First topic: Direct sums and quotient spaces

We start with a homework problem:

9. Let W_1 and W_2 be subspaces of a vector space V such that $W_1 + W_2 = V$ and $W_1 \cap W_2 = \{0\}$. Prove that for each vector α in V there are *unique* vectors α_1 in W_1 and α_2 in W_2 such that $\alpha = \alpha_1 + \alpha_2$.

In a bit more detail: Let V be a vector space and let X and Y be subspaces. Show that the following are equivalent:

1. Every vector in V can be written in exactly one way as $\vec{x} + \vec{y}$ for $\vec{x} \in X$ and $\vec{y} \in Y$.
2. Every vector in V can be written as $\vec{x} + \vec{y}$ for $\vec{x} \in X$ and $\vec{y} \in Y$, and $X \cap Y = \{0\}$.

We start with a homework problem: Let V be a vector space and let X and Y be subspaces. Show that the following are equivalent:

1. Every vector in V can be written in exactly one way as $\vec{x} + \vec{y}$ for $\vec{x} \in X$ and $\vec{y} \in Y$.
2. Every vector in V can be written as $\vec{x} + \vec{y}$ for $\vec{x} \in X$ and $\vec{y} \in Y$, and $X \cap Y = \{0\}$.

Proof: In either case, we are assuming that every vector can be written as $\vec{x} + \vec{y}$.

(1) $\implies$ (2): If $\vec{u}$ is in $X \cap Y$, then $\vec{x} + \vec{y} = (\vec{x} + \vec{u}) + (\vec{y} - \vec{u})$. This would give multiple formulas for the same vector unless $\vec{u} = \vec{0}$.

(2) $\implies$ (1): Suppose, to the contrary, that $\vec{x}_1 + \vec{y}_1 = \vec{x}_2 + \vec{y}_2$. Then $\vec{x}_1 - \vec{x}_2 = \vec{y}_2 - \vec{y}_1$, so assumption (2) tells us that $\vec{x}_1 - \vec{x}_2 = \vec{y}_2 - \vec{y}_1 = 0$, and we have $\vec{x}_1 = \vec{x}_2$ and $\vec{y}_1 = \vec{y}_2$. $\square$

In this case, we'll say that $V = X \oplus Y$.

For example, $\mathbb{R}^3 = \{\begin{bmatrix} x \\ y \\ z \end{bmatrix} : x + y + z = 0\} \oplus \{\begin{bmatrix} t \\ t \\ t \end{bmatrix}\}$.

In this case, we'll say that $V = X \oplus Y$.

For example, $\mathbb{R}^3 = \{\begin{bmatrix} x \\ y \\ z \end{bmatrix} : x + y + z = 0\} \oplus \{\begin{bmatrix} t \\ t \\ t \end{bmatrix}\}$.

In finite dimensional vector spaces over $\mathbb{R}$, we always have $\mathbb{R}^n = L \oplus L^\perp$, but this isn't our focus right now.

In this case, we'll say that $V = X \oplus Y$.

For example, $\mathbb{R}^3 = \{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} : x + y + z = 0\} \oplus \{ \begin{bmatrix} t \\ t \\ t \end{bmatrix} \}$.

In finite dimensional vector spaces over $\mathbb{R}$, we always have $\mathbb{R}^n = L \oplus L^\perp$, but this isn't our focus right now.

If $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m$ is a basis of X, and $\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_n$ is a basis of Y, then $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m, \vec{y}_1, \vec{y}_2, \ldots, \vec{y}_n$ is a basis of V. In particular, $\dim V = \dim X + \dim Y$.

In this case, we'll say that $V = X \oplus Y$.

For example, $\mathbb{R}^3 = \{\begin{bmatrix} x \\ y \\ z \end{bmatrix} : x + y + z = 0\} \oplus \{\begin{bmatrix} t \\ t \\ t \end{bmatrix}\}$.

