BASES, COORDINATES, LINEAR MAPS, AND MATRICES

Math 223

Bases and Coordinates

Let V be a vector space, and suppose that $\mathcal{B} = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ is a basis for V. This basis induces a linear map from $\mathbb{R}^n$ to V, namely, given $(x_1, \ldots, x_n)^T$ in $\mathbb{R}^n$ we use $x_1, \ldots, x_n$ as coefficients of $\mathbf{v}_1, \ldots, \mathbf{v}_n$ in a linear combination.[1] In symbols the map is

$$(1) \qquad \mathbb{R}^n \to V, \qquad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n.$$

This map defines a linear one-to-one correspondence between $\mathbb{R}^n$ and V. Although the map going from V back to $\mathbb{R}^n$ doesn't have a nice formula in general, it's a familiar process. Namely, given $\mathbf{v}$ in V, write $\mathbf{v}$ as a linear combination of $\mathbf{v}_1, \ldots, \mathbf{v}_n$. The coefficients define an element of $\mathbb{R}^n$. These coefficients are called the coordinates of $\mathbf{v}$ relative to the basis $\mathcal{B}$, and are denoted as $[\mathbf{v}]_{\mathcal{B}}$. More precisely, $[\mathbf{v}]_{\mathcal{B}}$ is the unique element of $\mathbb{R}^n$ that maps to $\mathbf{v}$ in (1).

Note that the formula in (1) can be written as

$$(2) \qquad x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n = [\mathbf{v}_1, \ldots, \mathbf{v}_n] \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

where we can think of $[\mathbf{v}_1, \ldots, \mathbf{v}_n]$ as a row matrix whose entries are elements of V (as opposed to numbers). If V is a subspace of $\mathbb{R}^k$ and we write $\mathbf{v}_1, \ldots, \mathbf{v}_n$ as column vectors, then $[\mathbf{v}_1, \ldots, \mathbf{v}_n]$ is a $k \times n$ matrix, and the product in (2) is simply the usual matrix multiplication. (Recall that if A is a matrix with n columns and $\mathbf{c}$ is a vector in $\mathbb{R}^n$, then $A\mathbf{c}$ is the linear combination of the columns of A using the entries of $\mathbf{c}$ as coefficients.) Using this notation, we get that

$$(3) \qquad \mathbf{v} = [\mathbf{v}_1, \ldots, \mathbf{v}_n][\mathbf{v}]_{\mathcal{B}} \quad \text{for all} \quad \mathbf{v} \in V$$

and

$$\left[[\mathbf{v}_1, \ldots, \mathbf{v}_n] \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \right]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \text{ for all } \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n.$$

---

[1] In this handout, all vectors in R^k will be denoted by column vectors.

(Pause for a moment to think about what these equations mean!) The relationship between the basis $\mathcal{B}$ and the coordinate system it induces on V can be summarized by the following diagram.

$$\begin{array}{c} V \\ [\ ]_{\mathcal{B}} \big\downarrow \big\uparrow {\scriptstyle \text{mult by } [\mathbf{v}_1, \ldots, \mathbf{v}_n]} \\ \mathbb{R}^n \end{array}$$

Example 1. Let V be the subspace of $\mathbb{R}^3$ given by $3x - 2y + 4z = 0$. The vectors $\mathbf{v}_1 = (2, 3, 0)^T$ and $\mathbf{v}_2 = (0, 2, 1)^T$ are in V and are independent. Since $\dim V$ is 2, then $\mathcal{B} = [\mathbf{v}_1, \mathbf{v}_2]$ is a basis for V. The vector $\mathbf{w} = (2, 15, 6)^T$ is in V, and so is a linear combination of $\mathbf{v}_1$ and $\mathbf{v}_2$. In fact,

$$\mathbf{w} \;=\; \mathbf{v}_1 + 6\mathbf{v}_2 \;=\; [\mathbf{v}_1, \mathbf{v}_2] \begin{pmatrix} 1 \\ 6 \end{pmatrix}, \quad \text{and so} \quad [\mathbf{w}]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 6 \end{pmatrix}.$$

