vectors

- we package a number of data values into one atomic object
- we will study basic properties of collections of such objects
- look at notions of "flatness" of such collections.

basics

- a *vector in* $\mathbb{R}^m$ is an ordered collection of m real numbers.
  - also called a *point*
- We will write a vector as a vertical colum, such as

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} := \begin{bmatrix} 1 \\ -5 \end{bmatrix}$$

- index the entries starting at the top, with the index 1.
- we will use a bold font for a vector $\mathbf{v}$ When we index an entry, such as $v_2 = -5$, we will use a regular font.

visualizing

- $\mathbb{R}^2$ for now.
- draw a dot in plane with coordinates (v_1, v_2).
  - This will be especially useful when we want to visualize lots of vectors, such as "those forming a line".
- draw an arrow, with its tail at the origin and its head at (v_1, v_2).
- draw a "free" arrow, with its tail at any point with coordinates (p_1, p_2), and its head at $(p_1 + v_1, p_2 + v_2)$.
  - This will be useful when we want to do things like visualize vector-vector addition.
- these are 3 visualizations. the mathematical object is the collection of numbers!

addition

- The *zero vector* $\mathbf{0}$ in $\mathbb{R}^m$ is the vector where all of its m entries are 0.
- We define *vector-vector addition* between two vectors in $\mathbb{R}^m$ by adding them together entry by entry, as in:

$$\begin{bmatrix} -17 \\ \pi \end{bmatrix} + \begin{bmatrix} 19 \\ .5 \end{bmatrix} = \begin{bmatrix} 2 \\ \pi + .5 \end{bmatrix}$$

- sums are written as $\mathbf{w} := \mathbf{u} + \mathbf{v}$.
- visualize: start at the origin, move along $\mathbf{u}$ taking us to the tip of $\mathbf{u}$, move along the free vector $\mathbf{v}$ taking us to a new point.
  - the coordinates of this new point give us $\mathbf{w}$.
  - we may also add an arrow from the origin as a visualization of $\mathbf{w}$.
- Clearly it is true that for all $\mathbf{v}$, we have $\mathbf{v} + \mathbf{0} = \mathbf{v}$.
- It is also clearly true that for all $\mathbf{v}$ and $\mathbf{u}$, $\mathbf{v} + \mathbf{u} = \mathbf{u} + \mathbf{v}$.

scalar mul

- Given a vector $\mathbf{v}$ and a real number, c, we call a *scalar*, we define the *scalar-vector multiply*, $c\mathbf{v}$ to be the vector with entries cv_i.

$$3 \begin{bmatrix} 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 6 \\ 15 \end{bmatrix}$$

- We will also write this product in reverse order, such as $\mathbf{v}c$.
- visualize scalar multiplication on a vector by lengthening or shortening a free-vector.
  - When c is negative, we need to flip the orientation of our visualized free-vector.
- Clearly, for any vector $\mathbf{v}$, we have $0\mathbf{v} = \mathbf{0}$.
- We define the unary negation operator as $-\mathbf{v} := -1\mathbf{v}$.
- We define the binary subtraction operator as $\mathbf{v} - \mathbf{u} := \mathbf{v} + (-\mathbf{u})$.

properties

- There are a bunch of properties that are true for vectors and these operations. Their correctness is easy to prove.

  Theorem (properties). *For all vectors* $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *in* R^m *and scalars* c *and* d *the following are true*

  - *Commutativity:* $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$
  - *Associativity:* $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$
  - *Zero:* $\mathbf{u} + \mathbf{0} = \mathbf{u}$
  - *Additive inverse:* $\mathbf{u} - \mathbf{u} = \mathbf{0}$
  - *Distributivity:* $c(\mathbf{u} + \mathbf{v}) = c\mathbf{u} + c\mathbf{v}$
  - *More distributivity:* $(c + d)\mathbf{u} = c\mathbf{u} + d\mathbf{u}$
  - $c(d(\mathbf{u})) = (cd)\mathbf{u}$
  - $1\mathbf{u} = \mathbf{u}$

Applications

- In physics, robotics, computer vision and computer graphics, we use vectors in $\mathbb{R}^3$ to represent points in a real or simulated 3D space.
- We can use an "rgb"-vector in $\mathbb{R}^3$ to represent a color.
- We might use a vector in $\mathbb{R}^{44,000}$ to represent one second of a digital audio clip.
- If we have a digital image, of pixel dimensions h-by-w,
- We can represent a vertical column of pixels by a vector in $\mathbb{R}^h$.
- We can represent a horizontal row of pixels by a vector in $\mathbb{R}^w$.
- We can also represent the whole image as a vector in $\mathbb{R}^{h \times w}$,
  - here we would need some rule to go from two-dimensional image coordinates to one-dimensional vector coordinates.
- Many times, a vector is just used to capture some data. If I rate 100 movies with numerical scores, then this can be stored as a vector in $\mathbb{R}^{100}$.

