Lecture 3k

Compositions and Combinations

(pages 139-141)

Since linear mappings are first and foremost types of functions, we can take sums, scalar multiples, and compositions of linear mappings just as we could any function. But here are those definitions, just to refresh your memory. (The definition will be given for linear mappings specifically, as that is what we are concerned with.)

Definition: Let L be a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$, and let $t \in \mathbb{R}$ be a scalar. We define (tL) to be the mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ such that

$$(tL)(\vec{x}) = t(L(\vec{x}))$$

Example: Let $L : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by $L(x_1, x_2, x_3) = (5x_2, x_1 + 3x_3)$. Then:

$$\begin{array}{rcl} L(1,1,1) = (5,4) & \Rightarrow & (2L)(1,1,1) = 2L(1,1,1) = 2(5,4) = (10,8) \\ L(1,2,3) = (10,10) & \Rightarrow & (\frac{1}{2}L)(1,2,3) = \frac{1}{2}L(1,2,3) = \frac{1}{2}(10,10) = (5,5) \\ L(-1,0,1) = (0,2) & \Rightarrow & (-4L)(-1,0,1) = -4L(-1,0,1) = -4(0,2) = (0,-8) \end{array}$$

Definition: Let L and M be linear mappings from $\mathbb{R}^n$ to $\mathbb{R}^m$. We define $(L+M)$ to be the mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ such that

$$(L + M)(\vec{x}) = L(\vec{x}) + M(\vec{x})$$

Example: Let L be as in the previous example, and let $M : \mathbb{R}^3 \to \mathbb{R}^2$ be defined by $M(x_1, x_2, x_3) = (0, x_1 + x_2 + x_3)$. Then:

$$\begin{array}{ll} L(1,1,1) = (5,4) & M(1,1,1) = (0,3) \Rightarrow \\ & (L+M)(1,1,1) = L(1,1,1) + M(1,1,1) = (5,4) + (0,3) = (5,7) \\ L(1,2,3) = (10,10) & M(1,2,3) = (0,6) \Rightarrow \\ & (L+M)(1,2,3) = L(1,2,3) + M(1,2,3) = (10,10) + (0,6) = (10,16) \\ L(-1,0,1) = (0,2) & M(-1,0,1) = (0,0) \Rightarrow \\ & (L+M)(-1,0,1) = L(-1,0,1) + M(-1,0,1) = (0,2) + (0,0) = (0,2) \end{array}$$

One thing to notice is that we can not take the sum of arbitrary linear mappings, but instead only those with the same domain and the same codomain. Things become more complicated when we look at compositions of linear mappings.

Definition: Let $L : \mathbb{R}^n \to \mathbb{R}^m$ and $N : \mathbb{R}^m \to \mathbb{R}^p$ be linear mappings. The **composition** $N \circ L : \mathbb{R}^n \to \mathbb{R}^p$ is defined by

$$(N \circ L)(\vec{x}) = N(L(\vec{x})) \qquad \text{for all } \vec{x} \in \mathbb{R}^n.$$

Example: Let L be as in the previous two examples, and let $N : \mathbb{R}^2 \to \mathbb{R}^4$ be defined by $N(x_1, x_2) = (x_1, -x_1, x_2, -2x_2)$. Then:

$$\begin{aligned}
&(N \circ L)(1,1,1) = N(L(1,1,1)) = N(5,4) = (5,-5,4,-8)\\
&(N \circ L)(1,2,3) = N(L(1,2,3)) = N(10,10) = (10,-10,10,-20)\\
&(N \circ L)(-1,0,1) = N(L(-1,0,1)) = N(0,2) = (0,0,2,-4)
\end{aligned}$$

Again, we cannot take the composition of arbitrary linear mappings, but only those with compatible domains and codomains. Specifically, we need the codomain of L to be the same as the domain of N. The domain of L and the codomain of N are not restricted.

Though we start with linear mappings, since these definitions are simply based on the fact that they are functions, we don't yet know that our resulting functions are in fact linear as well. But, of course, they are. We see the first part of this in the following theorem.

Theorem 3.2.4: If L and M are linear mappings from $\mathbb{R}^n$ to $\mathbb{R}^m$ and $t \in \mathbb{R}$, the $(L + M)$ and (tL) are linear mappings.

The next result is an exercise in the textbook:

Exercise 3.2.5: If $L : \mathbb{R}^n \to \mathbb{R}^m$ and $N : \mathbb{R}^m \to \mathbb{R}^p$ are linear mappings, then $N \circ L$ is a linear mapping.

