Math 115A - Week 5
Textbook sections: 1.1-2.5
Topics covered:

- Co-ordinate changes
- Stuff about the midterm

* * * * *

Changing the basis

- In last week's notes, we used bases to convert vectors to co-ordinate vectors, and linear transformations to matrices. We have already mentioned that if one changes the basis, then the co-ordinate vectors and matrices also change. Now we study this phenomenon more carefully, and quantify exactly how changing the basis changes these co-ordinate vectors and matrices.

- Let's begin with co-ordinate vectors. Suppose we have a vector space V and two ordered bases β, β' of that vector space. Suppose we also have a vector v in V. Then one can write v as a co-ordinate vector either with respect to β - thus obtaining $[v]^\beta$ - or with respect to $[v]^{\beta'}$. The question is now: how are $[v]^\beta$ and $[v]^{\beta'}$ related?

- Fortunately, this question can easily be resolved with the help of the identity operator $I_V : V \to V$ on V. By definition, we have

  $$I_V v = v.$$

  We now convert this equation to matrices, but with a twist: we measure the domain V using the basis β, but the range V using the basis β'! This gives

  $$[I_V]_\beta^{\beta'}[v]^\beta = [v]^{\beta'}.$$

  Thus we now know how to convert from basis β to basis β':

  $$[v]^{\beta'} = [I_V]_\beta^{\beta'}[v]^\beta. \quad (0.1)$$

  (If you like, the two bases β have "cancelled each other" on the right-hand side).

- **Example.** Let $V = \mathbf{R}^2$, and consider both the standard ordered basis $\beta := ((1,0),(0,1))$ and a non-standard ordered basis $\beta' := ((1,1),(1,-1))$. Let's pick a vector $v \in \mathbf{R}^2$ at random; say $v := (5,3)$. Then one can easily check that $[v]^\beta = \begin{pmatrix} 5 \\ 3 \end{pmatrix}$ and $[v]^{\beta'} = \begin{pmatrix} 4 \\ -1 \end{pmatrix}$ (why?).

- Now let's work out $[I_V]_\beta^{\beta'}$. Thus we are applying I_V to elements of β and writing them in terms of β'. Since

$$I_V(1,0) = (1,0) = \frac{1}{2}(1,1) + \frac{1}{2}(1,-1)$$

and

$$I_V(0,1) = (0,1) = \frac{1}{2}(1,1) - \frac{1}{2}(1,-1)$$

we thus see that

$$[I_V]_\beta^{\beta'} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

We can indeed verify the formula (0.1), which in this case becomes

$$\begin{pmatrix} 4 \\ -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{pmatrix} \begin{pmatrix} 5 \\ 3 \end{pmatrix}.$$

Note that $[I_V]_\beta^{\beta'}$ is different from $[I_V]_\beta^\beta$, $[I_V]_{\beta'}^{\beta'}$, or $[I_V]_{\beta'}^\beta$. For instance, we have

$$[I_V]_\beta^\beta = [I_V]_{\beta'}^{\beta'} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

(why?), while

$$[I_V]_{\beta'}^\beta = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

(why?). Note also that $[I_V]_{\beta'}^\beta$ is the inverse of $[I_V]_\beta^{\beta'}$ (can you see why this should be the case, without doing any matrix multiplication?)

- **A scalar example**. Let V be the space of all lengths; this has a basis $\beta := (yard)$, and a basis $\beta' := (foot)$. Since the identity I_V applied to a yard yields three feet, we have

$$[I_V]_{(yard)}^{(foot)} = (3)$$

(i.e. the identity on lengths is three feet per yard). Thus for any length v,

$$[v]^{(foot)} = [I_V]^{(foot)}_{(yard)}[v]^{(yard)} = (3)[v]^{(yard)}.$$

Thus for instance, if $[v]^{(yard)} = (4)$ (so v is four yards), then $[v]^{(foot)}$ must equal $(3)(4) = (12)$ (i.e. v is also twelve feet).

- Conversely, we have $[I_V]^{(yard)}_{(foot)} = (1/3)$ (i.e. we have 1/3 yards per foot).

- The matrix $[I_V]^{\beta'}_{\beta}$ is called the *change of co-ordinate matrix from β to β'*; it is the matrix we use to multiply by when we want to convert β co-ordinates to β'. Very loosely speaking, $[I_V]^{\beta'}_{\beta}$ measures how much of β' lies in β (just as $[I_V]^{(foot)}_{(yard)}$ measures how many feet lie in a yard).