General Vector Spaces (opt)

- In this class we will just deal with the vectors spaces $\mathbb{R}^m$, but there are many ways this can be abstracted generalized.
- One can work with the "complex" vectors spaces $\mathbb{C}^m$.
- One can work with vector spaces over other "fields".
- One can work with "infinite-dimensional" vector spaces.
- One can define an abstract vector space to be any set of objects, together with a notion of addition and scalar multiplication that satisfies all of the properties of vector algebra.
  - This approach lets us think of vectors as abstract atomic objects, and not as collections of numbers.
  - The theorems we are about to prove about $\mathbb{R}^m$ only depend on these properties, and so also hold for abstract vector spaces.

Linear combintation

- Given a finite list of n vectors $\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n$ in $\mathbb{R}^m$ and scalars $c_1, c_2 ... c_n$, the following expression is well defined

$$c_1\mathbf{a}_1 + ... c_n\mathbf{a}_n$$

  and it describes a vector $\mathbf{v} \in \mathbb{R}^m$.

  - We can write this in symbolically as $\mathbf{v} := \sum_j c_j\mathbf{a}_j$

- Since $\mathbf{v}$ can be written in this way, we say that $\mathbf{v}$ is a *linear combination* of the $\mathbf{a}_j$.
- For a linear combination, we will call the c_j the *weights* or the *coefficients*. The weights (even all of them) are allowed to be 0.

notation

- if i have two defined object a and b, then $a = b$ is a statement about their equality.
- if i have one defined object b, the $a := b$ is a definition for a new object a.

picture form

- We can also write a linear combination in picture form as

$$\mathbf{v} := \begin{bmatrix} \mathbf{a}_1 & ... & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} c_1 \\ ... \\ c_n \end{bmatrix}$$

- the lefter object is a "row" of n vectors in $\mathbb{R}^m$, and the righter object is a vector in $\mathbb{R}^n$.
- you use the "two finger rule",
  - left finger goes l2r, right finger goes t2b.
- Note the objects here for the left finger are vectors, not numbers.

span

Definition. If some vector $\mathbf{v}$ is a linear combination of a *set* of vectors $\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\}$ in $\mathbb{R}^m$, then we say that $\mathbf{v}$ is *in the span* of the $\mathbf{a}_i$.
The set of vectors that are in the span of the $\mathbf{a}_i$ is called *the span* of the $\mathbf{a}_i$, or $Span(\{\mathbf{a}_i\})$. As a corner case, we define the span of the empty set to be $\mathbf{0}$.
For some set S, the statement $\{\mathbf{a}_i\}$ *spans* S, means $S = Span(\{\mathbf{a}_i\})$. ■

basic examples

- in $\mathbb{R}^3$, the span of the empty set is just the zero vector.
- The span of a single, non-zero vector, is all of the vectors long a line through the origin.
- The span of two, non-zero vectors, where one is not a scalar multiple of the other, is the set of vectors along a plane through the origin.
- The span of three, non-zero vectors, will generally be all of $\mathbb{R}^3$.
- If I add a 4th vector, the span will still be all of $\mathbb{R}^3$.

funny examples

- The span of the single vector $\mathbf{0}$ is just $\{\mathbf{0}\}$.
- The span of two vectors, where the second vector is just a scalar multiple of the first vector, is just a set of vectors along a line.
- The span of three vectors, where the third vector is in the plane spanned by the first two vectors, is just that plane.
- in these cases, the span is smaller than would be suggested by the number of vectors.
  - there is some kind of reduncancy in our vector set.

