbf{0} \in V$ satisfying $\mathbf{v} + \mathbf{0} = \mathbf{v} = \mathbf{0} + \mathbf{v}$.

(d) *Additive Inverse*: For each $\mathbf{v} \in V$ there is an element $-\mathbf{v} \in V$ such that $\mathbf{v} + (-\mathbf{v}) = \mathbf{0} = (-\mathbf{v}) + \mathbf{v}$.

(e) *Distributivity*: $(c + d)\mathbf{v} = (c\mathbf{v}) + (d\mathbf{v})$, and $c(\mathbf{v} + \mathbf{w}) = (c\mathbf{v}) + (c\mathbf{w})$.

(f) *Associativity of Scalar Multiplication*: $c(d\mathbf{v}) = (cd)\mathbf{v}$.

(g) *Unit for Scalar Multiplication*: the scalar $1 \in \mathbb{R}$ satisfies $1\mathbf{v} = \mathbf{v}$.

The operations (i) & (ii) just tell us that if we start with vectors $\underline{v}, \underline{w} \in V$ and real scalars $c, d \in \mathbb{R}$, we are free to add scaled versions together & we will stay in V, i.e., $c\underline{v} + d\underline{w} \in V$ for any choices of $c, d, \underline{v}, \underline{w}$.

The axioms that follow are a formalization of properties we expect addition and multiplication to follow: these are true with ordinary numbers and ordinary vectors, and we want them to hold for generic vectors too. We will work through some familiar and not so familiar examples soon, but end with some additional important properties that can be deduced from the axioms (a)–(g).

(P1) The $\underline{0}$ vector in Axiom (c) is unique

(P2) The additive inverse $-\underline{v}$ in Axiom (d), called the negative of $\underline{v}$, is unique.

(P3) $0\underline{v} = \underline{0}$ [NOTE: 0 on LHS is a scalar, $\underline{0}$ on RHS is the $\underline{0}$ vector these are different!]

(P4) $-1(\underline{v}) = -\underline{v}$

(P5) $c\underline{0} = \underline{0}$ for any $c \in \mathbb{R}$

(P6) If $c\underline{v} = \underline{0}$ then either $c = 0$ or $\underline{v} = \underline{0}$.

Note that these are all properties that obviously hold for ordinary numbers and ordinary vectors. This tells us they also hold for our new abstract vectors.

Example 1: Euclidean Space $\mathbb{R}^n$

The prototypical example is $\mathbb{R}^n$, the space of n dim. column vectors:

$$\underline{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}, \quad \underline{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}.$$

Vector addition & scalar multiplication are defined as usual ($c \in \mathbb{R}$):

$$\underline{v} + \underline{w} = \begin{bmatrix} v_1 + w_1 \\ v_2 + w_2 \\ \vdots \\ v_n + w_n \end{bmatrix}, \quad c\,\underline{v} = \begin{bmatrix} cv_1 \\ cv_2 \\ \vdots \\ cv_n \end{bmatrix}.$$

The zero vector is $\underline{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$.

The axioms of Def'n 2.1 are satisfied by virtue of the laws of vector addition and scalar multiplication.

Tip: whenever you find yourself confused about vector spaces, try to imagine a familiar space like $\mathbb{R}^3$. We are using the same geometric ideas, but now the arrows (vectors) are encodings of different objects that we want to scale and add together.

Example 2: Space of mxn real matrices $\mathbb{R}^{m\times n}$

A new, but not so different example, is the space of mxn matrices, denoted by $\mathbb{R}^{m\times n}$ (the superscript mxn here denotes the matrix dimension)

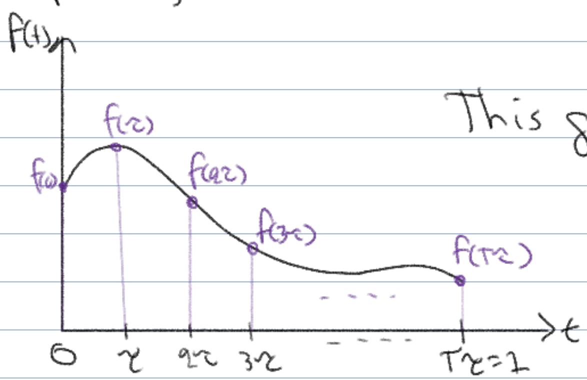
$\mathbb{R}^{m\times n}$ forms a vector space under the laws of matrix addition and scalar multiplication (the $\underline{0}$ vector is the all zeros matrix $0 \in \mathbb{R}^{m\times n}$).

