EECS 16A Designing Information Devices and Systems I
Spring 2020 Lecture Notes Note 3

3.1 Linear Dependence

Recall the simple tomography example from Note 1, in which we tried to determine the composition of a box of bottles by shining light at different angles and measuring light absorption. From what we've seen so far, we know that we need to take at least 9 measurements to uniquely identify the 9 bottles in a box. However, will taking *any* 9 measurements guarantee that we can find a solution? Answering this question requires an understanding of linear dependence. In this note, we will define linear dependence (and independence), and take a look at what it implies for systems of linear equations.

3.1.1 What is Linear Dependence?

Linear dependence is a very useful concept that is often used to characterize the "redundancy" of information in real world applications[1]. We will give two equivalent definitions of linear dependence.

Definition 3.1 (Linear Dependence (I)): A set of vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is linearly dependent if there exist scalars $\alpha_1, \ldots, \alpha_n$ such that $\alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n = \vec{0}$ and not all α_i's are equal to zero.

Definition 3.2 (Linear Dependence (II)): A set of vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is linearly dependent if there exist scalars $\alpha_1, \ldots, \alpha_n$ and an index i such that $\vec{v}_i = \sum_{j \neq i} \alpha_j \vec{v}_j$.[2]

In words, the second definition is saying that a set of vectors is linearly dependent if one of the vectors can be realized as a linear combination of the others.

Why did we introduce two equivalent definitions? They could be useful in different settings. For example, it is often easier mathematically to show linear dependence with definition (I). Can you see why? If we would like to prove linear dependence with definition (II), we need to first choose a vector $\vec{v}_i$ and show that it is a linear combination of the other vectors. However, with definition (I), we don't need to try to "single out" a vector to get started with the proof. We can blindly write down the equation $\alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n = \vec{0}$ and begin our proof from there. On the other hand, definition (II) gives us a more intuitive way to talk about redundancy. If a vector can be constructed from the rest of the vectors, then this vector does not contribute any information that is not already captured by the other vectors.

Now we will show that the two definitions are equivalent. This is the first formal proof in the course! We will walk you through it.[3] First, we ask the question, "What does it mean when we say two definitions are

[1] The term "redundancy" does not have a precise definition in this context; we are using it here to appeal to your sense of intuition. The precise terminology is that defined on this page: linear dependence.

[2] In case you are unfamiliar with this notation, the $\sum$ symbol is simply shorthand for addition. For instance, $\alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n$ can be written as $\sum_{i=1}^{n} \alpha_i \vec{x}_i$ or $\sum_i \alpha_i \vec{x}_i$, which is a sum over all possible i values. In this instance, $\sum_{j \neq i} \alpha_j \vec{v}_j$ is the sum over all $\alpha_j \vec{v}_j$ excluding the $\alpha_i \vec{v}_i$ term, which can also be calculated as $\alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n - \alpha_i \vec{v}_i$.

[3] Note 4 provides a more in depth treatment on how to approach proofs.

equivalent?” It means that when the condition in definition (I) holds, the condition in definition (II) must hold as well. And when the condition in definition (II) holds, the condition in definition (I) must also hold. So there are two directions that we have to show:

(i) To see how definition (II) implies definition (I), we start from the condition in definition (II) — suppose there exist scalars $\alpha_1, \ldots, \alpha_n$ and an index i such that $\vec{v}_i = \sum_{j \neq i} \alpha_j \vec{v}_j$. We want to somehow transform this equation into the form that appears in definition (I). How can we achieve that? We can move $\vec{v}_i$ to the right:

$$\vec{0} = -\vec{v}_i + \sum_{j \neq i} \alpha_j \vec{v}_j. \tag{1}$$

Now setting $\alpha_i = -1$, we have

$$\vec{0} = \alpha_i \vec{v}_i + \sum_{j \neq i} \alpha_j \vec{v}_j = \sum_{j} \alpha_j \vec{v}_j. \tag{2}$$

Since $\alpha_i = -1$, at least one of the α_j terms is not zero, and the condition in definition (I) is satisfied.