In finite dimensional vector spaces over $\mathbb{R}$, we always have $\mathbb{R}^n = L \oplus L^\perp$, but this isn't our focus right now.

If $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m$ is a basis of X, and $\vec{y}_1, \vec{y}_2, \ldots, \vec{y}_n$ is a basis of Y, then $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_m, \vec{y}_1, \vec{y}_2, \ldots, \vec{y}_n$ is a basis of V. In particular, $\dim V = \dim X + \dim Y$.

So, when we write vectors in the coordinates of this basis, the X-entries come first and then the Y-entries. Similarly, if $V_1 = X_1 \oplus Y_1$ and $V_2 = X_2 \oplus Y_2$, then linear transformations $V_1 \to V_2$ are given by block matrices.

$$\left[\begin{array}{c|c} X_1 \to X_2 & Y_1 \to X_2 \\ \hline X_1 \to Y_2 & Y_1 \to Y_2 \end{array}\right].$$

There is another, related but distinct, construction, called direct sum. Today, I'll write $\boxplus$ for that other construction, but people normally call them both $\oplus$, and I'll start doing that soon.

There is another, related but distinct, construction, called direct sum. Today, I'll write $\boxplus$ for that other construction, but people normally call them both $\oplus$, and I'll start doing that soon.

Let X and Y be two vector spaces over the same field F. We define the vector space $X \boxplus Y$ as follows:

- The elements of $X \boxplus Y$ are ordered pairs $(\vec{x}, \vec{y})$ with $\vec{x} \in X$ and $\vec{y} \in Y$.
- Addition is defined as $(\vec{x}_1, \vec{y}_1) + (\vec{x}_2, \vec{y}_2) = (\vec{x}_1 + \vec{x}_2, \vec{y}_1 + \vec{y}_2)$.
- Scalar multiplication is defined as $c(\vec{x}_1, \vec{y}_1) = (c\vec{x}_1, c\vec{y}_1)$.

There is another, related but distinct, construction, called direct sum. Today, I'll write $\boxplus$ for that other construction, but people normally call them both $\oplus$, and I'll start doing that soon.

Let X and Y be two vector spaces over the same field F. We define the vector space $X \boxplus Y$ as follows:

- The elements of $X \boxplus Y$ are ordered pairs $(\vec{x}, \vec{y})$ with $\vec{x} \in X$ and $\vec{y} \in Y$.
- Addition is defined as $(\vec{x}_1, \vec{y}_1) + (\vec{x}_2, \vec{y}_2) = (\vec{x}_1 + \vec{x}_2, \vec{y}_1 + \vec{y}_2)$.
- Scalar multiplication is defined as $c(\vec{x}_1, \vec{y}_1) = (c\vec{x}_1, c\vec{y}_1)$.

So, if X and Y are both subspaces of V and $V = X \oplus Y$, then $X \boxplus Y$ is isomorphic to $X \oplus Y$, by $(\vec{x}, \vec{y}) \mapsto \vec{x} + \vec{y}$. But we are allowed to talk about $X \boxplus Y$ without starting with a subspace that X and Y are both contained in.

Finally, quotient spaces. Let V be a vector space and let X be a subspace. Then there are many different subspaces Y with $V = X \oplus Y$.

Finally, quotient spaces. Let V be a vector space and let X be a subspace. Then there are many different subspaces Y with $V = X \oplus Y$.

Is there some natural way to talk about "the part of V which isn't X"?

Finally, quotient spaces. Let V be a vector space and let X be a subspace. Then there are many different subspaces Y with $V = X \oplus Y$.

