## Math 217 Worksheet 11: More About Bases (§3.3, §4.1)

IMPORTANT DEFINITIONS: Let $\mathcal{S}$ be a set of vectors in a vector space $V$.
The set $\mathcal{S}$ **spans** $V$ if every $v \in V$ is a linear combination of the elements of $\mathcal{S}$.
The set $\mathcal{S}$ is **linearly independent** if every relation on the elements of $\mathcal{S}$ is trivial.
The set $\mathcal{S}$ is a **basis** for $V$ if it is *both* linearly independent *and* spans $V$.
The **dimension** of $V$ is the number of elements in any (equivalently, every) basis.

**Theorem A:** Let $V$ be a vector space of finite dimension $n$, and let $\mathcal{S} = \{v_1, v_2, \dots, v_n\}$ be a subset of *exactly* $n$ *elements*. Then $\mathcal{S}$ is linearly independent if and only if $\mathcal{S}$ spans $V$.

**Problem 1.** For each vector space below, find two different bases, and determine the dimension. If you *already know* the dimension, how can you use Theorem A above to speed up your work?

(a) The **plane** defined by $x + y + z = 0$.

(b) The kernel of the matrix* $A = \begin{bmatrix} 1 & -21 & 0 \\ 0 & 0 & 1 \end{bmatrix}$. [HINT: How might rank-nullity speed up your work?]

(c) The image of the matrix $\begin{bmatrix} 1 & 2 & 3 & 0 & 1 \\ 2 & 4 & 6 & 0 & 2 \\ 3 & 6 & 9 & 1 & 4 \end{bmatrix}$.

(d) $\mathbb{C}$.

(e) The set of $3 \times 3$ *strictly* upper triangular matrices.

(f) The subset of $\mathcal{P}_4$ defined by $\{f(x) \mid f(0) = 0\}$. [How is this even a vector space?]

**Solution:** All the sets are vector spaces, which is easy to check. In each case, they are subsets of a bigger set we have already checked to be a vector space, so it's enough to check that they are subspaces.

(a) Basis $\left\{ \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ -1 \end{bmatrix} \right\}$. This is only one of many possible answers. Another is $\left\{ \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ -20 \\ 20 \end{bmatrix} \right\}$.

(b) The kernel of any linear transformation is always a subspace, hence it's a vector space! This is the solution set is $\left\{ \begin{bmatrix} 21y \\ y \\ 0 \end{bmatrix} \mid y \in \mathbb{R} \right\}$, so a basis is $\left\{ \begin{bmatrix} 21 \\ 1 \\ 0 \end{bmatrix} \right\}$. Another is $\left\{ \begin{bmatrix} -42 \\ -2 \\ 0 \end{bmatrix} \right\}$.

(c) We know the columns span the image, so we just need to remove redundant ones. Note that the second and third columns are multiples of the first, so we don't need them. Also, the fifth is the first plus the fourth. So a basis is $\left\{ \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$. Another is $\left\{ \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

*A common abuse of terminology is to write ker $A$ for "The kernel of the linear transformation $T_A$." And similarly, im $A$ means the image of $T_A$.

(d) $\{1, i\}$.

(e) A basis is $\left\{ \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \right\}$. Another is the same with $-1$ instead of one's everywhere.

(f) $\{x, x^2, x^3, x^4\}$. Or $\{x, x^2 + x, x^3 + x, x^4 + x\}$.

**Problem 2**. Let $\{v_1, \ldots, v_d\}$ be a basis for a vector space $V$.

(a) Prove that every element of $V$ can be written *uniquely* as a linear combination of $\{v_1, \ldots, v_d\}$.
[HINT: First sentence "Take an arbitrary...". Next show it *can* be written.... Then show uniqueness.]

(b) In the vector space $\mathcal{P}_3$, write $(x+1)^2$ uniquely as a linear combination of the elements in the basis $\{1, x, x^2, x^3\}$.