WARNING: Matrix-matrix multiplication is NOT a vector space operation. Vector spaces only allow scalar multiplication and addition.

Notice that the vector space $\mathbb{R}^n$ is really $\mathbb{R}^{n\times 1}$ (nx1 matrices, or column vectors), so we really could've just started with this example.

Example 3: Sampled functions over an interval

In digital signal processing, we work with sampled versions of functions $f(t)$ over a time interval, say $[0,1]$, because we need to store these on a computer. This is obtained by sampling $f(t)$ at times $\{0, \tau, 2\tau, \dots, T\tau = 1\}$ (here τ is our sampling period, and we assume that $\frac{1}{\tau} = T$ is the # of samples taken over $[0,1]$).



This gives us a vector of size $T+1$ defined as

$$\underline{f} = \begin{bmatrix} f(0) \\ f(\tau) \\ f(2\tau) \\ \vdots \\ f(T\tau) \end{bmatrix}$$

which is nothing but an ordinary vector of size $T+1$. Adding sampled versions of two functions $f(t)$ and $g(t)$ is the same as adding their vectors $\underline{f} + \underline{g}$.

Similarly scaling the sampled version of $f(t)$ by $c \in \mathbb{R}$ is the same as computing $c\underline{f}$ in the usual way. Sampling the zero function $z(t) = 0$ gives the usual zero vector $\underline{0}$.

We have just argued that the space of functions sampled at the same points over an interval is not only a vector space, but is in fact $\mathbb{R}^{T+1}$!

Example 4: Doubly infinite sequences of numbers $\mathbb{S}$

Let $\mathbb{S}$ be the space of all doubly infinite sequences of numbers:

$$\{y_k\} = (\dots, y_{-2}, y_{-1}, y_0, y_1, y_2, \dots)$$

This space arises in many areas of engineering whenever a signal is sampled over an undefined interval: this happens in control theory, signal processing, biology, and optics. We will also call $\mathbb{S}$ the space of discrete time signals.

If we define addition as $\{y_k\} + \{z_k\} = \{y_k + z_k\}$ (element wise) and scalar multiplication as $c\{y_k\} = \{cy_k\}$ (scale each entry) then the vector space axioms can be verified exactly as we did for $\mathbb{R}^n$.

This is our first example of a vector space where the vectors are "not just an arrow in $\mathbb{R}^n$." In fact, each vector in $\mathbb{S}$ has infinitely many elements! Nevertheless, we can still think of each vector $\{y_k\} \in \mathbb{S}$ as an "arrow" that adds according to the parallelogram picture above & gets stretched/shrunk when multiplied by a scalar.

Example 5: Real polynomials of degree $\leq n$ $P^{(n)}$

Let's venture further into unfamiliar territory! Consider the space:

$$P^{(n)} = \{p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0\}$$

consisting of all real polynomials of degree $\leq n$. In the above, the polynomial coefficients $a_0, \ldots, a_n$ are allowed to be any real numbers. For example,

$$P^{(1)} = \{p(x) = a_1 x + a_0\}$$

is the set of all linear polynomials, since given any linear polynomial $q(x) = mx+b$, setting $a_1 = m$ and $a_0 = b$ shows that $q(x) \in P^{(1)}$.

We claim that under the usual definition of polynomial addition:

$$p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0, \quad q(x) = b_n x^n + b_{n-1}x^{n-1} + \cdots + b_1 x + b_0$$

$$\begin{aligned} p(x) + q(x) &= (a_n + b_n)x^n + (a_{n-1}+b_{n-1})x^{n-1} + \cdots + (a_1+b_1)x + (a_0+b_0) \\ &= d_n x^n + d_{n-1}x^{n-1} + \cdots + d_1 x + d_0 \end{aligned}$$

(we defined $d_i = a_i + b_i$ in the above)

and scalar multiplication:

$$\begin{aligned} cp(x) &= ca_n x^n + ca_{n-1}x^{n-1} + \cdots + ca_1 x + ca_0 \qquad (\tilde{a}_i = ca_i) \\ &= \tilde{a}_n x^n + \tilde{a}_{n-1}x^{n-1} + \cdots + \tilde{a}_1 x + \tilde{a}_0 \end{aligned}$$

that $P^{(n)}$ is a vector space. The axioms of addition and scalar multiplications can be checked to be satisfied because we add coefficients elementwise, and scaling a polynomial by c is done by scaling each coefficient.