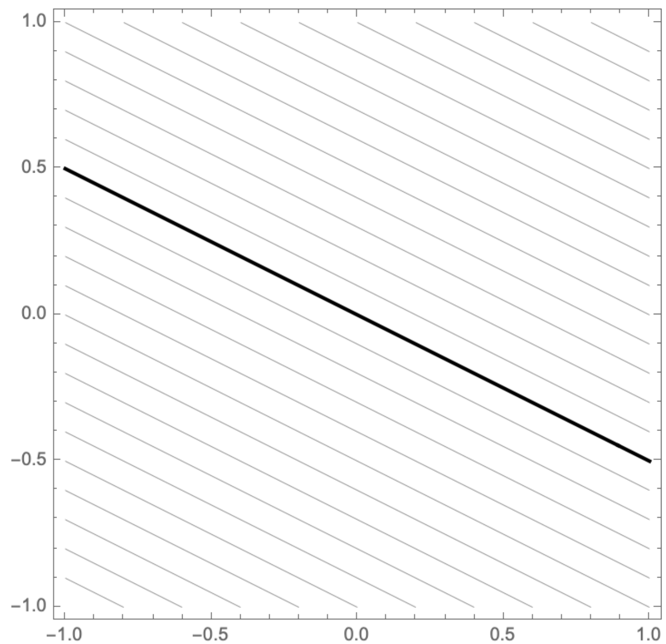
Is there some natural way to talk about "the part of V which isn't X"?

One answer, for finite dimensional real vector spaces, is $X^{\perp} = \{\vec{y} : \langle \vec{y}, \vec{x} \rangle = 0 \text{ for all } \vec{x} \in X\}$. But that isn't the answer I am talking about today.

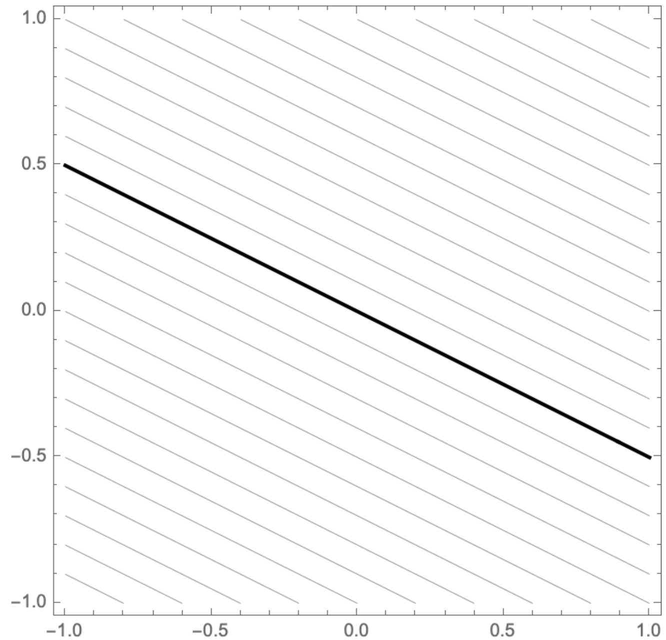
Finally, quotient spaces. Let V be a vector space and let X be a subspace. Then there are many different subspaces Y with $V = X \oplus Y$.

Is there some natural way to talk about "the part of V which isn't X"?

Define $\vec{v}_1 \equiv \vec{v}_2 \bmod X$ if $\vec{v}_1 - \vec{v}_2 \in X$.



Define $\vec{v}_1 \equiv \vec{v}_2 \mod X$ if $\vec{v}_1 - \vec{v}_2 \in X$.