**Solution:** Take arbitrary $v \in V$. Since $\{v_1, \ldots, v_d\}$ is a basis, we know it spans $V$, so we can write $v = c_1 v_1 + \cdots + c_d v_d$. This is unique, because if also $v = c_1' v_1 + \cdots + c_d' v_d$, then subtracting we have a relation $0 = (c_1 - c_1')v_1 + \cdots + (c_d - c_d')v_d$. Because $\{v_1, \ldots, v_d\}$ is linearly independent, we know the relation is trivial, so each coefficient $c_i - c_i' = 0$. This means $c_i = c_i'$ for all $i$.

For (b), we have $(x+1)^2 = x^2 + 2x + 1$ so it is $1\, v_1 + 2\, v_2 + 1\, v_3 \, + 0\, v_4$.

**Problem 3: Bases of $\mathbb{R}^n$.** Let $\vec{v}_1, \ldots, \vec{v}_m$ be vectors in $\mathbb{R}^n$.

(a) Prove that if $m > n$, then $\{\vec{v}_1, \ldots, \vec{v}_m\}$ is not linearly independent.
[Hint for (a) and (b): write the vectors as columns in an $n \times m$ matrix $A$, and consider $\mathrm{rref}(A)$.]

(b) Prove that if $m < n$, then $\{\vec{v}_1, \ldots, \vec{v}_m\}$ does not span $\mathbb{R}^n$.

(c) Use (a) and (b) to prove that every basis for $\mathbb{R}^n$ contains exactly $n$ elements.

(d) Discuss Theorem B below in the special case where $V$ is a *coordinate space.* In the case $V = \mathbb{R}^n$, show that (i) implies both (ii), (iii), and (iv).
[HINT: You've done most of the work already!]

**Solution:**

(a) Let $A = [\vec{v}_1 \ \cdots \ \vec{v}_m] \in \mathbb{R}^{n \times m}$. If $m > n$, then at least one column of $\mathrm{rref}(A)$ does not have a leading 1. By our knowledge for how to solve systems of linear equations from Chapter 1, this means that there is a free variable and the system of linear equations $A\vec{x} = \vec{0}$ has a non-zero solution, say $\vec{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_m \end{bmatrix} \in \mathbb{R}^m$. As we saw on WS 10, problem 4a, this means that $c_1 \vec{v}_1 + \cdots + c_m \vec{v}_m = \vec{0}$ is a non-trivial relation on the columns of $A$. So $\{\vec{v}_1, \ldots, \vec{v}_m\}$ is not linearly independent.

(b) Again let $A = [\vec{v}_1 \ \cdots \ \vec{v}_m] \in \mathbb{R}^{n \times m}$. If $m < n$, then at least one row of $\mathrm{rref}(A)$ does *not have* a leading 1 in it. In other words, the rank of $A$ is strictly less than $n$. From Worksheet 10, we know the rank of $A$ is the dimension of the image of the linear transformation $T_A$ given by left multiplication by $A$. So the dimension of the image of $T_A$ is strictly less than $n$, which means that $\mathrm{im} T_A$ is *properly* contained in $\mathbb{R}^n$. In other words, we can find a vector $\vec{b} \in \mathbb{R}^n$ that does not belong to $\mathrm{im}(T_A) = \mathrm{Span}(\vec{v}_1, \ldots, \vec{v}_m)$, so $\{\vec{v}_1, \ldots, \vec{v}_m\}$ does not span $\mathbb{R}^n$.

(c) Suppose $\{\vec{v}_1, \ldots, \vec{v}_m\}$ is a basis for $\mathbb{R}^n$. If $m < n$, then by (b), the set $\{\vec{v}_1, \ldots, \vec{v}_m\}$ does not span $\mathbb{R}^n$, contrary to its being a basis. So $m \geq n$. But if $m > n$, then by (a), the set $\{\vec{v}_1, \ldots, \vec{v}_m\}$ can not be linearly independent. So $m \leq n$. We conclude that $m = n$.