- These interesting cases, brings up our next idea.

dependence

Definition. Given a non-empty, finite set of n vectors $\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\}$ in $\mathbb{R}^m$, if there are a set of scalars $c_1, c_2 ... c_n$, **not all zero** such that

$$\mathbf{0} = c_1\mathbf{a}_1 + ...c_n\mathbf{a}_n$$

then we say that these vectors are *linearly dependent*. If there is no such set of scalars then say that they are *linearly independent*. ■

- it means we can start at the origin, walk in the n directions, and get back to the origin.
- Intuitively, linear dependence means that there is a certain redundancy in our set.
- in all of our funny examples above (and one not funny example), it turns out that there was linear dependence. (draw examples)
- BQ1: If $\{\mathbf{a}_1...\mathbf{a}_n\}$ are linearly dependent. What happens if I add another vector $\mathbf{a}_{n+1}$ to the set. Can the new set be enlinearly independent?
- BQ2: Suppose that $\mathbf{0}$ is one of the vectors in our set, can this set be linearly independent?

algorithms

- In easy cases, one can eyeball a set of vectors and tell if they are linearly independent or not.
- here is a simple example that is clearly dependendent

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

- More generally, there are various algorithms that can do this test.
  - But we will not delve into such algorithms now.
  - Ultimately, my favorite algorithm for this is based on something called a "singluar value decomposition", which we will learn much later.

logic terminology break

- consider the *statement* "A implies B".
- the *converse* statement is "B implies A"
  - just because a statement is true does not automatically mean that its converse is true.
  - ex. "if it is raining then i carry an umbrella" vs. "if i carry an umbrella then it must be raining"
- sometimes both a statement and its converse is true. In this case, i will need to say so.
  - i may use "iff"
- the *contrapositive* statement is "not B implies not A"
  - if a statement is true, then this automatically means that its contrapositive is true.
  - ex. "if it is raining then i carry an umbrella" vs. "if i do not have an umbrella then it must not be raining"
  - one way to prove that a statement is true is to prove that its contrapositive is true!

span and LD

Lemma (Hammer)**.** *If there is a linear dependence on a set of vectors $\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n, \mathbf{v}\}$ in $\mathbb{R}^m$, where the coefficient on $\mathbf{v}$ is non-zero, then $\mathbf{v}$ is in the span of $\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\}$. The converse is true as well.*

We will prove the lemma under the assumption that $n \geq 1$, though the lemma remains true for the corner case where the $\{\mathbf{a}_i\}$ is the empty set.

Proof. By assumption we have

$$\mathbf{0} = d_1\mathbf{a}_1 + ...d_n\mathbf{a}_n + c\mathbf{v}$$

for some d_i and c, where c is non-zero. Then the following manipulations are well defined

$$\begin{aligned}
\mathbf{0} &= -1/c(d_1\mathbf{a}_1 + ...d_n\mathbf{a}_n) - \mathbf{v} \\
\mathbf{v} &= -1/c(d_1\mathbf{a}_1 + ...d_n\mathbf{a}_n) \\
\mathbf{v} &= (-d_1/c)\mathbf{a}_1 - ...(-d_n/c)\mathbf{a}_n
\end{aligned}$$

The last statement means that $\mathbf{v}$ is in the span of $\{\mathbf{a}_1, \mathbf{a}_2, ..., \mathbf{a}_n\}$.
For the converse direction, the assumptions tell us that there exists e_i such that

$$\mathbf{v} = e_1\mathbf{a}_1 + ...e_n\mathbf{a}_n$$

which then means

$$\mathbf{0} = e_1\mathbf{a}_1 + ...e_n\mathbf{a}_n - 1\mathbf{v}$$

□

proof notes

- We start with the assumptions and make well defined steps until we get to our desired conclusion.
- This particular argument is the main kind of manipulation that we will use over and over again when reasoning about linear dependence.

....

- the converse tells us that if $\mathbf{v}$ is in the span of an initial set of $\mathbf{a}_i$, then the augmented set has a linear dependence.
- so if I have two vectors that span some plane, and i pick a third vector in that plane, then the set of these three vectors must be linearly dependent.
- Eventually, we will want to make the more numerical claim: "any three vectors in a plane must be linearly dependent"
- we are not quite ready to deal with this yet as we need to have a good defintion of a plane, and we haven't yet seen the Steinitz lemma.

set terminology

- Suppose that A and B are two sets.
- The statement $A \supseteq B$ means that the set A contains all of the elements the set B, and these two sets might be equal.
  - An equivalent statement is $B \subseteq A$.
- The statement $A \supsetneq B$ means that the set A strictly contains the set B, that is, it contains all of the elements of the set B and also contains some extra elements that are not in B.
  - An equivalent statement is $B \subsetneq A$.

