MATH 217 - LINEAR ALGEBRA
HOMEWORK 2, SOLUTIONS

Part A (10 points)

Solve the following problems from the book:
Section 1.3: 34, 50, 62
Section 2.1: 8, 40, 44

Solution.

1.3.34: (a) $A\vec{e}_1 = \begin{bmatrix} a \\ d \\ g \end{bmatrix}$, $A\vec{e}_2 = \begin{bmatrix} b \\ e \\ h \end{bmatrix}$, $A\vec{e}_3 = \begin{bmatrix} c \\ f \\ k \end{bmatrix}$.

(b) $B\vec{e}_1 = \vec{v}_1$, $B\vec{e}_2 = \vec{v}_2$, $B\vec{e}_3 = \vec{v}_3$.

1.3.50: This system has *no solutions*, that is, it is inconsistent. We know this because $[A \mid \vec{b}]$ is a 4×4 matrix, so if its rank is 4 then it has a pivot position in every row (i.e., its reduced row echelon form has a leading 1 in every row). This means that the bottom row of its reduced row echelon form corresponds to the (inconsistent) linear equation $0x_1 + 0x_2 + 0x_3 = 1$. Therefore the system has no solutions.

1.3.62 Given $a, b \in \mathbb{R}$. The vector $\begin{bmatrix} 1 \\ c \\ c^2 \end{bmatrix}$ is a linear combination of the vectors $\begin{bmatrix} 1 \\ a \\ a^2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ b \\ b^2 \end{bmatrix}$ for some $c \in \mathbb{R}$ if and only if there exists $x_1, x_2 \in \mathbb{R}$ such that

$$\begin{bmatrix} 1 \\ c \\ c^2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ a \\ a^2 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ b \\ b^2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ a & c \\ a^2 & b^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Row-reducing the associated augemented matrix, we get

$$\left[\begin{array}{cc|c} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & b-a & c-a \\ 0 & b^2-a^2 & c^2-a^2 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 0 & b-a & c-a \\ 0 & 0 & (c-a)(c-b) \end{array}\right],$$

which will lead to a solution to the original linear system provided $c = a$ or $c = b$. In other words, $\begin{bmatrix} 1 \\ c \\ c^2 \end{bmatrix}$ is a linear combination of the vectors $\begin{bmatrix} 1 \\ a \\ a^2 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ b \\ b^2 \end{bmatrix}$ if and only if $c = a$ or $c = b$.

2.1.8: Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear transformation given by

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + 7x_2 \\ 3x_1 + 20x_2 \end{bmatrix} \quad \text{for all } \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2.$$

In order to find the inverse of T, we solve the linear system

$$\begin{aligned} x_1 + 7x_2 &= y_1 \\ 3x_1 + 20x_2 &= y_2 \end{aligned}$$

for x_1 and x_2 in terms of y_1 and y_2. The augemented matrix of this linear system is

$$\begin{bmatrix} 1 & 7 & y_1 \\ 3 & 20 & y_2 \end{bmatrix}, \quad \text{with rref} \quad \begin{bmatrix} 1 & 0 & -20y_1 + 7y_2 \\ 0 & 1 & 3y_1 - y_2 \end{bmatrix}.$$

Thus the inverse T^{-1} of T is given by

$$T^{-1}\left(\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}\right) = \begin{bmatrix} -20y_1 + 7y_2 \\ 3y_1 - y_2 \end{bmatrix}.$$

Alternative Solution: Note that $T(\vec{x}) = A\vec{x}$ for all $\vec{x} \in \mathbb{R}^2$ where $A = \begin{bmatrix} 1 & 7 \\ 3 & 20 \end{bmatrix}$. Thus $T^{-1}(\vec{y}) = A^{-1}\vec{y}$ for all $\vec{y} \in \mathbb{R}^2$, where A^{-1} is the *inverse matrix* of A. This solution method is not really valid yet since we will not learn about matrix inverses until Section 2.4, but once we learn about them this will be the most efficient way to solve this problem since there is a nice formula for the inverse of an invertible 2×2 matrix:

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

You can check that this gives the correct solution in this case!