(d) This is the case of Theorem B where $V$ is a *coordinate space* $\mathbb{R}^n$. Assume $V = \mathbb{R}^n$. The equivalence of (i) and (iv) is already proved by Problem 2, so we focus on the other three statements. To see (i) implies (ii), suppose $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a basis (by (c), it contains exactly $n$ elements). If this set is *not a maximal linearly independent set*, then we can add some $v_{n+1}$ to it, to get a larger linearly independent set $\{\vec{v}_1, \ldots, \vec{v}_n, \vec{v}_{n+1}\}$. This contradicts (a), since then we'd have $n + 1$ linearly independent vectors in $\mathbb{R}^n$.

To see (i) implies (iii), again suppose $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a basis. If $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is not a minimal spanning set, we can remove some vector from it, say $\vec{v}_n$ (after reordering the $\vec{v}_i$) and still have a spanning set. But then $\{\vec{v}_1, \ldots, \vec{v}_{n-1}\}$ is a spanning set with $n - 1$ elements, contradicting (b).

**Theorem B:** For any subset $\mathcal{B}$ of a vector space $V$, the following are equivalent:

(i) $\mathcal{B}$ is a basis of $V$;

(ii) $\mathcal{B}$ is a **maximal linearly independent subset** of $V$, meaning that $\mathcal{B}$ is linearly independent and no strictly larger superset of $\mathcal{B}$ is linearly independent;

(iii) $\mathcal{B}$ is a **minimal spanning subset** for $V$, meaning that spans $V$ and no strictly smaller subset of $\mathcal{B}$ spans $V$;

(iv) Every element of $V$ can be written in **one and only one** way as a linear combination of elements in $\mathcal{B}$.

**Problem 4: A lemma on span.** Let $V$ be a vector space, and let $\mathcal{S}$ a subset that spans $V$. Suppose that some $w \in \mathcal{S}$ is in the Span of the set $\mathcal{S} \setminus \{w\}$. Prove that $\mathcal{S} \setminus \{w\}$ also spans $V$.

**Solution:** Take arbitrary $v \in V$. We know we can write $v$ as a linear combination of the vectors in $\mathcal{S}$, say $v = c_1 v_1 + \cdots + c_t v_t$ for some scalars $c_i$ and vectors $v_i \in \mathcal{S}$. If none of the $v_i$ is equal to $w$, we see that $v \in \text{Span}(\mathcal{S} \setminus \{w\})$. On the other hand, if some $v_i = w$, without loss of generality, say $v_1 = w$ (and no other $v_i$ in our sum is $w$). By assumption, we can write $w$ as a linear combination of the vectors in $\mathcal{S} \setminus \{w\}$. That is, $w = a_2 v_2 + \cdots + a_t v_t$, and so

$$v = c_1(a_2 v_2 + \cdots + a_t v_t) + c_2 v_2 + \ldots c_t v_t = (c_1 a_2 + c_2) v_2 + \cdots + (c_1 a_t + c_t) v_t \in \text{Span}(\mathcal{S} \setminus \{w\}).$$

This shows that every $v$ in $V$ is in $\text{Span}(\mathcal{S} \setminus \{w\})$, that is, that $\mathcal{S} \setminus \{w\}$ spans $V$.

**Problem 5: A lemma on linear independence**. Let $V$ be a vector space, and let $\mathcal{S}$ be a linearly independent subset[†] of $V$.

(a) For any $v \in V \setminus \mathcal{S}$, prove that $\mathcal{S} \cup \{v\}$ is linearly independent or $v \in \text{Span}(\mathcal{S})$.
[PROOF TECHNIQUE: To prove "P or Q", you can prove "Not P implies Q." Do you see why?]

[†]Assume $\mathcal{S}$ is a finite set. But the natural argument you are likely to write down works also if $\mathcal{S}$ is infinite.