- Change of co-ordinate matrices are always square (why?) and always invertible (why?).

- **Example** Let $V := P_2(\mathbf{R})$, and consider the two bases $\beta := (1, x, x^2)$ and $\beta' := (x^2, x, 1)$ of V. Then

$$[I_V]^{\beta'}_{\beta} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

(why?).

- **Example** Suppose we have a mixture of carbon dioxide CO_2 and carbon monoxide CO molecule. Let V be the vector space of all such mixtures, so it has an ordered basis $\beta := (CO_2, CO)$. One can also use just the basis $\beta' = (C, O)$ of carbon and oxygen atoms (where we are ignoring the chemical bonds, etc., and treating each molecule as simply the sum of its components, thus $CO_2 = 1 \times C + 2 \times O$ and $CO = 1 \times C + 1 \times O$). Then we have

$$[I_V]^{\beta'}_{\beta} = \begin{pmatrix} 1 & 1 \\ 2 & 1 \end{pmatrix};$$

this can be interpreted as saying that CO_2 contains 1 atoms of carbon and 2 atoms of oxygen, while CO contains 1 atom of carbon and 1 atom of oxygen. The inverse change of co-ordinate matrix is

$$[I_V]_{\beta'}^{\beta} = ([I_V]_{\beta}^{\beta'})^{-1} = \begin{pmatrix} -1 & 1 \\ 2 & -1 \end{pmatrix};$$

this can be interpreted as saying that an atom of carbon is equivalent to -1 molecules of CO_2 and $+2$ molecules of CO, while an atom of oxygen is equivalent to 1 molecule of CO_2 and -1 molecules of CO. Note how the inverse matrix does not quite behave the way one might naively expect; for instance, given the factor of 2 in the conversion matrix $[I_V]_{\beta}^{\beta'}$, one might expect a factor of $1/2$ in the inverse conversion matrix $[I_V]_{\beta'}^{\beta}$; instead we get strange negative numbers all over the place. (To put it another way, since CO_2 contains two atoms of oxygen, why shouldn't oxygen consist of $\frac{1}{2}$ of a molecule of CO_2? Think about it).

- **Example** (This rather lengthy example is only for physics-oriented students who have some exposure to special relativity; everyone else can safely skip this example). One of the fundamental concepts of special relativity is that space and time should be treated together as a single vector space, and that different observers use different bases to measure space and time.

- For simplicity, let us assume that space is one-dimensional; people can move in only two directions, which we will call right and left. Let's say that observers measure time in years, and space in light-years.

- Let's say there are two observers, Alice and Bob. Alice is an *inertial observer*, which means that she is not accelerating. Bob is another inertial observer, but travelling at a fixed speed $\frac{3}{5}c$ to the right, as measured by Alice; here c is of course the speed of light.

- An *event* is something which happens at a specific point in space and time. Any two events are separated by some amount of time and some amount of distance; however the amount of time and distance that separates them depends on which observer is perceiving. For instance, Alice might measure that event Y occurred 8 years later and 4 light-years to the right of event X, while Bob might measure the distance and duration between the two events differently. (In this case, it turns out that B measures Y as occurring 7 years later and 1 light-year to the *left* of event X).

- Let V denote the vector space of all possible displacements in both space and time; this is a two-dimensional vector space, because we are assuming space to be one-dimensional. (In real life, space is three dimensional, and so spacetime is four dimensional). To measure things in this vector space, Alice has a unit of length - let's call it the *Alice-light-year*, and a unit of time - the *Alice-year*. Thus in the above example, if we call v the vector from event X to event Y, then $v = 8 \times Alice-year + 4 \times Alice-light-year$. (We'll adopt the convention that a displacement of length in the right direction is positive, while a displacement in the left direction is negative). Thus Alice uses the ordered basis $(Alice-light-year, Alice-year)$ to span the space V.