Example 2. Let $V = \{p \in \mathcal{P}_3 \mid p(1) = 0\}$. This is a two-dimensional subspace of $\mathcal{P}_3$, and a basis for V is $\mathcal{B} = [x - 1, x(x - 1)]$. The polynomial $q(x) = 2x^2 + x - 3$ is in V, and so is a linear combination of $x - 1$ and $x(x - 1)$. We have

$$q(x) \;=\; 3(x - 1) + 2x(x - 1) \;=\; [x - 1, x(x - 1)] \begin{pmatrix} 3 \\ 2 \end{pmatrix}, \quad \text{and so} \quad [q]_{\mathcal{B}} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}.$$

Change of Basis

Now suppose that $\mathcal{B} = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$ and $\tilde{\mathcal{B}} = [\tilde{\mathbf{v}}_1, \ldots, \tilde{\mathbf{v}}_n]$ are two bases for V. We then have two coordinate systems on V, and it's good to have a way to get from one to the other, that is, if you have the coordinates of $\mathbf{v}$ relative to one basis, you should be able to compute the coordinates relative to the other basis. The advantage of representing coordinate systems as linear maps from V to $\mathbb{R}^n$ is that these maps will do the work for us.

The two coordinate maps from V to $\mathbb{R}^n$ are

$$[\ ]_{\mathcal{B}} \colon V \to \mathbb{R}^n, \quad [\mathbf{v}]_{\mathcal{B}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \qquad \text{and} \qquad [\ ]_{\tilde{\mathcal{B}}} \colon V \to \mathbb{R}^n, \quad [\mathbf{v}]_{\tilde{\mathcal{B}}} = \begin{pmatrix} \tilde{x}_1 \\ \vdots \\ \tilde{x}_n \end{pmatrix},$$

where $x_1, \ldots, x_n$ are the coordinates of $\mathbf{v}$ relative to $\mathcal{B}$ and $\tilde{x}_1, \ldots, \tilde{x}_n$ are the coordinates of $\mathbf{v}$ relative to $\tilde{\mathcal{B}}$. We can arrange these maps into a diagram as follows.

$$\begin{array}{ccc} & V & \\ {\scriptstyle [\ ]_{\mathcal{B}}} \swarrow & & \searrow {\scriptstyle [\ ]_{\tilde{\mathcal{B}}}} \\ \mathbb{R}^n & \underset{?}{\longrightarrow} & \mathbb{R}^n \end{array}$$

The map across the bottom is defined by going around the diagram: a vector $(x_1, \ldots, x_n)^T$ in the left $\mathbb{R}^n$ is taken up to V via $[\;]_{\mathcal{B}}^{-1}$, which is the map in (1), and then down to the right $\mathbb{R}^n$ via $[\;]_{\tilde{\mathcal{B}}}$. Now $(x_1, \ldots, x_n)^T$ is the coordinate vector relative to $\mathcal{B}$ of some vector $\mathbf{v}$ in V. The result of going up to V via $[\;]_{\mathcal{B}}^{-1}$ is just $\mathbf{v}$ itself! Going down to the right $\mathbb{R}^n$ results in the coordinates of $\mathbf{v}$ relative to $\tilde{\mathcal{B}}$. Thus, the map across the bottom changes the coordinates of $\mathbf{v}$ relative to $\mathcal{B}$ into the coordinates of $\mathbf{v}$ relative to $\tilde{\mathcal{B}}$! We denote it by $C_{\mathcal{B}\tilde{\mathcal{B}}}$. This map is linear, since it's the composition of linear maps, and so it is given by multiplication by some $n \times n$ matrix, also denoted by $C_{\mathcal{B}\tilde{\mathcal{B}}}$. Here is the completed diagram.

(4)
$$\begin{array}{ccc} & V & \\ {}^{[\;]_{\mathbf{B}}}\swarrow & & \searrow^{[\;]_{\tilde{\mathbf{B}}}} \\ \mathbb{R}^n & \xrightarrow[C_{\mathbf{B}\tilde{\mathbf{B}}}]{} & \mathbb{R}^n \end{array}$$

In this diagram there are two ways to take a vector $\mathbf{v}$ in V to the right $\mathbb{R}^n$, namely go directly there, or go the long way through the left $\mathbb{R}^n$. By the way the bottom map is defined, it follows that it doesn't matter which way you go, you get the same thing. The equation that says this is

(5)
$$C_{\mathcal{B}\tilde{\mathcal{B}}}[\mathbf{v}]_{\mathcal{B}} = [\mathbf{v}]_{\tilde{\mathcal{B}}}.$$

Note that the equation says that mulitplying the coordinates of $\mathbf{v}$ relative to $\mathcal{B}$ by the matrix $C_{\mathcal{B}\tilde{\mathcal{B}}}$ results in the coordinates of $\mathbf{v}$ relative to $\tilde{\mathcal{B}}$.

