Lecture 3v

Inverse Linear Mappings

(pages 170-3)

At the end of the previous lecture, we looked at using the matrix inverse $A^{-1}$ to find $\vec{x}$ such that $A\vec{x} = \vec{b}$. Specifically, we saw that $A^{-1}\vec{b}$ was the only such $\vec{x}$, which means that $A(A^{-1}\vec{b}) = \vec{b}$. Instead of thinking of this as a system of equations, or as matrix multiplication, let's interpret this as a linear mapping. Then we have shown that $(A \circ A^{-1})\vec{b} = \vec{b}$ for all $\vec{b} \in \mathbb{R}^n$. This means that $A \circ A^{-1}$ is the same as the identity mapping Id, defined by $\mathrm{Id}(\vec{x}) = \vec{x}$. And it is in this way that we consider the notion of the inverse of a linear mapping.

Definition: If $L : \mathbb{R}^n \to \mathbb{R}^n$ is a linear mapping and there exists another linear mapping $M : \mathbb{R}^n \to \mathbb{R}^n$ such that $M \circ L = \mathrm{Id} = L \circ M$, then $L$ is said to be invertible, and $M$ is called the inverse of $L$, usually denoted $L^{-1}$.

This definition parallels the definition of an invertible matrix. Note, in particular, that we only define the inverse of a linear operator (a linear mapping whose domain and codomain are the same), which parallels the fact that we only defined the inverse for square matrices.

But thinking of linear mappings as functions, our definition of inverse is the same as the usual definition of a function inverse. For, if $M$ is the inverse of $L$ and $L(\vec{x}) = \vec{y}$, then $M(\vec{y}) = M(L(\vec{x})) = (M \circ L)(\vec{x}) = \mathrm{Id}(\vec{x}) = \vec{x}$, so if $L(\vec{x}) = \vec{y}$, then $M(\vec{y}) = \vec{x}$. Thus, $M$ simply reverses the action of $L$. Moreover, if $M(\vec{v}) = \vec{w}$, then $L(\vec{w}) = L(M(\vec{v})) = (L \circ M)(\vec{v}) = \mathrm{Id}(\vec{v}) = \vec{v}$. So, if $M(\vec{v}) = \vec{w}$, then $L(\vec{w}) = \vec{v}$, and we see that $L$ reverses the action of $M$.

As all of our work with linear mappings simply ends up being work with matrices, we note the following (hopefully obvious) fact.

Theorem 3.5.5: Suppose that $L : \mathbb{R}^n \to \mathbb{R}^n$ is a linear mapping with standard matrix $[L] = A$, and that $M : \mathbb{R}^n \to \mathbb{R}^n$ is a linear mapping with standard matrix $[M] = B$. Then $M$ is the inverse of $L$ if and only if $B$ is the inverse of $A$.

Proof of Theorem 3.5.5: Since $[L \circ M] = [L][M] = AB$, we see that $L \circ M = \mathrm{Id}$ if and only if $[L \circ M] = I$, if and only if $AB = I$. And $AB = I$ if and only if $B$ is the inverse of $A$. Similarly, $M \circ L = \mathrm{Id}$ if and only if $BA = I$, which happens if and only if $B$ is the inverse of $A$. So, we see that $L \circ M = \mathrm{Id} = M \circ L$ if and only if $B$ is the inverse of $A$.

