# LINEAR DEPENDENCE, BASES, AND SUBSPACES

MATH 196, SECTION 57 (VIPUL NAIK)

**Corresponding material in the book**: Sections 3.2 and 3.3.

## Executive summary

(1) A *linear relation* between a set of vectors is defined as a linear combination of these vectors that is zero. The *trivial* linear relation refers to the trivial linear combination being zero. A nontrivial linear relation is any linear relation other than the trivial one.
(2) The trivial linear relation exists between any set of vectors.
(3) A set of vectors is termed *linearly dependent* if there exists a nontrivial linear relation between them, and *linearly independent* otherwise.
(4) Any set of vectors containing a linearly dependent subset is also linearly dependent. Any subset of a linearly independent set of vectors is a linearly independent set of vectors.
(5) The following can be said of sets of small size:
   - The empty set (the only possible set of size zero) is considered linearly independent.
   - A set of size one is linearly dependent if the vector is the zero vector, and linearly independent if the vector is a nonzero vector.
   - A set of size two is linearly dependent if either one of the vectors is the zero vector or the two vectors are scalar multiples of each other. It is linearly independent if both vectors are nonzero and they are not scalar multiples of one another.
   - For sets of size three or more, a *necessary* condition for linear independence is that no vector be the zero vector and no two vectors be scalar multiplies of each other. However, this condition is not sufficient, because we also have to be on the lookout for other kinds of linear relations.
(6) Given a nontrivial linear relation between a set of vectors, we can use the linear relation to write one of the vectors (any vector with a nonzero coefficient in the linear relation) as a linear combination of the other vectors.
(7) We can use the above to prune a spanning set as follows: given a set of vectors, if there exists a nontrivial linear relation between the vectors, we can use that to write one vector as a linear combination of the others, and then remove it from the set *without affecting the span*. The vector thus removed is termed a *redundant vector*.
(8) A *basis* for a subspace of $\mathbb{R}^n$ is a linearly independent spanning set for that subspace. Any finite spanning set can be pruned down (by repeatedly identifying linear relations and removing vectors) to reach a basis.
(9) The size of a basis for a subspace of $\mathbb{R}^n$ depends only on the choice of subspace and is *independent* of the choice of basis. This size is termed the *dimension* of the subspace.
(10) Given an ordered list of vectors, we call a vector in the list *redundant* if it is redundant relative to the preceding vectors, i.e., if it is in the span of the preceding vectors, and *irredundant* otherwise. The irredundant vectors in any given list of vectors form a basis for the subspace spanned by those vectors.
(11) Which vectors we identify as redundant and irredundant depends on how the original list was ordered. However, the *number* of irredundant vectors, insofar as it equals the dimension of the span, does not depend on the ordering.
(12) If we write a matrix whose column vectors are a given list of vectors, the linear relations between the vectors correspond to vectors in the kernel of the matrix. Injectivity of the linear transformation given by the matrix is equivalent to linear independence of the vectors.

(13) Redundant vector columns correspond to non-leading variables and irredundant vector columns correspond to leading variables if we think of the matrix as a coefficient matrix. We can row-reduce to find which variables are leading and non-leading, then look at the irredundant vector columns in the *original* matrix.
(14) *Rank-nullity theorem*: The nullity of a linear transformation is defined as the dimension of the kernel. The nullity is the number of non-leading variables. The rank is the number of leading variables. So, the sum of the rank and the nullity is the number of columns in the matrix for the linear transformation, aka the dimension of the domain. See Section 3.7 of the notes for more details.
(15) The problem of finding all the vectors orthogonal to a given set of vectors can be converted to solving a linear system where the rows of the coefficient matrix are the given vectors.

## 1. Linear relation

1.1. **Preliminaries.** Previously, we have seen the concepts of *linear combination*, *span*, and *spanning set*. We also saw the concept of the *trivial* linear combination: this is the linear combination where all the coefficients we use are zero. The trivial linear combination gives rise to the zero vector.