We know that if A is a matrix with n columns and if $\mathbf{e}_j$ is the j^{th} vector in the standard basis for $\mathbb{R}^n$, then $A\mathbf{e}_j$ is the j^{th} column of A. Since $[\mathbf{v}_j]_{\mathcal{B}} = \mathbf{e}_j$, equation (5) gives us

$$(j^{\text{th}} \text{ column of } C_{\mathcal{B}\tilde{\mathcal{B}}}) = C_{\mathcal{B}\tilde{\mathcal{B}}}\mathbf{e}_j = C_{\mathcal{B}\tilde{\mathcal{B}}}[\mathbf{v}_j]_{\mathcal{B}} = [\mathbf{v}_j]_{\tilde{\mathcal{B}}}.$$

In words, the j^{th} column of the matrix $C_{\mathcal{B}\tilde{\mathcal{B}}}$ consists of the coordinates of the j^{th} basis vector of $\mathcal{B}$ relative to the basis $\tilde{\mathcal{B}}$. We can write

$$C_{\mathcal{B}\tilde{\mathcal{B}}} = \left([\mathbf{v}_1]_{\tilde{\mathcal{B}}}, \ldots, [\mathbf{v}_n]_{\tilde{\mathcal{B}}}\right).$$

Here's another way to see this. If $\mathbf{v} \in V$, then

$$\mathbf{v} = [\mathbf{v}_1, \ldots, \mathbf{v}_n][\mathbf{v}]_{\mathcal{B}} = [\tilde{\mathbf{v}}_1, \ldots, \tilde{\mathbf{v}}_n][\mathbf{v}]_{\tilde{\mathcal{B}}} = [\tilde{\mathbf{v}}_1, \ldots, \tilde{\mathbf{v}}_n]C_{\mathcal{B}\tilde{\mathcal{B}}}[\mathbf{v}]_{\mathcal{B}},$$

where the last equality follows from (5). If we let $[\mathbf{v}]_{\mathcal{B}} = (x_1, \ldots, x_n)^T$, then we have

$$[\mathbf{v}_1, \ldots, \mathbf{v}_n]\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = [\tilde{\mathbf{v}}_1, \ldots, \tilde{\mathbf{v}}_n]C_{\mathcal{B}\tilde{\mathcal{B}}}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

for all $(x_1, \ldots, x_n)^T$ in $\mathbb{R}^n$. It follows that

$$[\mathbf{v}_1, \ldots, \mathbf{v}_n] = [\tilde{\mathbf{v}}_1, \ldots, \tilde{\mathbf{v}}_n]C_{\mathcal{B}\tilde{\mathcal{B}}}.$$

This last step is similar to the fact that if A and B are matrices of the same dimension and that $A\mathbf{x} = B\mathbf{x}$ for all vectors $\mathbf{x}$, then $A = B$.) Each side of this equation is a list of n vectors in V. The j^{th} vector on the left is $\mathbf{v}_j$. The j^{th} vector on the right is $[\tilde{\mathbf{v}}_1, \ldots, \tilde{\mathbf{v}}_n]$ times the j^{th} column of $C_{\mathcal{B}\tilde{\mathcal{B}}}$. Thus

$$\mathbf{v}_j = [\tilde{\mathbf{v}}_1, \ldots, \tilde{\mathbf{v}}_n](j^{\text{th}} \text{ column of } C_{\mathcal{B}\tilde{\mathcal{B}}}).$$

But this says that the j^{th} column of $C_{\mathcal{B}\tilde{\mathcal{B}}}$ is the coordinate vector of $\mathbf{v}_j$ relative to $\tilde{\mathcal{B}}$ by the $\tilde{\mathcal{B}}$ version of (3).