**Solution:** Assume that for some $v \in V$, the set $\mathcal{S} \cup \{v\}$ is linearly dependent. This means there are vectors $v_1, \ldots, v_n$ in $\mathcal{S}$ along with scalars $c_0, \ldots, c_n \in \mathbb{R}$, not all zero, such that

$$c_0 v + c_1 v_1 + \cdots + c_n v_n \;=\; 0_V. \tag{$*$}$$

Since $\mathcal{S}$ is linearly independent, we must have $c_0 \neq 0$, since otherwise $c_1 v_1 + \cdots + c_n v_n = 0_V$ would be a nontrivial relation on $\mathcal{S}$. Since $c_0 \neq 0$, we can rewrite $(*)$ as

$$v \;=\; (-c_0^{-1} c_1) v_1 + \cdots + (-c_0^{-1} c_n) v_n,$$

which shows that $v \in \mathrm{Span}(\mathcal{S})$.

(b) Prove that *only one* of the two possibilities in (a) can hold. [Hint: Try assuming both hold, and derive a contradiction.]

**Solution:** Assume *both* $v \in \mathrm{Span}(\mathcal{S})$ *and* $\mathcal{S} \cup \{v\}$ is linearly independent. Since $v \in \mathrm{Span}(\mathcal{S})$, we can write $v = c_1 v_1 + \cdots c_n v_n$ where $v_1, \ldots, v_n \in \mathcal{S}$ and $c_1, \ldots, c_n \in \mathbb{R}$. Then

$$c_1 x_1 + \cdots + c_n x_k - v \;=\; 0_V$$

is a nontrivial linear relation on $\mathcal{S} \cup \{v\}$ (the coefficient on $v$ is $-1 \neq 0$), which means that $\mathcal{S} \cup \{v\}$ is linearly dependent. This contradicts our assumption, proving (b).

(c) In the special case $V = \mathbb{R}^{2\times 2}$, $\mathcal{S} = \left\{ \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \right\}$, which of the two possibilities in (a) holds for the vector $v = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$? Find a vector $v' \in V$ so that the other possibility holds.

**Solution:** $\mathcal{S} \bigcup \{v\}$ is linearly independent. For $v' = \begin{bmatrix} 2 & 2 \\ -\pi & -\pi \end{bmatrix}$, we have $v' \in \mathrm{Span}\mathcal{S}$.

**Problem 6: The proof of Theorems A and B.** Let $\mathcal{B}$ be a subset of a vector space $V$.
[Hint: Scaffold carefully and use the Lemmas you proved!]

(a) Prove Theorem A follows from Theorem B.

**Solution:** Let $V$ be a vector space of dimension $n$. Assume $\{v_1, \ldots, v_n\}$ is linearly independent. If $\{v_1, \ldots, v_n\}$ does not span $V$, this means that there is some $w \in V \setminus \mathrm{Span}(v_1, \ldots, v_n)$. By the lemma in Problem 4, this means that $\{w, v_1, \ldots, v_n\}$ is linearly independent. But now we have a linearly independent set with $n+1$ elements in it, so a *maximal* linearly independent set has *at least* $n+1$ elements. Theorem B says that then $V$ has a basis with more at least $n+1$ vectors in it, contrary to $\dim V = n$. So $\{v_1, \ldots, v_n\}$ spans $V$.

For the other direction, say that $\{v_1, \ldots, v_n\}$ spans $V$. If $\{v_1, \ldots, v_n\}$ is not linearly independent, there is a relation $c_1 v_1 + \cdots + c_n v_n = 0$ with at least one $c_i$ not zero. Without loss of generality (by re-ordering the $v_i$), we can assume $c_n \neq 0$. But then $v_n = \frac{-c_1}{c_n} v_n - \cdots - \frac{c_{n-1}}{c_n} v_{n-1}$, so $v_n \in \mathrm{Span}\{v_1, \ldots, v_n\}$. So a minimal spanning set has at most $n-1$ elements, and by Theorem B, this means the dimension is at most $n-1$, again a contradiction.

