Chapter 1 Proofs

1 Chapter 1 Proofs

These are the theorems from Chapter 1 which I proved in class on Monday, September 11th. I have made this document to demonstrate what I believe is an acceptable level of detail when asked to prove something. It's not necessary to write as formally as I do; what you should focus on is how I clearly define all terms/symbols before using them; how I carefully justify each step in the proof; and the overall structure of a proof showing two sets are equal or two statements are equivalent.

Theorem 1.1. *range($\boldsymbol{A}$) is the space spanned by the columns of $\boldsymbol{A}$.*

Comments: This is the exact statement of the theorem as it appears in the textbook. First, in order to reason about the matrix $\boldsymbol{A}$ in our proof, we should use variables to denote the number of rows and columns. We should also specify the nature of the entries of $\boldsymbol{A}$. We'll let m denote the number of rows, n the number of columns, and allow for complex-valued entries. Put succintly:

$$\boldsymbol{A} \in \mathbb{C}^{m \times n}$$

After specifying the dimensions of $\boldsymbol{A}$ and what space its elements belong to, the definition of range($\boldsymbol{A}$) is:

$$\text{range}(\boldsymbol{A}) = \{\boldsymbol{y} \in \mathbb{C}^m : \boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \text{ for some } \boldsymbol{x} \in \mathbb{C}^n\}$$

Let's also introduce some notation for the object "the space spanned by the columns of $\boldsymbol{A}$". Translating words into mathematical symbols is a skill you should develop, but I will do the translation here. The "space spanned by the columns of $\boldsymbol{A}$" is usually called the column space of $\boldsymbol{A}$, and it is the set of linear combinations of the columns of $\boldsymbol{A}$. We can express it as follows: "Let $\boldsymbol{a}_j$ denote the jth column of $\boldsymbol{A}$. Then the column space of $\boldsymbol{A}$ is:

$$\text{col}(\boldsymbol{A}) = \text{span}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\} = \left\{ \sum_{j=1}^{n} \gamma_j \boldsymbol{a}_j : \gamma_1, \ldots, \gamma_n \in \mathbb{C} \right\}$$

If we were restricting attention to only $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, then we would consider $\gamma_1, \ldots, \gamma_n \in \mathbb{R}$ instead.

What this proof is stating is that the two sets range($\boldsymbol{A}$) and col($\boldsymbol{A}$) are the same, even though they are defined differently. If the proof of this theorem were on a homework assignment I would probably state the question as follows:

Question 1: Let $\boldsymbol{A} \in \mathbb{C}^{m \times n}$. Show that range($\boldsymbol{A}$) = col($\boldsymbol{A}$).

To prove set equality, we typically do two steps: Show that range($\boldsymbol{A}$) $\subset$ col($\boldsymbol{A}$) (every element of range($\boldsymbol{A}$) is also an element of col($\boldsymbol{A}$)), then show that col($\boldsymbol{A}$) $\subset$ range($\boldsymbol{A}$).

The proof on the next page would be an acceptable proof for this theorem. (You don't need to write as formally as I do, but it can make it easier to grade.) Red text denotes a comment from me which is not part of the proof

Proof. **Step 1:** Let $\boldsymbol{y}$ be an arbitrary element of range$(\boldsymbol{A})$. By the definition of range$(\boldsymbol{A})$, there exists a vector $\boldsymbol{x} \in \mathbb{C}^n$ such that

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x}. \tag{1}$$

By the column-based interpretation of matrix-vector multiplication (this phrase justifies my next equation), we know that

$$\boldsymbol{A}\boldsymbol{x} = \sum_{j=1}^{n} x_j \boldsymbol{a}_j \tag{2}$$

where x_j are the components of $\boldsymbol{x}$, and $\boldsymbol{a}_j$ are the columns of $\boldsymbol{A}$. (Notice how I explicitly said what x_j and $\boldsymbol{a}_j$ were, instead of having the grader figure out what I'm talking about).