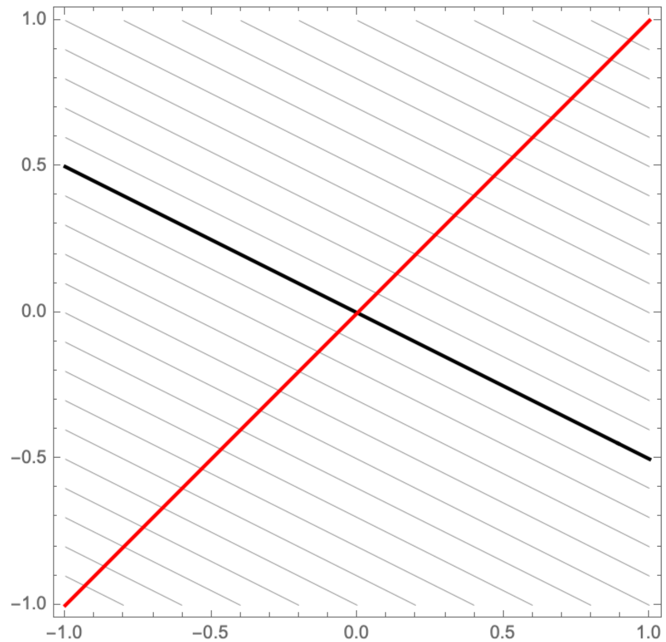


Check that

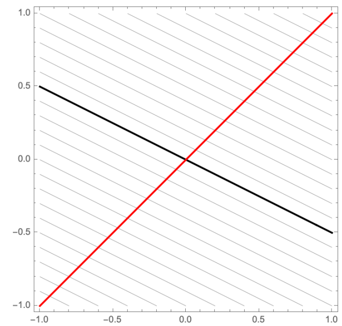
- If $\vec{v}_1 \equiv \vec{v}_2$ and $\vec{w}_1 \equiv \vec{w}_2$ then $\vec{v}_1 + \vec{w}_1 \equiv \vec{v}_2 + \vec{w}_2$.
- If $\vec{v}_1 \equiv \vec{v}_2$ and c is a scalar then $c\vec{v}_1 \equiv c\vec{v}_2$.

The elements of V/X are the equivalence classes for V/X, with addition and scalar multiplication defined as above.

If $V = X \oplus Y$, then $Y \longrightarrow V \longrightarrow V/X$ is an isomorphism.

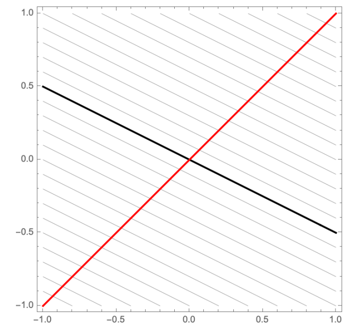


If $V = X \oplus Y$, then $Y \longrightarrow V \longrightarrow V/X$ is an isomorphism.



Proof: Since $Y \to V$ and $V \to V/X$ are linear maps, the composite map is linear; we need to check that it is bijective. In other words, we need to check that, for each equivalence class $\vec{z} + X$, there is exactly one $\vec{y} \in Y$ in $\vec{z} + X$.

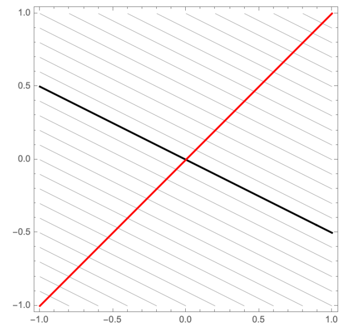
If $V = X \oplus Y$, then $Y \longrightarrow V \longrightarrow V/X$ is an isomorphism.



Proof: Since $Y \to V$ and $V \to V/X$ are linear maps, the composite map is linear; we need to check that it is bijective. In other words, we need to check that, for each equivalence class $\vec{z} + X$, there is exactly one $\vec{y} \in Y$ in $\vec{z} + X$.

Saying that $\vec{y}$ is in $\vec{z} + X$ is the same as saying that there is some $\vec{x} \in X$ with $\vec{y} = \vec{z} + \vec{x}$ or, in other words, $\vec{z} = -\vec{x} + \vec{y}$.

If $V = X \oplus Y$, then $Y \longrightarrow V \longrightarrow V/X$ is an isomorphism.



Proof: Since $Y \to V$ and $V \to V/X$ are linear maps, the composite map is linear; we need to check that it is bijective. In other words, we need to check that, for each equivalence class $\vec{z} + X$, there is exactly one $\vec{y} \in Y$ in $\vec{z} + X$.