2.1.40: Let $T : \mathbb{R} \to \mathbb{R}$ be a linear transformation. Then for all $x \in \mathbb{R}$ we have $T(x) = T(1x) = T(1)x$, so if we let $m = T(1)$ then $T(x) = mx$ for all $x \in \mathbb{R}$. Thus every liner transformation from $\mathbb{R}$ to $\mathbb{R}$ is just scaling by some constant, and has as its graph a (non-vertical) straight line through the origin.

2.1.44: Let $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$, $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ and

$$A = \begin{bmatrix} 0 & -v_3 & v_2 \\ v_3 & 0 & -v_1 \\ -v_2 & v_1 & 0 \end{bmatrix}.$$

Note that

$$T(\vec{x}) = \vec{v} \times \vec{x} = \begin{bmatrix} -v_3x_2 + v_2x_3 \\ v_3x_1 - v_1x_3 \\ -v_2x_1 + v_1x_2 \end{bmatrix} = A\vec{v}$$

so T is linear.

Part B (25 points)

Some of the problems below cover material that is in the "More Joy of Sets" handout, available on Canvas. Please read that handout carefully before attempting these problems. In particular, recall from that handout that a function $f : X \to Y$ is

- *injective* if for all $x_1, x_2 \in X$, if $f(x_1) = f(x_2)$ then $x_1 = x_2$;

- *surjective* if for all $y \in Y$ there is $x \in X$ such that $f(x) = y$;
- *bijective* if f is both injective and surjective; and
- *invertible* if there is a function $g : Y \to X$ such that $g \circ f = \mathrm{id}_X$ and $f \circ g = \mathrm{id}_Y$.

Problem 1. In parts (a) – (d) below, determine whether the given function is *injective*, *surjective*, *both*, or *neither*, and state whether the function is invertible. *No justification needed.*

(a) The function $p : \mathbb{R} \to [-9, \infty)$ defined by $p(x) = x^2 - 2x - 8$.
(b) The function $f : \mathbb{R} \to (0, \infty)$ defined by $f(x) = e^{-x^2}$.
(c) The function $q : \mathcal{P}(\mathbb{N}) \setminus \{\emptyset\} \to \mathbb{N}$ defined by $q(A) = \min A$.
(d) The function $g : \mathbb{R} \setminus \{3\} \to \mathbb{R} \setminus \{2\}$ defined by $g(x) = \frac{2x+7}{x-3}$.
(e) The function $F : \mathbb{R} \to \mathbb{R}$ defined by $F(x) = x(x^2 - 1)$.
(f) The function $h : \mathbb{R}^3 \to \mathbb{R}$ defined by $h(x, y, z) = e^{xyz}$.

Solution.

(a) The function p is surjective but not injective, so it is not bijective and hence not invertible.
(b) The function f is neither injective nor surjective, so it is not bijective and hence not invertible.
(c) The function q is surjective but not injective, so it is not bijective and hence not invertible.
(d) The function g is bijective, hence invertible. Its inverse is given by $g^{-1}(y) = \frac{3x+7}{13(x-2)}$.
(e) The function F is surjective but not injective, so it is not bijective and hence not invertible.
(f) The function h is neither injective nor surjective, so it is not bijective and hence not invertible.

Problem 2. Let X, Y, and Z be sets, and let $f : X \to Y$ and $g : Y \to Z$ be functions.

(a) Prove that if f and g are injective, then $g \circ f$ is injective.
(b) Suppose $g \circ f$ is injective. Does it necessarily follow that f is injective? How about g? Justify your claims with a proof or counterexample.
(c) Prove that if f and g are surjective, then $g \circ f$ is surjective.
(d) Suppose $g \circ f$ is surjective. Does it necessarily follow that f is surjective? How about g? Justify your claims with a proof or counterexample.

Solution.