(ii) Now let’s show the reverse — that definition (I) implies definition (II). Suppose the condition in definition (I) is true. Then there exist scalars $\alpha_1, \ldots, \alpha_n$ such that $\alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n = 0$ and not all α_i’s are equal to zero. Since at least one of the α_i’s is nonzero, let’s assume that α_1 is nonzero since we can always reorder terms in the summation because addition is commutative. Now how do we get the equation into the form identical to that in definition (II)? Observe that if we move $\alpha_1 \vec{v}_1$ to the other side of equation and divide both sides by α_1 (a valid operation, since $\alpha_1 \neq 0$ by assumption), we have

$$\vec{v}_1 = -\sum_{j \neq 1} \left(\frac{\alpha_j}{\alpha_1} \right) \vec{v}_j. \tag{3}$$

We see that this is identical to the second definition. (In our proof, we made the assumption that $\alpha_1 \neq 0$. However, notice that we could as well have supposed that $\alpha_2 \neq 0$, $\alpha_3 \neq 0$, or any index i so that $\alpha_i \neq 0$. The convention is to set the first index, in this case 1, to be nonzero. In mathematical texts, we typically write *“Without loss of generality (W.L.O.G.), we let $\alpha_1 \neq 0$.”*)

Now that we have introduced the notion of linear dependence, what does it mean to be linearly *independent*?

3.1.2 Linear Independence

Definition 3.3 (Linear Independence): A set of vectors is linearly independent if it is not linearly dependent. More specifically, from the first definition of linear dependence we can deduce that a set of vectors $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is linearly independent if $\alpha_1 \vec{v}_1 + \cdots + \alpha_n \vec{v}_n = \vec{0}$ implies $\alpha_1 = \cdots = \alpha_n = 0$.

Let’s see some simple examples of linear dependence and linear independence.

Example 3.1 (Linear dependence of 2 vectors): Consider vectors $\vec{a} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$. These vectors are linearly dependent because we can write $\vec{b}$ as a scaled version of $\vec{a}$:

$$\vec{b} = \begin{bmatrix} 4 \\ 2 \end{bmatrix} = 2 \times \begin{bmatrix} 2 \\ 1 \end{bmatrix} = 2 \times \vec{a}.$$

Example 3.2 (Linear independence of 2 vectors): Consider vectors $\vec{a} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}$. We will show that the two vectors are linearly independent. Consider scalars α_1, α_2 such that $\alpha_1 \vec{a} + \alpha_2 \vec{b} = \vec{0}$. We can write this vector equation as a system of linear equations:

$$\begin{aligned}
&\alpha_1 \vec{a} + \alpha_2 \vec{b} = \vec{0} \\
\Rightarrow &\alpha_1 \begin{bmatrix} 2 \\ 1 \end{bmatrix} + \alpha_2 \begin{bmatrix} 1 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
\Rightarrow &\begin{bmatrix} 2\alpha_1 \\ \alpha_1 \end{bmatrix} + \begin{bmatrix} \alpha_2 \\ 5\alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
\Rightarrow &\begin{bmatrix} 2\alpha_1 + \alpha_2 \\ \alpha_1 + 5\alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\
\Rightarrow &\begin{cases} 2\alpha_1 + \alpha_2 = 0 \\ \alpha_1 + 5\alpha_2 = 0 \end{cases}
\end{aligned}$$

Solving this system of linear equations with Gaussian elimination yields a unique solution, $\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. So, $\vec{a}$ and $\vec{b}$ are linearly independent by definition.

Additional Resources For more on the definition of linear dependence, read *Strang* pages 164-167 or read *Schuam's* pages 121-124. For additional practice with these ideas, try *Schuam's* Problems 4.17 to 4.22, and 4.89 to 4.96.

3.1.3 Linear Dependence and Systems of Linear Equations

Previously, we saw that a system of linear equations can have zero solutions, a unique solution, or infinitely many solutions. Is there a way to tell what kind of solution a system of linear equations has without running Gaussian elimination or explicitly solving for the solution? Yes! Recall that a system of linear equations can be written in matrix-vector form as $A\vec{x} = \vec{b}$, where A is a matrix of variable coefficients, $\vec{x}$ is a vector of variables, and $\vec{b}$ is a vector of values that these weighted sums must equal. We will show that just looking at the columns or rows of the matrix A can help tell us about the solutions to $A\vec{x} = \vec{b}$.