We now move to a disturbing observation: it is possible that a nontrivial linear combination of vectors give rise to the zero vector. For instance, consider the three vectors:

$$\vec{v_1} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \vec{v_2} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \vec{v_3} = \begin{bmatrix} 3 \\ 10 \end{bmatrix}$$

We note that:

$$\vec{v_3} = 3\vec{v_1} + 5\vec{v_2}$$

Thus, we get that:

$$\vec{0} = 3\vec{v_1} + 5\vec{v_2} + (-1)\vec{v_3}$$

In other words, the zero vector arises as a *nontrivial* linear combination of these vectors.

We will now codify and study such situations.

1.2. **Linear relation and nontrivial linear relation.** Suppose $\vec{v_1}, \vec{v_2}, \dots, \vec{v_r}$ are vectors. A *linear relation* between $\vec{v_1}, \vec{v_2}, \dots, \vec{v_r}$ involves a choice of (possibly equal, possibly distinct, and possibly some zero) real numbers $a_1, a_2, \dots, a_r$ such that:

$$a_1\vec{v_1} + a_2\vec{v_2} + \cdots + a_r\vec{v_r} = \vec{0}$$

The linear relation is termed a *trivial* linear relation if *all* of $a_1, a_2, \dots, a_r$ are 0. Note that for *any* collection of vectors, the trivial linear relation between them exists. Thus, the trivial linear relation is not all that interesting, but it is included for completeness' sake.

The more interesting phenomenon is that of a *nontrivial* linear relation. Note here that nontriviality requires that at least one of the coefficients be nonzero, but it does not require that *all* coefficients be nonzero.

The existence of nontrivial linear relations is not a given; there may be sets of vectors with *no* nontrivial linear relation. Let's introduce some terminology first, then explore the meaning of the ideas.

1.3. **Linear dependence and linear independence.** Consider a non-empty set of vectors. We say that the set of vectors is *linearly dependent* if there exists a nontrivial linear relation between the vectors. A non-empty set of vectors is called *linearly independent* if it is not linearly dependent, i.e., there exists *no* nontrivial linear relation between the vectors in the set.

We begin with this observation: If a subset of a set of vectors is linearly dependent, then the whole set of vectors is also linearly dependent. The justification is that a nontrivial linear relation within a subset gives a nontrivial linear relation in the whole set by extending to zero coefficients for the remaining vectors. For instance, suppose $\vec{v_1}$, $\vec{v_2}$, $\vec{v_3}$, and $\vec{v_4}$ are vectors and suppose we have a linear relation between $\vec{v_1}$, $\vec{v_2}$, and $\vec{v_3}$:

$$3\vec{v_1} + 4\vec{v_2} + 6\vec{v_3} = \vec{0}$$

Then, we also have a linear relation between the four vectors $\vec{v_1}$, $\vec{v_2}$, $\vec{v_3}$, and $\vec{v_4}$:

$$3\vec{v_1} + 4\vec{v_2} + 6\vec{v_3} + 0\vec{v_4} = \vec{0}$$

An obvious corollary of this is that any subset of a linearly *independent* set is linearly independent.

1.4. **Sets of size zero.** By convention, the empty set is considered linearly independent.

1.5. **Sets of size one: linear dependence and independence.** A set of size one is:

- linearly *dependent* if the vector is the zero vector
- linearly *independent* if the vector is a nonzero vector

In particular, this means that any set of vectors that contains the zero vector must be linearly dependent.

1.6. **Sets of size two: linear dependence and independence.** A set of size two is:

- linearly *dependent* if either one of the vectors is zero or both vectors are nonzero but the vectors are scalar multiples of each other, i.e., they are in the same line.
- linearly *independent* if both vectors are nonzero and they are not scalar multiples of each other.

Pictorially, this means that the vectors point in different directions.

A corollary is that if we have a set of two or more vectors and two vectors in the set are scalar multiples of each other, then the set of vectors is linearly dependent.