NOTE: this is very similar to how we add & scale ordinary vectors, this is NOT a coincidence at all! We'll see why later!

What is the zero vector here? It is the zero polynomial satisfying $a_0 = a_1 = \cdots = a_{n-1} = a_n = 0$. Vectors in $P^{(n)}$ are polynomial functions: you should think of these as "arrows", like ordinary vectors, living in the space of polynomials. We'll see soon that we can make this picture precise.

WARNING: The space of n-degree polynomials is NOT a vector space. For example, consider $p(x) = x^2+1$ and $q(x) = -x^2+x$, both of degree 2. Their sum, $p(x)+q(x) = x^2+1+(-x^2+x) = x+1$ has degree 1! This is why $P^{(n)}$ is defined as polynomials of degree $\leq n$, because we can't "leave" the space when we add elements.

WARNING: Even though they look similar, a scalar c and a constant polynomial $a_0 \in \mathbb{P}^{(n)}$ are different objects! You should think of a_0 as an "arrow" in $\mathbb{P}^{(n)}$, and not a real number. This takes some time to get used to.

Example 6: Real valued functions over an interval $\mathcal{F}(I)$

Our last example will be the most abstract example of a vector space we see today.

Let $I \subset \mathbb{R}$ be an interval (a common choice is $[0,1]$, the closed interval from 0 to 1). Let us define the function space $\mathcal{F}(I)$ whose elements are all real-valued functions $f(x)$ defined for all $x \in I$.

This is a very complicated and big collection of objects. Fortunately, it has the structure of a vector space!

Define • addition in the usual way: $(f+g)(x) = f(x)+g(x)$ for all $x \in I$
• Scalar mult. in the usual way: $(cf)(x) = c\,f(x)$

Note: $(f+g)$ denotes the "new" function obtained by adding f and g. We define the value of $f+g$ at $x \in I$ by $(f+g)(x) = f(x)+g(x)$. This is tricky! The vectors here are $(f+g)$, f, and g. The variable x is not related to the vector space: it is only used to define how $(f+g)$ is computed from f and g. A useful trick to remind yourself of this is to not write out the argument x when performing vector space operations.

Let's see a concrete example:

Let $f(x) = 1+\sin 2x$ and $g(x) = 2+0.5x$, & set $I = [0,1]$. Then $f, g \in \mathcal{F}(I)$.

To compute $f+g$, we remember that $(f+g)(x) = f(x)+g(x)$, so that

$$\begin{aligned}(f+g)(x) &= f(x)+g(x)\\ &= 1+\sin 2x+2+0.5x\\ &= 3+0.5x+\sin 2x\end{aligned}$$

We see that $f+g$ is a new function that is defined for all $x \in [0,1]$. Therefore $f+g \in \mathcal{F}(I)$.

Tip: If this is confusing, pretend that we are using sampled versions $\underline{f}$ and $\underline{g}$ of $f(t)$ and $g(t)$. Adding $\underline{f}+\underline{g}$ gives me a new vector, just like adding $f+g$ gives me a new function.

Subspaces

Vector spaces arising in applications typically consist of an appropriate subset of vectors from some larger vector space.

These vector spaces "inside" of other vector spaces are called subspaces.

A subspace of a vector space V is a subset $W \subset V$ that is a vector space itself under the same rules for vector addition and scalar mult., and the same zero element.

There is a simple way to check if a subset $W \subset V$ of a vector space is a subspace: W needs to be non-empty and satisfy

$$c\underline{v} + d\underline{w} \in W \qquad \circledast$$

for any vectors $\underline{v}, \underline{w} \in W$ and any scalars $c, d \in \mathbb{R}$.

Subspaces are said to be closed under addition and scalar multiplication because they satisfy $(\circledast)$.

Subspaces must always contain the $\underline{0}$ vector (why?).

Example 1: Subspaces in $\mathbb{R}^3$

Let's work out all of the kinds of subspaces we can have in $\mathbb{R}^3$

a) The trivial subspace $W = \{\underline{0}\}$ consisting of only the $\underline{0}$ vector. We check that $c\underline{0} + d\underline{0} = \underline{0} \in W$ for any scalars $c, d \in \mathbb{R}$. This isn't very interesting, but it gives us our first example (a point at zero).