Saying that $\vec{y}$ is in $\vec{z} + X$ is the same as saying that there is some $\vec{x} \in X$ with $\vec{y} = \vec{z} + \vec{x}$ or, in other words, $\vec{z} = -\vec{x} + \vec{y}$. Since $V = X \oplus Y$, there is exactly one way to write $\vec{z}$ as $\vec{x}' + \vec{y}$, and then we take $\vec{x} = -\vec{x}'$. $\square$

If $T : V \to W$ is a linear map, then $V/\mathrm{Ker}(T)$ is isomorphic to $\mathrm{Image}(T)$.

If $T : V \to W$ is a linear map, then $V/\mathrm{Ker}(T)$ is isomorphic to $\mathrm{Image}(T)$.

Proof: Notice that, if $\vec{v}_1 - \vec{v}_2$ is in $\mathrm{Ker}(T)$, then $T(\vec{v}_1) = T(\vec{v}_2)$, so we get a well defined function from $V/\mathrm{Ker}(T)$ to W.

If $T : V \to W$ is a linear map, then $V/\mathrm{Ker}(T)$ is isomorphic to $\mathrm{Image}(T)$.

Proof: Notice that, if $\vec{v}_1 - \vec{v}_2$ is in $\mathrm{Ker}(T)$, then $T(\vec{v}_1) = T(\vec{v}_2)$, so we get a well defined function from $V/\mathrm{Ker}(T)$ to W. Its image is $\mathrm{Image}(T)$.

If $T : V \to W$ is a linear map, then $V/\mathrm{Ker}(T)$ is isomorphic to $\mathrm{Image}(T)$.

Proof: Notice that, if $\vec{v}_1 - \vec{v}_2$ is in $\mathrm{Ker}(T)$, then $T(\vec{v}_1) = T(\vec{v}_2)$, so we get a well defined function from $V/\mathrm{Ker}(T)$ to W. Its image is $\mathrm{Image}(T)$. It is injective: If $T(\vec{v}_1) = T(\vec{v}_2)$ then $\vec{v}_1 - \vec{v}_2$ is in $\mathrm{Ker}(T)$, so $\vec{v}_1$ and $\vec{v}_2$ are the same point of $V/\mathrm{Ker}(T)$. And it is easy to check that this map is linear.

If $T : V \to W$ is a linear map, then $V/\text{Ker}(T)$ is isomorphic to $\text{Image}(T)$.

Proof: Notice that, if $\vec{v}_1 - \vec{v}_2$ is in $\text{Ker}(T)$, then $T(\vec{v}_1) = T(\vec{v}_2)$, so we get a well defined function from $V/\text{Ker}(T)$ to W. Its image is $\text{Image}(T)$. It is injective: If $T(\vec{v}_1) = T(\vec{v}_2)$ then $\vec{v}_1 - \vec{v}_2$ is in $\text{Ker}(T)$, so $\vec{v}_1$ and $\vec{v}_2$ are the same point of $V/\text{Ker}(T)$. And it is easy to check that this map is linear.

So we have a bijective linear map, and thus an isomorphism. $\square$

Topic Two: The rank nullity theorem, from two perspectives

The rank nullity theorem Let $T : V \to W$ be a linear transformation, with $\dim V < \infty$. Then

$$\dim V = \dim \operatorname{Ker}(T) + \dim \operatorname{Image}(T).$$

The rank nullity theorem Let $T : V \to W$ be a linear transformation, with $\dim V < \infty$. Then

$$\dim V = \dim \mathrm{Ker}(T) + \dim \mathrm{Image}(T).$$

First proof (matrices) Let's assume W is also finite dimensional, and identify V with F^m and W with F^n. So T is given by an $n \times m$ matrix A. Then

$$\begin{aligned} \dim V &= \text{number of columns of } A \\ \dim \mathrm{Ker}(T) &= \text{number of free columns of } \mathrm{rref}(A) \\ \dim \mathrm{Image}(T) &= \text{number of pivot columns of } \mathrm{rref}(A) \end{aligned}$$

$$\#(\text{columns}) = \#(\text{free columns}) + \#(\text{pivot columns}). \quad \square$$

The rank nullity theorem Let $T : V \to W$ be a linear transformation, with $\dim V < \infty$. Then

$$\dim V = \dim \operatorname{Ker}(T) + \dim \operatorname{Image}(T).$$

Second proof (abstract vector spaces) Choose a basis $\vec{v}_1$, $\vec{v}_2$, $\ldots$, $\vec{v}_k$ for $\operatorname{Ker}(T)$.