LD is a form of redundancy

Lemma (Pull-off Lemma)**.** *Let $\{\mathbf{a}_1..\mathbf{a}_n\}$ be a set of linearly dependent vectors in $\mathbb{R}^m$ with span S. Then there is at least one of the $\mathbf{a}_i$ that can be removed such that the span of the remaining $\mathbf{a}_i$ still equals S.*

Proof. By the assumption of linear dependence, we have

$$\mathbf{0} = c_1\mathbf{a}_1 + ...c_n\mathbf{a}_n$$

with not all $c_i = 0$. Without loss of generality, we can reorder our indices so that afterwards, we have $c_n \neq 0$. Then

$$\mathbf{a}_n = -1/c_n(c_1\mathbf{a}_1 + ...c_{n-1}\mathbf{a}_{n-1})$$

...

- From the definition of spanning, for any $\mathbf{v}$ in S, we have

$$\begin{aligned}\mathbf{v} &= d_1\mathbf{a}_1 + ...d_n\mathbf{a}_n \\ &= d_1\mathbf{a}_1 + ... + d_{n-1}\mathbf{a}_{n-1} - d_n/c_n(c_1\mathbf{a}_1 + ...c_{n-1}\mathbf{a}_{n-1}) \\ &= e_1\mathbf{a}_1 + ...e_{n-1}\mathbf{a}_{n-1}\end{aligned}$$

for appropriate e_i. This last expression shows that $\mathbf{v}$ is in the span of the remaining $\mathbf{a}_i$. So we have shown

$$span\{\mathbf{a}_1..\mathbf{a}_n\} \subseteq span\{\mathbf{a}_1..\mathbf{a}_{n-1}\}$$

...

- the other dierctional set inclusion is much easier:

$$span\{\mathbf{a}_1..\mathbf{a}_n\} \supseteq span\{\mathbf{a}_1..\mathbf{a}_{n-1}\}$$

- for any $\mathbf{v} \in span\{\mathbf{a}_1..\mathbf{a}_{n-1}\}$, we have

$$\mathbf{v} = d_1\mathbf{a}_1 + ...d_{n-1}\mathbf{a}_{n-1}$$

for some set of d_i

- and so

$$\mathbf{v} = d_1\mathbf{a}_1 + ...d_{n-1}\mathbf{a}_{n-1} + 0\mathbf{a}_n$$

  - so $\mathbf{v} \in span\{\mathbf{a}_1..\mathbf{a}_n\}$.

- so, together we get

$$span\{\mathbf{a}_1..\mathbf{a}_n\} = span\{\mathbf{a}_1..\mathbf{a}_{n-1}\}$$

□

proof notes

- Proof notes: we don't care which of the c_i is non-zero, just that there is at least one of them. For simplicity we assume that this is the nth coefficient, but nothing would change if it were a different coefficient.
- picked "any vector" in S and named it $\mathbf{v}$ and used its properties, and then named "some" d_i that we knew had to exist.
  - so it is true for all vectors in S.

adding more vectors

Lemma (Push-on Lemma)**.** *Let $\{\mathbf{a}_1..\mathbf{a}_n\}$ be a linearly independent set of vectors in $\mathbb{R}^m$, Let $\mathbf{a}_{n+1}$ be any vector in $\mathbb{R}^m$ that is not in the span of $\{\mathbf{a}_1..\mathbf{a}_n\}$. Then $\{\mathbf{a}_1..\mathbf{a}_n, \mathbf{a}_{n+1}\}$, is linearly independent.*

Proof. Suppose that $\{\mathbf{a}_1..\mathbf{a}_{n+1}\}$ were linearly dependent. Then this means that

$$\mathbf{0} = c_1\mathbf{a}_1 + ... + c_n\mathbf{a}_n + c_{n+1}\mathbf{a}_{n+1}$$

(with at least one non-zero coefficient). Claim: $c_{n+1} \neq 0$. Proof: If $c_{n+1} = 0$, then

$$\begin{aligned}\mathbf{0} &= c_1\mathbf{a}_1 + ... + c_n\mathbf{a}_n + 0\mathbf{a}_{n+1} \\ &= c_1\mathbf{a}_1 + ... + c_n\mathbf{a}_n\end{aligned}$$

(with at least one non-zero coefficient). which implies that $\{\mathbf{a}_1..\mathbf{a}_n\}$ is a linearly dependent set. This would contradict our assumption. Now that we have established $c_{n+1} \neq 0$, then from the hammer Lemma we would have $\mathbf{a}_{n+1}$ in the span of $\mathbf{a}_1..\mathbf{a}_n$. But this would contradict our assumption that $\mathbf{a}_{n+1}$ was not in the span of $\{\mathbf{a}_1..\mathbf{a}_n\}$. □

proof notes

- Here we see a proof by *contradiction*.
  - We add an assumption that our desired claim is false.

