Worksheet 7: Compositions and Inverses of Transformations (§2.3 §2.4)

Problem 1. Let $\mathbb{R}^2 \xrightarrow{T} \mathbb{R}^3$ and $\mathbb{R}^3 \xrightarrow{S} \mathbb{R}^2$ be linear transformations between coordinates space.

(a) Consider the composition $S \circ T$. What is its source vector space? What is its target?

(b) Suppose that T and S are given by the formulas

$$T(\begin{bmatrix} x \\ y \end{bmatrix}) = \begin{bmatrix} x+2y \\ -y \\ -x+3y \end{bmatrix} \qquad \text{and} \qquad S(\begin{bmatrix} x \\ y \\ z \end{bmatrix}) = \begin{bmatrix} x+z \\ -y \end{bmatrix}.$$

Compute an explicit formula for $S \circ T(\begin{bmatrix} x \\ y \end{bmatrix})$. Is $S \circ T$ also a linear transformation?

(c) Use the **Key Theorem** to find the standard matrices of S, T and $S \circ T$; call these A, B and C.

(d) Compute the matrix product AB, and compare to C. What do you notice? Why is this?

(e) Is $T \circ S = S \circ T$? What is the matrix for $T \circ S$ in terms of A and B?

(f) Prove Theorem 1 below. [Caution: Be careful about the order!]

Theorem 1: If $\mathbb{R}^n \xrightarrow{T_M} \mathbb{R}^m$ and $\mathbb{R}^m \xrightarrow{T_N} \mathbb{R}^p$ are linear transformations with standard matrices M and N respectively, then the composition $T_N \circ T_M$ is a linear transformation with standard matrix NM.

Solution:

(a) The source and target are both $\mathbb{R}^2$.

(b) $S \circ T(\begin{bmatrix} x \\ y \end{bmatrix}) = S(\begin{bmatrix} x+2y \\ -y \\ -x+3y \end{bmatrix}) = \begin{bmatrix} x+2y+(-x+3y) \\ y \end{bmatrix} = \begin{bmatrix} 5y \\ y \end{bmatrix}$

(c&d) $A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 \\ 0 & -1 \\ -1 & 3 \end{bmatrix}$, and $C = \begin{bmatrix} 0 & 5 \\ 0 & 1 \end{bmatrix}$. Note that $AB = C$.

(e) No! $T \circ S \neq S \circ T$—for one thing the sources are different: $\mathbb{R}^3$ versus $\mathbb{R}^2$. The matrix of $T \circ S$ is BA which is 3×3.

(f) For any $\vec{x} \in \mathbb{R}^n$, we have

$$T_N \circ T_M(\vec{x}) = T_N(T_M(\vec{x})) = T_N(M\vec{x}) = N(M\vec{x}) = (NM)\vec{x} = T_{NM}(\vec{x}).$$

Here, every equality uses a definition or basic property of matrix multiplication (the first is definition of composition, the second is definition of T_M, the third is definition of T_N, the fourth is the association property of matrix multiplication, and the last is the definition of T_{NM}. We have shown on a previous worksheet that functions of the form T_A for any matrix A are always linear, so this completes the proof. QED.

Definitions: Let X and Y be any sets, and $X \xrightarrow{\phi} Y$ any mapping.

- ϕ is **surjective** if for all $y \in Y$, there exists $x \in X$ such that $\phi(x) = y$.
- ϕ is **injective** if for all $y \in Y$, there is at most one $x \in X$ with $\phi(x) = y$. Equivalently, ϕ is injective if, whenever $\phi(x) = \phi(y)$ for $x, y \in X$, it follows that $x = y$.
- ϕ is **bijective** (or invertible) if it is *both* injective and surjective—that is for $\forall y \in Y$, there is a *unique* $x \in X$ such that $\phi(x) = y$. Equivalently, ϕ is bijective if it has an **inverse map** $\psi : Y \to X$, meaning that $\phi \circ \psi$ is the identity on Y and $\psi \circ \phi$ is the identity on X.

