Lecture 6:

Recap: • Defined matrix multiplication $\longleftrightarrow$ Composition of linear transformations

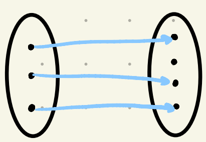
• Defined invertibility, gave an algorithm for finding A^{-1}.

Today: injectivity, surjectivity, kernel, image, subspaces, dimension (bases).

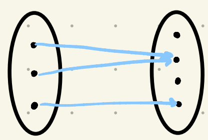
Injectivity and surjectivity

"Not losing information"

Definition 1: A function between sets $f: X \to Y$ is injective iff $f(x) = f(y)$ implies $x = y$ for any $x, y \in X$.

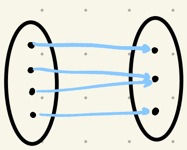


injective

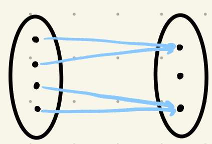


not injective.

Definition 2: A function between sets $f: X \to Y$ is surjective iff for all $y \in Y$, there exists $x \in X$ s.t. $f(x) = y$.



surjective



not surjective

Definition 3: A function between sets $f: X \to Y$ is bijective iff it is injective and surjective.

Remark: Let f be a bijective function. Then we can define an inverse $f^{-1}: Y \to X$ as follows: take $y \in Y$. Since f is surjective, there is some $x \in X$ mapping to y. Since f is injective, there is in fact only one. So set $f^{-1}(y) = x$.

Recall Example 7 from Lecture 5: the linear transformation T with matrix $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ did not have an inverse. We can see what fails:

- T is not injective: $f\begin{pmatrix} 0 \\ a \end{pmatrix} = f\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ for all $a \in \mathbb{R}$, yet $\begin{pmatrix} 0 \\ a \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$.
- T is not surjective: $f\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}$, so f never attains $\begin{pmatrix} a \\ b \end{pmatrix}$ for $b \neq 0$.

Definition 4: The image of a linear tranformation $T: \mathbb{R}^n \to \mathbb{R}^m$ is the set $\mathrm{Im}(T) = \{T(v) : v \in \mathbb{R}^n\}$.

Remark: the image "measures the surjectivity" of T: the closer $\mathrm{Im}\,T$ is to $\mathbb{R}^m$, the closer it is to being surjective.

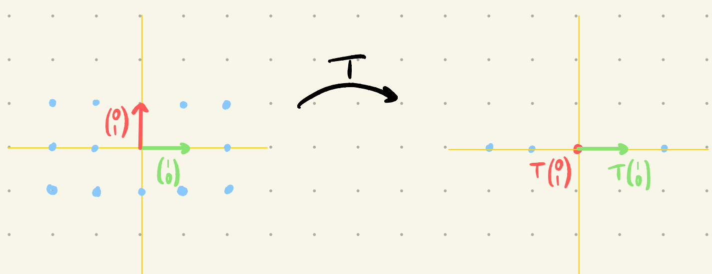
Computing the image of a linear tranformation would be hard if it wasn't for the following theorem.

<u>Theorem 1</u>: The image of a linear transformation is the span of $T\left(\begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}\right), \dots, T\left(\begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}\right)$.

Example continued: the image of $T \leftrightarrow \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ is

$$\left\{\begin{pmatrix} a \\ 0 \end{pmatrix} : a \in \mathbb{R}\right\} = \mathrm{Span}\left(T\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$$
$$= \mathrm{Span}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix}\right)$$
$$= \mathrm{Span}\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right).$$

Geometrically,



$\Rightarrow$ Image is $\mathrm{Span}\left(T\begin{pmatrix} 1 \\ 0 \end{pmatrix}\right)$.

<u>Proof of Thm 1</u>: We want to show $\mathrm{Im}\,T = \mathrm{Span}(T(v_1), \dots, T(v_n))$, so we prove $\mathrm{Im}\,T \subseteq \mathrm{Span}(T(v_1), \dots, T(v_n))$ and $\mathrm{Im}\,T \supseteq \mathrm{Span}(T(v_1), \dots, T(v_n))$

$\subseteq$) Write $e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$. Let $y \in \mathrm{Im}(T)$. Then by definition, $y = T(x)$ for some $x \in \mathbb{R}^n$.

Next, write $x = \lambda_1 e_1 + \dots + \lambda_n e_n$. Then $y = T(x) = T(\lambda_1 e_1 + \dots + \lambda_n e_n) \underset{T \text{ linear}}{=} \lambda_1 T(e_1) + \dots + \lambda_n T(e_n)$, so y is in the span of $T(e_1), \dots, T(e_n)$.

2) let $w = \lambda_1 T(v_1) + \dots + \lambda_n T(v_n) \in \mathrm{Span}(T(v_1), \dots, T(v_n))$, for some choices of $\lambda_1, \dots, \lambda_n \in \mathbb{R}$. Then,

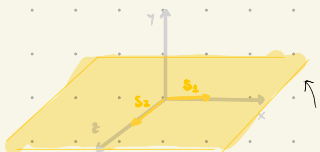
$$w \underset{T \text{ linear}}{=} T(\lambda_1 v_1 + \dots + \lambda_n v_n) \in \mathrm{Im}\, T, \text{ as desired.}$$

The subset $\mathrm{Im}\, T \subseteq \mathbb{R}^m$ is an example of a subspace:

Definition 5: A (linear) subspace S of $\mathbb{R}^n$ is a subset $S \subseteq \mathbb{R}^n$ such that: for all $s_1, s_2 \in S$ and $\lambda_1, \lambda_2 \in \mathbb{R}$, $\lambda_1 s_1 + \lambda_2 s_2 \in S$.