b) The entire space $\mathbb{R}^3$: clearly $c\underline{v} + d\underline{w} \in \mathbb{R}^3$ for any $\underline{v}, \underline{w} \in \mathbb{R}^3$ and $c, d \in \mathbb{R}$

c) The set of all vectors of the form $\begin{bmatrix} v \\ 2v \\ 3v \end{bmatrix}$. Pick two vectors $\underline{v} = \begin{bmatrix} v \\ 2v \\ 3v \end{bmatrix}$, $\underline{w} = \begin{bmatrix} w \\ 2w \\ 3w \end{bmatrix}$

and two scalars $c, d \in \mathbb{R}$. Then:

$$c\underline{v} + d\underline{w} = \begin{bmatrix} cv \\ 2cv \\ 3cv \end{bmatrix} + \begin{bmatrix} dw \\ 2dw \\ 3dw \end{bmatrix} = \begin{bmatrix} cv+dw \\ 2(cv+dw) \\ 3(cv+dw) \end{bmatrix} = \begin{bmatrix} x \\ 2x \\ 3x \end{bmatrix}$$

is also in our set. The subspace defined in this way is the line parallel to $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

d) The set of all vectors of the form $\begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$. Pick two vectors $\underline{v} = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}$, $\underline{w} = \begin{bmatrix} \hat{x} \\ \hat{y} \\ 0 \end{bmatrix}$ and any two scalars $c, d \in \mathbb{R}$. Then:

$$c\underline{v} + d\underline{w} = \begin{bmatrix} cx + d\hat{x} \\ cy + d\hat{y} \\ 0 \end{bmatrix} = \begin{bmatrix} \tilde{x} \\ \tilde{y} \\ 0 \end{bmatrix}$$

Is also in our set. This subspace is the xy-plane.

In general, subspaces of $\mathbb{R}^3$ are one of these 4 types: a point, a line, a plane, or all of $\mathbb{R}^3$. Key to all of these is that they must go through the origin! WARNING: lines & planes not passing through the origin are NOT subspaces!

Example 1(a): Subsets of $\mathbb{R}^3$ that are NOT subspaces

a) The set of vectors of the form $\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$, because the $\underline{0}$ vector is not in this set.

b) The non-negative orthant $\mathcal{O}^+ = \{x \geq 0, y \geq 0, z \geq 0\}$. Take any $\underline{x} \in \mathcal{O}^+$: then $-\underline{x} \notin \mathcal{O}^+$. So $\mathcal{O}^+$ is not closed under scalar multiplication.

c) The unit sphere $S^2 = \{x^2 + y^2 + z^2 = 1\}$ because $\underline{0} \notin S^2$. More generally, curved surfaces, like the paraboloid $P = \{z = x^2 + y^2\}$ are not subspaces (why not?)

Example 2: Subspaces of Discrete Time Signals

Here our base vector space is $\mathbb{S}$, the space of doubly infinite signals $\{y_k\} = (\ldots, y_{-2}, y_{-1}, y_0, y_1, y_2, \ldots)$.

a) The set $\mathbb{S}_{0:T}$ that are zero for all indices except $k = \{0, 1, 2, \ldots, T\}$: if $\{y_k\}, \{z_k\} \in \mathbb{S}_{0:T}$ then $c\{y_k\} + d\{z_k\} = \{cy_k + dz_k\}$ is also zero for all indices not in $\{0, 1, \ldots, T\}$ for any $c, d \in \mathbb{R}$.

b) The set Σ_0 of signals that sum to zero, i.e., $\{y_k\} \in \Sigma_0$ if and only if

$$\sum_{k=-\infty}^{\infty} y_k = 0$$

To check this, we compute the sum of $c\{y_k\} + d\{z_k\} = \{cy_k + dz_k\}$:

$$\sum_{k=-\infty}^{\infty} (cy_k + dz_k) = c\sum_{k=-\infty}^{\infty} y_k + d\sum_{k=-\infty}^{\infty} z_k = c \cdot 0 + d \cdot 0 = 0$$

since $\sum y_k = 0$ & $\sum z_k = 0$.

Example 3: Subspaces of Matrices

The following are easily verified to be subspaces of $\mathbb{R}^{n\times n}$.

a) The space Sym^n of symmetric matrices, i.e. $n\times n$ matrices M satisfying $m_{ij} = m_{ji}$ for all $i, j = 1, \dots, n$. For example, Sym^2 are matrices of the form

$$M = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

b) The space of diagonal matrices, i.e., $n\times n$ matrices M with $m_{ij} = 0$ if $i \neq j$.

c) The space of upper triangular matrices and the space of lower triangular matrices are both subspaces.