Problem 2. Examples. For each linear transformation below, consider whether it is injective, surjective, and/or bijective. If it is bijective, describe the inverse map.

(a) The mapping $\mathbb{R}^3 \to \mathbb{R}^3$ which scales every vector by 2.

(b) The mapping $\mathbb{R}^3 \to \mathbb{R}^3$ which rotates every vector by $37°$ counterclockwise around the x-axis.

(c) The mapping $\mathbb{R}^2 \to \mathbb{R}^2$ defined by projection onto a line L.

(d) The shear $\mathbb{R}^2 \to \mathbb{R}^2$ defined by multiplication by the matrix $\begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix}$.

(e) The map $\mathbb{R}^2 \to \mathbb{R}^3$ sending $\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} x \\ y \\ x+y \end{bmatrix}$.

Solution: (a) This is surjective, injective, and invertible. The inverse scales by $\frac{1}{2}$.

(b) Invertible (hence surjective and injective). The inverse rotates clockwise by $37°$.

(c) Not surjective, since the image is the line L, which is strictly smaller than the target $\mathbb{R}^2$. Not injective, since all points on a given line perpendicular to L have the same image. Not invertible.

(d) Invertible. The inverse is $\begin{bmatrix} 1 & -5 \\ 0 & 1 \end{bmatrix}$.

(e) Injective but not surjective. We can never hit $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$, for example, so it is not surjective.

Definition: A matrix $A \in \mathbb{R}^{n \times n}$ is **invertible** if there exists $B \in \mathbb{R}^{n \times n}$ s.t. $AB = BA = I_n$. In this case, the matrix B is called the **inverse** of A and denoted by A^{-1}.
[There is only one such B. This is not completely obvious but you will prove this in Problem 3.]

Problem 3: Prove that an invertible* matrix has a *unique* inverse.
[HINT: Use the standard proof technique to show uniqueness: assume B and C both work, then show $B = C$].

Solution: Suppose $AB = BA = I_n$ and also $AC = CA = I_n$. Then $B = I_nB = (CA)B = C(AB) = CI_n = C$.

Problem 4. Suppose that $A = \begin{bmatrix} 1 & 0 & 2 \\ -2 & 1 & -4 \\ 0 & -1 & 1 \end{bmatrix}$. Find the inverse of A. [HINT: Remember the 2.4 reading!]

*Don't forget that only square matrices can be invertible!

Solution: Use the technique in the book of row-reducing A simultaneously with an adjacent 3×3 identity matrix. We get $A^{-1} = \begin{bmatrix} -3 & -2 & -2 \\ 2 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix}$.

Problem 5: Invertible Linear Maps and Matrices. Let $T_A : \mathbb{R}^n \to \mathbb{R}^n$ be the linear transformation with standard matrix A.

(a) Prove T_A is bijective if and only if for all $\vec{b} \in \mathbb{R}^n$, there exists a unique solution to $A\vec{x} = \vec{b}$.

(b) How can you tell if T_A is bijective from the *rank* of A? from the row reduced echelon form of A?

(c) Assuming A is invertible, show that $T_A : \mathbb{R}^n \to \mathbb{R}^n$ is invertible with inverse $T_{A^{-1}}$.
[Hint: Compose T_A and $T_{A^{-1}}$ in both orders.]

(d) Scaffold the proof of Theorem 2 below. What parts of the proof are complete? [Don't complete it now.]

Theorem 2: A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is invertible if and only if its standard matrix A is invertible; In this case, the inverse of T is the linear transformation with standard matrix A^{-1}.