The rank nullity theorem Let $T : V \to W$ be a linear transformation, with $\dim V < \infty$. Then

$$\dim V = \dim \mathrm{Ker}(T) + \dim \mathrm{Image}(T).$$

Second proof (abstract vector spaces) Choose a basis $\vec{v}_1$, $\vec{v}_2$, $\ldots$, $\vec{v}_k$ for $\mathrm{Ker}(T)$. Complete it to a basis $\vec{v}_1$, $\vec{v}_2$, $\ldots$, $\vec{v}_k$, $\vec{v}_{k+1}$, $\vec{v}_{k+2}$, $\ldots$, $\vec{v}_m$. (See Homework Problem 4.)

The rank nullity theorem Let $T : V \to W$ be a linear transformation, with $\dim V < \infty$. Then

$$\dim V = \dim \mathrm{Ker}(T) + \dim \mathrm{Image}(T).$$

Second proof (abstract vector spaces) Choose a basis $\vec{v}_1$, $\vec{v}_2$, $\ldots$, $\vec{v}_k$ for $\mathrm{Ker}(T)$. Complete it to a basis $\vec{v}_1$, $\vec{v}_2$, $\ldots$, $\vec{v}_k$, $\vec{v}_{k+1}$, $\vec{v}_{k+2}$, $\ldots$, $\vec{v}_m$. (See Homework Problem 4.)

The rank nullity theorem Let $T : V \to W$ be a linear transformation, with $\dim V < \infty$. Then

$$\dim V = \dim \operatorname{Ker}(T) + \dim \operatorname{Image}(T).$$

Second proof (abstract vector spaces) Choose a basis $\vec{v}_1$, $\vec{v}_2$, $\ldots$, $\vec{v}_k$ for $\operatorname{Ker}(T)$. Complete it to a basis $\vec{v}_1$, $\vec{v}_2$, $\ldots$, $\vec{v}_k$, $\vec{v}_{k+1}$, $\vec{v}_{k+2}$, $\ldots$, $\vec{v}_m$. (See Homework Problem 4.)

Then the equivalence classes of $\vec{v}_{k+1}$, $\vec{v}_{k+2}$, $\ldots$, $\vec{v}_n$ form a basis of $V/\operatorname{Ker}(T)$. Or, without talking about quotients, $T(\vec{v}_{k+1})$, $T(\vec{v}_{k+2})$, $\ldots$, $T(\vec{v}_m)$ is a basis of $\operatorname{Image}(T)$.

In class, we proved the second way of framing this: $T(\vec{v}_{k+1})$, $T(\vec{v}_{k+2})$, $\ldots$, $T(\vec{v}_m)$ is a basis of Image(T).

Linear independence: Suppose that we have a linear relation $c_{k+1}T(\vec{v}_{k+1}) + \cdots + c_m T(\vec{v}_m) = \vec{0}$. Then $T(c_{k+1}\vec{v}_{k+1} + \cdots + c_m\vec{v}_m) = \vec{0}$, so $c_{k+1}\vec{v}_{k+1} + \cdots + c_m\vec{v}_m$ is in Ker(T). Since $\vec{v}_1$, $\vec{v}_2$, $\ldots$, $\vec{v}_k$ is a basis of Ker(T), we can write $c_{k+1}\vec{v}_{k+1} + \cdots + c_m\vec{v}_m = a_1\vec{v}_1 + \cdots + a_k\vec{v}_k$. Since the $\vec{v}_i$ are linearly independent, we have $a_1 = a_2 = \cdots = a_k = c_{k+1} = \cdots = c_m = 0$ and, in particular, $c_{k+1} = \cdots = c_m = 0$.