Example 4: Subspaces of function space $F(I)$

Here our base space is $F(I)$, the space of real-valued functions defined on the interval I.

a) $P^{(n)}$ the space of polynomials, defined over I. The 0 function is in $P^{(n)}$, and $P^{(n)}$ is a vector space contained in $F(I)$.

b) The space $C^0(I)$ of all continuous functions defined on I. Showing closure of this space relies on knowing that if $f(x)$ and $g(x)$ are continuous, then so is $cf(x) + dg(x)$ for any $c, d \in \mathbb{R}$, something you may have seen in Math 1400/1410.

Span & Linear Independence

A natural way of constructing a subspace is to start with some building blocks $\underline{v}_1, \dots, \underline{v}_k \in V$ from the vector space we are working in, and consider all possible linear combinations of them.

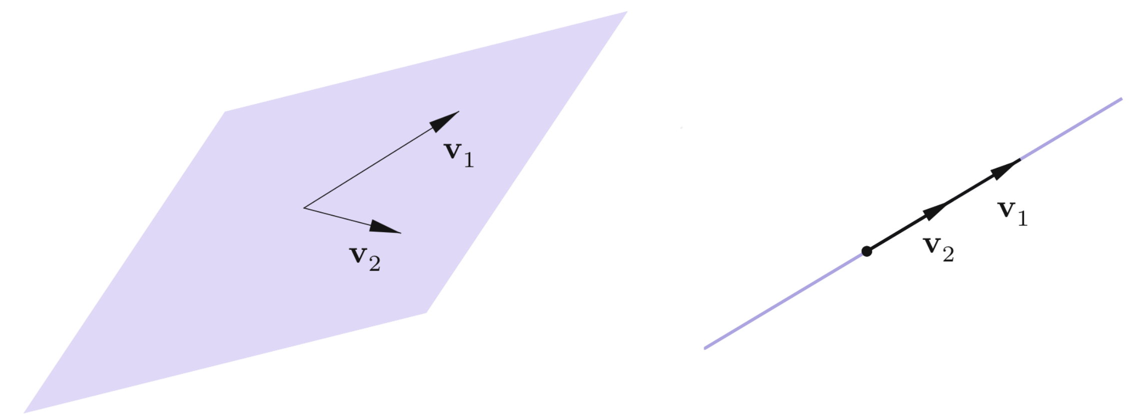
Let's make this idea formal: Let $\underline{v}_1, \dots, \underline{v}_k$ be elements of a vector space V. A sum of the form

$$c_1\underline{v}_1 + c_2\underline{v}_2 + \cdots + c_k\underline{v}_k = \sum_{i=1}^{k} c_i\underline{v}_i,$$

where the coefficients $c_1, c_2, \dots, c_k$ are any scalars, is a linear combination of $\{\underline{v}_1, \dots, \underline{v}_k\}$. Their span is the subset

$$W = \text{span}\{\underline{v}_1, \dots, \underline{v}_k\} \subset V$$

consisting of all possible linear combinations with scalars $c_1, \dots, c_k \in \mathbb{R}$.



A plane and a line spanned by two vectors. In the case of the line, $\underline{v}_1$ and $\underline{v}_2$ are parallel, so there is some scalar $c \in \mathbb{R}$ such that $\underline{v}_1 = c\underline{v}_2$.

Key fact: The span $W = \text{span}\{\underline{v}_1, \dots, \underline{v}_k\}$ of any finite collection of vector space elements $\underline{v}_1, \dots, \underline{v}_k \in V$ is a subspace of V.

This key fact is not hard to check using the properties of vector addition and scalar multiplication, but it is a surprisingly powerful tool for generating useful subspaces, and for checking if a vector $\underline{v} \in V$ also lives in the subspace W.

Example 1 If $\underline{v}_1$ and $\underline{v}_2$ are vectors in $\mathbb{R}^3$, then if $\underline{v}_1$ and $\underline{v}_2$ are not parallel, they define a plane composed of vectors of the form $c\underline{v}_1 + d\underline{v}_2$. If they are parallel, then they define a line composed of vectors of the form $c\underline{v}_1$. See the picture above.

