POINT, COORDINATE, AND CHANGE OF BASIS TRANSFORMATIONS

1. Point and Coordinate Transformations

We begin with the vector space $\mathcal{P}_2[x]$ of quadratic polynomials, along with the basis $\mathcal{B} = \{1, x, x^2\}$. We define the *point transformation* $T^P_{\mathcal{B}} : \mathbb{R}^3 \to \mathcal{P}_2[x]$ defined by $T^P_{\mathcal{B}}([a_1, a_2, a_3]^t) = a_1 + a_2 x + a_3 x^2$. This is a linear transformation from $\mathbb{R}^3$ to $\mathcal{P}_2[x]$ (The transformation is a function of a_1, a_2, a_3, and it is linear in these. It certainly is not linear in the variable x!) The point transformation applied to the point $[0, 1, 2]^t$ gives us the polynomial which has 0 of the first basis vector 1, 1 of the second basis vector x, and 2 of the third basis vector x^2. Let us remark that when we speak of basis here we are assuming we have fixed an ordering on the elements in the basis. Thus, if we let $\mathcal{B}' = \{1, x^2, x\}$, then the corresponding point transformation would be $T^P_{\mathcal{B}'}([a_1, a_2, a_3]^t) = a_1 + a_2 x^2 + a_3 x$. In particular,

$$T^P_{\mathcal{B}}([0, 1, 2]^t) = x + 2x^2$$

whereas

$$T^P_{\mathcal{B}'}([0, 1, 2]^t) = x^2 + 2x.$$

Given any basis for $\mathcal{P}_2[x]$ we get a corresponding point transformation. Let us give one more basis, $\mathcal{B}'' = \{1, x^2 + x, x^2 - x\}$. Then

$$T^P_{\mathcal{B}''}([a_1, a_2, a_3]^t) = a_1 + a_2(x^2 + x) + a_3(x^2 - x).$$

Next we wish to go the other way. Given any quadratic polynomial we can express it as a linear combination of basis vectors for $\mathcal{P}_2[x]$, and there is only one such linear combination. We define the *coordinate transformation of $\mathcal{P}_2[x]$ with respect to the basis $\mathcal{B}$* by $T^C_{\mathcal{B}} : \mathcal{P}_2[x] \to \mathbb{R}^3, T^C_{\mathcal{B}}(b_1 + b_2 x + b_2 x^2) = [b_1, b_2, b_3]^t$. Thus, given a polynomial, we find the coefficients for each basis vector, and only keep track of the coefficients. Again, this will depend on the basis we use. In particular $T^C_{\mathcal{B}'}(b_1 + b_2 x + b_2 x^2) = [b_1, b_3, b_2]^t$ and $T^C_{\mathcal{B}''}(b_1 + b_2 x + b_2 x^2) = [b_1, \frac{b_2+b_3}{2}, \frac{b_3-b_2}{2}]^t$. Why? For $\mathcal{B}'$ we just swapped the order of two of the basis elements. We will give an example to illustrate what's going on a little better, especially for $\mathcal{B}''$.

Let us use the polynomial $p(x) = 3x + 4x^2$ to illustrate. First we write this in the basis $\mathcal{B}$. We have $p(x) = 0 \cdot 1 + 3x + 4x^2$ and so $T^C_{\mathcal{B}}(p(x)) = [0, 3, 4]^t$. Next we write $p(x)$ in terms of the second bases. We have $p(x) = 0 \cdot 1 + 4x^2 + 3x$, and so $T^C_{\mathcal{B}'}(p(x)) = [0, 4, 3]^t$. Next we write $p(x)$ in terms of the third basis. We need to find c_1, c_2, c_3 so that

$$3x + 4x^2 = c_1 \cdot 1 + c_2(x^2 + x) + c_3(x^2 - x).$$

Equating like coefficients, we find that $c_1 = 0, c_2 - c_3 = 3$ and $c_2 + c_3 = 4$. Therefore, $c_2 = \frac{7}{2}$ and $c_3 = \frac{1}{2}$. $T^C_{\mathcal{B}''}(p(x))$ gives us the 3-tuple of c's, so $T^C_{\mathcal{B}''}(p(x)) = [0, \frac{7}{2}, \frac{1}{2}]^t$.

