Math 217 Worksheet 8: Kernel, Span and Image (§3.1)

Definitions: Given a linear transformation $V \xrightarrow{T} W$ between vector spaces, we have

- The **source** of T is V, and the **target** of T is W.
- The **image** of T is the subset of the target $\{w \in W \mid w = T(v) \text{ for some } v \in V\}$.
- The **kernel** of T is the subset of the source $\{v \in V \mid T(v) = 0_W\}$. Put differently, the kernel of T is the *pre-image* of 0_W.

We write $\operatorname{im} T$ and $\ker T$, respectively, for the image and kernel of T.

ADVICE: In encountering new definitions, please keep in mind concrete examples you already know—in this case, think about V as $\mathbb{R}^n$ and W as $\mathbb{R}^m$ the first time through. Model your future understanding on this case, but be aware that **there are other important examples** and there are important differences. Beware: it makes no sense to say **a linear map is "a matrix"** unless *source and target* are both *coordinate spaces* of column vectors.

Problem 1. Examples. For each linear transformation below, determine the source, target, image and kernel. Say whether or not T is injective, surjective and/or bijective.

(a) $T : \mathbb{R}^3 \to \mathbb{R}^2$ such that $T(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}) = \begin{bmatrix} x_1 + x_2 + x_3 \\ x_1 + x_2 + x_3 \end{bmatrix}$. Describe $\ker T$ and $\operatorname{im} T$ geometrically, and in set notation.

(b) $T : \mathbb{R} \to \mathbb{R}^3$ defined by $T(x) = \begin{bmatrix} x \\ x \\ x \end{bmatrix}$. For the image: describe $\operatorname{im} T$ geometrically, and in set notation.

(c) $\mathbb{R}^2 \xrightarrow{T} \mathbb{R}^2$ given by left multiplication by $\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$. [HINT: Think geometrically!]

(d) The transformation $d : \mathcal{P} \to \mathcal{P}$ sending f to its derivative df/dx.

(e) The trace map $\mathbb{R}^{2\times 2} \to \mathbb{R}$ sending $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ to $a + d$. Describe $\ker T$ in set notation.

Solution:

(a) $\ker T = \{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \mid x_1 + x_2 + x_3 = 0\}$. That's a plane through the origin with normal vector $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Not injective, since this whole plane collapses to the origin. The image $\operatorname{im} T \subseteq \mathbb{R}^2$ is the line $y = x$ going through the origin and through $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. Not surjective, since this line is not all of $\mathbb{R}^2$.

(b) Source is $\mathbb{R}$, target is $\mathbb{R}^3$, image $\operatorname{im} T = \{ \begin{bmatrix} x \\ x \\ x \end{bmatrix} \in \mathbb{R}^3 | x \in \mathbb{R}\}$ is a line through the origin and $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$. Also $\ker T = \{0\}$. Injective, not surjective.

(c) This is the map of rotation counterclockwise by θ. It is clearly both surjective and injective, since it has an inverse (rotation clockwise by θ). The image is thus all of $\mathbb{R}^2$ and the kernel is $\{\vec{0}\}$.

(d) Source and target are both the polynomials $\mathcal{P}$. It is surjective: The image is all of $\mathcal{P}$ since every polynomial has an anti-deriviative (which is also a polynomial). It is not injective, since x^2 and $x^2 + 1$ have the same image. The kernel is the set of constant functions.

(e) The image is all of $\mathbb{R}$. The kernel is the set $\{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \mid a + d = 0\}$. This is surjective but not injective.

Theorem A: A linear transformation $V \to W$ is injective if and only if its kernel is $\{0_V\}$.

Problem 2. Confirm that Theorem A holds in each of the 5 examples you studied in Problem 1.

Problem 3. Let $T_A : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear transformation defined by $\vec{x} \mapsto A\vec{x}$ where $A = \begin{bmatrix} \pi & 3 \\ \sqrt{17} & \frac{1}{2} \\ 0 & 1 \end{bmatrix}$.

Show $\left\{ \begin{bmatrix} \pi \\ \sqrt{17} \\ 0 \end{bmatrix}, \begin{bmatrix} 3 \\ \frac{1}{2} \\ 1 \end{bmatrix} \right\} \subseteq \operatorname{im} T_A$. [Technique: to show $w \in \operatorname{im} T$, find a vector v in the source such that $T(v) = w$.]

Solution: $\begin{bmatrix} \pi \\ \sqrt{17} \\ 0 \end{bmatrix} = A\vec{e}_1 = T_A(\vec{e}_1)$ so $\begin{bmatrix} \pi \\ \sqrt{17} \\ 0 \end{bmatrix} \in \operatorname{im} T_A$. Likewise, $\begin{bmatrix} 3 \\ \frac{1}{2} \\ 1 \end{bmatrix} = A\vec{e}_2 = T_A(\vec{e}_2)$. By the "Unusually Useful Lemma," we see in general, the columns of A are in the image of T_A since $T(\vec{e}_i) = A\vec{e}_i$ is the i-th column of A.

Problem 4. True or False: A map is surjective if and only if its image is equal to its target. Explain!