Example 2 Let $W = \text{span}\{\underline{v}_1, \underline{v}_2\} \subset \mathbb{R}^3$ be the subspace spanned by

$$\underline{v}_1 = \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \quad \text{and} \quad \underline{v}_2 = \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix}.$$

Is the vector $\underline{v} = \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix}$ an element of W? If $\underline{v} \in W$ then we can find scalars $c_1, c_2 \in \mathbb{R}$ such that $\underline{v} = c_1 \underline{v}_1 + c_2 \underline{v}_2$, that is:

$$\begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} = c_1 \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ -3 \\ 1 \end{bmatrix} = \begin{bmatrix} c_1 + 2c_2 \\ -2c_1 - 3c_2 \\ c_1 + c_2 \end{bmatrix},$$

That is, c_1 and c_2 must satisfy the linear system:

$$\begin{aligned} c_1 + 2c_2 &= 0 \\ -2c_1 - 3c_2 &= 1 \\ c_1 + c_2 &= -1 \end{aligned}$$

which we solve using Gaussian Elimination to get $c_1 = -2$, $c_2 = 1$, so that $\underline{v} = -2\underline{v}_1 + \underline{v}_2 \in W$. On the other hand, $\tilde{v} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$ is not in W: setting up a similar system of equations, we would find it has no solution and hence $\tilde{v} \notin W$.

Example 3 Let $V = \mathcal{F}(\mathbb{R})$ denote the space of scalar functions $f(x)$ defined on $\mathbb{R}$.

a) The span of the three monomials $f_1(x) = 1$, $f_2(x) = x$, $f_3(x) = x^2$ is the set of functions of the form:

$$f(x) = c_1 f_1(x) + c_2 f_2(x) + c_3 f_3(x) = c_1 + c_2 x + c_3 x^2$$

where $c_1, c_2, c_3 \in \mathbb{R}$ are arbitrary. But we've seen this before! $\text{span}\{1, x, x^2\} = P^{(2)}$ the subspace of all polynomials of degree ≤ 2.

Linear Independence and Dependence

Linear dependence captures a notion of "redundancy" in a collection of vectors: the elements $\underline{v}_1, \ldots, \underline{v}_k \in V$ are linearly dependent if there exist scalars $c_1, \ldots, c_k$ not all zero such that

$$c_1 \underline{v}_1 + \cdots + c_k \underline{v}_k = \underline{0}. \quad (1)$$

Elements that are not linearly dependent are called linearly independent.

The condition (1) says that we can write one of the $\underline{v}_i$ as a linear combination of the other vectors: hence it does not add anything new to the span of the collection.

<u>Example</u> $\underline{e}_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$, $\underline{e}_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$, $\underline{e}_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}$ are linearly independent because the only way $c_1\underline{e}_1 + c_2\underline{e}_2 + c_3\underline{e}_3 = 0$ is if $c_1 = c_2 = c_3 = 0$.

<u>Example</u> The collection $\{\underline{e}_1, \underline{e}_2, \underline{e}_3, \underline{v}\} \subset \mathbb{R}^3$ is linearly dependent because

$$\underline{v} = \begin{bmatrix}v_1\\v_2\\v_3\end{bmatrix} = v_1\begin{bmatrix}1\\0\\0\end{bmatrix} + v_2\begin{bmatrix}0\\1\\0\end{bmatrix} + v_3\begin{bmatrix}0\\0\\1\end{bmatrix} = v_1\underline{e}_1 + v_2\underline{e}_2 + v_3\underline{e}_3.$$

<u>Example</u> $\underline{v}_1 = \begin{bmatrix}1\\2\\-1\end{bmatrix}$ and $\underline{v}_2 = \begin{bmatrix}0\\3\\1\end{bmatrix}$ are linearly independent. To see this,

suppose that $c_1\underline{v}_1 + c_2\underline{v}_2 = \begin{bmatrix}c_1\\2c_1+3c_2\\-c_1+c_2\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}$

For this to happen $c_1 = 0$ (from first eqn), and hence $c_2 = 0$.

<u>Example</u> The polynomials $p_1(x) = x-2$, $p_2(x) = x^2-5x+4$, $p_3(x) = 3x^2-4x$, $p_4(x) = x^2-1$ are linearly dependent since:

$$p_1(x) + p_2(x) - p_3(x) + 2p_4(x) = 0 \text{ for all } x$$

On the other hand, the three polynomials $p_1(x)$, $p_2(x)$, and $p_3(x)$ are linearly independent. To see why, suppose

$$c_1p_1(x) + c_2p_2(x) + c_3p_3(x) = (c_2+3c_3)x^2 + (c_1-5c_2+4c_3)x - 2c_1 + 4c_2 = 0 \text{ for all } x.$$

This is true if $c_2+3c_3=0$, $c_1-5c_2-4c_3=0$, $-2c_1+4c_2=0$, which only has the trivial solution $c_1=c_2=c_3=0$.