There is a very nice relationship between the point and coordinate transformations for a *fixed* basis. Let us check in the case of the third basis.

$$T^C_{\mathcal{B}''} \circ T^P_{\mathcal{B}''}([a_1, a_2, a_3]^t) = T^C_{\mathcal{B}''}(T^P_{\mathcal{B}''}([a_1, a_2, a_3]^t))$$
$$= T^C_{\mathcal{B}''}(a_1 \cdot 1 + a_2(x^2 + x) + a_3(x^2 - x)) = [a_1, a_2, a_3]^t.$$

So T^C is un-doing T^P. This also works in the other direction, meaning T^P undoes T^C :

$$T^P_{\mathcal{B}''} \circ T^C_{\mathcal{B}''}(b_1 + b_2x + b_3x^2) = T^P_{\mathcal{B}''}(T^C_{\mathcal{B}''}(b_1 + b_2x + b_3x^2))$$
$$= T^P_{\mathcal{B}''}([b_1, \frac{b_2 + b_3}{2}, \frac{b_3 - b_2}{2}]^t) = b_1 + \frac{b_2 + b_3}{2}(x^2 + x) + \frac{b_3 - b_2}{2}(x^2 - x)$$
$$= b_1 + b_2x + b_3x^2.$$

There is nothing special about $\mathcal{P}_2[x]$ here. Given any finite dimensional vector space $\mathcal{V}$ and a basis $\mathcal{B} = \{v_1, v_2, \cdots, v_n\}$ we can define point and coordinate transformations $T^P_{\mathcal{B}} : \mathbb{R}^n \to \mathcal{V}$ by $T^P_{\mathcal{B}}([a_1, a_2, \cdots, a_n]^t) = a_1v_1 + a_2v_2 + \cdots + a_nv_n$, and $T^C_{\mathcal{B}} : \mathcal{V} \to \mathbb{R}^n$, by $T^C_{\mathcal{B}}(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = [a_1, a_2, \cdots, a_n]^t$. Notice we use $\mathbb{R}^n$ when our vector space is of dimension n. Again, these depend not only on the vector space $\mathcal{V}$ but on the basis $\mathcal{B}$, and in fact the ordering of the elements of the basis. Also, $T^P_{\mathcal{B}} \circ T^C_{\mathcal{B}}(a_1v_1 + a_2v_2 + \cdots + a_nv_n) = a_1v_1 + a_2v_2 + \cdots + a_nv_n$ and $T^C_{\mathcal{B}} \circ T^P_{\mathcal{B}}([a_1, a_2, \cdots, a_n]^t) = [a_1, a_2, \cdots, a_n]^t$, so the point and coordinate transformations undo each other again.

These transformations have a deep philosophical impact on finite dimensional vector spaces, namely that all vector spaces of dimension n look like $\mathbb{R}^n$, at least as far as *linear* algebra is concerned. (Actually, I should say once we choose a basis for our vector space, then it looks like $\mathbb{R}^n$, but the precise way in which it looks like $\mathbb{R}^n$ depends heavily on the basis we chose.) For instance, if I want to see if the polynomials $\{x+3, x-4, x+7\}$ are independent, then I first choose a basis $\mathcal{B}$, say $\{1, x, x^2\}$. Then take the coordinate transformations, to get the points $\{[3, 1, 0]^t, [-4, 1, 0]^t, [7, 1, 0]^t\}$. Then I check these for dependencies. Since the coordinate and point transformations are linear and they undo each other, then the set of polynomials $\{x + 3, x - 4, x + 7\}$ is dependent precisely if the set of vectors $\{[3, 1, 0]^t, [-4, 1, 0]^t, [7, 1, 0]^t\}$ is dependent. I should remark, however, that it is not always best to think of all n dimensional vector spaces as being $\mathbb{R}^n$. For instance, $\mathcal{M}(2, 2)$ is a 4-dimensional vector space, so after using coordinate transformations, it looks just like $\mathbb{R}^4$. But, this set of matrices has a richer algebraic structure than that of vector space since we can multiply matrices as well. This additional structure is lost under the coordinate transformation.

2. Change of Basis Matrices

Often times we will have some vector space with a good basis. Think of $\mathbb{R}^n$ with the standard basis or $\mathcal{P}_d[x]$ with the monomial basis $\{1, x, x^2, \cdots, x^d\}$. However, for the particular problem we are studying, it might be better to use some other basis. (Chapter 5 will be concerned with locating things called *eigenbases*, in which we will need to be able to go from the standard basis in $\mathbb{R}^n$ into one of these eigenbases.) Thus, we would like to have some systematic way of converting between two bases for the same space.