Theorem 3.1: **If the system of linear equations $A\vec{x} = \vec{b}$ has an infinite number of solutions, then the columns of A are linearly dependent.**

Let's see why this is the case: If the system has infinite number of solutions, it must have at least two distinct solutions. Let's call them $\vec{x}_1$ and $\vec{x}_2$; our assumption that these are distinct solutions can be rewritten as $\vec{x}_1 \neq \vec{x}_2$. (Note that $\vec{x}_1, \vec{x}_2$ are vectors, not entries of vectors.) Then $\vec{x}_1$ and $\vec{x}_2$ must satisfy

$$A\vec{x}_1 = \vec{b} \tag{4}$$

$$A\vec{x}_2 = \vec{b}. \tag{5}$$

Subtracting the first equation from the second equation, we have $A\left(\vec{x}_2 - \vec{x}_1\right) = \vec{0}$. Let $\vec{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \vec{x}_2 - \vec{x}_1$.

Because $\vec{x}_1 \neq \vec{x}_2$, not all α_i's are equal to zero. Let the columns of A be $\vec{a}_1, \ldots, \vec{a}_n$. Then, $A\vec{\alpha} = \sum_{i=1}^{n} \alpha_i \vec{a}_i = \vec{0}$. Hence, by definition, the columns of A are linearly dependent.

Note that in this proof, we used the property of matrix multiplication that $A\vec{\alpha} = \sum_{i=1}^{n} \alpha_i \vec{a}_i$. We scale each column and add them together. In other words, matrix-vector multiplication is a linear combination of columns. This property is often a useful way to think about matrix multiplication. The following example might help:

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} \alpha_1 a_1 + \alpha_2 b_1 + \alpha_3 c_1 \\ \alpha_1 a_2 + \alpha_2 b_2 + \alpha_3 c_2 \\ \alpha_1 a_3 + \alpha_2 b_3 + \alpha_3 c_3 \end{bmatrix} = \alpha_1 \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} + \alpha_2 \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} + \alpha_3 \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

Theorem 3.2: **If the columns of A in the system of linear equations $A\vec{x} = \vec{b}$ are linearly dependent, then the system does not have a unique solution.**

Remark. Note, the above statement does not guarantee that $A\vec{x} = \vec{b}$ is a consistent system of linear equations. It could be inconsistent (i.e., have no solution), in which case it also does not have a unique solution.

Let's walk through this proof step by step: we'll start by assuming we have a matrix A with linearly dependent columns, and then we will show that this means that the system does not have a unique solution.

Since we are interested in the columns of A, let's start by explicitly defining the columns of A:

$$A = \begin{bmatrix} | & | & & | \\ \vec{a}_1 & \vec{a}_2 & \ldots & \vec{a}_n \\ | & | & & | \end{bmatrix},$$

We've defined A to have linearly dependent columns, so by the definition of linear dependence, there exist scalars $\alpha_1, \ldots, \alpha_n$ such that $\alpha_1 \vec{a}_1 + \ldots + \alpha_n \vec{a}_n = \vec{0}$ where not all of the α_i's are zero. We can put these α_i's in a vector

$$\vec{\alpha} = \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

and by the definition of matrix-vector multiplication, we can compactly write the expression above:

$$A\vec{\alpha} = \vec{0}$$

where $\vec{\alpha} \neq \vec{0}$.

Recall that we are trying to show that the system of equations $A\vec{x} = \vec{b}$ does not have a unique solution. We know that systems of equations can have either zero, one, or infinite solutions. If our system of equations has zero solutions, then it cannot have a unique solution, so we don't need to consider this case. Now let's consider the case where we have at least one solution, $\vec{x}$:

$$\begin{aligned} A\vec{x} &= \vec{b} \\ A\vec{x} + \vec{0} &= \vec{b} \\ A\vec{x} + A\vec{\alpha} &= \vec{b} \\ A\left(\vec{x} + \vec{\alpha}\right) &= b \end{aligned}$$

Therefore, $\vec{x} + \vec{\alpha}$ is also a solution to the system of equations! Since both $\vec{x}$ and $\vec{x} + \vec{\alpha}$ are solutions, and $\vec{\alpha} \neq \vec{0}$, the system has more than one solution. We've now proven the theorem.

Note that we can add any scalar multiple of $\vec{\alpha}$ to $\vec{x}$ and it will still be a solution – therefore, if there is at least one solution to the system and the columns of A are linearly dependent, then there are infinite solutions. Let's think about why this makes sense intuitively. In an experiment, each column in matrix A represents the influence of each variable x_i on the measurements. If the columns are linearly dependent, this means that some of the variables influence the measurement in "the same" way (interpreted precisely in terms of linear dependence), and therefore cannot be disambiguated. The next example illustrates this idea.

