Review notes on coordinates, change of basis

1. Introduction

Vector spaces come from many applications, including some where they are not just vectors of numbers, for example solution spaces of differential equations. But to compute with them we need to have numbers! To get numbers, we need coordinates. To get coordinates, we need a basis.

In Math 115A, there are three kinds of situations in which we get numbers by using coordinates.

- In a vector space V with basis B, for each $\mathbf{v} \in V$ we get a *coordinate vector* $[\mathbf{v}]_B$ (a column vector).
- For a linear transformation $T : V \to W$, if B is a basis of V and C is a basis of W, then we get a matrix $[T : B, C]$.
- For two bases B, B' of one vector space V, we get a *transition* matrix $[B'/B]$ (a square matrix, also called a *change-of-basis* matrix).

Examples.

In the following, two main examples of vector spaces will be used:

- V is the solution space of $y'' = y$ (rather than $y'' = -y$). As you know, there are two good bases. One is e^x, e^{-x}. The other is $\cosh x, \sinh x$, where $\cosh x = \frac{1}{2}(e^x + e^{-x})$ (the average of e^x and e^{-x}) and $\sinh x = \frac{1}{2}(e^x - e^{-x})$. Both bases are useful in applications.

  The purpose of this example is to show why it is necessary to talk about vector spaces in a general context rather than just sticking to $\mathbf{R}^n$: The "vectors" in V are functions and not n-tuples of numbers.

  As an example of a linear transformation from V to V we can use differentiation, since you know it well and yet it doesn't come from a matrix to start with. Differentiation is a linear transformation because of the rules $(p + q)' = p' + q'$ and$(cp)' = cp'$, where c is any constant.

- V is $\mathbf{R}^n$ and the basis is the standard basis. In this kind of example you should have the intuition that "everything is going to come out right". But what does that mean in each application?

For linear transformations, watch to see whether we are talking about

- The general case $T : V \to W$, where the bases in V and W are assumed to be different, or
- the case $T : V \to V$, where $W = V$ and we usually use the same basis on both sides of the arrow. In this case, T is often called a *linear operator*, since it "operates" on V.

2. Coordinates

If V is a vector space with a basis $\mathbf{v}_1, \cdots \mathbf{v}_n$, then, as you know, each vector $\mathbf{v}$ in V can be written uniquely as

$$\mathbf{v} = r_1\mathbf{v}_1 + \cdots + r_n\mathbf{v}_n$$

In this case we say that $\mathbf{v}$ has *coordinates* $r_1, \ldots, r_n$ relative to this basis. The *coordinate vector* of $\mathbf{v}$ is a column vector:

$$[\mathbf{v}]_B = \begin{bmatrix} r_1 \\ \cdots \\ r_n \end{bmatrix},$$

where B denotes the basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$. (The book calls this an "ordered basis" because the order in which you list the basis vectors makes a difference.)

Example 2.1. For $V = \mathbf{R}^2$, find the coordinate vector of $(2,3)$ relative to the basis $(1,2)$ and $(3,4)$.

Solution. Write $(2,3) = r_1(1,2) + r_2(3,4)$ and treat r_1, r_2 as unknowns. This gives simultaneous equations in r_1 and r_2. Solving them, you get $r_1 = \frac{1}{2}, r_2 = \frac{1}{2}$, so the answer is $\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$.

Example 2.2. For $V = P(1)$, find the coordinate vector of $2 + 3x$ relative to the basis $1 + x, 3 + 4x$.

Solution. You can write $2 + 3x = r_1(1 + x) + r_2(3 + 4x)$, equate coefficients (remembering that x is for the polynomials, not the unknown), and solve to get the answer $\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$. Or, you can use the isomorphism $P(1) \cong \mathbf{R}^2$ to translate the problem into $\mathbf{R}^2$, where it is the same as Example 2.1.

Example 2.3. If $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$ is a basis of V, what are the coordinates of $\mathbf{v}_2$ relative to this basis?

Solution. Easy: $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$.

Example 2.4. If we think of $\mathbf{R}^3$ as the space of all column vectors of length 3, what is the coordinate vector of $\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$ relative to the standard basis $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$?

Solution. Easy: $\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = 2\mathbf{e}_1 + 3\mathbf{e}_2 + 4\mathbf{e}_3$, so we get the coordinate vector $\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$. As you see, if you use the standard basis in $\mathbf{R}^n$, any vector is its *own* coordinate vector, so everything works out consistently.

3. A higher point of view.

You can think of $\mathbf{R}^n$ as n-dimensional graph paper. Most other vector spaces are best thought of as blank paper with just an origin.

When you choose a basis for an n-dimensional vector space V, you are giving it coordinates, so you are essentially laying n-dimensional graph paper on V. Mathematically, this is the idea of *isomorphism*, a one-to-one correspondence that preserves the vector operations of addition and multiplication by scalars: $V \cong \mathbf{R}^n$, using $\mathbf{v} \leftrightarrow [\mathbf{v}]_B$.

