Lecture 1r

Change of Coordinates and Linear Mappings

(pages 240-242)

Let's take another look at an example from the previous lecture:

Example: Let $L : M(2,2) \to P_2$ be defined by

$$L\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = (a+b) + (a+c)x + (a+d)x^2$$

We found the matrix for L with respect to two different pairs of bases. If $\mathcal{B}_1$ is the standard basis for $M(2,2)$ and $\mathcal{C}_1$ is the standard basis for P_2, we have

$${}_{\mathcal{C}_1}[L]_{\mathcal{B}_1} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix}$$

and if $\mathcal{B}_2 = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\}$ and $\mathcal{C}_2 = \{1, 1+x, 1+x+x^2\}$, we have

$${}_{\mathcal{C}_2}[L]_{\mathcal{B}_2} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 2 \end{bmatrix}$$

Since it is reasonably easy to find the coordinates for a matrix with respect to the standard basis, we can easily use our first matrix to compute $L(\mathbf{x})$. But before we can use the second matrix to compute $L(\mathbf{x})$, we would first need to find the $\mathcal{B}_2$-components of $\mathbf{x}$. Now, if we only wanted to do that for one vector $\mathbf{x}$, we would do so by setting up an equation and row reducing a matrix. However, the more likely situation is that we would want to compute $L(\mathbf{x})$ for many (perhaps even all) matrices $\mathbf{x} \in M(2,2)$, and the faster way to do that would be to find the change of coordinates matrix from the standard basis ($\mathcal{B}_1$) to $\mathcal{B}_2$. We've already looked at how to do that, so the point now is that we can also use our change of coordinates matrix to find ${}_{\mathcal{C}_2}[L]_{\mathcal{B}_2}$ if we already have ${}_{\mathcal{C}_1}[L]_{\mathcal{B}_1}$.

Let R be the change of coordinates matrix from $\mathcal{B}_1$ to $\mathcal{B}_2$, and let Q be the change of coordinates matrix from $\mathcal{C}_1$ to $\mathcal{C}_2$. Then we have that

$$\begin{array}{llll} [L(\mathbf{x})]_{\mathcal{C}_2} & = Q[L(\mathbf{x})]_{\mathcal{C}_1} & = Q\,{}_{\mathcal{C}_1}[L]_{\mathcal{B}_1}[\mathbf{x}]_{\mathcal{B}_1} & \text{and} \\ [L(\mathbf{x})]_{\mathcal{C}_2} & = {}_{\mathcal{C}_2}[L]_{\mathcal{B}_2}[\mathbf{x}]_{\mathcal{B}_2} & = {}_{\mathcal{C}_2}[L]_{\mathcal{B}_2}R[\mathbf{x}]_{\mathcal{B}_1} & \end{array}$$

So, for every vector $[\mathbf{x}]_{\mathcal{B}_1}$ in $\mathbb{R}^n$, which will in fact be every vector in $\mathbb{R}^n$, we have that

$$Q_{\mathcal{C}_1}[L]_{\mathcal{B}_1}[\mathbf{x}]_{\mathcal{B}_1} = {}_{\mathcal{C}_2}[L]_{\mathcal{B}_2}R[\mathbf{x}]_{\mathcal{B}_1}$$

It's been a while, but if you recall Theorem 3.1.4, then you see that this means

$$Q_{\mathcal{C}_1}[L]_{\mathcal{B}_1} = {}_{\mathcal{C}_2}[L]_{\mathcal{B}_2}R$$

Since R is a change of coordinates matrix, R is invertible (as R^{-1} is the change of coordinates matrix from $\mathcal{B}_2$ to $\mathcal{B}_1$) and so we have

$${}_{\mathcal{C}_2}[L]_{\mathcal{B}_2} = Q_{\mathcal{C}_1}[L]_{\mathcal{B}_1}R^{-1}$$

Now, I started this explanation in regards to our example, but I finished using only general notation, so in fact we've shown that this holds in general.

Summary: Let $L : \mathbb{V} \to \mathbb{W}$ be a linear mapping, $\mathcal{B}_1$ and $\mathcal{B}_2$ be bases for $\mathbb{V}$, and $\mathcal{C}_1$ and $\mathcal{C}_2$ be bases for $\mathbb{W}$. Then if we let R be the change of coordinates matrix from $\mathcal{B}_1$ to $\mathcal{B}_2$ and Q be the change of coordinates matrix from $\mathcal{C}_1$ to $\mathcal{C}_2$, we have

$${}_{\mathcal{C}_2}[L]_{\mathcal{B}_2} = Q_{\mathcal{C}_1}[L]_{\mathcal{B}_1}R^{-1}$$

Example: Let's go ahead and apply this result to our example. To do so, we will first need to find R^{-1} and Q. We have

$$\begin{aligned}
R^{-1} &= \begin{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}_{\mathcal{B}_1} & \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}_{\mathcal{B}_1} & \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}_{\mathcal{B}_1} & \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}_{\mathcal{B}_1} \end{bmatrix} \\
&= \begin{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} \end{bmatrix} \\
&= \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}
\end{aligned}$$

and

$$
\begin{aligned}
Q &= \left[\ [1]_{\mathcal{C}_2} \quad [x]_{\mathcal{C}_2} \quad [x^2]_{\mathcal{C}_2}\ \right] \\
&= \left[\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right] \\
&= \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix}
\end{aligned}
$$

