MATH 317 NOTE: CHANGE OF BASIS

WILLIAM DEMEO

Abstract. This handout explains how to convert from one basis representation of a linear transformation to another. It assumes a lot of the reader. In particular, the reader is expected to already know about change-of-basis formulae, and this handout is intended primarily for reference when reviewing for the final exam.

1. Changing basis representations of vectors

Let V be an n-dimensional vector space and suppose $\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$ and $\mathcal{V}' = \{\mathbf{v}'_1, \mathbf{v}'_2, \dots, \mathbf{v}'_n\}$ are two bases for V. Recall that the $\mathcal{V}$-basis representation of a vector $\mathbf{x} \in V$ is defined to be the n-tuple $[\mathbf{x}]_{\mathcal{V}} = (c_1, c_2, \dots, c_n)$ of coefficients satisfying

$$\mathbf{x} = c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_n\mathbf{v}_n.$$

Similarly, the representation of $\mathbf{x}$ in the basis $\mathcal{V}'$ is the n-tuple $[\mathbf{x}]_{\mathcal{V}'} = (d_1, d_2, \dots, d_n)$ of coefficients satisfying

$$\mathbf{x} = d_1\mathbf{v}'_1 + d_2\mathbf{v}'_2 + \cdots + d_n\mathbf{v}'_n.$$

The first question we ask is, if we are given $[\mathbf{x}]_{\mathcal{V}}$, how to compute $[\mathbf{x}]_{\mathcal{V}'}$? And if we are given $[\mathbf{x}]_{\mathcal{V}'}$, how to compute $[\mathbf{x}]_{\mathcal{V}}$? In this first section, we review how to do this. In the next section we will take up the similar question for representations of linear transformations.

To change between different basis representations of vectors in $\mathcal{V}$, we construct the *change-of-basis matrix* P from $\mathcal{V}'$ to $\mathcal{V}$ as follows: write each $\mathbf{v}'_i \in \mathcal{V}'$ as a linear combination of the vectors in the first basis,

$$\mathbf{v}'_i = \alpha_{1i}\mathbf{v}_1 + \alpha_{2i}\mathbf{v}_2 + \cdots + \alpha_{ni}\mathbf{v}_n. \tag{1.1}$$

That is, we compute the coefficients, $[\mathbf{v}'_i]_{\mathcal{V}} = (\alpha_{1i}, \alpha_{2i}, \dots, \alpha_{ni})$. It is important to notice that equation (1.1) is equivalent to the following matrix vector multiplication:

$$\mathbf{v}'_i = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \alpha_{1i} \\ \vdots \\ \alpha_{ni} \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix} [\mathbf{v}'_i]_{\mathcal{V}}.$$

Date: December 12, 2015

The latest draft is available at github.com/williamdemeo .

If we do the same for each vector in the basis $\mathcal{V}'$, we have

$$\begin{bmatrix} | & | & & | \\ \mathbf{v}_1' & \mathbf{v}_2' & \cdots & \mathbf{v}_n' \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix} \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & & \\ \vdots & & \ddots & \\ \alpha_{n1} & & & \alpha_{nn} \end{bmatrix}. \tag{1.2}$$

The *change-of-basis matrix* P is the matrix on the right-hand side of equation (1.2). That is, P is formed by putting $[\mathbf{v}_1']_{\mathcal{V}}$ in the first column, $[\mathbf{v}_2']_{\mathcal{V}}$ in the second column, and so on. So

$$P = \begin{bmatrix} | & | & & | \\ [\mathbf{v}_1']_{\mathcal{V}} & [\mathbf{v}_2']_{\mathcal{V}} & \cdots & [\mathbf{v}_n']_{\mathcal{V}} \\ | & | & & | \end{bmatrix} = \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & & \\ \vdots & & \ddots & \\ \alpha_{n1} & & & \alpha_{nn} \end{bmatrix}.$$

If we define the matrices

$$B := \begin{bmatrix} | & | & & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_n \\ | & | & & | \end{bmatrix} \quad \text{and} \quad B' := \begin{bmatrix} | & | & & | \\ \mathbf{v}_1' & \mathbf{v}_2' & \cdots & \mathbf{v}_n' \\ | & | & & | \end{bmatrix}, \tag{1.3}$$

then Equation (1.2) becomes $B' = BP$. Since $\mathcal{V}$ is a basis, B is invertible, so we see that P can be computed as $P = B^{-1}B'$.

Theorem 1.1. *Multiplication by the matrix P converts the $\mathcal{V}'$-basis representation of a vector into the $\mathcal{V}$-basis representation of the same vector.*

Proof. The claim is that $[\mathbf{x}]_{\mathcal{V}} = P[\mathbf{x}]_{\mathcal{V}'}$. Indeed, if we define the matrices B and B' as in (1.3), then we have $\mathbf{x} = B[\mathbf{x}]_{\mathcal{V}}$ and $\mathbf{x} = B'[\mathbf{x}]_{\mathcal{V}'}$. Also, equation (1.2) can be written as $B' = BP$, so we have

$$BP[\mathbf{x}]_{\mathcal{V}'} = B'[\mathbf{x}]_{\mathcal{V}'} = \mathbf{x} = B[\mathbf{x}]_{\mathcal{V}}.$$