1.7. **Sets of size more than two.** For sets of size more than two, linear relations could be fairly elaborate. For instance, a linear relation involving three vectors may occur even if no individual vector is a multiple of any other. Such a linear relation relies on one vector being "in the plane" of the other two vectors. For instance, if one vector is the average of the other two vectors, that creates a linear relation. Explicitly, consider the case of three vectors:

$$\vec{v_1} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \vec{v_2} = \begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}, \vec{v_3} = \begin{bmatrix} 5 \\ 6 \\ 7 \end{bmatrix}$$

Notice that the middle vector is the average of the first and third vector. Thus, we get the linear relation:

$$\vec{v_2} = \frac{\vec{v_1} + \vec{v_3}}{2}$$

We can rearrange this as a linear relation:

$$\vec{0} = \frac{1}{2}\vec{v_1} - \vec{v_2} + \frac{1}{2}\vec{v_3}$$

Note that this is not the only linear relation possible. Any multiple of this also defines a linear relation, albeit, an *equivalent* linear relation. For instance, we also have the following linear relation:

$$\vec{0} = \vec{v_1} - 2\vec{v_2} + \vec{v_3}$$

1.8. **Rewriting a linear relation as one vector in terms of the others.** Given a *nontrivial* linear relation between vectors, we can rewrite that relation in the form of expressing one vector as a linear combination of the other vectors. Here's the reasoning:

- We can find a vector that is being "used" nontrivially, i.e., the coefficient in front of that vector is nonzero.
- We can move that vector to the other side of the equality.
- Divide both sides by its coefficient.

For instance, consider the linear relation:

$$3\vec{v_1} + 7\vec{v_2} + 0\vec{v_3} + 9\vec{v_4} = \vec{0}$$

Note that the coefficient on $\vec{v_3}$ is 0. So, we cannot use this linear relation to write $\vec{v_3}$ in terms of the other vectors. However, we can write $\vec{v_1}$ in terms of the other vectors, or we can write $\vec{v_2}$ in terms of the other vectors, or we can write $\vec{v_4}$ in terms of the other vectors. Let's take the example of $\vec{v_2}$.

We have:

$$3\vec{v_1} + 7\vec{v_2} + 9\vec{v_4} = \vec{0}$$

We can isolate the vector $\vec{v_2}$:

$$3\vec{v_1} + 9\vec{v_4} = -7\vec{v_2}$$

We can now divide both sides by $-7$ to get:

$$\frac{-3}{7}\vec{v_1} + \frac{-9}{7}\vec{v_4} = \vec{v_2}$$

or, written the other way around:

$$\vec{v_2} = \frac{-3}{7}\vec{v_1} + \frac{-9}{7}\vec{v_4}$$

## 2. Span, basis, and redundancy

2.1. **Span and basis: definitions.** Suppose $V$ is a vector subspace of $\mathbb{R}^n$. Let's recall what this means: $V$ contains the zero vector, and it is closed under addition and scalar multiplication.

Recall that a subset $S$ of $V$ is termed a *spanning set* for $V$ if the span of $S$ is $V$, i.e., if every vector in $V$, and no vector outside $V$, can be expressed as a linear combination of the vectors in $S$.

A *basis* for $V$ is a spanning set for $V$ that is linearly independent. Note that any linearly independent set is a basis for the subspace that it spans.

2.2. **Pruning our spanning set.** As before, suppose $V$ is a vector subspace of $\mathbb{R}^n$ and $S$ is a spanning set for $V$. Suppose that $S$ is not a basis for $V$, because $S$ is not linearly independent. In other words, there is at least one nontrivial linear relation between the elements of $S$.

Pick one nontrivial linear relation between the elements of $S$. As described in an earlier section, we can use this relation to write one vector as a linear combination of the others. Once we have achieved this, we can "remove" this vector without affecting the span, because it is *redundant* relative to the other vectors. In other words, removing a redundant vector (a vector in $S$ that is a linear combination of the other vectors in $S$) does not affect the span of $S$. This is because for any vector in the span of $S$ that can be expressed as a linear combination using the redundant vector, the redundant vector can be replaced by the linear combination of the other vectors that it is.