Combining the last two equations, it follows that

$$\boldsymbol{y} = \sum_{j=1}^{n} x_j \boldsymbol{a}_j \tag{3}$$

Since $\boldsymbol{x} \in \mathbb{C}^n$, we have $x_j \in \mathbb{C}$ for $j = 1, \ldots, n$. By the definition of col$(\boldsymbol{A})$, it follows that $\boldsymbol{y} \in$ col$(\boldsymbol{A})$. Since $\boldsymbol{y} \in$ range$(\boldsymbol{A})$ was arbitrary, we have shown that

$$\text{range}(\boldsymbol{A}) \subset \text{col}(\boldsymbol{A}) \tag{4}$$

Step 2: Let $\boldsymbol{y} \in$ col$(\boldsymbol{A})$ be arbitrary. By definition, there exist scalars $\gamma_1, \ldots, \gamma_n \in \mathbb{C}$ such that

$$\boldsymbol{y} = \sum_{j=1}^{n} \gamma_j \boldsymbol{a}_j. \tag{5}$$

Define a vector $\boldsymbol{x} \in \mathbb{C}^n$ by:

$$\boldsymbol{x} \equiv \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \vdots \\ \gamma_n \end{bmatrix} \tag{6}$$

By the column-based interpretation of matrix-vector products, we know that

$$\sum_{j=1}^{n} \gamma_j \boldsymbol{a}_j = \boldsymbol{A}\boldsymbol{x} \tag{7}$$

Combining eq. (5) and eq. (7), it follows that

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} \tag{8}$$

By definition, it follows that $\boldsymbol{y} \in$ range$(\boldsymbol{A})$. Since $\boldsymbol{y} \in$ col$(\boldsymbol{A})$ was arbitrary, it follows that

$$\text{col}(\boldsymbol{A}) \subset \text{range}(\boldsymbol{A}) \tag{9}$$

$\square$

Theorem 1.2. *A matrix $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ with $m \geq n$ has full rank if and only if it maps no two distinct vectors to the same vector.*

Comments: The phrase "$\boldsymbol{A}$ maps no two distinct vectors to the same vector" is a little confusing. I would rather say that "$\boldsymbol{A}$ is injective", which means the following:

$$\boldsymbol{x} \neq \boldsymbol{y} \implies \boldsymbol{A}\boldsymbol{x} \neq \boldsymbol{A}\boldsymbol{y}.$$

I.e. if $\boldsymbol{x}$ does not equal to $\boldsymbol{y}$, then $\boldsymbol{A}\boldsymbol{x}$ does not equal to $\boldsymbol{A}\boldsymbol{y}$.

The following is an equivalent definition of injectivity, which is the "contrapositive" of the previous statement:

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{A}\boldsymbol{y} \implies \boldsymbol{x} = \boldsymbol{y}$$

So this theorem states that for any $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ with $m \geq n$, the following two conditions are equivalent:

- $\boldsymbol{A}$ is full rank
- $\boldsymbol{A}$ is injective

To prove this, we need to show that whenever $m \geq n$, every full-rank $\boldsymbol{A}$ is also injective, and then show that every injective $\boldsymbol{A}$ is also full-rank.

The following is an acceptable proof:

Proof. **Step 1:** Let $\boldsymbol{A} \in \mathbb{C}^{m\times n}$ be full rank, where $m \geq n$. Since $m \geq n$, we have

$$\operatorname{rank}(\boldsymbol{A}) = \min\{m, n\} = n \tag{10}$$

By definition of rank, we have:

$$n = \operatorname{rank}(\boldsymbol{A}) = \dim[\operatorname{col}(\boldsymbol{A})] = \dim\left[\operatorname{span}\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\}\right] \tag{11}$$

where $\boldsymbol{a}_j$ is the jth column of $\boldsymbol{A}$.

Since the n column vectors span an n-dimensional space, they must be linearly independent, and thus form a basis for $\operatorname{span}\{\boldsymbol{a}_1, \ldots, \boldsymbol{a}_n\} = \operatorname{col}(\boldsymbol{A}) = \operatorname{range}(\boldsymbol{A})$. (Note that I am using Theorem 1.1 here. I'll order the homework questions in such a way that you can freely use facts from earlier problems, even if you aren't able to complete them successfully. You can also use facts I state in class.)