Example 3.3 (Intuition): Suppose we have a black and white image with two pixels. We cannot directly see the shade of each pixel, but we can measure how much light the two pixels absorb in total. Can we figure out the shade of each pixel? Let's model this as a system of linear equations. Suppose pixel 1 absorbs x_1 units of light and pixel 2 absorbs x_2 units of light. Our measurement indicates that total amount of light absorbed by the image is 10 units of light. Then we could write down the equation,

$$x_1 + x_2 = 10. \tag{6}$$

Written in matrix form, we have

$$\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 10 \end{bmatrix}. \tag{7}$$

We see that the columns are $\begin{bmatrix} 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \end{bmatrix}$. The total amount of light absorbed is influenced by 1 unit of x_1 and 1 unit of x_2. However, we cannot pin down the exact influence by x_1 and x_2 because if pixel 1 absorbs c units less, we can just have pixel 2 absorb c units more. This is connected with the fact that the two columns are linearly dependent — if one pixel absorbs less, it is possible to find a way such that the other pixel absorbs more to make up for the loss (the column of that pixel can be written as a linear combination of the columns of the other pixels).

This result has important implications to the design of engineering experiments. Often times, we can't directly measure the values of the variables we're interested in. However, we can measure the total weighted contribution of each variable. The hope is that we can fully recover each variable by taking several of such measurements. Now we can ask: "What is the minimum number of measurements we need to fully recover the solution?" and "How do we design our experiment so that we can fully recover our solution with the minimum number of measurements?" Consider the tomography example. We are confident that we can figure out the configuration of the stack when the columns of the lighting pattern matrix A in $A\vec{x} = \vec{b}$ are linearly independent. On the other hand, if the columns of the lighting pattern matrix are linearly dependent, we know that we don't yet have enough information to figure out the configuration. Checking whether the columns are linearly independent gives us a way to validate whether we've effectively designed our experiment.

3.2 Row Perspective

(This section is optional for the course.)

So far, we have seen a number of results relating the columns of a matrix to its corresponding system of linear equations. But what about the rows? Intuitively, each row represents some measurement: for example, if our linear system is

$$\begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix},$$

then the variables we want to measure are α_1, α_2, and α_3, and the second row represents the measurement $a_2\alpha_1 + b_2\alpha_2 + c_2\alpha_3$. If we take less than 3 measurements, then of course we cannot recover all three variables. So suppose we take 3 or more measurements. If the number of measurements taken is at least the number of variables and we still cannot completely determine the variables, then at least one of our measurements must be redundant (it doesn't give us any new information). This intuition suggests that the number of variables we can recover is equal to the number of unique measurements, or the number of linearly independent rows.

While this is an intuitive argument, we need a formal proof to be sure of the reasoning. This formal proof will come in a later note when we talk about rank.

Note that we now have two perspectives: in the matrix, each row represents a measurement, while each column corresponds to a variable. Therefore, if the columns are linearly dependent, then we have at least one redundant variable. From the perspective of rows, linear dependency tells us that we have one or more redundant measurements.

3.3 Span

Let's introduce **span**, a concept closely related to linear dependence that will be used throughout this course.

Definition 3.4 (Span): The span of a set of vectors $\{v_1, \ldots, v_n\}$ is the set of *all* linear combinations of $\{v_1, \ldots, v_n\}$. We can write this mathematically as

$$span(v_1, \ldots, v_n) = \left\{ \sum_{i=1}^{n} \alpha_i \vec{v}_i \mid \alpha_i \in \mathbb{R} \right\}$$

We can now rephrase our second definition of linear dependence: A set of vectors is linearly dependent if any one of the vectors is in the *span* of the remaining vectors.

When given a matrix A, the **range** and **column space** of A both refer to the span of the columns of A! Some people also colloquially refer to this as the *span of* A, but this terminology is not widely accepted and should be avoided.

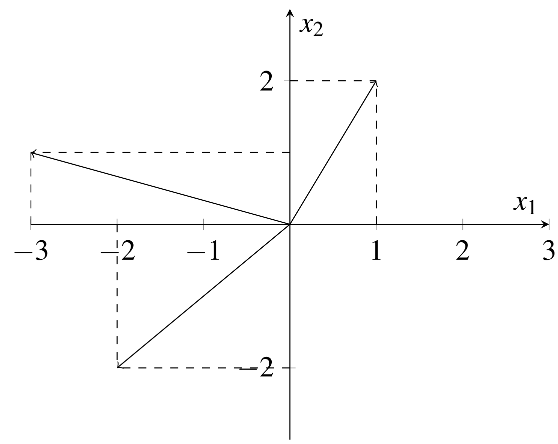
Additional Resources For more on linear span, read *Schaum's* pages 119-121. For additional practice with these ideas, try Problems 4.13 to 4.16, 4.66, 4.69, and 4.83 to 4.88.