Now it is a simple case of matrix multiplication:

$$
\begin{aligned}
{}_{\mathcal{C}_2}[L]_{\mathcal{B}_2} &= Q\,{}_{\mathcal{C}_1}[L]_{\mathcal{B}_1} R^{-1} \\
&= \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 2 & 2 \\ 1 & 1 & 2 & 2 \\ 1 & 1 & 1 & 2 \end{bmatrix} \\
&= \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 2 \end{bmatrix}
\end{aligned}
$$

which is the same as our previous result.

Course Author's Note: I have deviated from the text again here, in order to provide the most general result. So let's see how to take our result, and make it the same as the result in the text. For starters, the text only looks at the case where L is a linear operator, so $\mathbb{V}$ and $\mathbb{W}$ are the same. It further makes the restriction that $\mathbb{V}$ is $\mathbb{R}^n$, although that is rather extreme. Why does the text bother with the notation $[\vec{x}]_S$, when $[\vec{x}]_S = \vec{x}$? But there is value in assuming that one of our bases is the standard basis S. The other significant restriction is that the book sets $\mathcal{B}_1 = \mathcal{C}_1$ and $\mathcal{B}_2 = \mathcal{C}_2$. That is, not only does it have the domain and codomain equal to each other, but we are looking at the matrix with respect to the same basis for the domain and codomain. So, to get the result in the book, you have to set $\mathcal{B}_1 = \mathcal{C}_1 = S$ and $\mathcal{B}_2 = \mathcal{C}_2 = \mathcal{B}$. Then R is the change of coordinates matrix from S to $\mathcal{B}$, and Q is also the change of coordinates matrix from S to $\mathcal{B}$. As such, $R = Q$, and our result is now

$$[L]_{\mathcal{B}} = R[L]_{\mathcal{S}}R^{-1}$$

This almost what the text writes, except we need to replace R with P^{-1} and R^{-1} with P. And that is because they are defined as inverses. While I found it more straightforward to use the change of coordinates matrix from $\mathcal{S}$ to $\mathcal{B}$, the text based its result on the change of coordinates matrix from $\mathcal{B}$ to $\mathcal{S}$.

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 will first find $[L]_S$:

$$\begin{aligned} [L]_S &= \begin{bmatrix} L\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) & L\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) & L\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) \end{bmatrix} \\ &= \begin{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \\ 4 \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix} \end{aligned}$$

Next, we need to find P:

$$\begin{aligned} P &= \begin{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}_S & \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}_S & \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}_S \end{bmatrix} \\ &= \begin{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} & \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} & \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \end{aligned}$$

Now we need to find P^{-1}, using the matrix inverse algorithm:

$$\left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right] R_3 - R_1 \sim \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & -1 & 1 & -1 & 0 & 1 \end{array}\right] R_3 + R_2$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 2 & -1 & 1 & 1 \end{array}\right] (1/2)R_3 \sim \left[\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & -1/2 & 1/2 & 1/2 \end{array}\right] \begin{array}{l} R_1 + R_3 \\ R_2 - R_3 \\ \end{array}$$

$$\sim \left[\begin{array}{ccc|ccc} 1 & 2 & 0 & 1/2 & 1/2 & 1/2 \\ 0 & 1 & 0 & 1/2 & 1/2 & -1/2 \\ 0 & 0 & 1 & -1/2 & 1/2 & 1/2 \end{array}\right] \begin{array}{c} R_1 - 2R_2 \\ \\ \\ \end{array} \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & -1/2 & -1/2 & 3/2 \\ 0 & 1 & 0 & 1/2 & 1/2 & -1/2 \\ 0 & 0 & 1 & -1/2 & 1/2 & 1/2 \end{array}\right]$$

So $P^{-1} = \begin{bmatrix} -1/2 & -1/2 & 3/2 \\ 1/2 & 1/2 & -1/2 \\ -1/2 & 1/2 & 1/2 \end{bmatrix}$.

At long last, we can compute $[L]_{\mathcal{B}}$:

$$\begin{aligned} [L]_{\mathcal{B}} &= P^{-1}[L]_{\mathcal{S}}P \\ &= \begin{bmatrix} -1/2 & -1/2 & 3/2 \\ 1/2 & 1/2 & -1/2 \\ -1/2 & 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 4 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & -1 \\ 0 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} -1/2 & -1/2 & 3/2 \\ 1/2 & 1/2 & -1/2 \\ -1/2 & 1/2 & 1/2 \end{bmatrix} \begin{bmatrix} 2 & 4 & -2 \\ 1 & 3 & 0 \\ 1 & 5 & 4 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 4 & 7 \\ 1 & 1 & -3 \\ 0 & 2 & 3 \end{bmatrix} \end{aligned}$$

This is, of course, the same matrix we found in the previous lecture.