Explicitly, suppose we have a relation of the form:

$$\vec{v_1} + 2\vec{v_2} - \vec{v_3} = \vec{0}$$

We use this to write $\vec{v_3}$ in terms of $\vec{v_1}$ and $\vec{v_2}$:

$$\vec{v_3} = \vec{v_1} + 2\vec{v_2}$$

Now, consider an arbitrary vector $\vec{v}$ expressible in terms of these:

$$\vec{v} = a_1\vec{v_1} + a_2\vec{v_2} + a_3\vec{v_3}$$

Using the expression above, replace $\vec{v_3}$ by $\vec{v_1} + 2\vec{v_2}$ to get:

$$\vec{v} = a_1\vec{v_1} + a_2\vec{v_2} + a_3(\vec{v_1} + 2\vec{v_2})$$

This simplifies to:

$$\vec{v} = (a_1 + a_3)\vec{v_1} + (a_2 + 2a_3)\vec{v_2}$$

In other words, getting rid of $\vec{v_3}$ doesn't affect the span: if something can be written as a linear combination using $\vec{v_3}$, it can also be written as a linear combination without using $\vec{v_3}$. So, we can get rid of $\vec{v_3}$.

2.3. **We can get rid of vectors only one at a time!** In the example above, we noted that it is possible to get rid of the vector $\vec{v_3}$ based on the linear relation that nontrivially uses $\vec{v_1}, \vec{v_2}, \vec{v_3}$. Thus, $\vec{v_3}$ is redundant relative to $\vec{v_1}$ and $\vec{v_2}$, so we can remove $\vec{v_3}$ from our spanning set.

We could have similarly argued that $\vec{v_2}$ is redundant relative to $\vec{v_1}$ and $\vec{v_3}$, and therefore, that $\vec{v_2}$ can be removed while preserving the span.

We could also have similarly argued that $\vec{v_1}$ is redundant relative to $\vec{v_2}$ and $\vec{v_3}$, and therefore, that $\vec{v_1}$ can be removed while preserving the span.

In other words, we could remove *any* of the vectors $\vec{v_1}, \vec{v_2}, \vec{v_3}$ that are involved in a nontrivial linear relation.

However, we *cannot* remove them together. The reason is that once one of the vectors is removed, that destroys the linear relation as well, so the other two vectors are no longer redundant based on this particular linear relation (they may still be redundant due to other linear relations). In a sense, every time we use a linear relation to remove one redundant vector, we have "used up" the linear relation and it cannot be used to remove any other vectors.

This suggests something: it is not so much an issue of *which* vectors are redundant, but rather a question of *how many*. At the core is the idea of dimension as a size measure. We now turn to that idea.

2.4. **Repeated pruning, and getting down to a basis.** As before, suppose $V$ is a vector subspace of $\mathbb{R}^n$ and $S$ is a *finite* spanning set for $V$. Our goal is to find a subset of $S$ that is a basis for $V$.

If $S$ is already linearly independent, that implies in particular that it is a basis for $V$, and we are done.

If $S$ is *not* already linearly independent, there exists a nontrivial linear relation in $S$. Then, by the method discussed in the preceding section, we can get rid of one element of $S$ and get a smaller subset that still spans $V$.

If this new subset is linearly independent, then we have a basis. Otherwise, repeat the process: find a nontrivial linear relation within this smaller spanning set, and use that to get rid of another vector.

The starting set $S$ was finite, so we can perform the process only finitely many times. Thus, after a finite number of steps, we will get to a subset of $S$ that is linearly independent, and hence a basis for $V$.

In the coming section, we will discuss various computational approaches to this pruning process. Understanding the process conceptually, however, is important for a number of reasons that shall become clear later.

2.5. **Basis and dimension.** Recall that we had defined the *dimension* of a vector space as the minimum possible size of a spanning set for the vector space.

The following are equivalent for a subset $S$ of $\mathbb{R}^n$:

- $S$ is a linearly independent set.
- $S$ is a basis for the subspace that it spans.
- The size of $S$ equals the minimum possible size of a spanning set for the span of $S$.
- The size of $S$ equals the dimension of the span of $S$.

Now, given a subspace $V$ of $\mathbb{R}^n$, there are many different possibilities we can choose for a basis of $V$. For instance, if $V$ has a basis comprising the vectors $\vec{v_1}$ and $\vec{v_2}$, we could choose another basis comprising the vectors $\vec{v_1}$ and $\vec{v_1} + \vec{v_2}$. Even for one-dimensional spaces, we have many different choices for a basis of size one: any nonzero vector in the space will do.

Although there are many different possibilities for the basis, the *size* of the basis is an invariant of the subspace, namely, it is the dimension. The specific vectors used can differ, but the number needed is determined.