(a) Suppose f and g are injective. Let $x, x' \in X$, and suppose $x \neq x'$. Then $f(x) \neq f(x')$ since f is injective, which implies $g(f(x)) \neq g(f(x'))$ since g is injective. We have shown that for all $x, x' \in X$, if $x \neq x'$ then $(g \circ f)(x) \neq (g \circ f)(x')$, so we conclude that $g \circ f$ is injective as desired.
(b) If $g \circ f$ is injective, then f must be injective but g need not be. To prove the first claim, we will show the contrapositive; that is, we prove that if f is not injective then neither is $g \circ f$. Assume f is not injective, and fix $x_1, x_2 \in X$ such that $x_1 \neq x_2$ but $f(x_1) = f(x_2)$. Then $g(f(x_1)) = g(f(x_2))$, which shows $g \circ f$ is not injective, as claimed. To prove the second claim, we give a counterexample. Let $X = Y = \mathbb{R}$, let f be the exponential function defined by $f(x) = e^x$, and let g be the squaring function defined by $g(x) = x^2$. Then g is not injective, since for instance $g(1) = 1 = g(-1)$. However, to see that $g \circ f$ is injective, note that for all real numbers $x < y$ we have $0 < e^x < e^y$, which implies $(e^x)^2 < (e^y)^2$. Thus $g \circ f$ is strictly increasing and is therefore injective even though g is not.

(c) Suppose f and g are surjective, and let $z \in Z$. Using surjectivity of g, fix $y \in Y$ such that $g(y) = z$, and then using surjectivity of f, fix $x \in X$ such that $f(x) = y$. Then $(g \circ f)(x) = g(f(x)) = g(y) = z$. We have shown that for all $z \in Z$ there is $x \in X$ such that $(g \circ f)(x) = z$, so we conclude that $g \circ f$ is surjective as desired.

(d) If $g \circ f$ is surjective, then g must be surjective but f need not be. To prove the first claim, suppose $g \circ f$ is surjective and let $z \in Z$ be arbitrary. Fix $x \in X$ such that $g(f(x)) = z$. Let $y = f(x)$, so $g(y) = z$. This shows that g is surjective. To prove the second claim, we give a countexample: let $X = Y = \mathbb{R}$, let $Z = [0, \infty)$, and let f and g both be the squaring function. Then f is not surjective, but $g \circ f$ is the function $x \mapsto x^4$ from $\mathbb{R}$ to $[0, \infty)$ so $g \circ f$ *is* surjective.

Problem 3. Let X and Y be sets, and let $f : X \to Y$ be a function. Recall that for every $A \subseteq X$, we define the *forward image*, or *direct image* of A under f to be the set $f[A] = \{f(x) : x \in A\}$.

(a) Prove that $f[A \cap B] \subseteq f[A] \cap f[B]$ for all $A, B \subseteq X$.
(b) Prove that $f[A] \setminus f[B] \subseteq f[A \setminus B]$ for all $A, B \subseteq X$.
(c) Show by example[1] that we might have $f[A] \cap f[B] \not\subseteq f[A \cap B]$ for some $A, B \subseteq X$.
(d) Show by example that we might have $f[A \setminus B] \not\subseteq f[A] \setminus f[B]$ for some $A, B \subseteq X$.
(e) Prove that if f is injective, then $f[A \cap B] \supseteq f[A] \cap f[B]$ for all $A, B \subseteq X$.
(f) Prove that if f is injective, then $f[A] \setminus f[B] \supseteq f[A \setminus B]$ for all $A, B \subseteq X$.

Solution.

(a) Let X and Y be sets, let $f : X \to Y$ be a function, and let $A, B \subseteq X$. Let $x \in A \cap B$. Then $x \in A$, so $f(x) \in f[A]$, and $x \in B$, so $f(x) \in f[B]$, which means $f(x) \in f[A] \cap f[B]$. Since $x \in A \cap B$ was arbitrary, this shows $f[A \cap B] \subseteq f[A] \cap f[B]$.