  - We do some valid reasoning under this assumption that leads to a contradiction (like a violation of one of our assumptions).
  - The fact that we arrived at a contradiction means that our desired claim must actually be true.
- we stared by stating and establishing a subclaim.
  - In the above proof, we used contradiction twice, once for the whole lemma, and once for the subclaim in its proof.
- we used and cited a previous lemma

Steinitz exchange

- The next lemma gives us a *numerical* relationship between spanning and lin ind.
- It looks a bit odd, but this lemma will key in the next section.

  Lemma (Steinitz Exchange Lemma)**.** *Let* $\{\mathbf{b}_1 ....\mathbf{b}_k\}$ *be a set of vectors in* $\mathbb{R}^m$ *with span* S*. Let* $\{\mathbf{a}_1 ....\mathbf{a}_n\}$ *be a linearly independent set of vectors in* S*. Then* $k \geq n$*.*

- part of proof in pset.
- multiple applications of reasoning we have already used.
- this is getting us closer to our idea that any three vectors in a plane must be linearly dependent.
  - we still don't have a definition of a plane, but we now know that if two vectors span some set S , then any three vectors in S must be lin. dep.

Subspace

- a subspace is something that will ultimately turn out to be exactly the notion of a span of vectors in $\mathbb{R}^m$.
- But here our definition does not use any generating vectors!

  Definition. We say that a subset V of $\mathbb{R}^m$ is *closed under addition* if for any two vectors $\mathbf{u}$ and $\mathbf{v}$ in V, we have $\mathbf{u}+\mathbf{v}$ in V. We say that a subset V of $\mathbb{R}^m$ is *closed under scalar multiplication* if for any vector $\mathbf{v}$ in V and scalar c, we have $c\mathbf{v}$ in V.

  A non-empty subset V of $\mathbb{R}^m$ that is closed under addition and scalar multiplication is called a *subspace* of $\mathbb{R}^m$. ■

examples

- The smallest subspace of $\mathbb{R}^m$ is $\{\mathbf{0}\}$.
- The largest subpsace is $\mathbb{R}^m$ itself.
- In $\mathbb{R}^2$ or $\mathbb{R}^3$, we can verify that vectors along a single line that goes through the origin form a subspace.
- In $\mathbb{R}^3$, we can verify that vectors along a single plane that goes through the origin form a subspace.
- Q: why must $\mathbf{0}$ be in every subspace?

span is susbspace

Lemma. *The span* S *of a finite set of vectors* $\{\mathbf{a}_1..\mathbf{a}_n\}$*, in* $\mathbb{R}^m$ *is a subspace.*

Proof. Let $\mathbf{u}$ and $\mathbf{v}$ be any two vectors in S. Then they must be expressible as

$$\mathbf{u} = d_1\mathbf{a}_1 + ...d_n\mathbf{a}_n$$

and

$$\mathbf{v} = e_1\mathbf{a}_1 + ...e_n\mathbf{a}_n$$

for some d_i and e_i. To verify that S is closed under addition, we need to show that $\mathbf{u}+\mathbf{v}$ is in S. When we do this addition, and use the properties of vector algebra we find that

$$\mathbf{u}+\mathbf{v} = (d_1+e_1)\mathbf{a}_1 + ...(d_n+e_n)\mathbf{a}_n.$$

This expression puts the sum in S. A similar argument can be used to show closedness under scalar multiplication □

- We will see later that every subspace of $\mathbb{R}^m$ is the span of some finite set of vectors, making these concepts identical.

proof notes

- we need to work with any pair of vectors in our proof.
  - Specific examples of vectors will not do the trick.

basis

- we describe sets of vectors that are neither redundant, or insufficient to span a given subspace.

  Definition. A *basis* for a subspace V is a set of linearly independent vectors that spans V. ■

  Corner case: if V is $\{\mathbf{0}\}$, then the empty set is its basis.

simplest example

- The following three vectors form a basis for $\mathbb{R}^3$.

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

- It is easy to verify that these are linearly independent and that any vector in $\mathbb{R}^3$ can be obtained through a linear combination of these three vectors.
- Vectors of this form, with all zero entries, except for a single 1 in the $i$the position are called the *indicator* vectors, $\mathbf{e}_i$.
- the following is another basis (less obvious)

$$\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

unique rep

- this is an aside, but it is a super useful property of a basis.
- Q: are there two etch a sketch knob-settings that put the cursor at the same point.

