1. Inverting linear transformations and matrices

Let $\mathrm{id}_X : X \to X$ denote the 'identify function' on a set X, given by $\mathrm{id}_X(x) = x$ for all elements $x \in X$.

Definition 1.1. *A linear tranformation $T : \mathbf{R}^n \to \mathbf{R}^m$ is called* invertible *if there exists another linear tranformation $S : \mathbf{R}^m \to \mathbf{R}^n$ such that $T \circ S = \mathrm{id}_{\mathbf{R}^m}$ and $S \circ T = \mathrm{id}_{\mathbf{R}^n}$. We call S the* inverse *of T, and we write $T^{-1} = S$.*

Theorem 1.2. *Let $T : \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation with standard matrix $A \in \mathcal{M}_{m\times n}$. Then the following are equivalent*

1. *T is invertible.*
2. *There is a matrix $B \in \mathcal{M}_{n\times m}$ such that $AB = I_{m\times m}$ and $BA = I_{n\times n}$.*
3. *The equation $A\mathbf{x} = \mathbf{b}$ has a unique solution for every $\mathbf{b} \in \mathbf{R}^m$.*
4. *A is square and non-singular.*

Proof. My strategy will be to show that (1) $\iff$ (2) and (3) $\iff$ (4) and then finally (2) $\iff$ (3).

That (1) and (2) are equivalent is more or less immediate: i.e. if $S : \mathbf{R}^m \to \mathbf{R}^n$ if some other linear transformation and B be its matrix. From class we know that the matrix for $T \circ S$ is AB and the matrix for $\mathrm{id}_{\mathbf{R}^m}$ is $I_{m\times m}$. Hence $T \circ S = \mathrm{id}_{\mathbf{R}^m}$ is equivalent to $AB = I_{m\times m}$. Similarly, $S \circ T = \mathrm{id}_{\mathbf{R}^n}$ is the same as $BA = I_{n\times n}$.

To see that (3) and (4) are equivalent, note first that if (3) holds, then A must be row equivalent to a matrix $\tilde{A}$ in reduced echelon form with a pivot in every column (because $A\mathbf{x} = \mathbf{b}$ has at *most* one solution) and a pivot in every row (because $A\mathbf{x} = \mathbf{b}$ has at *least* one solution, no matter what $\mathbf{b}$ is). Since there is at most one pivot in each row and column, it follows that the number of rows and columns of $\tilde{A}$ are the same. That is, $\tilde{A}$ and therefore also A are square matrices. From here, the equivalence between (3) and (4) is part of Proposition 4.1.6 in Shifrin (which I stated in class, too).

To see that (2) $\implies$ (3), let $B \in \mathcal{M}_{n\times m}$ be the matrix in (2) and $\mathbf{b} \in \mathbf{R}^m$ be any given vector. If $\mathbf{x} \in \mathbf{R}^m$ solves $A\mathbf{x} = \mathbf{b}$, then

$$B\mathbf{b} = B(A\mathbf{x}) = (BA)\mathbf{x} = I\mathbf{x} = \mathbf{x}.$$

That is, $\mathbf{x} = B\mathbf{b}$ is the only possible solution of $A\mathbf{x} = \mathbf{b}$. On the other hand, I can plug back in to check that it really is a solution:

$$A\mathbf{x} = A(B\mathbf{b}) = (AB)\mathbf{b} = I\mathbf{b} = \mathbf{b}.$$

In short $A\mathbf{x} = \mathbf{b}$ has the unique solution $\mathbf{x} := B\mathbf{b}$, i.e. (3) holds.

To see that (3) $\implies$ (2), recall from above that (3) implies that $m = n$, so that $A \in \mathcal{M}_{n\times n}$. I let $\mathbf{e}_j \in \mathbf{R}^n$ be the jth standard basis vector and apply (3) to get a vector $\mathbf{b}_j \in \mathbf{R}^n$ satisfying $A\mathbf{b}_j = \mathbf{e}_j$. Then I define $B := \begin{bmatrix}\mathbf{b}_1 & \dots & \mathbf{b}_n\end{bmatrix} \in \mathcal{M}_{n\times n}$. It follows that

$$AB = \begin{bmatrix}A\mathbf{b}_1 & \dots & A\mathbf{b}_n\end{bmatrix} = \begin{bmatrix}\mathbf{e}_1 & \dots & \mathbf{e}_n\end{bmatrix} = I.$$

That $BA = I$, too, follows from the lemma below. $\square$

Given the last item in this theorem, we can restrict out discussion of invertibility to linear transformations with the same source and target.

Definition 1.3. *A square matrix $A \in \mathcal{M}_{n\times n}$ is invertible if and only if there exists $B \in \mathcal{M}_{n\times n}$ such that $AB = BA = I$. We then call B the* inverse *of A and write $A^{-1} := B$.*

Lemma 1.4. *Given two square matrices $A, B \in \mathcal{M}_{n\times n}$, we have $AB = I$ if and only if $BA = I$.*

I follow the argument given in Shifrin, which is rather clever.

Proof. Since $AB = I$ and I is a non-singular square matrix, it follows from our homework problem 4.2.17b that A and B are both non-singular. Hence by Proposition 4.1.6 again the linear system $B\mathbf{x} = \mathbf{b}$ has a solutions for any $\mathbf{b} \in \mathbf{R}^n$. In particular, for each standard basis vector $\mathbf{e}_j \in \mathbf{R}^n$ there is a vectors $\mathbf{c}_j \in \mathbf{R}^n$ such that $B\mathbf{c}_j = \mathbf{e}_j$. And as in the proof of the Theorem, if I set $C = \begin{bmatrix}\mathbf{c}_1 & \dots & \mathbf{c}_n\end{bmatrix}$, then it follows that $BC = I$. This allows me to compute the product ABC in two different ways. On the one hand $ABC = A(BC) = AI = A$. On the other hand $ABC = (AB)C = IC = C$. Comparing the answers, I see that $A = C$. So $I = BC = BA$, which is what I aimed to show. $\square$