Let $\mathcal{V}$ be a vector space with two bases $\mathcal{B} = \{v_1, v_2, \cdots, v_n\}$ and $\mathcal{B}' = \{w_1, w_2, \cdots, w_n\}$. We define the change of basis transformation $C_{\mathcal{B}'\leftarrow\mathcal{B}} : \mathbb{R}^n \to \mathbb{R}^n$ as follows: First, start with $[a_1, a_2, \cdots, a_n]^t \in \mathbb{R}^n$. Apply the point transformation for the basis $\mathcal{B}$. This gives the vector space element $v = a_1v_1 + a_2v_2 + \cdots + a_nv_n$. Next, since $\mathcal{B}'$ is another basis for $\mathcal{V}$ there exists exactly one choice of constants $b_1, b_2, \cdots, b_n$ so that $v = b_1w_1 + b_2w_2 + \cdots + b_nv_n$. Then this string of b's is the output, $[b_1, b_2, \cdots, b_n]^t$. Another way to say all this is that we first apply the point transformation $T^P_{\mathcal{B}}$, then we apply the coordinate transformation $T^C_{\mathcal{B}'}$. Observe that this is not necessarily the identity operation. The point and coordinate transformations only undo each other if they are relative to the same basis! Here we are using different bases for the two transformations.

We will mostly be interested in doing this when our vector space is $\mathbb{R}^n$, and one of the bases is the standard basis. For example, consider $\mathbb{R}^2$ with $\mathcal{B} = \{[1, 0]^t, [1, 1]^t\}$ and $\mathcal{B}' = \{[1, 0]^t, [0, 1]^t\}$. Then let us find what $C_{\mathcal{B}'\leftarrow\mathcal{B}}$ does to the vector $[x, y]^t$. At first this is going to seem confusing, but bear with me! The vector $[x, y]^t$ that I input into C does not denote a point in the vector space $\mathbb{R}^2$ that I am interested in. It gives the coordinates for a point in that vector space relative to the basis $\mathcal{B}$. Thus, the input vector $[x, y]^t$ represents the point $x[1, 0]^t + y[1, 1]^t$ in the vector space I am in. $x[1, 0]^t + y[1, 1]^t$ is a point in my vector space. Combining like entries, I find $x[1, 0]^t + y[1, 1]^t = [x + y, y]^t$. Now I want to find the coordinates of the point $[x + y, y]^t$ relative to the standard basis. So, I write $[x + y, y]^t = (x + y)[1, 0]^t + y[0, 1]^t$. Finally, I interpret these coefficients as a coordinate vector, $[x + y, y]^t$. So,

$$C_{\mathcal{B}'\leftarrow\mathcal{B}} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+y \\ y \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix},$$

and so

$$C_{\mathcal{B}'\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$

Notice the first column of the matrix is the first basis vector in $\mathcal{B}$ and the second column is the second basis vector in $\mathcal{B}$. *This only works out this way if we are going from some basis into the standard basis!*

Suppose now I want to go from the basis $\mathcal{B}'$ to the basis $\mathcal{B}$. That is, I now want to find $C_{\mathcal{B}\leftarrow\mathcal{B}'}$. I could try to go through all of this again. However, I don't have to if I already know how to go from $\mathcal{B}$ to $\mathcal{B}'$. There is a simple relation;

$$C_{\mathcal{B}\leftarrow\mathcal{B}'} = (C_{\mathcal{B}'\leftarrow\mathcal{B}})^{-1}.$$

Now suppose I want to go between two arbitrary bases. Let $\mathcal{B}'' = \{[1,0]^t, [-1,1]^t\}$, and suppose I want to go from $\mathcal{B}$ to $\mathcal{B}''$. I could try to compute this matrix as I did above. However, there is a simpler way, based on the fact that it is easy to find these change of basis matrices when we are going from some basis into the standard basis.

First I find $C_{\mathcal{B}'\leftarrow\mathcal{B}}$ and $C_{\mathcal{B}'\leftarrow\mathcal{B}''}$, that is, the matrices that change whatever basis into the standard basis. These are

$$C_{\mathcal{B}'\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

and

$$C_{\mathcal{B}'\leftarrow\mathcal{B}''} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}.$$

(Remember, since I am going into the standard basis, I just need to put the columns of my basis into the matrix.)

Second, I compute (using the formula for 2×2 inverses)

$$C_{\mathcal{B}''\leftarrow\mathcal{B}'} = (C_{\mathcal{B}'\leftarrow\mathcal{B}''})^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

which is only coincidentally the same as $C_{\mathcal{B}'\leftarrow\mathcal{B}}$.