  Theorem (Unique Representation)**.** *Let V be a subspace and $\{\mathbf{a}_1..\mathbf{a}_n\}$ a basis for V. Then given any $\mathbf{v} \in V$ there is exactly one way that $\mathbf{v}$ can be written as a linear combination of the basis vectors.*

  Proof. Since the basis spans V, any vector V can be written, at least one way, as a linear combination of the basis vectors. Next we show that there is exactly one way.
  Let

$$\mathbf{v} = d_1\mathbf{a}_1 + ...d_n\mathbf{a}_n = e_1\mathbf{a}_1 + ...e_n\mathbf{a}_n$$

  Then

$$\begin{aligned} \mathbf{0} = \mathbf{v} - \mathbf{v} &= d_1\mathbf{a}_1 + ...d_n\mathbf{a}_n - (e_1\mathbf{a}_1 + ...e_n\mathbf{a}_n) \\ &= (d_1 - e_1)\mathbf{a}_1 + ...(d_n - e_n)\mathbf{a}_n \end{aligned}$$

  Since a basis must be linearly independent, we must have $d_j - e_j = 0$, for all j. □

size

Theorem (Basis size)**.** *Let V be a subspace and $\{\mathbf{a}_1..\mathbf{a}_n\}$ a basis and $\{\mathbf{b}_1..\mathbf{b}_k\}$ another basis. Then $k = n$.*

Proof. By the definition of a basis, $\{\mathbf{a}_1..\mathbf{a}_n\}$ is linearly independent. By the definition of a basis, $\{\mathbf{b}_1..\mathbf{b}_k\}$ spans V. Then from the Steintiz exchange lemma, we must have $k \geq n$.
Next, we reverse the argument: By the definition of a basis, $\{\mathbf{b}_1..\mathbf{b}_k\}$ is linearly independent. By the definition of a basis, $\{\mathbf{a}_1..\mathbf{a}_n\}$ spans V. Then from the Steinitz exchange lemma, we must have $n \geq k$.
Together, this means that $n = k$ □

proof notes

- This is a nifty proof method, where we show equality by proving two inequalities.

dimension

- Because all bases for V have the same size, we can use that size as a characteristic of V.

  Definition. Let V be a subspace of $\mathbb{R}^m$, with a basis made up n vectors. Then we say that V has *dimension* n. (ie. $Dim(V) = n$). ∎

  Q: what is the dimension of $\mathbb{R}^m$.

- dimension of V tells us how many coefficients we need in order to specify a vector in the space.
- what we called a "line through the origin" is a one-dimensional subspace. What we called a "plane through the origin" is a two-dimensional subspace.
- this gives us a clear defintion of things like lines and planes through the origin.
- BQ: Prove that that any three vectors in a plane through the origin. must be lin. dep.

small spanners

Lemma. *Let V be an n dimensional subspace and suppose that the set of n vectors $\{\mathbf{b}_1..\mathbf{b}_n\}$ (in V) spans V. Then $\{\mathbf{b}_1..\mathbf{b}_n\}$ is a basis*

Proof. We just need to show that $\{\mathbf{b}_1..\mathbf{b}_n\}$ is linearly independent.
Suppose that $\{\mathbf{b}_1..\mathbf{b}_n\}$ was linearly dependent. then from the pull-off lemma I could remove one $\mathbf{b}_i$ from the set without changing the span. This would give me a set of $n-1$ vectors that span V.
Meanwhile, from dimension, we know there is a basis $\mathbf{a}_1..\mathbf{a}_n$ of size n, which by definition must be linearly independent. This would contradict the Steinitz exchange lemma. □

large independent

Lemma. *Let V be n dimensional subspace and suppose if $\{\mathbf{a}_1..\mathbf{a}_n\}$ is a linearly independent set of vectors in V. Then $\{\mathbf{a}_1..\mathbf{a}_n\}$ is a basis.*