**Example**: Consider the linear mapping $T$ that is a dilation by a factor of 10 in $\mathbb{R}^2$. We saw in Section 3.3 that the standard matrix for this linear mapping is $\begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}$. But thinking of the mapping as a geometrical transformation, we easily see that it is invertible, since we can reverse the action of dilating by

10 by contracting by 1/10. The standard matrix for the mapping $C$ that is a contraction by 1/10 in $\mathbb{R}^2$ is $\begin{bmatrix} 1/10 & 0 \\ 0 & 1/10 \end{bmatrix}$. And Theorem 3.5.5 tells us that since $C = T^{-1}$, then $[C] = [T]^{-1}$. That is

$$\begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix}^{-1} = \begin{bmatrix} 1/10 & 0 \\ 0 & 1/10 \end{bmatrix}$$

which we can easily verify by multiplying: $\begin{bmatrix} 10 & 0 \\ 0 & 10 \end{bmatrix} \begin{bmatrix} 1/10 & 0 \\ 0 & 1/10 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

We can use the facts we've already discovered about inverse matrices, and now apply them to the study of linear mappings:

Theorem 3.5.6 (Invertible Matrix Theorem cont.): Suppose that $L : \mathbb{R}^n \to \mathbb{R}^n$ is a linear mapping with standard matrix $A = [L]$. Then the following statements are equivalent to each other and to the statements of Theorem 4.

(7) $L$ is invertible
(8) $\mathrm{Range}(L) = \mathbb{R}^n$
(9) $\mathrm{Null}(L) = \{\vec{0}\}$.

Proof of Theorem 3.5.6: To link this with the previous results in the Invertible Matrix Theorem, I will prove the chain (1) $\Rightarrow$ (7) $\Rightarrow$ (8) $\Rightarrow$ (9) $\Rightarrow$ (1), where $P \Rightarrow Q$ means "if $P$, then $Q$".

if (1), then (7): Suppose $A$ is invertible. Then, by Theorem 3.5.5, $L$ is invertible.

if (7), then (8): Suppose $L$ is invertible, and let $\vec{x} \in \mathbb{R}^n$. Then $L^{-1}(\vec{x})$ is such that $L(L^{-1}(\vec{x})) = \vec{x}$. Thus, $\vec{x}$ is in the range of $L$. And since we have shown this for all $\vec{x} \in \mathbb{R}^n$, we see that $\mathrm{Range}(L) = \mathbb{R}^n$.

if (8), then (9): Suppose $\mathrm{Range}(L) = \mathbb{R}^n$. Then $\mathrm{Col}(A) = \mathbb{R}^n$. This means that the rank of $A$ is $n$, and thus, by the rank-nullity theorem, the nullity of $A$ is 0. Thus, $\mathrm{Null}(A) = \{\vec{0}\}$, and so $\mathrm{Null}(L) = \{\vec{0}\}$.

if (9), then (1): Suppose $\mathrm{Null}(L) = \{\vec{0}\}$. Then the nullity of $A$ is 0, and thus, by the rank-nullity theorem, the rank of $A$ is $n$. From our previous result with the Invertible Matrix Theorem, we know that this means that $A$ is invertible.

You'll notice that my proof of the Invertible Matrix Theorem is not the same as the one presented in the book. For this part in particular, I choose to stick with the more standard cycle structure, while the book's choices were chosen for convenience. I also find that the book gives a more "hands on" proof of these facts, while my proofs use previous results more. I hope that you will take the time to read the proof in the book, as I think both proofs give insight into these connections. The textbook also notes that you should be able to link any of the properties (1)-(9) with each other, and you might want to spend some time considering which of these connections seem obvious, and which seem more surprising.

**Example:** Consider the linear mapping $L : \mathbb{R}^5 \to \mathbb{R}^5$ defined by

$$L(x_1, x_2, x_3, x_4, x_5) = (x_1 + x_2, x_3 + x_4, x_5, x_1 + x_2, x_3 + x_4)$$

$L$ is not invertible. Instead of finding the standard matrix for $L$ (which would be $5 \times 5$) and row reducing, we can instead notice that $L(1, -1, 1, -1, 0) = (1 - 1, 1 - 1, 0, 1 - 1, 1 - 1) = (0, 0, 0, 0, 0)$. Thus, $\mathrm{Null}(L) \neq \{\vec{0}\}$, and so by the Invertible Matrix Theorem $L$ is not invertible.