Let $\boldsymbol{x}$ and $\boldsymbol{y}$ be arbitrary distinct vectors in $\mathbb{C}^n$; denote their entries by x_j and y_j respectively. Define a vector $\boldsymbol{b} \in \mathbb{C}^m$ as:

$$\boldsymbol{b} \equiv \boldsymbol{A}\boldsymbol{x} = \sum_{j=1}^{n} x_j \boldsymbol{a}_j, \tag{12}$$

where we used the column-based interpretation of matrix-vector multiplication. Thus, $\boldsymbol{b} \in \operatorname{range}(\boldsymbol{A})$ and eq. (12) is the representation of $\boldsymbol{b}$ in the basis $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n\}$.

Since a vector's representation in a basis is unique (standard result from MATH 0520 or other linear algebra course), and $\boldsymbol{x} \neq \boldsymbol{y}$ it follows that

$$\boldsymbol{A}\boldsymbol{y} = \sum_{j=1}^{n} y_j \boldsymbol{a}_j \neq \boldsymbol{b} = \boldsymbol{A}\boldsymbol{x} \tag{13}$$

Since $\boldsymbol{x}$ and $\boldsymbol{y}$ were arbitrary, $\boldsymbol{A}$ is injective. So we have shown

$$\boldsymbol{A} \text{ is full-rank} \implies \boldsymbol{A} \text{ is injective} \tag{14}$$

Step 2: Let $\boldsymbol{A} \in \mathbb{C}^{m \times n}$ be injective, where $m \geq n$. Let $c_1, c_2, \ldots, c_n$ be scalars in $\mathbb{C}$ which satisfy:

$$\mathbf{0} = \sum_{j=1}^{n} c_j \boldsymbol{a}_j \tag{15}$$

Define a vector $\boldsymbol{c} \in \mathbb{C}^n$ as:

$$\boldsymbol{c} \equiv \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \tag{16}$$

Again, by the column-based definition of matrix-vector products (could probably stop saying this, the reader would get it by now) we can write eq. (15) as:

$$\mathbf{0} = \boldsymbol{A}\boldsymbol{c} \tag{17}$$

However, it is also true that

$$\mathbf{0} = \boldsymbol{A}\mathbf{0} \tag{18}$$

In the above equation there are two different zero vectors. The one on the left is the zero vector in $\mathbb{C}^m$, while the one on the right is the zero vector in $\mathbb{C}^n$. But this can be inferred from the dimensions of $\boldsymbol{A}$ so we can use the same symbol for both.

Combining eq. (17) and eq. (18), we have

$$\boldsymbol{A}\boldsymbol{c} = \boldsymbol{A}\mathbf{0} \tag{19}$$

Since $\boldsymbol{A}$ is injective, we must have $\boldsymbol{c} = \mathbf{0}$, which implies that $c_1 = c_2 = \cdots = c_n = 0$. It follows that the vectors $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_n$ are linearly independent. Thus,

$$\operatorname{rank}(\boldsymbol{A}) = \dim\left[\operatorname{span}\{\boldsymbol{a}_1, \cdots, \boldsymbol{a}_n\}\right] = n = \min\{m, n\} \tag{20}$$

so $\boldsymbol{A}$ is full-rank. So we have shown that

$$\boldsymbol{A} \text{ is injective} \implies \boldsymbol{A} \text{ is full-rank} \tag{21}$$

$\square$

Comment on Theorem 1.2: In the textbook's proof, Step 2 instead shows that

$$\boldsymbol{A} \text{ is not full-rank} \implies \boldsymbol{A} \text{ is not injective} \tag{22}$$

This is logically equivalent to eq. (21), and is called the "contrapositive" of eq. (21). This is known as "proof by contraposition". It is sometimes easier to prove things this way, but for this theorem, they're about the same level of difficulty.

Presentation of Proof: If you look at the textbook's proofs, you will notice they are much shorter than mine. I wrote the above proofs to a) carefully show all the steps in case the book's proofs are not clear to you, and b) demonstrate the flow of a proof for future theorems.

If you are inexperienced writing proofs, you are more likely to make an error if you try to write a terse proof like in the book. If you have written proofs before, you can probably get away with something quicker, but the person reading your work should be able to follow it (and errors are still more likely!)