(b) Prove Theorem $B$.

**Solution:**

We first show (i) and (ii) are equivalent. First assume (i), so $\mathcal{B}$ is a basis of $V$. In particular $\mathcal{B}$ is linearly independent. We need to show every larger set is linearly DEPENDENT. If $v \in V \setminus \mathcal{B}$, then since $\mathcal{B}$ spans $V$ we have $v \in \mathrm{Span}(\mathcal{B})$, which implies that $\mathcal{B} \cup \{v\}$ is not linearly independent by Problem 5.

Next assume (ii). suppose $\mathcal{B}$ is a maximal linearly independent set in $V$. We need to show that $\mathcal{B}$ spans $V$. If it does not, then there is some vector $v \in V$ which is NOT in $\mathrm{Span}(\mathcal{B})$. By Problem 5, this means $\mathcal{B} \cup \{v\}$ is linearly independent, contrary to our assumption that $\mathcal{B}$ is a *maximal* linearly independent set.

Now we know (i) and (ii) are equivalent. We next show (i) and (iii) are equivalent.

Assume (i), so $\mathcal{B}$ is a basis of $V$. Then $\mathcal{B}$ spans $V$, by definition of basis. Now let $\mathcal{S}$ be a proper subset of $\mathcal{B}$, which means we can find some vector $b \in \mathcal{B} \setminus \mathcal{S}$. Note that $\mathcal{S} \cup \{\vec{b}\}$ is linearly independent since it is a subset of the linearly independent set $\mathcal{B}$. By Problem 5, then $\vec{b} \notin \mathrm{Span}(\mathcal{S})$. This says no smaller subset of $\mathcal{B}$ can span $V$, so $\mathcal{B}$ is a minimal spanning set.

Assume (ii), let $\mathcal{B}$ be a minimal spanning set for $V$. We need to show that $\mathcal{B}$ is linearly independent. But if $\mathcal{B}$ is linearly *dependent,* we can we can find some $b \in \mathcal{B}$ which is a linear combination of the vectors in $\mathcal{B} \setminus \{b\}$. This implies that $\mathcal{B} \setminus \{b\}$ spans $V$ by Problem 4, contradicting the assummption that $\mathcal{B}$ is a *minimal* spanning set for $V$. Thus $\mathcal{B}$ is linearly independent, but it also spans $V$, so it is a basis of $V$. This shows (i) andd (iii) are equivalent.

We showed (i) implies (iv) are equivalent in Problem 2 (technically, we only did this when $V$ is finite dimensional, but the argument is the same). To see (iv) implies (i), assume that every element of $V$ can be written uniquely as a linear combination of the elements the basis $\mathcal{B}$. This immediately says that $\mathcal{B}$ spans $V$. However, if $\mathcal{B}$ is not linearly independent, then there are vectors $v_1, \ldots, v_n$ in $\mathcal{B}$ with a non-trivial relation $c_1 v_1 + \cdots + c_n v_n = 0$. As in the proof of (a), we can assume without loss of generality that $c_n \neq 0$ and see that $v_n \in \mathrm{Span}\{v_1, \ldots, v_{n-1}\}$. But then $\mathcal{B}$ is not a minimal spanning set, as we can remove $v_n$ from it. Hence $\mathcal{B}$ could not a basis (since (i) and (iii) are equivalent). This contradicts our hypothesis.

**Problem 7.** Let $\mathcal{S}$ be any subset of a vector space $V$. Prove that $\mathrm{Span}(\mathcal{S})$ is the *smallest subspace* of $V$ containing the set $\mathcal{S}$. [Hint: First show $\mathrm{Span}(\mathcal{S})$ is a subspace. Then suppose $\exists W$ subspace s.t. $\mathcal{S} \subset W \subsetneq \mathrm{Span}(\mathcal{S})$....]

