Linear Independence
How do we prove it?

Consider the following if-then statement:

"If A is an $m \times n$ matrix and $A\mathbf{v}_1, A\mathbf{v}_2, \ldots, A\mathbf{v}_p$ are linearly independent, then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ are linearly independent."

This is a true statement, but how do we go about proving it? To figure this out, let's analyze the it a little more closely.

The hypothesis of the implication is the statement

$$\mathcal{A}: \text{"}A \text{ is an } m \times n \text{ matrix and } A\mathbf{v}_1, A\mathbf{v}_2, \ldots, A\mathbf{v}_p \text{ are linearly independent."}$$

and the conclusion is

$$\mathcal{D}: \text{"}\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p \text{ are linearly independent."}$$

As usual, to prove "If $\mathcal{A}$ then $\mathcal{D}$", we start by assuming that $\mathcal{A}$ is true and then deduce that $\mathcal{D}$ is also true. What makes things challenging in this case is the nature of the statement $\mathcal{D}$: it is also an if-then statement! Indeed, the definition of linear independence tells us that $\mathcal{D}$ is really the statement

$$\mathcal{D}: \text{"If } x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0} \text{ then } x_1 = x_2 = \cdots = x_p = 0.\text{"}$$

In other words, the conclusion of our original if-then statement is another if-then statement with its own hypothesis and conclusion! In this case the hypothesis of $\mathcal{D}$ is

$$\mathcal{B}: \text{"}x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}.\text{"}$$

and the conclusion of $\mathcal{D}$ is

$$\mathcal{C}: \text{"}x_1 = x_2 = \cdots = x_p = 0.\text{"}$$

So the statement we are really trying to prove is the "compound" statement

"If $\mathcal{A}$ then (if $\mathcal{B}$ then $\mathcal{C}$)"

or, expressed entirely in symbols,

$$\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C}).$$

How in the world do we prove something like this?[1] Very carefully.

The easiest thing to do is to replace the expression $\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})$ with one that is logically equivalent but easier to understand. It turns out that the statement $\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})$ is logically equivalent to

$$(\mathcal{A} \text{ and } \mathcal{B}) \Rightarrow \mathcal{C}.$$

You can prove this rigorously using a truth table, but it should be intuitively clear that this makes sense. We can read $\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})$ as "Whenever $\mathcal{A}$ is true, then whenever $\mathcal{B}$ is true $\mathcal{C}$ is true." This certainly seems to be the same as the statement "Whenever $\mathcal{A}$ and $\mathcal{B}$ are true, then $\mathcal{C}$ is true," which is how we would read $(\mathcal{A} \text{ and } \mathcal{B}) \Rightarrow \mathcal{C}$.

So now we see that the statement $\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})$ can be proven by assuming $\mathcal{A}$ *and* $\mathcal{B}$ are true and then deducing that $\mathcal{C}$ is true. This is because that is exactly how we would prove $(\mathcal{A} \text{ and } \mathcal{B}) \Rightarrow \mathcal{C}$, which is a logically equivalent statement.

Summary: To prove the statement $\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})$, start by assuming that $\mathcal{A}$ *and* $\mathcal{B}$ are true, and then deduce that $\mathcal{C}$ is true.

Let's see how we would apply this to our original problem. We want to prove that

[1] The statement $\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})$ is *not* the same as the string of implications $\mathcal{A} \Rightarrow \mathcal{B} \Rightarrow \mathcal{C}$ because there is no "associative law" in logic that would allow us to shift parentheses (or omit them). In otherwords, we *are not* trying to prove that $\mathcal{A}$ implies $\mathcal{B}$ and that $\mathcal{B}$ implies $\mathcal{C}$.