Checking linear independence in $\mathbb{R}^n$

For now, let's focus on checking whether a collection $\underline{v}_1, \dots, \underline{v}_k \in \mathbb{R}^n$ is linearly dependent or not. We start by forming the $n \times k$ matrix $A = [\underline{v}_1 \ \underline{v}_2 \cdots \underline{v}_k]$ with columns defined by our vectors $\underline{v}_i$.

We start with the column interpretation of matrix-vector multiplication:

$$A\underline{c} = c_1\underline{v}_1 + c_2\underline{v}_2 + \cdots + c_k\underline{v}_k \quad \text{where } \underline{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_k \end{bmatrix},$$

That is, we write any linear combination in terms of matrix multiplication. This lets us make some beautiful connections between the geometry of the span of vectors and linear algebraic systems:

Theorem: Let $\underline{v}_1, \dots, \underline{v}_k \in \mathbb{R}^n$ and $A = [\underline{v}_1 \cdots \underline{v}_k] \in \mathbb{R}^{n \times k}$. Then

a) The vectors $\underline{v}_1, \dots, \underline{v}_k \in \mathbb{R}^n$ are linearly dependent if and only if there is a nonzero $\underline{c} \neq 0$ to $A\underline{c} = \underline{0}$.

b) The vectors are linearly independent if and only if the only solution to $A\underline{c} = \underline{0}$ is $\underline{c} = \underline{0}$.

c) A vector $\underline{b}$ lies in the span of $\underline{v}_1, \dots, \underline{v}_k$ if and only if $A\underline{c} = \underline{b}$ has at least one solution.

The proof of this theorem is not hard and consists of generalizing the examples we've seen so far: see online notes if curious!

This theorem allows us to make some important observations:

a) Any collection of $k > n$ vectors in $\mathbb{R}^n$ is linearly dependent, because we then have more variables (columns) than equations (rows), so we must have at least one free variable.

b) A collection of k vectors in $\mathbb{R}^n$ is linearly independent if and only if the equation $A\underline{c} = \underline{0}$ has no free variables, i.e., $\operatorname{rank} A = \#$ of pivots $= k$. This requires $k \leq n$.

c) A collection of k vectors spans $\mathbb{R}^n$ if and only if their $n \times k$ matrix has rank n. This requires $k \geq n$.

Basis and Dimension

The previous section was admittedly quite abstract, but it was necessary to set us to the extremely practical notion of a basis of a vector space.

This section is where the magic happens: we will show that any n-dimensional vector space doesn't just look like, but "behaves the same", as $\mathbb{R}^n$!

A basis of a vector space V is a finite collection $\underline{v}_1, \ldots, \underline{v}_n \in V$ that a) spans V, and b) linearly independent.

Another way of thinking about a basis is we are looking for the smallest collection of vectors that allow us to express any vector in V as a linear combination from our collection.

Example The standard basis of $\mathbb{R}^n$ consists of the n vectors

$$\underline{e}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \underline{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \ldots, \quad \underline{e}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix},$$

These clearly span $\mathbb{R}^n$ ($\underline{x} = x_1\underline{e}_1 + x_2\underline{e}_2 + \cdots + x_n\underline{e}_n$) and are linearly indep.

That we need n vectors to define a basis of $\mathbb{R}^n$ is no coincidence, in fact any basis of $\mathbb{R}^n$ consists of exactly n vectors.

Example The set $S = \{1, x, x^2, \ldots, x^n\}$ is a basis for $\mathcal{P}^{(n)}$, the space of polynomials of degree $\leq n$. This called the standard basis for $\mathcal{P}^{(n)}$.

Notice that $\mathcal{P}^{(n)}$ can have up to $n+1$ terms ($a_0, a_1x, a_2x^2, \ldots, a_nx^n$), and that its basis S has $n+1$ elements. Not a coincidence!

Fact: Suppose the vector space V has a basis composed of n elements $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n$. Then any other basis for V has the same number of elements, n, in it. This number is called the dimension of V, $\dim V = n$.

Example Both $\mathbb{R}^n$ and $\mathcal{P}^{(n-1)}$ have dimension n (note $\dim \mathcal{P}^{(n)} = n+1$ because of the constant term a_0 in $p(x) = a_0 + a_1x + \cdots + a_nx^n$).

Coordinate Systems

An important reason for specifying a basis for a vector space V is to impose a "coordinate system" on V.

This section will show that if $\dim V = n$, i.e., if the basis has n elements, then the coordinate system makes V behave exactly like $\mathbb{R}^n$!