- Similarly, Bob has the ordered basis $(Bob-light-year, Bob-year)$. These bases are related by the *Lorentz transformations*

$$Alice-light-year = \frac{5}{4}Bob-light-year - \frac{3}{4}Bob-year$$

$$Alice-year = -\frac{3}{4}Bob-light-year + \frac{5}{4}Bob-year$$

(this is because of Bob's velocity $\frac{3}{5}c$; different velocities give different transformations, of course. A derivation of the Lorentz transformations from Einstein's postulates of relativity is not too difficult, but is beyond the scope of this course). In other words, we have

$$[I_V]^{(Bob-light-year,Bob-year)}_{(Alice-light-year,Alice-year)} = \begin{pmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{pmatrix}.$$

- Some examples. Suppose Alice emits a flash of light (event X), waits for one year without moving, and emits another flash of light (event Y). Let v denote the vector from X to Y. Then from Alice's point of view, v consists of one year and 0 light-years (because she didn't move between events):

$$[v]^{(Alice-light-year,Alice-year)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

and thus

$$[v]^{(Bob-light-year,Bob-year)} = \begin{pmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 5/4 \\ -3/4 \end{pmatrix};$$

in other words, Bob perceives event Y occurring $5/4$ years afterward and $3/4$ light-years to the left of event X. This is consistent with the assumption that Bob was moving to the right at $\frac{3}{5}c$ with respect to Alice, so that from Bob's point of view Alice is receding to the left at $-\frac{3}{5}c$. This also illustrates the phenomenon of *time dilation* - what appears to be a single year from Alice's point of view becomes $5/4$ years when measured by Bob.

- Another example. Suppose a beam of light was emitted by some source (event A) in a left-ward direction and absorbed by some receiver (event B) some time later. Suppose that Alice perceives event B as occurring one year after, and one light-year to the left of, event A; this is of course consistent with light traveling at $1c$. Thus if w is the vector from A to B, then

$$[v]^{(Alice-light-year,Alice-year)} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

and thus

$$[v]^{(Bob-light-year,Bob-year)} = \begin{pmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 2 \\ -2 \end{pmatrix}.$$

Thus Bob views event B as occurring *two* years after and *two* light-years to the left of event A. Thus Bob still measures the speed of light as $1c$ (indeed, one of the postulates of relativity is that the speed of light is always a constant c to all inertial observers), but the light is "stretched out" over two years instead of one, resulting in Bob seeing the light at half the frequency that Alice would. This is the famous *Doppler red shift* effect in relativity (receding light has lower frequency and is red-shifted; approaching light has higher frequency and is blue-shifted).

- It may not seem like it, but this situation is symmetric with respect to Alice and Bob. We have

$$[I_V]^{(Bob-light-year,Bob-year)}_{(Alice-light-year,Alice-year)} = ([I_V]^{(Bob-light-year,Bob-year)}_{(Alice-light-year,Alice-year)})^{-1}$$

$$= \begin{pmatrix} 5/4 & -3/4 \\ -3/4 & 5/4 \end{pmatrix}^{-1} = \begin{pmatrix} 5/4 & 3/4 \\ 3/4 & 5/4 \end{pmatrix},$$

as one can verify by multiplying the above two matrices together (we will discuss matrix inversion in more detail next week). In other words,

$$Bob - light - year = \frac{5}{4} Alice - light - year + \frac{3}{4} Alice - year$$

$$Bob - year = \frac{3}{4} Alice - light - year + \frac{5}{4} Alice - year.$$

Thus for instance, just as Alice's years are time-dilated when measured by Bob, Bob's years are time-dilated when measured by Bob: if Bob emits light (event X'), waits for one year without moving (though of course still drifting at $\frac{3}{5}c$ as measured by Alice), and emits more light (event Y'), then Alice will perceive event Y' as occurring 5/4 years after and 3/4 light-years to the right of event X'; this is consistent with Bob travelling at $\frac{3}{5}c$, but Bob's year has been time dilated to $\frac{5}{4}$ years. (Why is it not contradictory for Alice's years to be time dilated when measured by Bob, and for Bob's years to be time dilated when measured by Alice? This is similar to the (CO_2, CO) versus (C, O) example: one molecule CO_2 contains two atoms of oxygen (plus some carbon), while an atom of oxygen consists of one molecule of CO_2 (minus some CO), and this is not contradictory. Vectors behave slightly different from scalars sometimes).

* * * * *

Co-ordinate change and matrices

- Let $T : V \to V$ be a linear transformation from a vector space V to itself. (Such transformations are sometimes called *automorphisms*, because they map onto themselves). Given any basis β of V, we can form a matrix $[T]_\beta^\beta$ representing T in the basis β. Of course, if we change the basis, from β to a different basis, say β', then the matrix changes also, to $[T]_{\beta'}^{\beta'}$. However, the two are related.