The concept of dimension can be understood in other related ways. For instance, the dimension is the number of independent parameters we need in a parametric description of the space. The natural parameterization of the subspace is by specifying a basis and using the coefficients for an arbitrary linear combination as the parameters. For instance, if $\vec{v}_1$ and $\vec{v}_2$ form a basis for a subspace $V$ of $\mathbb{R}^n$, then any vector $\vec{v} \in V$ can be written as:

$$\vec{v} = a_1\vec{v}_1 + a_2\vec{v}_2, \qquad a_1, a_2 \in \mathbb{R}$$

We can think of the coefficients $a_1, a_2$ (which we will later call the *coordinates*, but that's for next time) as the parameters in a parametric description of $V$. Different choices of these give different vectors in $V$, and as we consider all the different possible choices, we cover everything in $V$.

*Note*: We have *not* proved all parts of the statement above. Specifically, it is not *prima facie* clear why every basis should have the minimum possible size. In other words, we have not ruled out the *prima facie* possibility that there is a basis of size two and also a basis of size three. The abstract proof that any two bases must have the same size follows from a result called the "exchange lemma" that essentially involves a gradual replacement of the vectors in one basis by the vectors in the other. The proof uses the same sort of reasoning as our pruning idea. There are also other concrete proofs that rely on facts you have already seen about linear transformations and matrices.

Another way of framing this is that the dimension is something *intrinsic* to the subspace rather than dependent on how we parameterize it. It is an intrinsic geometric invariant of the subspace, having to do with the innards of the underlying linear structure.

## 3. Finding a basis based on a spanning set

3.1. **Redundant vectors in order.** The method above gives an *abstract* way of concluding that any spanning set can be trimmed down to a basis. The version stated above, however, is not a *practical* approach. The problem is that we don't yet know how to find a nontrivial linear relation. Or at least, we know it ... but not consciously. Let's make it conscious.

First, let's introduce a new, more computationally relevant notion of the redundancy of a vector. Consider an *ordered* list of vectors. In other words, we are given the vectors in a particular order. A vector in this list is termed *redundant* if it is redundant relative to the vectors that appear *before* it. Intuitively, we can think of it as follows: we are looking at our vectors one by one, reading from left to right along our list. Each time, we throw in the new vector, and *potentially* expand the span. In fact, one of these two cases occurs:

- The vector is redundant relative to the set of the preceding vectors, and therefore, it contributes nothing new to the span. Therefore, we do not actually need to add it in.
- The vector is irredundant relative to the set of the preceding vectors, and therefore, it adds a new dimension (literally and figuratively) to the span.

If we can, by inspection, determine whether a given vector is redundant relative to the vectors that appear before it, we can use that to determine the span. Basically, each time we encouter a redundant vector, we don't add it.

Thus, the sub-list comprising those vectors that are irredundant in the original ordered list gives a basis for the span of the original list.

For instance, suppose we have a sequence of vectors:

$$\vec{v_1}, \vec{v_2}, \vec{v_3}, \vec{v_4}, \vec{v_5}, \vec{v_6}, \vec{v_7}$$

Let's say that $\vec{v_1}$ is the zero vector. Then, it is redundant, so we don't add in. Let's say $\vec{v_2}$ is nonzero. So it is irredundant relative to what's appeared before (which is nothing), so we have so far built:

$$\vec{v_2}$$

Now, let's say $\vec{v_3}$ is a scalar multiple of $\vec{v_2}$. In that case, $\vec{v_3}$ is redundant and will not be added. Let's say $\vec{v_4}$ is again a scalar multiple of $\vec{v_2}$. Then, $\vec{v_4}$ is also redundant, and should not be added. Suppose now that $\vec{v_5}$ is not a scalar multiple of $\vec{v_2}$. Then, $\vec{v_5}$ is irredundant relative to the vectors that have appeared so far, so it deserves to be added:

$$\vec{v_2}, \vec{v_5}$$

We now consider the sixth vector $\vec{v_6}$. Suppose it is expressible as a linear combination of $\vec{v_2}$ and $\vec{v_5}$. Then, it is redundant, and should not be included. Now, let's say $\vec{v_7}$ is not a linear combination of $\vec{v_2}$ and $\vec{v_5}$. Then, $\vec{v_7}$ is irredundant relative to the preceding vectors, so we get:

$$\vec{v_2}, \vec{v_5}, \vec{v_7}$$

This forms a basis for the span of the original list of seven vectors. Thus, the original list of seven vectors spans a three-dimensional space, with the above as one possible basis.

