EECS 16A Designing Information Devices and Systems I
Spring 2023 Lecture Notes Note 3

3.1 Linear Dependence

Recall the 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. The Gaussian elimination algorithm implied that we needed to take at least 9 measurements to properly identify the 9 bottles in a box so that we had at least one equation per variable. 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. We will give two equivalent definitions of linear dependence, and formally prove their equivalence in Note 4.

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 $\vec{v}_i = \sum_{j \neq i} \alpha_j \vec{v}_j$.[1] In words, a set of vectors is linearly dependent if any one of the vectors could be written as a linear combination of the rest of the vectors.

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

[1] 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 $\sum_{i=1}^{n} \alpha_i\vec{v}_i$ or $\sum_i \alpha_i\vec{v}_i$, 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 $\alpha_1\vec{v}_1 + \cdots + \alpha_n\vec{v}_n - \alpha_i\vec{v}_i$.

3.1.2 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, $\begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. By definition, $\vec{a}$ and $\vec{b}$ are linearly independent.

> **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 unknown variables, and $\vec{b}$ is a vector of values that these weighted sums must equal. The following two theorems state that just looking at the columns or rows of the matrix A can tell us about the solutions to $A\vec{x} = \vec{b}$ (we will see the formal proofs of both theorems in Note 4).

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.**

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

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, and therefore cannot be disambiguated. Hence, there won't be enough information to determine a unique solution. 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 are 10 units of light. Then we could write down the equation,

$$x_1 + x_2 = 10. \tag{1}$$

Written in matrix form, we have

$$\begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 10 \end{bmatrix}. \tag{2}$$

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. The main idea here is that if we have n unknowns, we need n or more unique and linearly independent measurements (equations) to solve for each unknown.

3.2 Row Perspective

(Note that 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 and is linearly dependent on another measurement (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 closely related concept 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

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

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.

We should also be aware of the following annoying bit of jargon: that when given a matrix A, the *range* and *column space* of A are synonymous both refer to the span of the columns of A!

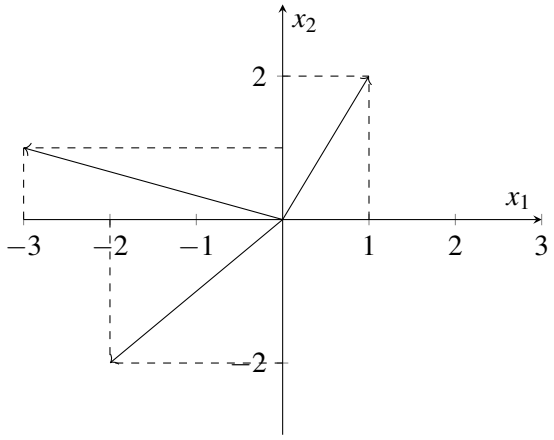
Additional Resources For more on linear span, read *Schaum's* pages 119-121. For additional practice with these ideas, try Problems 4.13 to 4.16, 4.66, 4.69, and 4.83 to 4.88.

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 simplify 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. We see here that there are two nonzero rows, no zero rows equating to a nonzero constant, and three unknowns. Thus, we have a consistent system with fewer equations than unknowns, so there are an *infinite number* of solutions, no matter how we choose a and b. The concept of "free variables" as it relates to span and number of solutions will be discussed later in this course.

In other words, no matter how a and b are chosen, there is not just one, but an *infinite number* of ways to compute a linear combination of our three vectors to reach $\begin{bmatrix} a \\ b \end{bmatrix}$. Thus, we have shown that our three vectors span the entirety of two-dimensional space - expressed mathematically,

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

3.4 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 $\text{span}\{\vec{v}_1, \vec{v}_2\}$?
   (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)