Solution: TRUE! Surjective means *every* $\vec{y}$ in the target is hit by the map—for every $\vec{y}$ in the target, there exists an $\vec{x}$ in the source such that $T(\vec{x}) = \vec{y}$. This simply says that every $\vec{y}$ in the target is in the image. So Im T = Target.

Definition: The **span** of a set $\{v_1, v_2, \ldots, v_n\}$ of vectors (in some vector space V) is the set of *all* linear combinations of the vectors $v_1, v_2, \ldots, v_n$.
That is, $\text{Span}\{v_1, v_2, \ldots, v_n\} = \{c_1v_1 + c_2v_2 + \cdots + c_nv_n \mid c_i \in \mathbb{R}\}$.
We also say the vectors $\{v_1, v_2, \ldots, v_n\}$ **span W** to mean that $W = \text{Span}\{v_1, v_2, \ldots, v_n\}$.

Problem 5. Span

(a) Describe the **span** of the set $\{\vec{e}_1\}$ in $\mathbb{R}^3$ in simple geometric language. Do the same for $\{\vec{e}_1, \vec{e}_2\}$ in $\mathbb{R}^3$. What about the span of $\{\vec{0}\}$ in $\mathbb{R}^3$? $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$?

(b) Describe the span of the set $\{\begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 0 \\ 4 \end{bmatrix}, \begin{bmatrix} 5 \\ 0 \\ 10 \end{bmatrix}\}$ in a simple way as a subset of $\mathbb{R}^3$. Write it in set-builder notation. Find the simplest/nicest **two** vectors that also span this subset.

(c) What is the *smallest* number of vectors that can span the subset of $\mathbb{R}^{10}$ spanned by $\vec{e}_1 + \vec{e}_2$, $\vec{e}_2 + \vec{e}_3$, $\vec{e}_3 + \vec{e}_4$, $\vec{e}_4 + \vec{e}_5$, $\vec{e}_5 + \vec{e}_1$, $\vec{e}_5 - \vec{e}_4$, $\vec{e}_4 - \vec{e}_3$, $\vec{e}_3 - \vec{e}_2$, $\vec{e}_2 - \vec{e}_1$?

(d) Describe the span of the polynomials $x, x^2, x^3 + x$ in the vector space $\mathcal{P}$ of all polynomials.

(e) Describe the span of the matrices $E_{11} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$ and $E_{22} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ in the vector space $\mathbb{R}^{2\times 2}$.

Solution:

(a) The span of $\{\vec{e}_1\}$ is the x-axis $\{\begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix} \mid x \in \mathbb{R}\}$. The span of $\{\vec{e}_1, \vec{e}_2\}$ is the xy-plane in $\mathbb{R}^3$. The span of $\{\vec{0}\}$ is the origin in $\mathbb{R}^3$. And Span $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\} = \mathbb{R}^3$.

(b) This span is the xz plane in $\mathbb{R}^3$. It is spanned by $\vec{e}_1$ and $\vec{e}_3$. In set-builder notation it is $\{\begin{bmatrix} x \\ 0 \\ z \end{bmatrix} \mid x, z \in \mathbb{R}\}$. Another correct answer is $\{c_1 \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} + c_2 \begin{bmatrix} 2 \\ 0 \\ 2 \end{bmatrix} \mid c_1, c_2 \in \mathbb{R}\}$.

(c) The smallest number of vectors that spans the given space is five. The set $\{\vec{e}_1, \vec{e}_2, \vec{e}_3, \vec{e}_4, \vec{e}_5\}$ is one such minimal set of spanning elements but not the only one.

(d) $\text{Span}\{x, x^2, x^3 + x\} = \{ax + bx^2 + cx^3 \mid a, b, c \in \mathbb{R}\}$, the set of polynomials of degree three or less with constant term zero.

(e) $\text{Span}\{E_{11}, E_{22}\}$ is the set of 2×2 diagonal matrices.

Theorem B: Let $\mathbb{R}^n \xrightarrow{T_A} \mathbb{R}^d$ be the linear transformation given by left multiplication by A. Then the image of T_A is equal to the span of the columns of A.

Problem 6. For parts (a) and (b) in Problem 1, confirm that Theorem B holds by examining the standard matrix of each transformation. Explain why Theorem B does not apply to parts (d) and (e).

Solution: For (a), the standard matrix is $\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$, which spans the line $y = x$ in $\mathbb{R}^2$. For (b), the standard matrix is $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, which spans the image line. For (d) and (e), the source and target are not coordinates spaces, so there is no standard matrix.

Problem 7: Proof of Theorem A. Let $T : V \to W$ be a linear transformation.

(a) Explain why $0_V \in \ker T$.

(b) Prove that if T is injective, then the kernel of T is $\{0_V\}$. [First Line: If not, suppose $v \in \ker T$ is non-zero...]

(c) State and prove the converse to (b). [Hint: Use the standard technique to show injectivity— assume two elements v_1, v_2 of the source have the same image under T, then "do some math" to show $v_1 = v_2$.]

(d) Explain how (b) and (c) together prove the following important theorem:

Solution:

(a) We proved earlier that $T(0_V) = 0_W$ for any linear transformation. So $0_V \in \ker T$.