- **Lemma 1.** Let V be a vector space with two bases β' and β, and let $Q := [I_V]_\beta^{\beta'}$ be the change of co-ordinates matrix from β to β'. Let

$T : V \to V$ be a linear transformation. Then $[T]_\beta^\beta$ and $[T]_{\beta'}^{\beta'}$ are related by the formula

$$[T]_{\beta'}^{\beta'} = Q[T]_\beta^\beta Q^{-1}.$$

- **Proof** We begin with the obvious identity

$$T = I_V T I_V$$

  and take bases (using Theorem 1 from last week's notes) to obtain

$$[T]_{\beta'}^{\beta'} = [I_V]_\beta^{\beta'} [T]_\beta^\beta [I_V]_{\beta'}^\beta.$$

  Substituting $[I_V]_\beta^{\beta'} = Q$ and $[I_V]_{\beta'}^\beta = ([I_V]_\beta^{\beta'})^{-1} = Q^{-1}$ we obtain the Lemma.

- **Example** Let's take a very simple portfolio model, where the portfolio consists of one type of stock (let's say GM stock), and one type of cash (let's say US dollars, invested in a money market fund). Thus a portfolio lives in a two-dimensional space V, with a basis $\beta := (Stock, Dollar)$. Let's say that over the course of a year, a unit of GM stock issues a dividend of two dollars, while a dollar invested in the money market fund would earn 2 percent, so that 1 dollar becomes 1.02 dollars. We can then define the linear transformation $T : V \to V$, which denotes how much a portfolio will appreciate within one year. (If one wants to do other operations on the portfolio, such as buy and sell stock, etc., this would require other linear transformations; but in this example we will just analyze plain old portfolio appreciation). Since

$$T(1 \times Stock) = 1 \times Stock + 2 \times Dollar$$

$$T(1 \times Dollar) = 1.02 \times Dollar$$

  we thus see that

$$[T]_\beta^\beta = \begin{pmatrix} 1 & 0 \\ 2 & 1.02 \end{pmatrix}.$$

  Now let's measure T in a different basis. Suppose that GM's stock has split, so that each old unit of Stock becomes two units of Newstock (so $Newstock = 0.5Stock$). Also, suppose for some reason (decimalization?) we wish to measure money in cents instead of dollars. So we now have a new basis $\beta' = (Newstock, Cent)$. Then we have

$$Newstock = 0.5Stock + 0Dollar; \quad Cent = 0Stock + 0.01Dollar,$$

so

$$Q = [I_V]_\beta^{\beta'} = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.01 \end{pmatrix}$$

while similar reasoning gives

$$Q^{-1} = [I_V]_{\beta'}^{\beta} = \begin{pmatrix} 2 & 0 \\ 0 & 100 \end{pmatrix}.$$

Thus

$$[T]_{\beta'}^{\beta'} = Q[T]_\beta^\beta Q^{-1} = \begin{pmatrix} 0.5 & 0 \\ 0 & 0.01 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2 & 1.02 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 100 \end{pmatrix}$$

which simplifies to

$$[T]_{\beta'}^{\beta'} = \begin{pmatrix} 1 & 0 \\ 100 & 1.02 \end{pmatrix}.$$

Thus

$$T(Newstock) = 1 \times Newstock + 100 \times Cent$$

$$T(Cent) = 0 \times Newstock + 1.02 \times Cent.$$

This can of course be deduced directly from our hypotheses; it is instructive to do so and to compare that with the matrix computation.

- **Definition** Two $n \times n$ matrices A, B are said to be *similar* if one has $B = QAQ^{-1}$ for some invertible $n \times n$ matrix Q.
- Thus the two matrices $[T]_\beta^\beta$ and $[T]_{\beta'}^{\beta'}$ are similar. Similarity is an important notion in linear algebra and we will return to this property later.

* * * * *

Stuff about the midterm

- The midterm will be on Nov 1, at the usual time and place (MS 5127, 1pm to 1:50 pm).

- There are five questions, some of which have two parts. The questions will be a mix of *computational questions* (compute the dimension of this vector space; compute the matrix of this linear transformation; etc.), *verification questions* (prove that this transformation is linear; prove that this set is a subspace; prove that this set is a basis; etc.), *short proofs* (where some hypotheses about vectors, linear transformations, etc. are given, and some conclusion must be deduced), and *example questions* (find a linear transformation, basis, etc. which obeys certain properties). In short, the questions will be much like those in the homework (though a little bit shorter, due to time constraints).