"If A is an $m \times n$ matrix and $A\mathbf{v}_1, A\mathbf{v}_2, \ldots, A\mathbf{v}_p$ are linearly independent, then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ are linearly independent."

is a true statement. We showed above that this is of the form $\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})$ where $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are

$$\begin{array}{lcl} \mathcal{A} & : & \text{"}A \text{ is an } m \times n \text{ matrix and } A\mathbf{v}_1, A\mathbf{v}_2, \ldots, A\mathbf{v}_p \text{ are linearly independent."} \\ \mathcal{B} & : & \text{"}x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}.\text{"} \\ \mathcal{C} & : & \text{"}x_1 = x_2 = \cdots = x_p = 0.\text{"} \end{array}$$

So, according to what we discussed above, we start by assuming that A is an $m \times n$ matrix and $A\mathbf{v}_1, A\mathbf{v}_2, \ldots, A\mathbf{v}_p$ are linearly independent, *and* $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}$. We need to conclude that $x_1 = x_2 = \cdots = x_p = 0$. Let's get to it!

Since

$$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}$$

if we multiply both sides by A we get, by linearity of matrix/vector multiplication,

$$A(x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p) = A\mathbf{0}x_1A\mathbf{v}_1 + x_2A\mathbf{v}_2 + \cdots + x_pA\mathbf{v}_p = \mathbf{0}.$$

Linear independence of $A\mathbf{v}_1, A\mathbf{v}_2, \ldots, A\mathbf{v}_p$ tells us that we must have $x_1 = x_2 = \cdots = x_p = 0$. This is what we needed to show.

Most linear independence proofs follow this same pattern. We are usually asked to prove a statement of the form

"If $\mathcal{A}$, then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ are linearly independent."

The same reasoning as above shows that if $\mathcal{B}$ is the statement "$x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}$" and $\mathcal{C}$ is the statement $x_1 = x_2 = \cdots = x_p = 0$, then we are really being asked to prove $\mathcal{A} \Rightarrow (\mathcal{B} \Rightarrow \mathcal{C})$. And now we know how to do this!

Summary: To prove the statement

"If $\mathcal{A}$, then $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ are linearly independent."

start by assuming that $\mathcal{A}$ and $x_1\mathbf{v}_1 + x_2\mathbf{v}_2 + \cdots + x_p\mathbf{v}_p = \mathbf{0}$ are true. Then conclude from this that $x_1 = x_2 = \cdots = x_p = 0$.

Let's finish by applying this technique to another example. In the proof I'm not going to explicitly point out that we're using this method, so be sure you can identify it.

Example. Let V be a finite-dimensional vector space with basis $\mathcal{B}$ and let $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p \in \mathcal{V}$. Show that if the $\mathcal{B}$-coordinate vectors $[\mathbf{v}_1]_{\mathcal{B}}, [\mathbf{v}_2]_{\mathcal{B}}, \ldots, [\mathbf{v}_p]_{\mathcal{B}}$ are linearly independent then so are the original vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$.

Proof. Suppose that $[\mathbf{v}_1]_{\mathcal{B}}, [\mathbf{v}_2]_{\mathcal{B}}, \ldots, [\mathbf{v}_p]_{\mathcal{B}}$ are linearly independent and that

$$c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p = \mathbf{0}. \tag{1}$$

Applying the coordinate map to this expression and using linearity gives

$$\begin{aligned} [c_1\mathbf{v}_1 + c_2\mathbf{v}_2 + \cdots + c_p\mathbf{v}_p]_{\mathcal{B}} &= [\mathbf{0}]_{\mathcal{B}} \\ c_1[\mathbf{v}_1]_{\mathcal{B}} + c_2[\mathbf{v}_2]_{\mathcal{B}} + \cdots + c_p[\mathbf{v}_p]_{\mathcal{B}} &= \mathbf{0}. \end{aligned}$$

Since $[\mathbf{v}_1]_{\mathcal{B}}, [\mathbf{v}_2]_{\mathcal{B}}, \ldots, [\mathbf{v}_p]_{\mathcal{B}}$ are linearly independent, this can only happen if $c_1 = c_2 = \cdots = c_p = 0$. That is, (1) can only hold if $c_1 = c_2 = \cdots = c_p = 0$, which shows that $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_p$ are linearly independent. $\square$