Example 3.4 (Span Practice): Let's see how to solve problems involving the span of a set of vectors. Consider the three vectors:

$$\begin{bmatrix}1\\2\end{bmatrix}, \begin{bmatrix}-3\\1\end{bmatrix}, \begin{bmatrix}-2\\-2\end{bmatrix}.$$

How can we compute and express their span?

First, let's try to gain some intuition for these vectors, by plotting them on a set of axes:



Intuitively, it seems like linear combinations of these three vectors can reach any point on the plane. Let's see if we can justify this rigorously. Consider an arbitrary point $\begin{bmatrix}a\\b\end{bmatrix}$ on the plane. We'd like to see if we can write this point as a linear combination of our three vectors - in other words, we'd like to show that no matter what a and b we pick, we can choose scalars c_1, c_2, and c_3 such that

$$c_1 \begin{bmatrix}1\\2\end{bmatrix} + c_2 \begin{bmatrix}-3\\1\end{bmatrix} + c_3 \begin{bmatrix}-2\\-2\end{bmatrix} = \begin{bmatrix}a\\b\end{bmatrix}.$$

Applying the rules of vector algebra that we presented in the previous note to simply the summation on the left-hand-side, this equation is equivalent to

$$\begin{bmatrix}(1)c_1 + (-3)c_2 + (-2)c_3\\(2)c_1 + (1)c_2 + (-2)c_3\end{bmatrix} = \begin{bmatrix}a\\b\end{bmatrix}$$

Observe that the above equation is essentially a system of linear equations with three unknowns - c_1, c_2, and c_3. Writing it in the standard "$Ax = b$" form, we obtain

$$\begin{bmatrix}1 & -3 & -2\\2 & 1 & -2\end{bmatrix} \begin{bmatrix}c_1\\c_2\\c_3\end{bmatrix} = \begin{bmatrix}a\\b\end{bmatrix}.$$

Recall our original goal - to show that no matter how a and b are chosen, we can solve for constants c_1, c_2, and c_3 such that the above equation is satisfied. But now that we have written it in the above form, we know how to solve it - Gaussian elimination! Remember that since we are trying to solve for the c_i as functions of a and b, a and b should not be treated as unknowns for the purpose of Gaussian elimination, but rather as arbitrarily chosen constants.

Observe that the pivot of the first row is already 1, so we can immediately eliminate the coefficient of c_1 in the second row by subtracting twice the first row from the second, to obtain

$$\begin{bmatrix} 1 & -3 & -2 \\ 0 & 7 & 2 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} a \\ b-2a \end{bmatrix}.$$

Now, we should scale the second row by a factor of $1/7$ in order to get a 1 in the coefficient for c_2 in the second row, so we obtain

$$\begin{bmatrix} 1 & -3 & -2 \\ 0 & 1 & 2/7 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} a \\ (1/7)b-(2/7)a \end{bmatrix}.$$

Observe that we have now placed our matrix of coefficients in row echelon form, so we can now determine whether a solution exists. It is straightforward to see that this system of equations is consistent for any choice of a, b, so that our three vectors span all vectors of dimension 2. In mathematical terms,

$$\text{span}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} -3 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ -2 \end{bmatrix} \right) = \mathbb{R}^2.$$

3.4 Span, Range, and Consistency of Linear Equations

In this section, let's connect the concept of span (or, more specifically, range of a matrix) to consistency of a given system of linear equations. First, let's dust off some terminology you presumably learned when you first encountered the concept of a *function* (probably way back in middle or high school). Associated with any function f are two sets: $\text{dom}(f)$, which we call the *domain* of f; and $\text{range}(f)$, which we call the *range* of f.

Recall that the domain $\text{dom}(f)$ is the set of points on which the function f is defined. Specifically, the function f evaluates to some quantity $f(\omega)$ for each choice of $\omega \in \text{dom}(f)$. Now, the range is defined as the set

$$\text{range}(f) = \{f(\omega) | \omega \in \text{dom}(f)\}.$$

In words, the output of the function f 'ranges' over all points in $\text{range}(f)$ as we let the input vary over the entire domain.