**Problem 8.** Let $W$ be a subspace of the vector space $V$. Prove or disprove:

(a) Every basis of $V$ contains a basis of $W$. FALSE! Say $V = \mathbb{R}^2$ and $W = \mathrm{Span}(\vec{e}_1 + \vec{e}_2)$. Observe that $\{\vec{e}_1, \vec{e}_2\}$ is basis for $\mathbb{R}^2$, but no subset is a basis for $W$.

(b) Every basis of $W$ is contained in a basis of $V$. TRUE! The basis for $W$ consists of linearly indepdent vectors in $W$, and also in $V$. So this set can be expanded to a maximal linearly independent set in $V$, or a basis of $V$.

**Problem 9. Avoid a Mental Trap.** Many vector spaces have an "obvious" or a "natural"—or as a mathematician might say, a **canonical** basis. For example, $\mathbb{R}^n$ has a canonical basis (What is it?) However, not all vector spaces have a natural choice of basis. Does the plane $2x - y + z = 0$ have a canonical basis, for example? If each student in your group writes down what they think is the most obvious basis, will you all agree? Does $\mathcal{P}_n$ have a canonical basis? Does $\mathbb{R}^{n \times m}$?

Moreover, we are often interested in **ordered bases**–meaning bases with a fixed ordering of the basis elements. For example, $(\vec{e}_1, \vec{e}_2)$ is a different **ordered basis** for $\mathbb{R}^2$ than $(\vec{e}_2, \vec{e}_1)$. Does $\mathbb{R}^n$ have a canonical *ordered* basis? Does $\mathcal{P}_n$? Does $\mathbb{R}^{n \times m}$? What about the other vector spaces in Problem 1?

**Solution:** The word "canonical" should not be interpreted as a precise term—understanding what is canonical or not depends a bit on the context and audience. But usually everyone agrees that that $\mathbb{R}^n$ has a canonical basis $\{\vec{e}_1, \ldots, \vec{e}_n\}$ as well as a canonical *ordered basis* $(\vec{e}_1, \ldots, \vec{e}_n)$, because it would be silly to write the $\vec{e}_i$ in some other order if we are trying to do everything as naturally as possible.

The vector spaces $\mathcal{P}_n$ and $\mathbb{R}^{m\times n}$, many would agree, also have canonical bases, namely $\{1, x, \ldots, x^n\}$ and $\{E_{ij} \mid 1 \le i \le m,\ 1 \le j \le n\}$. However, it is less agreed upon that these are canonical as *ordered* bases: some might prefer $(1, x, \ldots, x^n)$ and others $(x^n, \ldots, x, 1)$. For the basis $\{E_{ij} \mid 1 \le i \le m,\ 1 \le j \le n\}$ it feels even less clear. Should we list rows or columns first? Or maybe order them in some other way, moving along diagonals or something? At least I would argue, and I think most mathematicians would agree, that $\mathcal{P}_n$ and $\mathbb{R}^{m\times n}$ have canonical bases but not canonical ordered bases.

The subspace given by the plane with equation $2x - y + z = 0$ definitely *does not* have a canonical basis, as you surely realized if each in your group wrote down a basis. Some are nicer (for example, *sparser*–meaning more zeros) than others, but it's hard to argue that there is a clear "most obvious" choice of basis.

Of the vector spaces in Problem 1, $\mathbb{C}$ has an obviously canonical basis $\{1, i\}$, and some would argue that $(1, i)$ is a natural ordering. Likewise, strictly upper triangular $3 \times 3$ matrices have the basis $\{E_{ij} \mid 1 \le i < j \le 3\} = \{E_{12}, E_{13}, E_{23}\}$, and someone might argue there's a pretty natural way to order in this case. For the vector space in Problem 1 (f), $\{x^4, x^3, x^2, x\}$ is a decently obvious (canonical) basis but some might disagree.