Finally, note that $C_{\tilde{\mathcal{B}}\mathcal{B}}$, the matrix that converts coordinates relative to $\tilde{\mathcal{B}}$ into coordinates relative to $\mathcal{B}$, should be the inverse of $C_{\mathcal{B}\tilde{\mathcal{B}}}$, since it represents the linear transformation along the bottom of (4) from right to left.

Example 1 (cont). A second basis for $V = \{(x, y, z) \mid 3x - 2y + 4z = 0\}$ is $\tilde{\mathcal{B}} = [\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2]$, where $\tilde{\mathbf{v}}_1 = (2, 5, 1)^T$ and $\tilde{\mathbf{v}}_2 = (-4, 0, 3)^T$. You can show that $\mathbf{v}_1 = \frac{3}{5}\tilde{\mathbf{v}}_1 - \frac{1}{5}\tilde{\mathbf{v}}_2$ and $\mathbf{v}_2 = \frac{2}{5}\tilde{\mathbf{v}}_1 + \frac{1}{5}\tilde{\mathbf{v}}_2$. It follows that $C_{\mathcal{B}\tilde{\mathcal{B}}} = \begin{pmatrix} 3/5 & 2/5 \\ -1/5 & 1/5 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} 3 & 2 \\ -1 & 1 \end{pmatrix}$.

Recall that the coordinate vector of $\mathbf{w} = (2, 15, 6)^T$ relative to $\mathcal{B}$ is $[\mathbf{w}]_{\mathcal{B}} = (1, 6)^T$. Its coordinate vector relative to $\tilde{\mathcal{B}}$ should then be

$$[\mathbf{w}]_{\tilde{\mathcal{B}}} = C_{\mathcal{B}\tilde{\mathcal{B}}}[\mathbf{w}]_{\mathcal{B}} = \frac{1}{5}\begin{pmatrix} 3 & 2 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 6 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

Indeed, we have

$$[\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2]\begin{pmatrix} 3 \\ 1 \end{pmatrix} = 3\tilde{\mathbf{v}}_1 + \tilde{\mathbf{v}}_2 = \mathbf{w}.$$

It's important to remember here that the vector $\mathbf{w}$ doesn't change. What changes is the way we refer to it. The situation is similar to measuring a distance in feet and in meters. You get different results, not because the distance changes, but because you are measuring it differently.

The matrix converting coordinates relative to $\tilde{\mathcal{B}}$ into coordinates relative to $\mathcal{B}$ is $C_{\tilde{\mathcal{B}}\mathcal{B}} = C_{\mathcal{B}\tilde{\mathcal{B}}}^{-1} = \begin{pmatrix} 1 & -2 \\ 1 & 3 \end{pmatrix}$.

Example 2 (cont). A second basis for $V = \{p \in \mathcal{P}_3 \mid p(1) = 0\}$ is $\tilde{\mathcal{B}} = [x - 1, x^2 - 1]$. Since $x - 1$ is the first element of both bases, its coordinate vector relative to either basis is $(1, 0)^T$. We also have that

$$x(x-1) = -(x-1) + (x^2 - 1) = [x - 1, x^2 - 1]\begin{pmatrix} -1 \\ 1 \end{pmatrix},$$

and so $[x(x-1)]_{\tilde{\mathcal{B}}} = (-1, 1)^T$. The change of coordinates matrix is then $C_{\mathcal{B}\tilde{\mathcal{B}}} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$. From above we have $[q]_{\mathcal{B}} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$, and so

$$[q]_{\tilde{\mathcal{B}}} = C_{\mathcal{B}\tilde{\mathcal{B}}}[q]_{\mathcal{B}} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

Checking this,

$$[x-1, x^2-1]\begin{pmatrix}1\\2\end{pmatrix} = (x-1)+2(x^2-1) = 2x^2+x-3 = q(x),$$

as needed.

Linear Maps and Matrices

Let W be another vector space, and suppose $\hat{\mathcal{B}} = [\mathbf{w}_1, \ldots, \mathbf{w}_m]$ is a basis for W. As above, we get the linear map $[\ ]_{\hat{\mathcal{B}}}\colon W \to \mathbb{R}^m$ so that the entries of $[\mathbf{w}]_{\hat{\mathcal{B}}}$ are the coordinates of $\mathbf{w} \in W$ relative to $\hat{\mathcal{B}}$.