Example 3.1. An easy isomorphism: $P(2) \cong \mathbf{R}^3$ using $a + bx + cx^2 \leftrightarrow \begin{bmatrix} a \\ b \\ c \end{bmatrix}$.

Here $x \leftrightarrow \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, for example.

Example 3.2. A different, harder isomorphism: $P(1) \cong \mathbf{R}^2$ using the basis $x+1, x-1$ for $P(1)$. In this case, what does x correspond to? Notice that $x = \frac{1}{2}(x+1) + \frac{1}{2}(x-1)$ (you could solve for this as in Example 3.2), so $x \leftrightarrow \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$.

4. Matrices for linear transformations.

The easiest kind of linear transformation is a *matrix transformation*: Given an $m \times n$ matrix A, we make a linear transformation $T_A : \mathbf{R}^n \to \mathbf{R}^m$ by matrix multiplication:

$T_A(\mathbf{x}) = A\mathbf{x}$.

Here $\mathbf{x}$ *must* be written as a column vector for the multiplication to make sense. You saw how matrix transformations can be used to do rotations, reflections, projections, and other kinds of transformations. In fact, we proved in class that *any* linear transformation $\mathbf{R}^n \to \mathbf{R}^m$ has the form T_A for a suitable $m \times n$ matrix A. It is easy to find the matrix A, because the columns of A are the values $T(\mathbf{e}_1), \ldots, T(\mathbf{e}_n)$.

Now consider any linear transformation $T : V \to W$, where V has dimension n and W has dimension m. To start with, there is no matrix in sight.

Choose a basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ for V and call it B. Choose a basis $\mathbf{w}_1, \ldots, \mathbf{w}_m$ for W and call it C.

Now we can get a matrix of numbers to describe $T : V \to W$.

Start with $\mathbf{v}_1$ and look at $T(\mathbf{v}_1)$, which is in W. The coordinate vector of $T(\mathbf{v}_1)$ relative to the basis C is a column vector.

Now take $\mathbf{v}_2$; the coordinate vector of $T(\mathbf{v}_2)$ relative to the basis C is another column vector.

Now take $\mathbf{v}_3$, and so on. You get n column vectors in all, each of height m, so it seems natural to put them side-by-side to make an $m \times n$ matrix. This is the *matrix of T relative to the bases B* and C, denoted $[T : B, C]$ in the text. It describes T using numbers.

Example 4.1. Let V be the solution space of $y'' = y$ and let $T : V \to V$ be differentiation. On both sides of the arrow, use the basis B consisting of e^x, e^{-x}. Then we get the matrix $[T : B, B] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$. (Why?)

Example 4.2. Suppose we *start* with a matrix A, look at $T_A : \mathbf{R}^n \to \mathbf{R}^m$, and then take the matrix of T_A with respect to the standard bases in $\mathbf{R}^n$, $\mathbf{R}^m$. Do we get A back? *Yes!*, because the first column of the matrix of T_A is $T_A(\mathbf{e}_1) = A\mathbf{e}_1 =$ the first column of A, and so on.

Example 4.3. Take a matrix A again but look at the matrix of A relative to nonstandard bases. We expect to get some new matrix.

Specifically, take $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$, consider $T_A : \mathbf{R}^2 \to \mathbf{R}^2$, and use the basis $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ for both B and C.

We need to compute $T_A(\mathbf{v}_1) = A\mathbf{v}_1 = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$ and express it relative to the basis $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}$; we get $\begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4 \begin{bmatrix} 1 \\ 1 \end{bmatrix} + 0 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, so the coordinate vector is $\begin{bmatrix} 4 \\ 0 \end{bmatrix}$.

Next we need to compute $T_A(\mathbf{v}_2) = A\mathbf{v}_2 = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \begin{bmatrix} -2 \\ 2 \end{bmatrix}$ and express it relative to our basis; we get the coordinate vector $\begin{bmatrix} 0 \\ 2 \end{bmatrix}$.

Therefore the answer is the matrix $\begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$. Notice that in this example we get a diagonal matrix, which is always good. For this particular matrix A, then, the basis $\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ is actually a better basis than the standard basis!

Problem S-1. Let $V = W = P(3)$ and let $T(p) = p'$ (the derivative). Choose the same basis B consisting of $1, x, x^2, x^3$ for both V and W. Find $[T : B, B]$.

5. A higher point of view

We started with a linear transformation $T : V \to W$, chose bases for V and W, and got a resulting matrix for T. Is this really a new concept or does it really tie in to our previous idea of a matrix transformation?

When we chose a basis B for V, we were laying graph paper on V using an isomorphism. When we chose a basis C for W, we were laying graph paper on W using an isomorphism. The linear transformation $T : V \to W$ then essentially becomes a linear transformation from graph paper to graph paper (in other words, $\mathbf{R}^n$ to $\mathbf{R}^m$), and we know that such a linear transformation is a matrix transformation T_A for some matrix A. Is this the same matrix as $[T : B, C]$? Yes!