- Calculators are not allowed (and would not be much use anyway; you will not have to do any arithmetic more complicated than manipulation of one-digit numbers. For instance, one would not have to do any extensive Gaussian elimination computations). However, a single 5×7 index card will be allowed.

- Any theorem or lemma which is stated in the notes or in the textbook can be used without proof in the midterm; an exact citation is not required, saying something like "from the lecture notes we know that..." will be sufficient. (The statements of these results are probably a good thing to put on your index card). You will *not* need to memorize any of the proofs of these results, although some of the short proof questions may be similar (but not identical) to some of the shorter proofs given in the lecture notes.

- A practice midterm, with solutions, has been provided on the class web page.

* * * * *

Material covered by the midterm

- The midterm covers everything up to the Week 4 notes, i.e. Sections 1.1-2.4 from the textbook (excluding Section 1.7, which was not covered). In other words, you should be able to understand and work with the following concepts:

- Scalar; vector; vector space; subspace; span; spanning set; linear dependence; linear independence; basis; dimension; linear transformation; null space; nullity; range; rank; ordered basis; co-ordinate vector; matrix representation of a linear transformation; sum, scalar product, and product of linear transformations; sum, scalar product, and product of matrices; one-to-one transformations; onto transformations; invertible transformations (isomorphisms); isomorphic vector spaces.

- You will *not* need to know about the Lagrange interpolation formula, incidence matrices, or change-of-co-ordinate matrices. There will be very little need for Gaussian elimination in the midterm (you can use it, but one can also get by just by using high-school algebra as a substitute). You will *not* need to know the eight axioms (I)-(VIII) of a vector space, and you will *not* need to know any notation which is covered only in the exercises of the textbook. You will *not* need to memorize the proofs of any of the theorems in the text or the notes.

* * * * *

Common sources of confusion

- This course is very much about concepts, and on thinking clearly and precisely about these concepts. It is particularly important not to confuse two concepts which are similar but not identical, otherwise this can lead to one getting hopelessly lost when trying to work one's way through a problem. (This is true not just of mathematics, but also of other languages, such as English. If one confuses similar words - e.g. the adjective "happy" with the noun "happiness" - then one still might be able to read and write simple sentences and still be able to communicate (although you may sound unprofessional while doing so), but complex sentences will become very difficult to comprehend). Here I will list some examples of similar concepts that should be distinguished. These points may appear pedantic, but an inability to separate these concepts is usually a sign of some more fundamental problem in comprehending the material, and should be addressed as quickly as possible.

- **"Vector" versus "Vector space"**. A vector space consists of vectors, but is not actually a vector itself. Thus questions like "What is the dimension of $(1, 2, 3, 4)$?" are meaningless; $(1, 2, 3, 4)$ is a vector, not a vector space, and only vector spaces have a concept of dimension. A question such as "What is the dimension of (x_1, x_1, x_1)?" or "What is the dimension of $x_1 + x_2 + x_3 = 0$?" is also meaningless for the same reason, although "What is the dimension of $\{(x_1, x_1, x_1) : x_1 \in \mathbf{R}\}$?" or "What is the dimension of $\{(x_1, x_2, x_3) : x_1 + x_2 + x_3 = 0\}$?" are not.

- In a similar spirit, the zero vector 0 is distinct from the zero vector space $\{0\}$, and is in turn distinct from the zero linear transformation T_0. (And then there is also the zero scalar 0).

- **A set S of vectors, versus the span** $\text{span}(S)$ **of that set**. This is a similar problem. A statement such as "What is the dimension of $\{(1,0,0), (0,1,0), (0,0,1)\}$?" is meaningless, because the set $\{(1,0,0), (0,1,0), (0,0,1)\}$ is not a vector space. It is true that this set *spans* a vector space - $\mathbf{R}^3$ in this example - but that is a different object. Similarly, it is not correct to say that the sets $\{(1,0,0), (0,1,0), (0,0,1)\}$ and $\{(1,1,0), (1,0,1), (0,0,1)\}$ are equal - they may span the same space, but they are certainly not the same set.