Lastly, I need to observe one more formula, which says that if I want to go from basis 1 to basis 3, then I may do so by going from basis 1 to basis 2 and then go from basis 2 to basis 3. The formula is

$$C_{\mathcal{B}''\leftarrow\mathcal{B}} = C_{\mathcal{B}''\leftarrow\mathcal{B}'} C_{\mathcal{B}'\leftarrow\mathcal{B}}.$$

Using this, we find

$$C_{\mathcal{B}''\leftarrow\mathcal{B}} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}.$$

What does this all mean? Let us start with the vector $2[1,0]^t + 3[1,1]^t = [5,3]^t$, and suppose we want to express it in terms of the basis $\mathcal{B}''$. Thus, we need to find coefficients a, b so that $2[1,0]^t + 3[1,1]^t = a[1,0]^t + b[-1,1]^t$. Of course, one way is to explicitly solve a system of equations to find a, b but that will defeat the whole purpose of introducing change of basis transformations. Instead, I take the coordinates of $[5,3]^t$ with respect to the basis $\mathcal{B}$. The coordinate vector is $[2,3]^t$. Now I simply multiply it by $C_{\mathcal{B}''\leftarrow\mathcal{B}}$:

$$\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 8 \\ 3 \end{bmatrix}.$$

Let us check this. This is supposed to tell us that the desired a, b are 8 and 3, so we should have $2[1,0]^t + 3[1,1]^t = 8[1,0]^t + 3[-1,1]^t$. We can easily compute that both sides are in fact equal to $[5,3]^t$.

3. An Eigenbasis Example

I now want to illustrate a problem in which it is advantageous to use some other basis of $\mathbb{R}^2$ as opposed to the standard basis. Start with the matrix $A = \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix}$ and let $\mathcal{B} = \{[1,1]^t, [-1,1]^t\}$ and let $\mathcal{B}'$ denote the standard basis. Then we find

$$C_{\mathcal{B}' \leftarrow \mathcal{B}} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

In order to find $C_{\mathcal{B} \leftarrow \mathcal{B}'}$ we need only invert $C_{\mathcal{B}' \leftarrow \mathcal{B}}$, so

$$C_{\mathcal{B} \leftarrow \mathcal{B}'} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix}.$$

Now let us compute a rather random looking product of matrices:

$$\begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{-1}{2} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 3 & -1 \\ -1 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}.$$

So, by sandwiching A between these two change of basis matrices we reduced A from some arbitrary looking matrix to a diagonal one. Let D denote this diagonal matrix. Another way of writing this equality is then

$$(C_{\mathcal{B}' \leftarrow \mathcal{B}})^{-1} A C_{\mathcal{B}' \leftarrow \mathcal{B}} = D,$$

or equivalently

$$C_{\mathcal{B}' \leftarrow \mathcal{B}} D (C_{\mathcal{B}' \leftarrow \mathcal{B}})^{-1} = A.$$

Essentially, this says that A has a very simple form, we just need to look at it from the right perspective. The reason why all of this works is seen by looking at what A does on the basis $\mathcal{B}$. We compute

$$A \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad A \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 4 \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

These say that A doesn't change the directions of either $[1,1]^t$ or $[-1,1]^t$, it only stretches along those directions. Notice the stretching factors are exactly the entries of the diagonal matrix as well!

Given any square matrix B, we say a vector $\mathbf{x}$ is an eigenvector for B, with corresponding eigenvalue λ if $B\mathbf{x} = \lambda\mathbf{x}$. This means that B doesn't change the direction of the vector $\mathbf{x}$, it only stretches it, and it does so by a factor of λ. So in the above example $[1,1]^t$ is an eigenvector of A with eigenvalue 2, and $[-1,1]^t$ is an eigenvector with eigenvalue 4. The above example should illustrate that it is very useful look at a matrix A using a basis of eigenvectors. It leaves open many fundamental questions like can we always find eigenvectors, when can we find a basis of eigenvectors, how do we go about finding eigenvectors, and so on. These issues will be dealt with in Chapter 5.

Let me make one last remark about factorizations of matrices like the one for A. Often times in applications one is concerned with long term behavior of some process. For instance, if I continue playing poker will I get progressively more in debt? Often these questions are governed by linear models, which will mean that mathematically the process is described by a matrix. Perhaps the entries of A are telling me how much I stand to gain/lose in a single hand of poker. Then the tenth power of A will tell me how much I stand to gain/lose in 10 hands of poker. Thus, to understand long term behavior of this process, I would need to be able to compute higher and higher powers of A. This can be difficult. However, using

$$A = C_{\mathcal{B}' \leftarrow \mathcal{B}} D (C_{\mathcal{B}' \leftarrow \mathcal{B}})^{-1},$$

and a result from a recent homework set, we know

$$A^m = C_{\mathcal{B}' \leftarrow \mathcal{B}} D^m (C_{\mathcal{B}' \leftarrow \mathcal{B}})^{-1}.$$

We also know that it is easy to take powers of diagonal matrices,

$$D^m = \begin{bmatrix} 2^m & 0 \\ 0 & 4^m \end{bmatrix}.$$

Thus, this greatly simplifies the task of finding powers of A. Of course, if we only needed A^2 or A^3, it would be easier to compute it directly, but computing $A^1 0$, or $A^1 000$ are a different story.