Proof. We just need to show that $\{\mathbf{a}_1..\mathbf{a}_n\}$ spans V.
By construction, each $\mathbf{a}_i \in V$ so

$$V \supseteq Span(\{\mathbf{a}_1..\mathbf{a}_n\})$$

This last step uses the fact that a subspace must be closed under addition and scalar multiplication.
Suppose that V was a strict superset of $Span(\{\mathbf{a}_i\})$. I could pick any vector in V and not in $Span(\{\mathbf{a}_i\})$, and call this $\mathbf{a}_{n+1}$. Then from the push-on lemma, if I add this to my $\mathbf{a}_i$, I will obtain a linearly independent set of size $n+1$. This set will be in V. Meanwhile, from dimension, we know that V has a basis $\{\mathbf{b}_1..\mathbf{b}_n\}$ of size n, which by definition must span V. This would contradict the Steinitz exchange lemma. Thus, V is not a strict superset, and we have

$$V = Span(\{\mathbf{a}_i\})$$

□

basis thm

- Combining the small and large Lemmas we obtain the following

  Theorem (Basis theorem)**.** *Let V be an n dimensional subspace of $\mathbb{R}^m$. Then the following are equivalent:*

  - $\{\mathbf{a}_1..\mathbf{a}_n\}$ *is a basis for V.*
  - $\{\mathbf{a}_1..\mathbf{a}_n\}$ *are in V and are linearly independent.*
  - $\{\mathbf{a}_1..\mathbf{a}_n\}$ *spans V.*

- In short, once we know the dimension, we only need to check any one of the two basis-properties to know that we have a basis!
- BQ: Suppose that $V := Span(\{\mathbf{a}_1..\mathbf{a}_n\})$. We do not assume linear independence. What do we know about $Dim(V)$?

- BQ: What Lemma would you use to prove this?
- BQ: show that our non-obvious set above is a basis for $\mathbb{R}^3$.

Greed

- The Basis theorem tells us that we can be greedy when looking for a basis.

  Theorem (Basis extension theorem)**.** *Let V be an n-dimensional subspace of $\mathbb{R}^m$. Let $\{\mathbf{a}_1..\mathbf{a}_k\}$ be a linearly independent set of vectors in V, with $k < n$ Then we can extend this set with some set of vectors $\{\mathbf{a}_{k+1}..\mathbf{a}_n\}$ so that $\{\mathbf{a}_1..\mathbf{a}_n\}$ forms a basis.*

  Proof. By construction, each $\mathbf{a}_i \in V$ so

  $$V \supseteq Span(\{\mathbf{a}_1..\mathbf{a}_k\})$$

  (using the fact that a subspace must be closed under addition and scalar multiplication.)

  If $V = Span(\{\mathbf{a}_1..\mathbf{a}_k\})$, then we have a basis, and we are done. Otherwise, from the push-on lemma, we can add any vector that is in V but not in $Span(\{\mathbf{a}_1..\mathbf{a}_k\})$ to our set, while maintaining its linear independence.

  We can keep on doing this until the span of our linearly indepdent set equals V, giving us a basis. From the basis Theorem, once we reach n vectors, we must have a basis, so this process must terminate at n. $\square$

basis build

- So we don't have to be clever to find a basis for V.
- We can just start with the empty set as our "working-set", and keep on adding vectors from V while our working-set remains linearly independent.
- When hit n, we have a basis.

... not given dimension

- Suppose that V was defined as the span of some finite set of vectors of size N: $\{\mathbf{v}_1..\mathbf{v}_N\}$.
- And suppose we do not know the dimension of V.
- Then we can find a basis for V using the greedy expansion process.
- We iterate over all of the $\mathbf{v}_i$, adding each vector to our working-set if it does not render the working-set linearly dependent.
- When we are done, we know that each $\mathbf{v}_i$ is in the span of our linearly independent working-set.
- thus V is in the span of our working set.
- so we have a basis!

subspace has a basis (opt)

- now we can prove that every subspace has a basis

  Lemma. *Let V be a subspace of $\mathbb{R}^m$. Then there must be a finite set of vectors that spans V. In particular, there is a basis for V.*

- So every subspace of $\mathbb{R}^m$ has a dimension!

  Proof. We start with the empty set as a working-set. We iteratively expand our working set using vectors from V, (using the push-on Lemma) as long as our lin. ind. working set does not have a span that includes V.

  This process must terminate in at most m steps. For otherwise, we would have more than m linearly independent vectors in $\mathbb{R}^m$ (violating Steinitz).

  Thus, this process ends, with a lin. ind. working set with size $r \leq m$ and such that

  $$V \subseteq Span(\{\mathbf{a}_1..\mathbf{a}_r\})$$

Meanwhile, by construction, each $\mathbf{a}_i \in V$ so

$$V \supseteq Span(\{\mathbf{a}_1..\mathbf{a}_r\})$$

(This last step uses the fact that a subspace must be closed under addition and scalar multiplication.)