Now suppose $L\colon V \to W$ is linear. We get the following diagram of spaces and maps.

$$\begin{array}{ccc} V & \xrightarrow{L} & W \\ {\scriptstyle [\ ]_{\mathcal{B}}}\big\downarrow & & \big\downarrow{\scriptstyle [\ ]_{\hat{\mathcal{B}}}} \\ \mathbb{R}^n & \xrightarrow[?]{} & \mathbb{R}^m \end{array}$$

The map across the bottom is defined by going around the diagram from $\mathbb{R}^n$ up to V via $[\ ]_{\mathcal{B}}^{-1}$, which is the map in (1), then across to W via L, then down to $\mathbb{R}^m$ via $[\ ]_{\hat{\mathcal{B}}}$. This map is linear, and so is given by multiplication by some matrix, namely the matrix of L relative to the bases $\mathcal{B}$ and $\hat{\mathcal{B}}$, denoted by $[L]_{\mathcal{B}\hat{\mathcal{B}}}$. This can actually be used as the definition of this matrix! The diagram then becomes

(6)
$$\begin{array}{ccc} V & \xrightarrow{L} & W \\ {\scriptstyle [\ ]_{\mathcal{B}}}\big\downarrow & & \big\downarrow{\scriptstyle [\ ]_{\hat{\mathcal{B}}}} \\ \mathbb{R}^n & \xrightarrow[{[L]_{\mathcal{B}\hat{\mathcal{B}}}}]{} & \mathbb{R}^m \end{array}$$

With the bottom map defined by going around the other three sides of the diagram, it follows that if you start with a vector $\mathbf{v}$ in V and follow the maps to $\mathbb{R}^m$, it doesn't matter which way you go, you get the same thing. The equation that says this is

(7)
$$[L]_{\mathcal{B}\hat{\mathcal{B}}}[\mathbf{v}]_{\mathcal{B}} = [L(\mathbf{v})]_{\hat{\mathcal{B}}}.$$

The left side of this is the result of going from V down to $\mathbb{R}^n$ by taking the coordinates of $\mathbf{v}$ relative to $\mathcal{B}$, and then across to $\mathbb{R}^m$ by multiplying by the matrix $[L]_{\mathcal{B}\hat{\mathcal{B}}}$. The right side is the result of going across to W by L, and then down to $\mathbb{R}^m$ by taking coordinates of $L(\mathbf{v})$ relative to $\hat{\mathcal{B}}$.

We know that if A is a matrix with n columns and if $\mathbf{e}_j$ is the j^{th} vector in the standard basis for $\mathbb{R}^n$, then $A\mathbf{e}_j$ is the j^{th} column of A. Since $[\mathbf{v}_j]_{\mathcal{B}} = \mathbf{e}_j$ equation (7) gives us

$$[L]_{\mathcal{B}\hat{\mathcal{B}}}\mathbf{e}_j = [L]_{\mathcal{B}\hat{\mathcal{B}}}[\mathbf{v}_j]_{\mathcal{B}} = [L(\mathbf{v}_j)]_{\hat{\mathcal{B}}},$$

which says that the j^{th} column of $[L]_{\mathcal{B}\hat{\mathcal{B}}}$ is $[L(\mathbf{v}_j)]_{\hat{\mathcal{B}}}$. In words, the j^{th} column of the matrix of L consists of the coordinates of $L(\mathbf{v}_j)$ relative to the basis $\hat{\mathcal{B}}$. This is, of course, how we have computed matrices all along. The point is that this form for the matrix is forced upon us as soon as we say that the map across the bottom in diagram (6) is defined by going around the other three sides.

In many applications we work with a linear map from a vector space to itself, that is, $W = V$ in (6). In this case we generally take the bases $\mathcal{B}$ and $\hat{\mathcal{B}}$ to be the same, and the diagram becomes

$$(6') \qquad \begin{array}{ccc} V & \xrightarrow{\;L\;} & V \\ {\scriptstyle [\;]_{\mathcal{B}}}\downarrow & & \downarrow{\scriptstyle [\;]_{\mathcal{B}}} \\ \mathbb{R}^n & \xrightarrow[{[L]_{\mathcal{B}}}]{} & \mathbb{R}^n \end{array}$$

Note that we write $[L]_{\mathcal{B}}$ instead of $[L]_{\mathcal{B}\mathcal{B}}$, as the extra $\mathcal{B}$ is now redundant. The equation that goes around the diagram both ways is

$$(7') \qquad [L]_{\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} = [L(\mathbf{v})]_{\mathcal{B}},$$

which replaces equation (7).