Example 2: The span of a bunch of vectors is by definition a subspace. Take for instance the span of $\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ in $\mathbb{R}^3$. This is $S = \left\{ \begin{pmatrix} a \\ 0 \\ b \end{pmatrix} : a, b \in \mathbb{R} \right\}$. To see that this is a subspace, take $s_1 = \begin{pmatrix} a_1 \\ 0 \\ b_1 \end{pmatrix}$, $s_2 = \begin{pmatrix} a_2 \\ 0 \\ b_2 \end{pmatrix}$, $\lambda_1, \lambda_2 \in \mathbb{R}$. Then $\lambda_1 s_1 + \lambda_2 s_2 = \begin{pmatrix} \lambda_1 a_1 + \lambda_2 a_2 \\ 0 \\ \lambda_1 b_1 + \lambda_2 a_2 \end{pmatrix} \in S$.

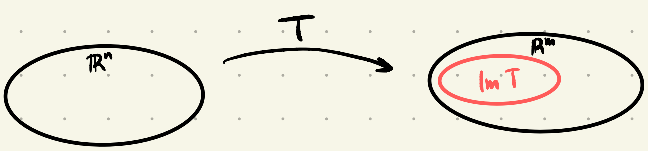
Geometrically, this is just the xz-plane in $\mathbb{R}^3$:



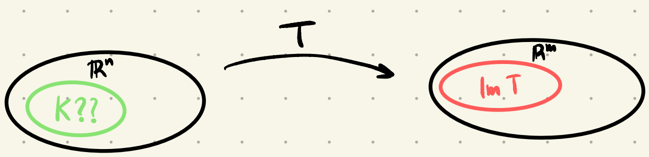
xz-plane

Remark: Theorem 1 says: $\mathrm{Im}\, T$ is a linear subspace of $\mathbb{R}^m$.

Discussion: We have a subspace "measuring" the surjectivity of T:



Idea: it would be nice to have a subspace $K \subseteq \mathbb{R}^n$ measuring the injectivity of T as well.



That subspace will be the kernel of T.

Definition 6: let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then the kernel of T is

$$\mathrm{Ker}(T) = \{ v \in \mathbb{R}^n : T(v) = 0 \}.$$

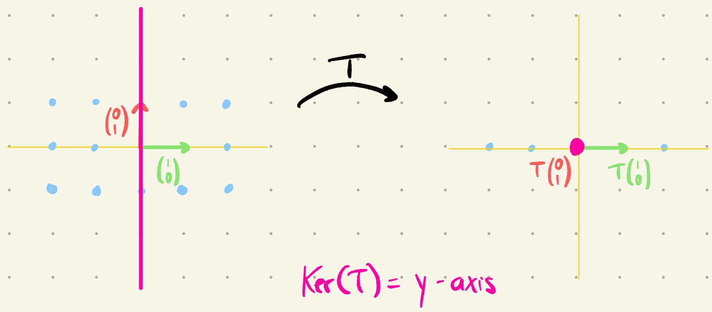
Theorem 2: Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then $\mathrm{Ker}(T)$ is a subspace of $\mathbb{R}^n$.

Proof: In-class exercise.

Question: how does this measure injectivity? Let's go back to our example: $T: \mathbb{R}^2 \to \mathbb{R}^2$ given by $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$. We have

$$\begin{aligned} \mathrm{Ker}(T) &= \left\{ \begin{pmatrix} a \\ b \end{pmatrix} : T\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} a \\ b \end{pmatrix} : \begin{pmatrix} a \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\} \\ &= \left\{ \begin{pmatrix} a \\ b \end{pmatrix} : a = 0 \right\}. \\ &= \left\{ \begin{pmatrix} 0 \\ b \end{pmatrix} : b \in \mathbb{R} \right\} \end{aligned}$$

Geometrically,



Notice that $\mathrm{Ker}(T)$ formalizes the idea that T "collapses a one dimensional subspace".

The larger $\mathrm{Ker}(T)$ is, the more T "collapses" and the less injective T is.

Question: could it be that T is not injective but does not collapse any subspace?

The following theorem says the answer is no.

Theorem 3: Suppose $T: \mathbb{R}^n \to \mathbb{R}^m$ is not injective. Then $\mathrm{Ker}(T) \neq \{0\}$.

Proof: If T is not injective, there exist distinct vectors $v, w \in \mathbb{R}^n$ such that $T(v) \neq T(w)$.

But then $T(v) - T(w) \neq 0 \Rightarrow T(v-w) = 0$, so $v - w \in \mathrm{Ker}(T)$. Finally note that $v - w \neq 0$ since v and w are distinct.

To summarize:

- The larger $\mathrm{Im}(T)$, the more surjective T is.
- The larger $\mathrm{Ker}(T)$, the less injective T is.

In-class exercises:

1. Find the image and kernel of the transformations associated to the following matrices:

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \end{pmatrix}, \quad \begin{pmatrix} 0 & 2 & -1 \\ 1 & -1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

2. Prove Theorem 2.

3. Find a linear transformation T with $\operatorname{Ker}(T) = \operatorname{Span}\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\}$, $\operatorname{Im}(T) = \operatorname{Span}\left\{\begin{pmatrix} 3 \\ 4 \end{pmatrix}\right\}$.