It is a bit unfortunate that Exercise 3.2.5 did not rate theorem status in the eyes of the author, but I will afford it that status. That is to say, that you should always take it for granted that sum, scalar multiples, and compositions of linear mappings are themselves linear mappings. I will give the proof of Theorem 3.2.4 now, and will leave the proof of Exercise 3.2.5 as a practice problem.

Proof of Theorem 3.2.4: Let L and M be linear mappings from $\mathbb{R}^n$ to $\mathbb{R}^m$ and let $t \in \mathbb{R}$. We show that $(L + M)$ preserves addition as follows:

$$\begin{aligned}
(L+M)(\vec{x}+\vec{y}) &= L(\vec{x}+\vec{y}) + M(\vec{x}+\vec{y}) && \text{(definition of function addition)}\\
&= L(\vec{x}) + L(\vec{y}) + M(\vec{x}) + M(\vec{y}) && \text{(because } L, M \text{ are linear)}\\
&= L(\vec{x}) + M(\vec{x}) + L(\vec{y}) + M(\vec{y}) && \text{(commutivity of addition in } \mathbb{R}^m\text{)}\\
&= (L+M)(\vec{x}) + (L+M)(\vec{y}) && \text{(definition of function addition)}
\end{aligned}$$

We show that $(L + M)$ preserves scalar multiplication as follows:

$$\begin{aligned}
(L+M)(t\vec{x}) &= L(t\vec{x}) + M(t\vec{x}) && \text{(definition of function addition)}\\
&= tL(\vec{x}) + tM(\vec{x}) && \text{(because } L, M \text{ are linear)}\\
&= t(L(\vec{x}) + M(\vec{x})) && \text{(distributive property of } \mathbb{R}^m\text{)}\\
&= t(L+M)(\vec{x}) && \text{(definition of function addition)}
\end{aligned}$$

Thus, we see that $(L+M)$ is a linear mapping. And now, we show that (tL) preserves addition as follows:

$$\begin{aligned}(tL)(\vec{x}+\vec{y}) &= t(L(\vec{x}+\vec{y})) && \text{(definition of scalar times function)}\\ &= t(L(\vec{x})+L(\vec{y})) && \text{(because } L \text{ is linear)}\\ &= tL(\vec{x})+tL(\vec{y}) && \text{(distributive property of } \mathbb{R}^m\text{)}\\ &= (tL)(\vec{x})+(tL)(\vec{y}) && \text{(definition of scalar times function)}\end{aligned}$$

Lastly, we show that (tL) preserves scalar multiplication as follows:

$$\begin{aligned}(tL)(s\vec{x}) &= t(L(s\vec{x})) && \text{(definition of scalar times function)}\\ &= t(sL(\vec{x})) && \text{(because } L \text{ is linear)}\\ &= (ts)(L(\vec{x})) && \text{(associativity of scalar multiplication in } \mathbb{R}^m\text{)}\\ &= (st)(L(\vec{x})) && \text{(commutivity of multiplication in } \mathbb{R}\text{)}\\ &= s(tL(\vec{x})) && \text{(associativity of scalar multiplication in } \mathbb{R}^m\text{)}\\ &= s((tL)(\vec{x})) && \text{(definition of scalar times function)}\end{aligned}$$

And since (tL) preserves addition and scalar multiplication, we see that (tL) is a linear mapping.

I choose to show that these mappings preserved addition and scalar multiplication separately, while the textbook (which only gives the proof of (tL)) used the combined method. Given that some of the properties used in the proof are back from section 1.2, I wanted to make sure that you remembered why we could do each and every operation. In my solution for Exercise 3.2.5, I use the combined method, so you can see an example of that technique there. Both are fine, and generally should depend on your comfort level with the various properties of $\mathbb{R}^n$.

So now that we know that $L+M$, tL, and $N \circ L$ are linear mappings, we would like to be able to find their standard matrices. Luckily, this is straightforward to do once you have $[L]$, $[M]$, and $[N]$.

<u>Theorem 3.2.5</u>: Let $L : \mathbb{R}^n \to \mathbb{R}^m$, $M : \mathbb{R}^n \to \mathbb{R}^m$, and $N : \mathbb{R}^m \to \mathbb{R}^p$ be linear mappings and $t \in \mathbb{R}$. Then

$$[L+M] = [L]+[M], \qquad [tL] = t[L], \qquad [N \circ L] = [N][L]$$

Proof of Theorem 5: To prove these, we will use Theorem 3.1.4, which states that two matrices A and B are equal if $A\vec{x} = B\vec{x}$ for all $\vec{x} \in \mathbb{R}^m$. With that in mind, we see that $[L+M] = [L]+[M]$ as follows:

$[L+M]\vec{x} = (L+M)(\vec{x}) = L(\vec{x}) + M(\vec{x}) = [L]\vec{x} + [M]\vec{x} = ([L]+[M])\vec{x}$, where this last equality is due to a distributive property of matrix multiplication.

