Lecture 1q

The Matrix of a Linear Mapping

(pages 235-239)

When we first studied linear mappings in Math 106, they were only between $\mathbb{R}^n$ and $\mathbb{R}^m$. In this setting, we were always able to find a matrix A such that our linear mapping $L(\vec{x})$ was the same as $A\vec{x}$. That's because, while $\vec{x} \in \mathbb{R}^n$ isn't technically a matrix, we could temporarily think of $\vec{x}$ like a matrix, and the matrix product $A\vec{x}$ had the properties we wanted. In general, we made great use of the matrix A, and would like to be able to do this for any linear mapping, not just ones between $\mathbb{R}^n$ and $\mathbb{R}^m$. But the product $A(2 + 3x - x^2)$ (for example) doesn't make any sense at all! So before we can find a matrix for a linear mapping whose domain is not $\mathbb{R}^n$, we first need to find some way to make our random vector look like a vector from $\mathbb{R}^n$. But wait–we've already done that! Once we fix a basis $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ for a vector space $\mathbb{V}$, we can look at the $\mathcal{B}$-coordinates for $\mathbf{x}$. That's only the first step in finding the matrix of a general linear mapping $L : \mathbb{V} \to \mathbb{W}$ though. Because our product $A[\mathbf{x}]_\mathcal{B}$ will be a vector from $\mathbb{R}^m$, not $\mathbb{W}$. So we will also need to fix a basis $\mathcal{C} = \{\mathbf{w}_1, \ldots, \mathbf{w}_m\}$ for $\mathbb{W}$, and we will have our output be in $\mathcal{C}$-coordinates. So while we can't find a matrix for L directly, we can find a matrix for L relative to the $\mathcal{B}$ and $\mathcal{C}$ coordinates. As such, instead of simply denoting the matrix for the linear mapping as $[L]$, we denote it as ${}_\mathcal{C}[L]_\mathcal{B}$, so that we remember what coordinates we are using. And this matrix ${}_\mathcal{C}[L]_\mathcal{B}$ will be such that

$$[L(\mathbf{x})]_\mathcal{C} =_\mathcal{C} [L]_\mathcal{B}[\mathbf{x}]_\mathcal{B}$$

We know that the matrix ${}_\mathcal{C}[L]_\mathcal{B}$ must exist, since the equation above defines a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$. Now our only problem is to find it. We go about this process the same way we went about finding $[L]$ in Math 106. The first thing we want to note is that there is a vector $[\mathbf{x}]_\mathcal{B} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ such that $\mathbf{x} = x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n$. Then we have that

$$\begin{aligned} L(\mathbf{x}) &= L(x_1\mathbf{v}_1 + \cdots + x_n\mathbf{v}_n) \\ &= x_1L(\mathbf{v}_1) + \cdots + x_nL(\mathbf{v}_n) \end{aligned}$$

But, recalling Theorem 4.4.1, we then have that

$$\begin{aligned}
[L(\mathbf{x})]_{\mathcal{C}} &= [x_1 L(\mathbf{v}_1) + \cdots + x_n L(\mathbf{v}_n)]_{\mathcal{C}} \\
&= x_1 [L(\mathbf{v}_1)]_{\mathcal{C}} + \cdots + x_n [L(\mathbf{v}_n)]_{\mathcal{C}} \\
&= \begin{bmatrix} [L(\mathbf{v}_1)]_{\mathcal{C}} & \cdots & [L(\mathbf{v}_n)]_{\mathcal{C}} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \\
&= \begin{bmatrix} [L(\mathbf{v}_1)]_{\mathcal{C}} & \cdots & [L(\mathbf{v}_n)]_{\mathcal{C}} \end{bmatrix} [\mathbf{x}]_{\mathcal{B}}
\end{aligned}$$

From this, we see that

$${}_{\mathcal{C}}[L]_{\mathcal{B}} = \begin{bmatrix} [L(\mathbf{v}_1)]_{\mathcal{C}} & \cdots & [L(\mathbf{v}_n)]_{\mathcal{C}} \end{bmatrix}$$

<u>Definition</u>: Let $\mathbb{V}$ be a vector space with basis $\mathcal{B} = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$, let $\mathbb{W}$ be a vector space with basis $\mathcal{C}$, and let $L : \mathbb{V} \to \mathbb{W}$ be a linear mapping. We define the **matrix of L with respect to the bases $\mathcal{B}$ and $\mathcal{C}$** to be the matrix

$${}_{\mathcal{C}}[L]_{\mathcal{B}} = \begin{bmatrix} [L(\mathbf{v}_1)]_{\mathcal{C}} & \cdots & [L(\mathbf{v}_n)]_{\mathcal{C}} \end{bmatrix}$$