(b) Let X and Y be sets, let $f : X \to Y$ be a function, and let $A, B \subseteq X$. Let $y \in f[A] \setminus f[B]$. Then $y \in f[A]$, so we may fix $x \in A$ such that $y = f(x)$, and since $y \notin f[B]$ we know $x \notin B$. Thus $x \in A \setminus B$, so $y = f(x) \in f[A \setminus B]$. Since $y \in f[A] \setminus f[B]$ was arbitrary, this shows $f[A] \setminus f[B] \subseteq f[A \setminus B]$.

(c) For instance, let $X = Y = \mathbb{R}$, let f be the squaring function given by $f(x) = x^2$, and let $A = (0, 1)$ and $B = (-1, 0)$. Then $f[A] = (0, 1) = f[B]$ but $A \cap B = \emptyset$, so $f[A] \cap f[B] = (0, 1) \not\subseteq \emptyset = f[A \cap B]$.

(d) For instance, again let $X = Y = \mathbb{R}$ and let f be the squaring function given by $f(x) = x^2$, and this time let $A = (-1, 1)$ and $B = (-1, 0]$, so $A \setminus B = (0, 1)$ and $f[A] = f[B] = [0, 1)$. Then $f[A \setminus B] = f[(0, 1)] = (0, 1) \not\subseteq \emptyset = f[A] \setminus f[B]$.

(e) Let X and Y be sets, let $f : X \to Y$ be a function, let $A, B \subseteq X$, and suppose now that f is injective. Let $y \in f[A] \cap f[B]$. Then $y \in f[A]$, so we may fix $x \in A$ such that $y = f(x)$, and also $y \in f[B]$, so we may fix $x' \in B$ such that $y = f(x')$. Since f is injective, $x = x'$, so $x \in A \cap B$ and thus $f(x) \in f[A \cap B]$. Since $y \in f[A] \cap f[B]$ was arbitrary, we have shown $f[A] \cap f[B] \subseteq f[A \cap B]$, as desired.

(f) Again let X and Y be sets, let $f : X \to Y$ be a function, let $A, B \subseteq X$, and suppose f is injective. Let $x \in A \setminus B$. Then $x \in A$, so $f(x) \in f[A]$. Since $x \notin B$ and f is injective, there is no $x' \in B$ such that $f(x) = f(x')$, so $f(x) \notin f[B]$. Thus $f(x) \in f[A] \setminus f[B]$. This completes the proof that $f[A \setminus B] \subseteq f[A] \setminus f[B]$.

[1] Remember that you also have to specify X, Y, and f in your example! The same goes for parts (e) and (f).

We call a function $T : \mathbb{R}^m \to \mathbb{R}^n$ a **linear transformation** if it satisfies:

(i) $T(\vec{x} + \vec{y}) = T(\vec{x}) + T(\vec{y})$ for all vectors $\vec{x}, \vec{y} \in \mathbb{R}^m$; and

(ii) $T(k\vec{x}) = kT(\vec{x})$ for all vectors $\vec{x} \in \mathbb{R}^m$ and all scalars $k \in \mathbb{R}$.

We say that a function $T : \mathbb{R}^m \to \mathbb{R}^n$ *preserves vector addition* if it satisfies (i) above, and that it *preserves scalar multiplication* if it satisfies (ii) above. So a function $T : \mathbb{R}^m \to \mathbb{R}^n$ is a *linear transformation* if it preserves both vector addition and scalar multiplication.[2]

Problem 4. Let m and n be positive integers, and let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a function. Prove that f is a linear transformation if and only if

$$f(a\vec{x} + b\vec{y}) \;=\; af(\vec{x}) + bf(\vec{y})$$

for all $a, b \in \mathbb{R}$ and $\vec{x}, \vec{y} \in \mathbb{R}^n$.