Spanning: Consider any vector $\vec{w}$ in Image(T). Then there is some $\vec{v} \in V$ with $T(\vec{v}) = \vec{w}$. Since the $\vec{v}_i$ are a basis of V, we can write $\vec{V} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k + c_{k+1}\vec{v}_{k+1} + \cdots + c_m\vec{v}_m$. But then, applying T, we have

$$\vec{w} = T(\vec{v}) = c_1T(\vec{v}_1) + \cdots + c_kT(\vec{v}_k) + c_{k+1}T(\vec{v}_{k+1}) + \cdots + c_mT(\vec{v}_m)$$

$$= \vec{0} + \cdot + \vec{0} + c_{k+1}T(\vec{v}_{k+1}) + \cdots + c_mT(\vec{v}_m) = c_{k+1}T(\vec{v}_{k+1}) + \cdots + c_mT(\vec{v}_m).$$

The rank nullity theorem Let $T : V \to W$ be a linear transformation, with $\dim V < \infty$. Then

$$\dim V = \dim \mathrm{Ker}(T) + \dim \mathrm{Image}(T).$$

Second proof (abstract vector spaces) Choose a basis $\vec{v}_1$, $\vec{v}_2$, $\ldots$, $\vec{v}_k$ for $\mathrm{Ker}(T)$. Complete it to a basis $\vec{v}_1$, $\vec{v}_2$, $\ldots$, $\vec{v}_k$, $\vec{v}_{k+1}$, $\vec{v}_{k+2}$, $\ldots$, $\vec{v}_m$. (See Homework Problem 4.)

Then the equivalence classes of $\vec{v}_{k+1}$, $\vec{v}_{k+2}$, $\ldots$, $\vec{v}_n$ form a basis of $V/\mathrm{Ker}(T)$. Or, without talking about quotients, $T(\vec{v}_{k+1})$, $T(\vec{v}_{k+2})$, $\ldots$, $T(\vec{v}_m)$ is a basis of $\mathrm{Image}(T)$.

So $\dim \mathrm{Ker}(T) = k$, $\dim \mathrm{Image}(T) = m - k$ and $\dim V = m$. $\square$

The rank nullity theorem Let $T : V \to W$ be a linear transformation, with $\dim V < \infty$. Then

$$\dim V = \dim \operatorname{Ker}(T) + \dim \operatorname{Image}(T).$$

What did we gain from this abstraction?

The rank nullity theorem Let $T : V \to W$ be a linear transformation, with $\dim V < \infty$. Then

$$\dim V = \dim \operatorname{Ker}(T) + \dim \operatorname{Image}(T).$$

What did we gain from this abstraction?

- Unimportant bonus: We didn't assume that $\dim W < \infty$, or even that W had a basis.

The rank nullity theorem Let $T : V \to W$ be a linear transformation, with $\dim V < \infty$. Then

$$\dim V = \dim \operatorname{Ker}(T) + \dim \operatorname{Image}(T).$$

What did we gain from this abstraction?

- Unimportant bonus: We didn't assume that $\dim W < \infty$, or even that W had a basis.
- Structural result: We get a basis of $\operatorname{Image}(T)$ by completing a basis of $\operatorname{Ker}(T)$.

The rank nullity theorem Let $T : V \to W$ be a linear transformation, with $\dim V < \infty$. Then

$$\dim V = \dim \operatorname{Ker}(T) + \dim \operatorname{Image}(T).$$

What did we gain from this abstraction?

- Unimportant bonus: We didn't assume that $\dim W < \infty$, or even that W had a basis.
- Structural result: We get a basis of $\operatorname{Image}(T)$ by completing a basis of $\operatorname{Ker}(T)$.
- Perhaps, some intuition?