- For a similar reason, it is not true that if you add $(0,0,1)$ to $\mathbf{R}^2$, that you "get" $\mathbf{R}^3$. First of all, $\mathbf{R}^2$ is not even contained in $\mathbf{R}^3$, although the xy-plane $V := \{(x, y, 0) : x, y, \in \mathbf{R}\}$, which is isomorphic to $\mathbf{R}^2$, *is* contained in $\mathbf{R}^3$. But also, if you take the union of V with $\{(0,0,1)\}$, one just gets the set $V \cup \{(0,0,1)\}$, which is a plane with an additional point added to it. The correct statement is that V and $(0,0,1)$ together *span* or *generate* $\mathbf{R}^3$.

- **"Finite" versus "finite-dimensional"**. A set is *finite* if it has finitely many elements. A vector space is *finite-dimensional* if it has a finite basis. The two notions are distinct. For instance, $\mathbf{R}^2$ is infinite (there are infinitely many points in the plane), but is finite-dimensional because it has a finite basis $\{(1,0), (0,1)\}$. The zero vector space $\{0\}$ contains just one element, but is zero-dimensional. The set $\{(1,0,0), (0,1,0), (0,0,1)\}$ is finite, but it does not have a dimension because it is not a vector space.

- **"Subset" versus "subspace"** Let V be a vector space. A *subset* U of V is any collection of vectors in V. If this subset is also closed under addition and scalar multiplication, we call U a *subspace*. Thus every subspace is a subset, but not conversely. For instance, $\{(1,0,0),(0,1,0),(0,0,1)\}$ is a subset of $\mathbf{R}^3$, but is not a subspace (so it does not have a dimension, for instance).

- **"Nullity" versus "null space"** The *null space* $N(T)$ of a linear transformation $T: V \to W$ is defined as the space $N(T) = \{v \in V : Tv = 0\}$; it is a subspace of V. The *nullity* nullity(T) is the dimension of $N(T)$; it is a *number*. So statements such as "the null space of T is three" or "the nullity of T is $\mathbf{R}^3$" are meaningless. Similarly one should distinguish "range" and "rank"

- **"Range $R(T)$" versus "final space"** This is an unfortunate notation problem. If one has a transformation $T: V \to W$, which maps elements in V to elements in W, V is sometimes referred to as the *domain* and W is referred to as the *range*. However, this notation is in conflict with the range $R(T)$, defined by $R(T) = \{Tv : v \in V\}$. If the transformation T is *onto*, then $R(T)$ and W are the same, but otherwise they are not. To avoid this confusion, I will try to refer to V as the *initial space* of T, and W as the *final space*. Thus in the map $T: \mathbf{R}^2 \to \mathbf{R}^3$ defined by $T(x_1, x_2) := (x_1, x_2, 0)$, the final space is $\mathbf{R}^3$, but the range is only the xy-plane, which is a subspace of $\mathbf{R}^3$.

- To re-iterate, in order to be able to say that a transformation T maps V to W, it is not required that every element of W is actually covered by T; if that is the case, we say that T maps V *onto* W, and not just *to* W. It is also not required that different elements of V have to map to different elements of W; that is true for *one-to-one* transformations, but not for general transformations. So in particular, if you are given that $T: V \to W$ is a linear transformation, you cannot necessarily assume that T is one-to-one or onto unless that is explicitly indicated in the question.

- **"Vector" versus "Co-ordinate vector"** A vector v in a vector space V is not the same thing as its co-ordinate vector $[v]^\beta$ with respect to some basis β of V. For one thing, v needn't be a traditional vector (a row or column of numbers) at all; it may be a polynomial, or a matrix, or a function - these are all valid examples of vectors in a vector space.

Secondly, the co-ordinate vector $[v]^\beta$ depends on the basis β as well as on the vector v itself; change β and you change the co-ordinate vector $[v]^\beta$, even if v itself is kept fixed. For instance, if $V = P_2(\mathbf{R})$ and $v = 3x^2+4x+5$, then v is a vector (even though it's also a polynomial). If $\beta := (1, x, x^2)$, then $[v]^\beta = \begin{pmatrix} 5 \\ 4 \\ 3 \end{pmatrix}$, but this is not the same object as $3x^2 + 4x + 5$; one is a column vector and the other is a polynomial. Calling them "the same" will lead to trouble if one then switches to another basis, such as $\beta' := (x^2, x, 1)$, since the co-ordinate vector is now $[v]^{\beta'} = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix}$, which is clearly different from $\begin{pmatrix} 5 \\ 4 \\ 3 \end{pmatrix}$.