Example 1 (cont). Let $P\colon \mathbb{R}^3 \to \mathbb{R}^3$ be orthogonal projection onto the xy-plane, and let $Q\colon \mathbb{R}^3 \to \mathbb{R}^3$ be orthogonal projection onto the subspace V. As you computed in a homework problem, the matrix for $Q \circ P$ is

$$\frac{1}{29}\begin{pmatrix} 20 & 6 & 0 \\ 6 & 25 & 0 \\ -12 & 8 & 0 \end{pmatrix}.$$

We define L to be this composition restricted to V, that is, $L\colon V \to V$, $L(\mathbf{v}) = (Q \circ P)(\mathbf{v})$. Applying this to the basis vectors of $\mathcal{B}$ we get

$$L(\mathbf{v}_1) = (Q \circ P)(\mathbf{v}_1) = \begin{pmatrix} 2 \\ 3 \\ 0 \end{pmatrix} = \mathbf{v}_1 + 0\mathbf{v}_2 = \frac{2}{29}(3\mathbf{v}_1 + 8\mathbf{v}_2)$$

and

$$L(\mathbf{v}_2) = (Q \circ P)(\mathbf{v}_2) = \frac{2}{29}\begin{pmatrix} 6 \\ 25 \\ 8 \end{pmatrix} = \frac{2}{29}(3\mathbf{v}_1 + 8\mathbf{v}_2) = \frac{6}{29}\mathbf{v}_1 + \frac{16}{29}\mathbf{v}_2.$$

(Note that $\mathbf{v}_1$ is in both V and the xy-plane, so it is unchanged by either projection.) The matrix for L relative to $\mathcal{B}$ is then $[L]_{\mathcal{B}} = \begin{pmatrix} 1 & 6/29 \\ 0 & 16/29 \end{pmatrix}$.

Example 2 (cont). Define $L\colon V \to V$ by $L(p) = (1-x)p'(x)$. You should verify that this is linear. Applying L to the elements of the basis $\mathcal{B}$ we get

$$L(x-1) = (1-x)(x-1)' = 1 - x = -(x-1) + 0x(x-1) = [x-1, x(x-1)]\begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

and

$$L\big(x(x-1)\big) = (1-x)(x^2-x)' = -2x^2+3x-1 = (x-1)-2x(x-1) = [x-1, x(x-1)]\begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

The matrix for L relative to $\mathcal{B}$ is then $[L]_{\mathcal{B}} = \begin{pmatrix} -1 & 1 \\ 0 & 2 \end{pmatrix}$.

How the Matrix for a Linear Map Changes Under a Change of Basis

If $\mathcal{B} = [\mathbf{v}_1, \dots, \mathbf{v}_n]$ and $\tilde{\mathcal{B}} = [\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_n]$ are bases for V, and if $L\colon V \to V$, we get matrices $C_{\mathcal{B}\tilde{\mathcal{B}}}$, $C_{\tilde{\mathcal{B}}\mathcal{B}}$, $[L]_{\mathcal{B}}$, and $[L]_{\tilde{\mathcal{B}}}$. How are these matrices related? We already know that the first two are inverses of each other, so the real question is how the matrix for L changes when the basis is changed from $\mathcal{B}$ to $\tilde{\mathcal{B}}$. The answer lies in putting together a complicated but natural diagram of spaces and maps by combining the previous diagrams (4) and (6′) along with their analogs for $C_{\tilde{\mathcal{B}}\mathcal{B}}$ and $[L]_{\tilde{\mathcal{B}}}$. Here they are.