Now, since B is clearly invertible (because $\mathcal{V}$ is a basis), we can apply B^{-1} to both sides of the equation $BP[\mathbf{x}]_{\mathcal{V}'} = B[\mathbf{x}]_{\mathcal{V}}$ to get the desired result. □

2. Changing basis representations of linear transformations

Let V and W be linear transformations and suppose $T : V \to W$ is a linear transformation, so that $T\mathbf{x} \in W$ for each $\mathbf{x} \in V$. Assume we have the following bases:

$$\mathcal{V} = \{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\} \quad \text{and} \quad \mathcal{V}' = \{\mathbf{v}_1', \mathbf{v}_2', \dots, \mathbf{v}_n'\} \quad \text{(two bases for } V\text{)}$$

$$\mathcal{W} = \{\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_n\} \quad \text{and} \quad \mathcal{W}' = \{\mathbf{w}_1', \mathbf{w}_2', \dots, \mathbf{w}_n'\} \quad \text{(two bases for } W\text{)}$$

Suppose P is the change-of-basis matrix from $\mathcal{V}'$ to $\mathcal{V}$ of the last section. Suppose Q is the change-of-basis matrix from $\mathcal{W}'$ to $\mathcal{W}$.

The $\mathcal{V}, \mathcal{W}$-basis representation of T is a matrix, denoted by $[T]_{\mathcal{V},\mathcal{W}}$, that when applied to $[\mathbf{x}]_{\mathcal{V}}$ computes the value of $T\mathbf{x}$ with respect to the basis $\mathcal{W}$. More precisely, if we have a vector $\mathbf{x} \in V$, then we can apply the matrix $[T]_{\mathcal{V},\mathcal{W}}$ to the $\mathcal{V}$-basis representation $[\mathbf{x}]_{\mathcal{V}}$ and obtain the $\mathcal{W}$-basis representation of $T\mathbf{x}$. In other terms,

$$[T\mathbf{x}]_{\mathcal{W}} = [T]_{\mathcal{V},\mathcal{W}}[\mathbf{x}]_{\mathcal{V}}.$$

Suppose we are given the matrix $[T]_{\mathcal{V},\mathcal{W}}$ and we wish to find the matrix $[T]_{\mathcal{V}',\mathcal{W}'}$. The matrix $[T]_{\mathcal{V}',\mathcal{W}'}$ should give $[T\mathbf{x}]_{\mathcal{W}'}$ when applied to the vector $[\mathbf{x}]_{\mathcal{V}'}$, that is,

$$[T\mathbf{x}]_{\mathcal{W}'} = [T]_{\mathcal{V}',\mathcal{W}'}[\mathbf{x}]_{\mathcal{V}'}.$$

In order to use $[T]_{\mathcal{V},\mathcal{W}}$ to effect this transformation, think about what must happen to $[\mathbf{x}]_{\mathcal{V}'}$ before we can apply $[T]_{\mathcal{V},\mathcal{W}}$ to it. We must first apply the matrix P from the previous section to get $[\mathbf{x}]_{\mathcal{V}} = P[\mathbf{x}]_{\mathcal{V}'}$, the $\mathcal{V}$-basis representation of $\mathbf{x}$. Only then can we apply the original matrix $[T]_{\mathcal{V},\mathcal{W}}$. But the result of this is the $\mathcal{W}$-basis representation of $T\mathbf{x}$; that is,

$$[T]_{\mathcal{V},\mathcal{W}}P[\mathbf{x}]_{\mathcal{V}'} = [T]_{\mathcal{V},\mathcal{W}}[\mathbf{x}]_{\mathcal{V}} = [T\mathbf{x}]_{\mathcal{W}},$$

whereas we initially set out to compute $[T\mathbf{x}]_{\mathcal{W}'}$. So we simply apply the change-of-basis matrix that goes from $\mathcal{W}$ to $\mathcal{W}'$, which is the inverse of Q. We now arrive at the final result:

$$[T\mathbf{x}]_{\mathcal{W}'} = [T]_{\mathcal{V}',\mathcal{W}'}[\mathbf{x}]_{\mathcal{V}'} = (Q^{-1}[T]_{\mathcal{V},\mathcal{W}}P)[\mathbf{x}]_{\mathcal{V}'}.$$

Since this holds for an arbitrary vector $\mathbf{x} \in V$, we see that

$$[T]_{\mathcal{V}',\mathcal{W}'} = Q^{-1}[T]_{\mathcal{V},\mathcal{W}}P, \tag{2.1}$$

and equation (2.1) is the *change-of-basis formula*.

To be clear, the matrices P and Q that appear in formula (2.1) are

$$P = \begin{bmatrix} | & | & & | \\ [\mathbf{v}_1']_{\mathcal{V}} & [\mathbf{v}_2']_{\mathcal{V}} & \cdots & [\mathbf{v}_n']_{\mathcal{V}} \\ | & | & & | \end{bmatrix} \quad \text{and} \quad Q = \begin{bmatrix} | & | & & | \\ [\mathbf{w}_1']_{\mathcal{W}} & [\mathbf{w}_2']_{\mathcal{W}} & \cdots & [\mathbf{w}_n']_{\mathcal{W}} \\ | & | & & | \end{bmatrix}.$$