Example: Let $L : M(2,2) \to P_2$ be defined by $L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (a+b) + (a+c)x + (a+d)x^2$, and let $\mathcal{B}$ be the standard basis for $M(2,2)$ (so $\mathcal{B} = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right\}$) and $\mathcal{C}$ be the standard basis for P_2 (so $\mathcal{C} = \{1, x, x^2\}$). To find ${}_{\mathcal{C}}[L]_{\mathcal{B}}$, we need to compute the $\mathcal{C}$-coordinates of the image under L of the $\mathcal{B}$ basis vectors.

$$\begin{aligned}
L\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) &= 1 + x + x^2, \text{ so} \\
\left[L\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right)\right]_{\mathcal{C}} &= \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \\
L\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right) &= 1, \text{ so} \\
\left[L\left(\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}\right)\right]_{\mathcal{C}} &= \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \\
L\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right) &= x, \text{ so} \\
\left[L\left(\begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}\right)\right]_{\mathcal{C}} &= \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \\
L\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right) &= x^2, \text{ so} \\
\left[L\left(\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right)\right]_{\mathcal{C}} &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}
\end{aligned}$$

And thus we have that

$${}_{\mathcal{C}}[L]_{\mathcal{B}} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

Most of the time we will be using the standard bases, but let's go ahead and look at the same linear mapping under different bases.

Example: Let $L : M(2,2) \to P_2$ be defined by $L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (a+b) + (a+c)x + (a+d)x^2$, and let $\mathcal{B} = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right\}$ and $\mathcal{C} = \{1, 1+x, 1+x+x^2\}$. To find ${}_{\mathcal{C}}[L]_{\mathcal{B}}$, we need to compute the $\mathcal{C}$-coordinates of the image under L of the $\mathcal{B}$ basis vectors.

$$\begin{aligned}
L\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) &= 1 + x + x^2, \text{ so} \\
\left[L\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right)\right]_{\mathcal{C}} &= \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \\
L\left(\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\right) &= 2 + x + x^2, \text{ so} \\
\left[L\left(\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}\right)\right]_{\mathcal{C}} &= \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \\
L\left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\right) &= 2 + 2x + x^2, \text{ so} \\
\left[L\left(\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}\right)\right]_{\mathcal{C}} &= \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} \\
L\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right) &= 2 + 2x + 2x^2, \text{ so} \\
\left[L\left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right)\right]_{\mathcal{C}} &= \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}
\end{aligned}$$

And thus we have that

$${}_{\mathcal{C}}[L]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 2 \end{bmatrix}$$

If our domain and codomain are the same vector space, then we might use the same basis $\mathcal{B}$ for both. In these situations, we simply write ${}_{\mathcal{C}}[L]_{\mathcal{B}}$ as $[L]_{\mathcal{B}}$.

Example: Let $L : \mathbb{R}^3 \to \mathbb{R}^3$ be defined by $L(a, b, c) = (2a, a + b, 4b + c)$, and let $\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \right\}$. To find $[L]_{\mathcal{B}}$, we need to compute the $\mathcal{B}$-coordinates of the image under L of the $\mathcal{B}$ basis vectors. First, let's just find the image under L of the $\mathcal{B}$ basis vectors.

$$\begin{aligned}
L(1, 0, 1) &= (2, 1, 1) \\
L(2, 1, 1) &= (4, 3, 5) \\
L(-1, 1, 0) &= (-2, 0, 4)
\end{aligned}$$

Now let's try to find the $\mathcal{B}$-coordinates for our results. We can quickly see that $\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$. But the other two are harder to see. So, let's solve for them. That is, we need to find $a_1, b_1, c_1 \in \mathbb{R}$ such that

$$a_1 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + b_1 \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + c_1 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix}$$

and $a_2, b_2, c_2 \in \mathbb{R}$ such that

$$a_2 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} + b_2 \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} + c_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 0 \\ 4 \end{bmatrix}$$

Our first equations is equivalent to the system

$$\begin{array}{cccl} a_1 & +2b_1 & -c_1 & = 4 \\ & b_1 & +c_1 & = 3 \\ a_1 & +b_1 & & = 5 \end{array}$$

while our second equation is equivalent to the system

$$\begin{array}{cccl} a_2 & +2b_2 & -c_2 & = -2 \\ & b_2 & +c_2 & = 0 \\ a_2 & +b_2 & & = 4 \end{array}$$