Thus

$$V = Span(\{\mathbf{a}_1..\mathbf{a}_r\})$$

and we are done. □

strict subspaces are smaller

Lemma (Small strict)**.** *Let V, and W be two subspaces of $\mathbb{R}^m$, with $V \subsetneq W$. Then $Dim(V)$ is strictly less than $Dim(W)$.*

- this makes perfect sence.
- you should be able to prove it.

examples

- for the following sets: is it lin ind?, what is the dim of its span?

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 10 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 10 \\ 0 \end{bmatrix} \begin{bmatrix} 5 \\ 10 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 5 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ 10 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

Affine Set

- Linear subspaces must include the vector $\mathbf{0}$.
  - Q: why?
- But we would also like to generalize the low dimensional notion of a line or plane that does not pass through the origin.
  - We have already seen such subsets arising as solutions to linear systems of equations.
- This will be easy. All we need to do is start with some n-dimensional subspace of $\mathbb{R}^m$, and "translate" it away from the origin.

formal version

Definition. Let V be an n-dimensional subspace of $\mathbb{R}^m$. Let $\mathbf{t}$ be a fixed "translation" vector in $\mathbb{R}^m$ Let S be the set of vectors $\mathbf{p}$ that can be written in the form $\{\mathbf{t} + \mathbf{v} | \mathbf{v} \in V\}$. Then we call S an n-dimensional *affine set*. ■

We allow for $\mathbf{t} \in V$, which would make $S = V$ and V itself an affine set.

- In an affine set, there is no notion of scalar multiplication, and no notion of adding elements together.
  - if one adds two vectors in S together, using vector addition in $\mathbb{R}^m$, the result will generally not remain in S
- We will visualize affine sets as the collection of points with coordinates p_i.

hyperplane

Definition. We refer to an $m-1$ dimensional affine subset of $\mathbb{R}^m$ as a *hyperplane*. We refer to an $m-1$ dimensional subspace of $\mathbb{R}^m$ as a *hyperspace*. ■

affine solution sets

- we saw in 3D, that each linear equation limited us to a (hyper)plane (affine), and that intersecting these planes gave us lines, which are smaller affine sets.
- in higher dimensions, the same will turn out to be true. a non-empty solution set to a linear system of equations must be an affine set of some dimension.
- one way to see this is to look at a linear system in echelon form

$$\begin{aligned} 2x_1 + 4x_2 - 3x_3 + 2x_4 + 10x_5 &= -2 \\ 0x_1 + 1x_2 - 5x_3 + 2x_4 - 5x_5 &= 4 \\ 0x_1 + 0x_2 + 0x_3 + 1x_4 - 2x_5 &= 3 \end{aligned}$$

- Using backsubstitution we concluded that the solutions were of the form

$$\begin{aligned} x_4 &= 3 + 2x_5 \\ x_2 &= -2 + 5x_3 + 1x_5 \\ x_1 &= -8 - 17/2x_3 - 9x_5 \end{aligned}$$

  with x_3 and x_5 free.

...

- Let us now express this solution set in vector notation. Setting $x_3 = x_5 = 0$ we find that one solution is

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -8 \\ -2 \\ 0 \\ 3 \\ 0 \end{bmatrix}$$

- More generally, we see that the solution space is the set of vectors of the form

$$\begin{aligned} \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} &= \begin{bmatrix} -8 \\ -2 \\ 0 \\ 3 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -17/2 \\ 5 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_5 \begin{bmatrix} -9 \\ 1 \\ 0 \\ 2 \\ 1 \end{bmatrix} \\ &=: \mathbf{t} + x_3\mathbf{a}_1 + x_5\mathbf{a}_2 \end{aligned}$$

- This is an affine set where V is the space spanned by the $\mathbf{a}_i$.
- Since we said that every linear system can be put into echelon form, this means that any solvable linear system has an affine solution set.

MATLAB

```
clear all;

b = [-2; 4; 3];
bRow = [-2, 4, 3];
A = [2, 4, -3, 2, 10; 0, 1, -5, 2, -5; 0, 0, 0, 1,- 2];
m = 3;
x = zeros(m,1);
c = 2*b + x;
r=rank(A);
```

- In MATLAB you usually do not need to declare variables and types before their use. In case you need to, the best way is by creating a `zero` matrix with the requested number of rows an columns.
- If we want to work with a set of vectors, we can store them as the columns of some matrix `A`. The command `rank` will return the dimension of the space spanned by these vectors.