Solution:

(a) Assume T_A is bijective. By definition, then, for all $\vec{b}$ in the target (which is $\mathbb{R}^n$), there is a unique $\vec{a}$ in the source (which is also $\mathbb{R}^n$) such that $T_A(\vec{a}) = \vec{b}$. In other words, $A\vec{a} = \vec{b}$, so $\vec{a}$ is the unique solution to $A\vec{x} = \vec{b}$. Conversely, suppose that for all $\vec{b} \in \mathbb{R}^n$, there exists a unique solution to $A\vec{x} = \vec{b}$. To see that T_A is surjective, take arbitrary $\vec{b}$ in the target (that is, in $\mathbb{R}^n$). Let $\vec{a}$ be the unique solution to $A\vec{x} = \vec{b}$. Note that $\vec{a}$ is the unique element in the source of T_A such that $T_A(\vec{a}) = \vec{b}$. Since $\vec{b}$ was arbitrary, we conclude that T_A is bijective.

(b) Thinking about a linear system with n variables and n equations, we know it has a solution for all $\vec{b}$ if and only if there are exactly n leading 1's in row reduced echelon form. This means $\text{rank}(A) = n$, or equivalently, $\text{rref}(A) = I_n$. So T_A is invertible if and only if its standard matrix A has rank n. This is the same as saying that $\text{rref}(A) = I_n$.

(c) The map $T_{A^{-1}} : \mathbb{R}^n \to \mathbb{R}^n$ is the inverse, because $T_{A^{-1}} \circ T_A = T_{A^{-1}A} = T_{I_n}$, which is the identity transformation on $\mathbb{R}^n$ (the first equality uses Theorem 1). And likewise, $T_A \circ T_{A^{-1}} = T_{AA^{-1}} = T_{I_n}$ is also the identity. This is the definition of invertible with inverse $T_{A^{-1}}$.

(d) Your scaffold should have two main parts:

  (a) If A is invertible, then T_A is invertible with inverse $T_{A^{-1}}$.

  (b) If T_A is invertible, then A is invertible.

We did the first part in (c). The second part is harder. For completeness, we give the full proof here (though it is Problem 8 b on the worksheet). In fact, we'll give two proofs here, one using Problem 8 (which is more rigorous) and one in the computational style of the book.

For the first proof: Assume T_A is invertible, and let S be its inverse mapping. *Because S is a linear transformation* (see Problem 8), we know that $S = T_B$ for some matrix B by the Key Theorem. By definition of invertible mapping, the compositions $T_A \circ T_B = T_B \circ T_A$ are both the identity map T_{I_n} on $\mathbb{R}^n$. By Theorem 1, then

$$T_{AB} = T_{BA} = I_n,$$

so by the uniqueness of the matrix in the Key theorem, we conclude $AB = BA = I_n$. This says that A is invertible.

For the more computational proof: Assume T_A is invertible. We know from (a) that the linear system $A\vec{x} = \vec{b}$ has a unique solution for all $\vec{b} \in \mathbb{R}^n$. From the theory in Chapter 1, and because the number of equations is equal to the number of variables, this means that the rank of A is n and the row reduced echelon form of A is the identity. Then, using the process in the book, we can compute the inverse of A by row-reducing $[A|I_n]$ (as in Problem 4 above).

Problem 6.

(a) If $f : X \to Y$ and $g : Y \to Z$ are invertible mappings (not necessarily linear), is their composition $g \circ f$ invertible? If so, what is its inverse?

(b) If A and B are invertible $n \times n$ matrices, is AB invertible? If so, what is its inverse?

(c) If $A_1, \ldots, A_k$ are invertible $n \times n$ matrices, what is $(A_1 \cdots A_k)^{-1}$? Prove it!

Solution:

(a) If $f : X \to Y$ and $g : Y \to Z$ are invertible, then so is $g \circ f$, and $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

(b) If A and B are invertible $n \times n$ matrices, then AB is invertible and $(AB)^{-1} = B^{-1}A^{-1}$.

(c) If $A_1, \ldots, A_k$ are invertible $n \times n$ matrices, then $A_1 \cdots A_k$ is invertible and

$$(A_1 \cdots A_k)^{-1} = A_k^{-1} \cdots A_1^{-1}.$$

We can prove this by induction on k. When $k = 2$, this is (b). Assume the statement for some fixed k. Then

$$(A_1 \cdots A_{k+1})^{-1} = ((A_1 \cdots A_k)A_{k+1})^{-1} = A_{k+1}^{-1}(A_1 \cdots A_k)^{-1} = A_{k+1}^{-1}A_k^{-1} \cdots A_1^{-1}.$$

The final equality is from our inductive assumption.