Since these systems have the same coefficient matrix, we can solve them simultaneously by row reducing the following doubly augmented matrix:

$$\left[\begin{array}{rrr|r|r} 1 & 2 & -1 & 4 & -2 \\ 0 & 1 & 1 & 3 & 0 \\ 1 & 1 & 0 & 5 & 4 \end{array}\right] R_3 - R_1 \quad \sim \left[\begin{array}{rrr|r|r} 1 & 2 & -1 & 4 & -2 \\ 0 & 1 & 1 & 3 & 0 \\ 0 & -1 & 1 & 1 & 6 \end{array}\right] R_3 + R_2$$

$$\sim \left[\begin{array}{rrr|r|r} 1 & 2 & -1 & 4 & -2 \\ 0 & 1 & 1 & 3 & 0 \\ 0 & 0 & 2 & 4 & 6 \end{array}\right] (1/2)R_3 \quad \sim \left[\begin{array}{rrr|r|r} 1 & 2 & -1 & 4 & -2 \\ 0 & 1 & 1 & 3 & 0 \\ 0 & 0 & 1 & 2 & 3 \end{array}\right] \begin{array}{l} R_1 + R_3 \\ R_2 - R_3 \\ \end{array}$$

$$\sim \left[\begin{array}{rrr|r|r} 1 & 2 & 0 & 6 & 1 \\ 0 & 1 & 0 & 1 & -3 \\ 0 & 0 & 1 & 2 & 3 \end{array}\right] \begin{array}{l} R_1 - 2R_2 \\ \\ \end{array} \quad \sim \left[\begin{array}{rrr|r|r} 1 & 0 & 0 & 4 & 7 \\ 0 & 1 & 0 & 1 & -3 \\ 0 & 0 & 1 & 2 & 3 \end{array}\right]$$

And so we see that $a_1 = 4$, $b_1 = 1$, $c_1 = 2$, $a_2 = 7$, $b_2 = -3$, and $c_2 = 3$. This means that $\begin{bmatrix} 4 \\ 3 \\ 5 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} 4 \\ 1 \\ 2 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ 0 \\ 4 \end{bmatrix}_{\mathcal{B}} = \begin{bmatrix} 7 \\ -3 \\ 3 \end{bmatrix}$. And all of this means that

$$[L]_{\mathcal{B}} = \begin{bmatrix} 0 & 4 & 7 \\ 1 & 1 & -3 \\ 0 & 2 & 3 \end{bmatrix}$$

We've spent a lot of time looking at how to find ${}_{\mathcal{C}}[L]_{\mathcal{B}}$, but what do we do with it when we have it? We use it to compute $[L(\mathbf{x})]_{\mathcal{C}}$ from $[\mathbf{x}]_{\mathcal{B}}$ of course!

Example: Let $L : M(2,2) \to P_2$ be defined by $L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (a+b) + (a+c)x + (a+d)x^2$, and let $\mathcal{B} = \left\{\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}\right\}$ and $\mathcal{C} = \{1, 1+x, 1+x+x^2\}$, as in our earlier example. We already found that

$${}_{\mathcal{C}}[L]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 2 \end{bmatrix}$$

Now, if we have $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 4 \end{bmatrix}$, then we get that

$$\begin{aligned} [L(\mathbf{x})]_{\mathcal{C}} &= {}_{\mathcal{C}}[L]_{\mathcal{B}}[\mathbf{x}]_{\mathcal{B}} \\ &= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ -1 \\ 4 \end{bmatrix} \\ &= \begin{bmatrix} 2 \\ -1 \\ 10 \end{bmatrix} \end{aligned}$$

This means that $L(\mathbf{x}) = 2(1) - 1(1+x) + 10(1+x+x^2) = 11 + 9x + 10x^2$.

We can also compute $L(\mathbf{x})$ by first finding $\mathbf{x}$: since $[\mathbf{x}]_{\mathcal{B}} = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 4 \end{bmatrix}$, we have $\mathbf{x} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + 2\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} - \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} + 4\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 6 & 5 \\ 3 & 4 \end{bmatrix}$. And we can then compute that $L(\mathbf{x}) = L\left(\begin{bmatrix} 6 & 5 \\ 3 & 4 \end{bmatrix}\right) = (6+5) + (6+3)x + (6+4)x^2 = 11 + 9x + 10x^2$, same as before.