(b) Assume T injective. Assume on the contrary, that $\ker(T)$ is not zero. Take any non-zero $\vec{v}$ in the kernel. By definition $T(\vec{v}) = 0$. But also $T(\vec{0}) = \vec{0}$. So by the injectivity assumption, $\vec{v} = \vec{0}$. Thus $\ker(T) = \{0\}$.

(c) Assume that $\ker(T) = \{0\}$. Let $v, w \in V$. Suppose that $T(v) = T(w)$. Then by linearity of T

$$0 = T(v) - T(w) = T(v - w) = 0,$$

so $v - w$ is in the kernel of T. But then $v - w = 0$, which means $v = w$. So T is injective. Thus T injective.

(d) This is an if and only if statement, so we need to prove two implications. (b) is one implication and (c) is the other.

Problem 8: Proof of Theorem B. Fix $A \in \mathbb{R}^{d\times n}$. Let $\mathbb{R}^n \xrightarrow{T_A} \mathbb{R}^d$ be the linear transformation defined by $T_A(\vec{x}) = A\vec{x}$.

(a) Let $\vec{C}_1, \ldots \vec{C}_n$ be the columns of A. Show that each $\vec{C}_i$ is in the image of T_A.
[Hint: You did this in Problem 3. Remember the unusually useful lemma: $A\vec{e}_j =??$].

(b) Prove that $\mathrm{Span}\{\vec{C}_1, \ldots, \vec{C}_n\} \subseteq \mathrm{im}\ T_A$. [First Line: "Take arbitrary $a_1\vec{C}_1 + \cdots + a_n\vec{C}_n \in \mathrm{Span}\{\vec{C}_1, \ldots, \vec{C}_n\}$...]

(c) Prove that $\mathrm{im}\ T_A \subseteq \mathrm{Span}\ \{\vec{C}_1, \ldots, \vec{C}_n\}$.
[First line: "Take arbitrary $\vec{y} = T_A(\vec{x}) \in \mathrm{im}\, T_A$. Write $A = \begin{bmatrix} \vec{C}_1 & \ldots & \vec{C}_n \end{bmatrix}$... "]

(d) Put together the previous steps to prove Theorem B.

Solution:

(a) We know $T_A(\vec{e}_j) = A\vec{e}_j = \vec{C}_j$ by the "Unusually Useful Lemma," so each column is in the image of T_A.

(b) Take arbitrary $a_1\vec{C}_1 + \cdots + a_n\vec{C}_n \in \mathrm{Span}\{\vec{C}_1, \ldots, \vec{C}_n\}$. From (a), we have $a_1\vec{C}_1 + \cdots + a_n\vec{C}_n = a_1T_A(\vec{e}_1) + \cdots + a_nT_A(\vec{e}_n)$, which equals $T_A(a_1\vec{e}_1) + \cdots + T_A(a_n\vec{e}_n) = T_A(a_1\vec{e}_1 + \cdots + a_n\vec{e}_n)$ by linearity. This shows that $a_1\vec{C}_1 + \cdots + a_n\vec{C}_n$ is in the image of T_A.

(c) Take arbitrary $\vec{y} \in \mathrm{im}\, T_A$. Then there exists $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ such that $\vec{y} = T_A(\vec{x}) = A\vec{x} = \begin{bmatrix} \vec{C}_1 & \ldots & \vec{C}_n \end{bmatrix} \vec{x} = x_1\vec{C}_1 + \ldots x_n\vec{C}_n \in \mathrm{Span}\ \{\vec{C}_1, \ldots, \vec{C}_n\}$.

(d) In (b), we showed that $\mathrm{Span}\{\vec{C}_1, \ldots, \vec{C}_n\} \subset \mathrm{im}T_A$. In (c), we showed that $\mathrm{im}T_A \subset \mathrm{Span}\ \{\vec{C}_1, \ldots, \vec{C}_n\}$. So $\mathrm{im}T_A = \mathrm{Span}\ \{C_1, \ldots, C_n\}$.

Problem 9. The **span** of an infinite subset $\mathcal{S}$ of a vector space V can also be defined. Try to write down a definition in proper set-builder notation. Caution! There is no such thing as an infinite sum of vectors! Find a spanning set for the vector space $\mathcal{P}$ of polynomials. Does $\mathcal{P}$ have a finite spanning set?

Solution: The **span** of a set $\mathcal{S}$ is the set of all (finite!) linear combinations of elements of $\mathcal{S}$, or $\text{Span}\mathcal{S} = \{\sum_{i=1}^{n} c_i x_i \mid n \in \mathbb{N},\ x_i \in \mathcal{S},\ c_i \in \mathbb{R},\ \text{for all } i = 1, \ldots, n\}$. The infinite set $\{1, x, x^2, x^3, \ldots\}$ spans $\mathcal{P}$ since every polynomial is a linear combination of these. No finite set spans $\mathcal{P}$: given any finite set $\{g_1, \ldots, g_t\}$ of polynomials, there is a maximal degree d among them. No polynomial of degree greater than d can be written as a linear combination of the g_i's since they all have degree less than or equal to d.