Theorem: Let $\underline{v}_1, \ldots, \underline{v}_n$ be a basis for a vector space V. Then for each $\underline{v} \in V$, there exists a unique set of coefficients $c_1, \ldots, c_n$ such that

$$\underline{v} = c_1 \underline{v}_1 + c_2 \underline{v}_2 + \cdots + c_n \underline{v}_n. \quad (2)$$

Proof: Since $\underline{v}_1, \ldots, \underline{v}_n$ is a basis for V, there exists at least one set of coefficients such that (2) holds. Suppose $\underline{v}$ also has the representation

$$\underline{v} = d_1 \underline{v}_1 + d_2 \underline{v}_2 + \cdots + d_n \underline{v}_n.$$

Then $\underline{0} = \underline{v} - \underline{v} = (d_1 - c_1)\underline{v}_1 + (d_2 - c_2)\underline{v}_2 + \cdots + (d_n - c_n)\underline{v}_n. \quad (*)$

But since $\underline{v}_1, \ldots, \underline{v}_n$ form a basis, they are linearly independent, meaning $(*)$ is only satisfied for $d_i - c_i = 0$, or $d_i = c_i$, for $i = 1, \ldots, n$.

For a given basis $\mathcal{B} = \{\underline{v}_1, \ldots, \underline{v}_n\}$, we can therefore define the vector $\underline{c} \in \mathbb{R}^n$ of coordinates for $\underline{x}$ relative to $\mathcal{B}$ by the weights in its representation:

$$\underline{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}, \quad \text{with } \underline{x} = c_1 \underline{v}_1 + c_2 \underline{v}_2 + \cdots + c_n \underline{v}_n.$$

This is very exciting!

Example: Consider the standard basis $\mathcal{B} = \{\underline{e}_1, \underline{e}_2\}$ for $\mathbb{R}^2$. Then we have $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \underline{e}_1 + x_2 \underline{e}_2$, and so the $\mathcal{B}$-coordinates of $\underline{x}$ are x_1 and x_2, as expected.

What if we instead use the basis $\mathcal{B}' = \left\{ \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix} \right\}$. We need to find coordinates c_1 and c_2 such that

$$c_1 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} c_1 + c_2 \\ c_1 - c_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

This linear system has a solution (it must!), and it is $c_1 = \frac{x_1 + x_2}{2}$, $c_2 = \frac{x_1 - x_2}{2}$

In the coordinate system defined by B', the coordinates for $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ are $\frac{1}{2}\begin{bmatrix} x_1 + x_2 \\ x_1 - x_2 \end{bmatrix}$. Moving from basis B to B' is called a change of basis. Can you pose this as finding the solution to $A\underline{c} = \underline{x}$?

<u>Example</u> Let $B = \{1, x, x^2, x^3\}$ be the standard basis for $P^{(3)}$. A typical element $p(x) \in P^{(3)}$ has the form $p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3$.

We can read off the coordinates of $p(x)$ with respect to B, which we encode in the vector $\underline{p}$:

$$\underline{p} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix}.$$

Notice that $\underline{p}$ lives in $\mathbb{R}^4$! Notice even further that the coefficients of the sum of polynomials:

$$p(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3, \quad q(x) = b_0 + b_1 x + b_2 x^2 + b_3 x^3$$
$$p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + (a_3 + b_3)x^3$$

can be obtained by adding their B-coordinate vectors in $\mathbb{R}^4$:

$$\underline{p} + \underline{q} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} + \begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} a_0 + b_0 \\ a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \end{bmatrix}$$

You should check that the same connection between $c\,p(x)$ and $c\underline{p}$ also holds. By setting a suitable coordinate system (via a basis), working with elements of $P^{(3)}$, which are polynomial functions, can be turned into working with ordinary vectors in $\mathbb{R}^4$!

This idea of going back and forth between two vector spaces is captured in terms of a vector space isomorphism. We do not yet have all of the tools needed to define this rigorously, but for now, we will interpret it as meaning that every vector space calculation in V is accurately reproduced in W, and vice versa.

In the above example, we used that $P^{(3)}$ and $\mathbb{R}^4$ are isomorphic, so we can add and scale either the polynomials directly, or work with their coefficient vectors in $\mathbb{R}^4$. Even though they are closely related, they are not the same thing. Rather, $P^{(3)}$ and $\mathbb{R}^4$ are different ways of representing polynomials of degree ≤ 3, connected via the chosen basis $B = \{1, x, x^2, x^3\}$.