You can see this in the diagram of Figure 1.

$$\begin{tikzcd} \mathbf{R}^n \arrow[r, dashed, "T_A"] \arrow[d, leftrightarrow] & \mathbf{R}^m \arrow[d, leftrightarrow] \\ V \arrow[r, "T"] & W \end{tikzcd}$$

Figure 1: A higher point of view

In fact, by following the arrows from $\mathbf{R}^n$ to V to W to $\mathbf{R}^m$, the composition is a linear transformation from $\mathbf{R}^n$ to $\mathbf{R}^m$, and that is T_A for the appropriate A.

Problem. Check that just the first column of A is the same as the first column of $[T : B, C]$. To do this, think how you get the first column of $[T; B, C]$ and compare it to $T_A(\mathbf{e}_1) = A\mathbf{e}_1$, which is the first column of A.

This explains why we had to use column vectors for coordinate vectors, instead of row vectors. Row vectors would give a transposed matrix.

6. Change of basis

Suppose we have *one* vector space V and *two* bases, an "old basis" B and a "new basis" B'. What is a good way of describing B' relative to B? Easy: Just take each vector in B' and find its coordinate vector relative to B, a column vector; put these column vectors as the columns of a matrix. This is the *transition matrix* or *change-of-basis matrix* that describes B' relative to B.

Example 6.1. Let V be the solution space of the DE $y'' = y$. Let B be the basis e^x, e^{-x} and let the "new" basis B' be $\cosh x, \sinh x$.

As we can see, $[B'/B] = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$, since the first column tells how to express $\cosh x$ in terms of e^x, e^{-x}, and the second column tells how to express $\sinh x$. What is $[B/B']$, the old in terms of the new? It's the inverse of $[B'/B]$, which comes out $\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$. This says $e^x = \cosh x + \sinh x$ and $e^{-x} = \cosh x - \sinh x$, which is true.

Example 6.2. Let $V = \mathbf{R}^n$, let the "old" basis be the standard basis, and let the "new" basis be some other basis; it doesn't matter what. Let P be the transition matrix. Then the columns of P *are* the new basis. (Why? Try a 2×2 example with some actual numbers.)

7. How a change of basis affects the matrix of a linear transformation

Suppose we have a linear transformation $T : V \to W$ and a basis B for V and a basis C for W. Then suppose we change to a new basis B' for V and a new basis C' for W. We have

- a matrix $A = [T : B, C]$ representing T relative to the "old" bases B and C.

- a matrix $A' = [T : B', C']$ representing T relative to the "new" bases B and C.

We also have

- a transition matrix $P = [B'/B]$ giving the new basis of V relative to the old basis,
- a transition matrix $Q = [C'/C]$ giving the new basis of W relative to the old basis.

There must be some way to find A' from A, P, and Q. What is it? It turns out that we have this formula:

$$A' = Q^{-1}AP.$$

In many cases, $V = W$ and $B = C$, $B' = C'$. In this case, we get

$$A' = P^{-1}AP.$$

Example 7.1. Let V be the solution space of $y'' = y$ and let $T : V \to V$ be differentiation. On both sides of the arrow, have the "old" basis B be e^x, e^{-x} and the "new" basis B' be $\cosh x, \sinh x$.

Then $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$ and $P = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & -\frac{1}{2} \end{bmatrix}$. (Why?) The formula gives $A' = P_{-1}AP = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, which is correct. (Why?)

Example 7.2. Consider $T_A : \mathbf{R}^2 \to \mathbf{R}^2$ for $A = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$. Have the "old" basis B be the standard basis and the "new" basis B' be $\mathbf{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\mathbf{v}_2 = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Then the "old" matrix for T_A is A and $P = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$. The formula gives $A' = P^{-1}AP = \begin{bmatrix} 4 & 0 \\ 0 & 2 \end{bmatrix}$, which is correct, since $T_A(\mathbf{v}_1) = A\mathbf{v}_1 = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4\mathbf{v}_1 = 4\mathbf{v}_1 + 0\mathbf{v}_2$, and similarly, $T_A(\mathbf{v}_2) = 2\mathbf{v}_2$.

In this example, notice that just one transformation T_A is used the whole time, but the matrix describing the transformation does change when the transformation is described relative to a new basis.

Note. There is a "higher point of view" that explains the change-of-basis formula, but since the diagram is more complicated we won't do it. (Ask if you are interested.)

8. Test yourself

Problem S-2. What are the ways we get numbers that describe (i) a vector in a vector space, (ii) a linear transformation, (iii) a change of basis? (Be able to be specific about each.)

Problem S-3. Be able to explain the $y'' = y$ example for each of (i), (ii), (iii).

Problem S-4. Explain why you get an $m \times n$ matrix for $T : V \to W$ when V has dimension n and W has dimension m, rather than the other way around (m and n).

Problem S-5. How does a change of basis affect the matrix of a linear transformation?

Problem S-6. Write the first formula from §7 in the notation of the text, $[T : B', C'] = \cdots$.