$$
(8)\qquad
\begin{array}{ccc}
 & V & \\
{}^{[\,]_{\mathbf{B}}}\swarrow & & \searrow{}^{[\,]_{\tilde{\mathbf{B}}}} \\
\mathbb{R}^n & \xrightarrow[C_{\mathbf{B}\tilde{\mathbf{B}}}]{} & \mathbb{R}^n
\end{array}
\qquad
\begin{array}{ccc}
 & V & \\
{}^{[\,]_{\tilde{\mathbf{B}}}}\swarrow & & \searrow{}^{[\,]_{\mathbf{B}}} \\
\mathbb{R}^n & \xrightarrow[C_{\tilde{\mathbf{B}}\mathbf{B}}]{} & \mathbb{R}^n
\end{array}
\qquad
\begin{array}{ccc}
V & \xrightarrow{L} & V \\
{\scriptstyle [\,]_{\mathbf{B}}}\big\downarrow & & \big\downarrow{\scriptstyle [\,]_{\mathbf{B}}} \\
\mathbb{R}^n & \xrightarrow[{[L]_{\mathbf{B}}}]{} & \mathbb{R}^n
\end{array}
\qquad
\begin{array}{ccc}
V & \xrightarrow{L} & V \\
{\scriptstyle [\,]_{\tilde{\mathbf{B}}}}\big\downarrow & & \big\downarrow{\scriptstyle [\,]_{\tilde{\mathbf{B}}}} \\
\mathbb{R}^n & \xrightarrow[{[L]_{\tilde{\mathbf{B}}}}]{} & \mathbb{R}^n
\end{array}
$$

The latter two diagrams can be attached along their common edge $V \xrightarrow{L} V$.

$$
\begin{array}{ccccc}
\mathbb{R}^n & \xleftarrow{[\,]_{\mathbf{B}}} & V & \xrightarrow{[\,]_{\tilde{\mathbf{B}}}} & \mathbb{R}^n \\
{\scriptstyle [L]_{\mathbf{B}}}\big\downarrow & & \big\downarrow{\scriptstyle L} & & {\scriptstyle [L]_{\tilde{\mathbf{B}}}}\big\downarrow \\
\mathbb{R}^n & \xleftarrow[{[\,]_{\mathbf{B}}}]{} & V & \xrightarrow{[\,]_{\tilde{\mathbf{B}}}} & \mathbb{R}^n
\end{array}
$$

We now consider the map across the top from the right $\mathbb{R}^n$ to the left $\mathbb{R}^n$. It is $C_{\tilde{\mathcal{B}}\mathcal{B}}$! Similarly, the map across the bottom from left to right is $C_{\mathcal{B}\tilde{\mathcal{B}}}$. Filling these in, the resulting diagram contains all four of those in (8)!

$$
(9)\qquad
\begin{array}{ccccc}
 & & C_{\tilde{\mathcal{B}}\mathcal{B}} & & \\
\big\downarrow & & & & \big\uparrow \\
\mathbb{R}^n & \xleftarrow{[\,]_{\mathbf{B}}} & V & \xrightarrow{[\,]_{\tilde{\mathbf{B}}}} & \mathbb{R}^n \\
{\scriptstyle [L]_{\mathbf{B}}}\big\downarrow & & \big\downarrow{\scriptstyle L} & & {\scriptstyle [L]_{\tilde{\mathbf{B}}}}\big\downarrow \\
\mathbb{R}^n & \xleftarrow[{[\,]_{\mathbf{B}}}]{} & V & \xrightarrow{[\,]_{\tilde{\mathbf{B}}}} & \mathbb{R}^n \\
\big\uparrow & & & & \big\downarrow \\
 & & C_{\mathcal{B}\tilde{\mathcal{B}}} & &
\end{array}
$$

Multiplying by the matrix $[L]_{\tilde{\mathcal{B}}}$ goes from the upper right $\mathbb{R}^n$ to the lower right $\mathbb{R}^n$. Another, equivalent, way to go between these corners is to go around the outside of the diagram counter clockwise, taking, in turn, the arrows marked $C_{\tilde{\mathcal{B}}\mathcal{B}}$, $[L]_{\mathcal{B}}$, and $C_{\mathcal{B}\tilde{\mathcal{B}}}$. The product of these matrices, in the appropriate order, must equal $[L]_{\tilde{\mathcal{B}}}$. We have

$$[L]_{\tilde{\mathcal{B}}} = C_{\mathcal{B}\tilde{\mathcal{B}}}[L]_{\mathcal{B}}C_{\tilde{\mathcal{B}}\mathcal{B}} = C_{\mathcal{B}\tilde{\mathcal{B}}}[L]_{\mathcal{B}}C_{\mathcal{B}\tilde{\mathcal{B}}}^{-1}.$$