Now, consider a matrix $A \in \mathbb{R}^{m\times n}$. Using the definition of matrix-vector multiplication, the quantity $A\vec{x}$ is a vector in $\mathbb{R}^m$ for any choice of $\vec{x} \in \mathbb{R}^n$. In this way, we can think of A as representing a function from $\mathbb{R}^n$ to $\mathbb{R}^m$ (i.e., for any "input" $\vec{x} \in \mathbb{R}^n$, the function "output" is $A\vec{x} \in \mathbb{R}^m$). In terms of the above discussion, the matrix A — thought of as a function — has domain $\mathbb{R}^n$, and range contained in $\mathbb{R}^m$. However, this seems to have potential for conflict, since we have already defined the range of a matrix as the span of its columns. You can rest easy, though, since the latter definition is in complete agreement with the general notion of range of a function described above. Indeed, let $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n$ denote the columns of A, and observe the following identities

$$\text{range}(A) = \{A\vec{x} \mid \vec{x} \in \mathbb{R}^n\} = \left\{ \sum_{i=1}^{n} x_i \vec{a}_i \,|\, x_i \in \mathbb{R} \right\} = \text{span}(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n).$$

Hence, the range of a matrix A (defined as the span of its columns) is equal to range(A) (defined in the sense of A being a function from $\mathbb{R}^n$ to $\mathbb{R}^m$). This sounds like a tautology, but it is actually a precise justification for why we refer to the range of a matrix as the span of its columns.

Having said all of this, consider the system of equations

$$A\vec{x} = \vec{b},$$

where $A \in \mathbb{R}^{m\times n}$. Consistency of this system is succinctly characterized by checking whether $\vec{b}$ is in the range of A. Indeed,

Theorem 3.3: The system of linear equations $A\vec{x} = \vec{b}$ is consistent if and only if $\vec{b} \in \text{range}(A)$.

Proof. The proof is really just an exercise using definitions. To start, let us assume $\vec{b} \in \text{range}(A)$. By definition, this means that

$$\vec{b} \in \text{range}(A) = \text{span}(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) = \{A\vec{x} \mid \vec{x} \in \mathbb{R}^n\}.$$

Hence, there is evidently $\vec{x} \in \mathbb{R}^n$ such that $A\vec{x} = \vec{b}$.

Now, let us assume the system is consistent. This implies existence of a solution $\vec{x}$ such that $A\vec{x} = \vec{b}$. Letting $x_1, \ldots, x_n$ denote the entries of the solution $\vec{x}$, the definition of matrix-vector multiplication implies $\vec{b}$ can be written as

$$\vec{b} = \sum_{i=1}^{n} x_i \vec{a}_i,$$

and the latter is an element of span$(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n)$ by definition. Hence, $\vec{b} \in \text{range}(A)$, completing the proof. $\square$

3.5 Practice Problems

These practice problems are also available in an interactive form on the course website.

1. Are the vectors $\vec{a} = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$ and $\vec{b} = \begin{bmatrix} -1 \\ 2 \end{bmatrix}$ linearly independent?
2. Are the vectors $\vec{a} = \begin{bmatrix} 2 \\ 2 \\ -1 \end{bmatrix}$, $\vec{b} = \begin{bmatrix} 1 \\ 6 \\ 2 \end{bmatrix}$, $\vec{c} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ linearly independent?
3. Suppose for some matrix $\mathbf{A}$, $\mathbf{A}\vec{x}_1 = \vec{b}$ and $\mathbf{A}\vec{x}_2 = \vec{b}$, where $\vec{x}_1 \neq \vec{x}_2$. Are the columns of $\mathbf{A}$ linearly independent?
4. Is $\vec{v} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$ in the *span* $\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 4 \\ -1 \end{bmatrix} \right)$?
5. Which of the following are equivalent to span$\{\vec{v}_1, \vec{v}_2\}$?

(a) $\text{span}\{\vec{v}_1, \vec{v}_1 + \vec{v}_2\}$

(b) $\text{span}\{\vec{v}_1\}$

(c) $\text{span}\{\vec{v}_1, \vec{v}_2 - \alpha\vec{v}_1\}$

(d) (a) and (b)

(e) (a) and (c)