Note that which vectors we identity as redundant and which vectors we identity as irredundant depends heavily on the manner in which we sequence our vectors originally. Consider the alternative arrangement of the original sequence:

$$\vec{v_4}, \vec{v_1}, \vec{v_3}, \vec{v_2}, \vec{v_7}, \vec{v_5}, \vec{v_6}$$

The irredundant vectors here are:

$$\vec{v_4}, \vec{v_7}, \vec{v_5}$$

Note that, because we ordered our original list differently, the *list* of irredundant vectors differs, so we get a different basis. But the *number* of irredundant vectors, i.e., the *size* of the basis, is the same. After all, this is the *dimension* of the space, and as such, is a geometric invariant of the space.

3.2. **The matrix and linear transformation formulation.** The problem we want to explicitly solve is the following:

> Given a collection of $m$ vectors in $\mathbb{R}^n$, find which of the vectors are redundant and which are irredundant, and use the irredundant vectors to construct a basis for the spanning set for that collection of vectors.

Consider the $n \times m$ matrix $A$ for a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$. We know that the columns of $A$ are the images of the standard basis vectors under $T$, and thus, the columns of $A$ form a spanning set for the image of $T$.

The problem that we are trying to solve is therefore equivalent to the following problem:

> Given a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$ with matrix $A$, consider the columns of $A$, which coincide with the images of the standard basis vectors. Find the irredundant vectors there, and use those to get a basis for the image of $T$.

3.3. **Linear relations form the kernel.** We make the following observation regarding linear relations:

> Linear relations between the column vectors of a matrix $A$ correspond to vectors in the kernel of the linear transformation given by $A$.

Let's understand this. Suppose $A$ is the matrix for a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$, so that $A$ is a $n \times m$ matrix. The columns of $A$ are the vectors $T(\vec{e_1})$, $T(\vec{e_2})$, $\dots$, $T(\vec{e_m})$. These also form a spanning set for the image of $T$.

Now, suppose there is a linear relation between the vectors, namely a relation of the form:

$$x_1 T(\vec{e_1}) + x_2 T(\vec{e_2}) + \cdots + x_m T(\vec{e_m}) = \vec{0}$$

Then, this is equivalent to saying that:

$$T(x_1 \vec{e_1} + x_2 \vec{e_2} + \cdots + x_m \vec{e_m}) = \vec{0}$$

or equivalently:

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix}\right) = \vec{0}$$

In other words, the vector:

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{bmatrix}$$

is in the kernel of $T$.

All these steps can be done in reverse, i.e., if a vector $\vec{x}$ is in the kernel of $T$, then its coordinates define a linear relation between $T(\vec{e_1})$, $T(\vec{e_2})$, $\dots$, $T(\vec{e_m})$.

3.4. **Special case of injective linear transformations.** Consider a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$ with matrix $A$, which is a $n \times m$ matrix. The following are equivalent:

- $T$ is an injective linear transformation.
- If you think of solving the linear system with coefficient matrix $A$, all variables are leading variables.
- $A$ has full column rank $m$.
- The kernel of $T$ is zero-dimensional, i.e., it comprises only the zero vector.
- The images of the standard basis vectors are linearly independent.
- The images of the standard basis vectors form a basis for the image.
- The image of $T$ is $m$-dimensional.

In particular, all these imply that $m \leq n$.

3.5. **Back to finding the irredundant vectors.** Recall that when we perform row reductions on the coefficient matrix of a linear system, we *do not change the solution set*. This is exactly why we can use row reduction to solve systems of linear equations, and hence, also find the kernel.

In particular, this means that when we row reduce a matrix, we *do not change the pattern of linear relations between the vectors*. This means that the information about which columns are redundant and which columns are irredundant does not change upon row reduction.