We could have discovered this formula by combining the formulas (5) and (7′) with their analogs for $C_{\tilde{\mathcal{B}}\mathcal{B}}$ and $[L]_{\tilde{\mathcal{B}}}$, since these formulas contain the same information as the corresponding diagrams. In retrospect it's not too difficult, and it emphasizes that, despite the complexity of the forest of arrows in (9), the matrix multiplications $[L]_{\mathcal{B}}$ and $[L]_{\tilde{\mathcal{B}}}$ down the left and right sides of the diagram are really just representations of the map L down the middle. If $\mathbf{v} \in V$, then

$$[L]_{\tilde{\mathcal{B}}}[\mathbf{v}]_{\tilde{\mathcal{B}}} = [L(\mathbf{v})]_{\tilde{\mathcal{B}}} = C_{\mathcal{B}\tilde{\mathcal{B}}}[L(\mathbf{v})]_{\mathcal{B}} = C_{\mathcal{B}\tilde{\mathcal{B}}}[L]_{\mathcal{B}}[\mathbf{v}]_{\mathcal{B}} = C_{\mathcal{B}\tilde{\mathcal{B}}}[L]_{\mathcal{B}}C_{\tilde{\mathcal{B}}\mathcal{B}}[\mathbf{v}]_{\tilde{\mathcal{B}}}.$$

We then have $[L]_{\tilde{\mathcal{B}}}[\mathbf{v}]_{\tilde{\mathcal{B}}} = C_{\mathcal{B}\tilde{\mathcal{B}}}[L]_{\mathcal{B}}C_{\tilde{\mathcal{B}}\mathcal{B}}[\mathbf{v}]_{\tilde{\mathcal{B}}}$ for all $\mathbf{v} \in V$. This is the same as

$$[L]_{\tilde{\mathcal{B}}}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = C_{\mathcal{B}\tilde{\mathcal{B}}}[L]_{\mathcal{B}}C_{\tilde{\mathcal{B}}\mathcal{B}}\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

for all $(x_1, \ldots, x_n)^T \in \mathbb{R}^n$. It follows that $[L]_{\tilde{\mathcal{B}}} = C_{\mathcal{B}\tilde{\mathcal{B}}}[L]_{\mathcal{B}}C_{\tilde{\mathcal{B}}\mathcal{B}}$. This should make sense. What it says is that computing $L(\mathbf{v})$ by using the coordinates of $\mathbf{v}$ and matrix of L relative to $\tilde{\mathcal{B}}$ is the same as converting to the $\mathcal{B}$ coordinate system, doing the computation in $\mathcal{B}$ coordinates, and then switching back to $\tilde{\mathcal{B}}$ coordinates.

Can you figure out which corners of the diagram the coordinate vectors $[\mathbf{v}]_{\tilde{\mathcal{B}}}$, $[L(\mathbf{v})]_{\tilde{\mathcal{B}}}$ $[L(\mathbf{v})]_{\mathcal{B}}$, and $[\mathbf{v}]_{\mathcal{B}}$ in this computation belong to?

Example 1 (cont). The matrix for L relative to $\tilde{\mathcal{B}}$ is

$$[L]_{\tilde{\mathcal{B}}} = C_{\mathcal{B}\tilde{\mathcal{B}}}[L]_{\mathcal{B}}C_{\mathcal{B}\tilde{\mathcal{B}}}^{-1} = \frac{1}{5}\begin{pmatrix} 3 & 2 \\ -1 & 1 \end{pmatrix} \frac{1}{29}\begin{pmatrix} 29 & 6 \\ 0 & 16 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 1 & 3 \end{pmatrix} = \frac{1}{145}\begin{pmatrix} 137 & -24 \\ -19 & 88 \end{pmatrix}.$$

Example 2 (cont). The matrix for L relative to $\tilde{\mathcal{B}}$ is

$$[L]_{\tilde{\mathcal{B}}} = C_{\mathcal{B}\tilde{\mathcal{B}}}[L]_{\mathcal{B}}C_{\mathcal{B}\tilde{\mathcal{B}}}^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -1 & 1 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 0 & -2 \end{pmatrix}.$$

Robert L. Foote, November 1997