Solution. First suppose that f is a linear transformation, so f preserves vector addition and scalar multiplication. Then for all $a, b \in \mathbb{R}$ and $\vec{x}, \vec{y} \in \mathbb{R}^n$ we have

$$f(a\vec{x} + b\vec{y}) \;=\; f(a\vec{x}) + f(b\vec{y}) \;=\; af(\vec{x}) + bf(\vec{y}).$$

Conversely, suppose $f(a\vec{x} + b\vec{y}) = af(\vec{x}) + bf(\vec{y})$ for all $a, b \in \mathbb{R}$ and $\vec{x}, \vec{y} \in \mathbb{R}^n$. Let $\vec{x}, \vec{y} \in \mathbb{R}^n$ and $c \in \mathbb{R}$ be arbitrary. Then by assumption we have

$$f(\vec{x} + \vec{y}) \;=\; f(1\vec{x} + 1\vec{y}) \;=\; 1f(\vec{x}) + 1f(\vec{y}) \;=\; f(\vec{x}) + f(\vec{y})$$

and

$$f(c\vec{x}) \;=\; f(c\vec{x} + 0\vec{y}) \;=\; cf(\vec{x}) + 0f(\vec{y}) \;=\; cf(\vec{x}),$$

which shows that f is linear as desired.

Problem 5. Let $\vec{a} \in \mathbb{R}^m$ and $\vec{b} \in \mathbb{R}^n$, and suppose that T is a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$ such that $T(\vec{a}) = \vec{b}$. Prove that

$$\{\vec{x} \in \mathbb{R}^m : T(\vec{x}) = \vec{0}\} \;=\; \{\vec{x} \in \mathbb{R}^m : T(\vec{a} + \vec{x}) = \vec{b}\}.$$

Solution. Assume the given hypotheses. Let $\vec{x} \in \mathbb{R}^m$ and suppose $T(\vec{x}) = \vec{0}$. Then

$$T(\vec{a} + \vec{x}) \;=\; T(\vec{a}) + T(\vec{x}) \;=\; T(\vec{a}) + \vec{0} \;=\; T(\vec{a}) \;=\; \vec{b}.$$

This shows $\{\vec{x} \in \mathbb{R}^m : T(\vec{x}) = \vec{0}\} \subseteq \{\vec{x} \in \mathbb{R}^m : T(\vec{a} + \vec{x}) = \vec{b}\}$. Conversely, let $\vec{x} \in \mathbb{R}^m$ and suppose $T(\vec{a} + \vec{x}) = \vec{b}$. Then

$$T(\vec{x}) \;=\; T(\vec{x} + \vec{0}) \;=\; T(\vec{x} + (\vec{a} - \vec{a})) \;=\; T((\vec{x} + \vec{a}) - \vec{a}) \;=\; T(\vec{x} + \vec{a}) - T(\vec{a}) \;=\; \vec{b} - \vec{b} \;=\; \vec{0}.$$

This shows $\{\vec{x} \in \mathbb{R}^m : T(\vec{x}) = \vec{0}\} \supseteq \{\vec{x} \in \mathbb{R}^m : T(\vec{a} + \vec{x}) = \vec{b}\}$, and we conclude that the two sets are equal as desired.

[2]Note that this definition differs from the one given in Section 2.1 of the textbook, although for functions from $\mathbb{R}^m$ to $\mathbb{R}^n$ the two definitions are equivalent. *In this class we will always prefer to use the definition given above, including for Problem 4.*

Solution. Assume the hypotheses, and let $\vec{x} \in \mathbb{R}^m$. Since T is linear and $T(\vec{a}) = \vec{b}$, we have

$$T(\vec{a} + \vec{x}) \;=\; T(\vec{a}) + T(\vec{x}) \;=\; \vec{0} + T(\vec{x}) \;=\; T(\vec{x}).$$

Thus for all $\vec{x} \in \mathbb{R}^m$ we have $T(\vec{x}) = \vec{0}$ if and only if $T(\vec{a} + \vec{x}) = \vec{b}$, so

$$\{\vec{x} \in \mathbb{R}^m : T(\vec{x}) = \vec{0}\} \;=\; \{\vec{x} \in \mathbb{R}^m : T(\vec{a} + \vec{x}) = \vec{b}\}.$$