- There is of course a similar distinction between a linear transformation T, and the matrix representation $[T]_\beta^\gamma$ of that transformation T.

- **"Closed under addition" versus "preserves addition"**. A set U is *closed under addition* if, whenever x and y are elements of U, the sum $x + y$ is also in U. A transformation $T : V \to W$ *preserves addition* if, whenever x and y are elements in V, that $T(x + y) = Tx + Ty$. The two concepts refer to two different things - one has to do with a *set*, and the other with a *transformation* - but surprisingly it is still possible to confuse the two. For instance, in one of the homework questions, one has to show that the *set* $T(U)$ is closed under addition, and someone thought this meant that one had to show that $T(x+y) = Tx+Ty$ for all x in U; probably what happened was that the presence of the letter T in the set $T(U)$ caused one to automatically think of the preserving addition property rather the closed-under-addition property. Of course, there is a similar distinction between "closed under scalar multiplication" and "preserves scalar multiplication".

- **A vector v, versus the image Tv of that vector**. A linear transformation $T : V \to W$ can take any vector v in V, and transform it to a new vector Tv in W. However, it is dangerous to say things like v "becomes" Tv, or v "is now" Tv. If one is not careful, one may soon write that v "is" Tv, or think that every property that v has, automatically also holds for Tv. The correct thing to say is that Tv is the *image* of v under T; this image may preserve some properties of the original vector v, but it may distort or destroy others. In general, Tv is not equal to v (except in special cases, such as when T is the identity transformation).

- **Hypothesis versus Conclusion**. This is not a confusion specific to linear algebra, but nevertheless is an important distinction to keep in mind when doing any sort of "proof" type question. You should always know what you are currently *assuming* (the hypotheses), and what you are trying to prove (the conclusion). For instance, if you are trying to prove

  Show that if $T : V \to W$ is linear, then
  $T(cx + y) = cTx + Ty$ for all $x, y \in V$ and all scalars c

  then your hypothesis is that $T : V \to W$ is a linear transformation (so that $T(x + y) = Tx + Ty$ and $T(cx) = cTx$ for all $x, y \in V$ and all scalars c, and your objective is to prove that for any vectors $x, y \in V$ and any scalar c, we have $T(cx + y) = cTx + Ty$.

- On the other hand, if you are trying to prove that

  Show that if $T : V \to W$ is such that $T(cx + y) = cTx + Ty$
  for all $x, y \in V$ and all scalars c, then T is linear.

  then your hypotheses are that T maps V to W, and that $T(cx + y) = cTx + Ty$ for all $x, y \in V$ and all scalars c, and your objective is to prove that T is linear, i.e. that $T(x + y) = Tx + Ty$ for all vectors $x, y \in V$, and that $T(cx) = cTx$ for all vectors $x \in V$ and scalars c. This is a completely different problem from the previous one. (Part 2. of Question 7 of 2.2 requires you to prove *both* of these implications, because it is an "if and only if" question).

- In your proofs, it may be a good idea to identify which of the statements that you write down are things that you *know* from the hypotheses, and which ones are those that you *want*. Little phrases like "We are given that", "It follows that", or "We need to show that" in the proof are very helpful, and will help convince the grader that you actually know what you are doing! (provided that those phrases are being used correctly, of course).

- **"For all" versus "For some"**. Sometimes it is really important to read the "fine print" of a question - it is all to easy to jump to the equations without reading all the English words which surround those equations. For instance, the statements

  "Show that $Tv = 0$ for some non-zero $v \in V$"

  is completely different from

  "Show that $Tv = 0$ for all non-zero $v \in V$"

  In the first case, one only needs to exhibit a *single* non-zero vector v in V for which $Tv = 0$; this is a statement which could be proven by an example. But in the second case, no amount of examples will help; one has to show that $Tv = 0$ for *every* non-zero vector v. In the second case, probably what one would have to do is start with the hypotheses that $v \in V$ and that $v \neq 0$, and somehow work one's way to proving that $Tv = 0$.

- Because of this, it is very important that you read the question carefully before answering. If you don't understand exactly what the question is asking, you are unlikely to write anything for the question that the grader will find meaningful.