Similarly, we see that $[tL] = t[L]$ as follows:

$[tL]\vec{x} = (tL)(\vec{x}) = tL(\vec{x}) = t([L]\vec{x}) = (t[L])\vec{x}$, where this last equality is due to the associativity of scalar multiplication in matrices.

Finally, we see that $[N \circ L] = [N][L]$ as follows:

$[N \circ L]\vec{x} = (N \circ L)(\vec{x}) = N(L(\vec{x})) = N([L]\vec{x}) = [N]([L]\vec{x}) = ([N][L])\vec{x}$, where this last equality is due to the associativity of matrix multiplication.

Example: Let's find $[L]$, $[M]$, and $[N]$ for the linear mappings in our earlier examples, and use them to find $[2L]$, $[L+M]$, and $[N \circ L]$.

To find $[L]$, we note that $L(1,0,0) = (0,1)$, $L(0,1,0) = (5,0)$, and $L(0,0,1) = (0,3)$. Thus,

$$[L] = \begin{bmatrix} 0 & 5 & 0 \\ 1 & 0 & 3 \end{bmatrix}$$

To find $[M]$, we note that $M(1,0,0) = (0,1)$, $M(0,1,0) = (0,1)$, and $M(0,0,1) = (0,1)$. Thus,

$$[M] = \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

To find $[N]$, we note that $N(1,0) = (1,-1,0,0)$ and $N(0,1) = (0,0,1,-2)$. Thus,

$$[N] = \begin{bmatrix} 1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & -2 \end{bmatrix}$$

This means that

$$[2L] = 2[L] = 2\begin{bmatrix} 0 & 5 & 0 \\ 1 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 10 & 0 \\ 2 & 0 & 6 \end{bmatrix}$$

We can verify that this coincides with our earlier result, by seeing that $[2L]\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 & 10 & 0 \\ 2 & 0 & 6 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 10 \\ 8 \end{bmatrix}$. And we have that

$$[L+M] = [L] + [M] = \begin{bmatrix} 0 & 5 & 0 \\ 1 & 0 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 5 & 0 \\ 2 & 1 & 4 \end{bmatrix}$$

We can verify that this coincides with our earlier results by seeing that

$$[L+M]\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}0&5&0\\2&1&4\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}5\\7\end{bmatrix}$$

$$[L+M]\begin{bmatrix}1\\2\\3\end{bmatrix}=\begin{bmatrix}0&5&0\\2&1&4\end{bmatrix}\begin{bmatrix}1\\2\\3\end{bmatrix}=\begin{bmatrix}10\\16\end{bmatrix}$$

$$[L+M]\begin{bmatrix}-1\\0\\1\end{bmatrix}=\begin{bmatrix}0&5&0\\2&1&4\end{bmatrix}\begin{bmatrix}-1\\0\\1\end{bmatrix}=\begin{bmatrix}0\\2\end{bmatrix}$$

Lastly, we have that

$$[N\circ L]=[N][L]=\begin{bmatrix}1&0\\-1&0\\0&1\\0&-2\end{bmatrix}\begin{bmatrix}0&5&0\\1&0&3\end{bmatrix}=\begin{bmatrix}0&5&0\\0&-5&0\\1&0&3\\-2&0&-6\end{bmatrix}$$

We can verify that this coincides with our earlier results by seeing that

$$[N\circ L]\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}0&5&0\\0&-5&0\\1&0&3\\-2&0&-6\end{bmatrix}\begin{bmatrix}1\\1\\1\end{bmatrix}=\begin{bmatrix}5\\-5\\4\\-8\end{bmatrix}$$

$$[N\circ L]\begin{bmatrix}1\\2\\3\end{bmatrix}=\begin{bmatrix}0&5&0\\0&-5&0\\1&0&3\\-2&0&-6\end{bmatrix}\begin{bmatrix}1\\2\\3\end{bmatrix}=\begin{bmatrix}10\\-10\\10\\-20\end{bmatrix}$$

$$[N\circ L]\begin{bmatrix}-1\\0\\1\end{bmatrix}=\begin{bmatrix}0&5&0\\0&-5&0\\1&0&3\\-2&0&-6\end{bmatrix}\begin{bmatrix}-1\\0\\1\end{bmatrix}=\begin{bmatrix}0\\0\\2\\-4\end{bmatrix}$$