For a matrix in reduced row-echelon form, the columns corresponding to the leading variables are irredundant and the columns corresponding to the non-leading variables are redundant. The leading variable columns are irredundant relative to the preceding columns because each leading variable columns uses a new row for the first time. The non-leading variable columns are redundant because they use only the existing rows for which standard basis vectors already exist in preceding leading variables. Consider, for instance:

$$\begin{bmatrix} 1 & 2 & 7 & 0 & 4 \\ 0 & 0 & 0 & 1 & 6 \end{bmatrix}$$

The first and fourth variable here are leading variables. The second, third, and fifth variable are non-leading variables. The column vectors are:

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \begin{bmatrix} 7 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 4 \\ 6 \end{bmatrix}$$

- The first column vector corresponds to a leading variable and is irredundant. In fact, it is the first standard basis vector.
- The second column vector corresponds to a non-leading variable and is redundant: Note that it does not use any new rows. It uses only the first row, for which we already have the standard basis vector. In fact,the second column vector is explicitly twice the first column vector.
- The third column vector corresponds to a non-leading varable and is redundant: The reasoning is similar to that for the second column vector. Explicitly, this third column vector is 7 times the first column vector.
- The fourth column vector corresponds to a leading variable and is irredundant: Note that it is the first vector to use the second row. Hence, it is not redundant relative to the preceding vectors.
- The fifth column vector corresponds to a non-leading variable and is redundant: It uses both rows, but we already have standard basis vectors for both rows from earlier. Hence, it is a linear combination

of those. Explicitly, it is 4 times the first column vector plus 6 times the fourth column vector. Thus, it is redundant.

The detailed example above hopefully illustrates quite clearly the *general* statement that the column vectors corresponding to leading variables are irredundant whereas the column vectors corresponding to non-leading variables are redundant. Note that all this is being said for a matrix that is already in reduced row-echelon form. But we already noted that the linear relations between the columns are invariant under row reductions. So whatever we conclude after converting to rref about which columns are redundant and irredundant *also* applies to the original matrix.

Thus, the following algorithm works:

- Convert the matrix to reduced row-echelon form. Actually, it suffices to convert the matrix to row-echelon form because all we really need to do is identify which variables are leading variables and which variables are non-leading variables.
- The columns in the *original* matrix corresponding to the leading variables are the irredundant vectors, and form a basis for the image. Please note that the actual column vectors we use are the column vectors of the original matrix, not of the rref.

3.6. **Procedural note regarding the kernel.** We had earlier seen a procedure to find a spanning set for the kernel of a linear transformation. It turns out that the spanning set obtained that way, providing one vector for each non-leading variable, is actually linearly independent, and hence, gives a basis for the kernel. The dimension of the kernel is thus equal to the number of non-leading variables, or equivalently, equals the total number of columns minus the rank.

3.7. **Rank and nullity.** We define the *nullity* of a linear transformation $T : \mathbb{R}^m \to \mathbb{R}^n$ as the dimension of the kernel of $T$. We will return a while later to the concept of nullity in more gory detail. For now, we state a few simple facts about rank and nullity that will hopefully clarify much of what will come later.

Suppose $A$ is the matrix of $T$, so $A$ is a $n \times m$ matrix. The following are true:

- The nullity of $A$ is the dimension of the kernel of $T$.
- The rank of $A$ is the dimension of the image of $T$.
- The sum of the rank of $A$ and the nullity of $A$ is $m$.
- The nullity of $A$ is at most $m$.
- The rank of $A$ is at most $\min\{m, n\}$.
- The nullity of $A$ is 0 (or equivalently, the rank of $A$ is $m$, so full column rank) if and only if $T$ is injective. See the preceding section on injective transformations for more on this. Note that this forces $m \leq n$.
- The rank of $A$ is $n$ (so full row rank) if and only if $T$ is surjective.

3.8. **Finding all the vectors orthogonal to a given set of a vectors.** Suppose we are given a bunch of vectors in $\mathbb{R}^n$. We want to find all the vectors in $\mathbb{R}^n$ whose dot product with any vector in this collection is zero. This process is relatively straightforward: set up a matrix whose *rows* are the given vectors, and find the kernel of the linear transformation given by that matrix.