Problem 7. Prove or disprove:

(a) If A and B are both 2×2 matrices, then $AB = BA$. [Hint: Find an explicit counterexample.]

(b) If S and T are both linear transformations $\mathbb{R}^n \to \mathbb{R}^n$, then $S \circ T = T \circ S$.

Solution: (a) is False! Counterexample: Let $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, and $B = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. Then $AB = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \neq BA = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$.

Also (b) is false. For an explicit counterexample, we can take the maps T_A and T_B of $\mathbb{R}^2$ given by multiplication by A and B, respectively. We have $T_A \circ T_B = T_{AB}$ is multiplication by zero, so it sends every element of its source to the zero vector in $\mathbb{R}^2$. But $T_B \circ T_A = T_{BA}$ is multiplication by $BA = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$. This map sends $\vec{e}_2$ to $\vec{e}_1$, not zero. So $T_A \circ T_B \neq T_B \circ T_A$.

Problem 8.

(a) Let V and W be vector spaces, and suppose that $T : V \to W$ is a bijective linear transformation[†] from V to W. Prove that T^{-1} is linear.

(b) Use (a) to complete the proof of Theorem 2. [HINT: Assume $T_A : \mathbb{R}^n \to \mathbb{R}^n$ is invertible. Part (a) says the inverse function has the form T_B for some $B \in \mathbb{R}^{n\times n}$. Now compose T_A and T_B in both orders.]

Solution: Let $w, z \in W$, and write $x = T^{-1}(w)$ and $y = T^{-1}(z)$, so that $T(x) = w$ and $T(y) = z$. Since T is linear, we have $T(x + y) = T(x) + T(y) = w + z$, which, after applying T^{-1} to each side, gives us

$$T^{-1}(w + z) \;=\; T^{-1}(T(x + y)) \;=\; x + y \;=\; T^{-1}(w) + T^{-1}(z).$$

Similarly, given $w \in W$ and $c \in \mathbb{R}$, write $x = T^{-1}(w)$ so that $T(x) = w$, and observe that since $T(cx) = cT(x) = cw$ by linearity of T, we have

$$T^{-1}(cw) \;=\; T^{-1}(T(cx)) \;=\; cx \;=\; cT^{-1}(w).$$

The solution to part (b) appears in the solutions to Problem 5 (d).

Problem 9: A bogus proof. Suppose $A \in \mathbb{R}^{2\times 3}$ and $B \in \mathbb{R}^{3\times 2}$, so that both products AB and BA make sense. Consider the following "proof" of the statement: "*if* $AB = I_2$, *then* $BA = I_3$."

Proof. Suppose $AB = I_2$.

$\Rightarrow B(AB) = BI_2$ (multiplying both sides by B on the left)
$\Rightarrow B(AB) = B$ (by properties of I_2)
$\Rightarrow (BA)B = B$ (by the associative property of matrix mutliplication)
$\Rightarrow (BA)B = I_3B$ (by properties of I_3)
$\Rightarrow BA = I_3$ (canceling B from both sides). □

Is this proof correct? Is the statement even true? What if $A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ and $B = A^\top$? Where is the error in the "proof"? What can you conclude about "canceling matrices?"

Solution: The statement is not true, as can be seen using the counterexample

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}.$$

So the proof cannot be correct! The invalid step is the last one, where B is "canceled" from both sides to obtain $BA = I_3$ from $(BA)B = I_3B$. The problem is that you can only cancel a matrix from both sides of an equation *if* that matrix is invertible, which B is not; what you are really doing when you "cancel" a matrix is multiplying by its inverse and then simplifying.

[†] A bijective linear transformation between vector spaces is called an *isomorphism*.