Lecture 9: Linear transformations

Reading: 

2.3  4.3–4.4

Today: Linear transformations $\mathbb{R}^n \to \mathbb{R}^m$

Goal: Understand the geometry of linear maps.

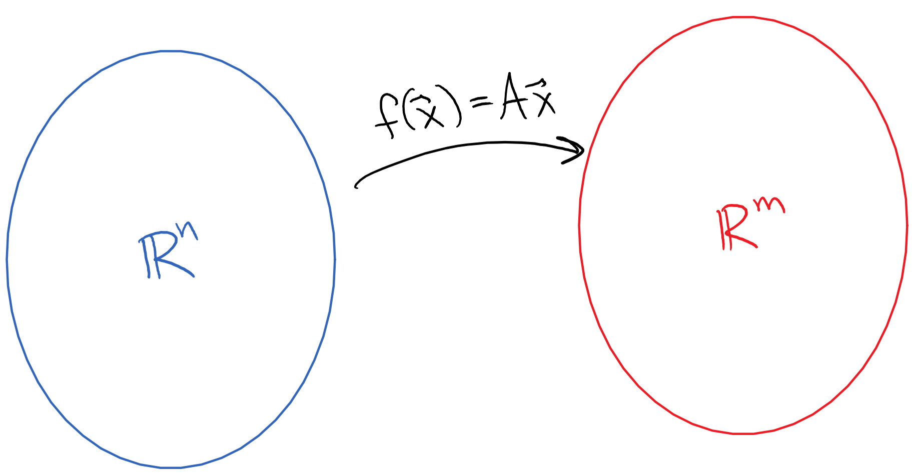
Recall:

Definition: For an $m \times n$ matrix A,

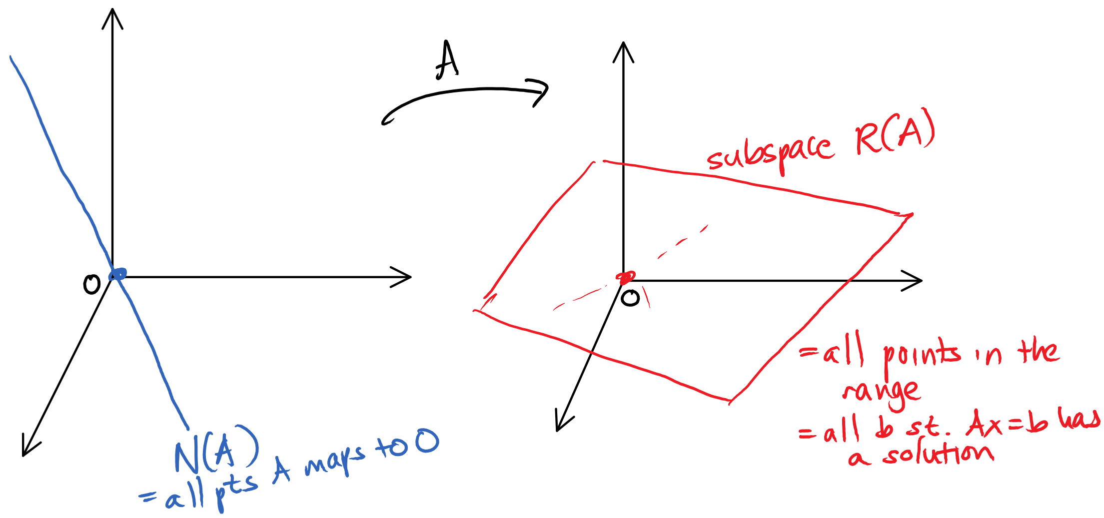
- The nullspace of A is
$$N(A) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = 0\}.$$
- The range (or column space) of A is
$$R(A) = \{A\vec{x} \mid \vec{x} \in \mathbb{R}^n\}.$$
$$= \text{Span(columns of } A)$$
- The rowspace of A is $R(A^T)$

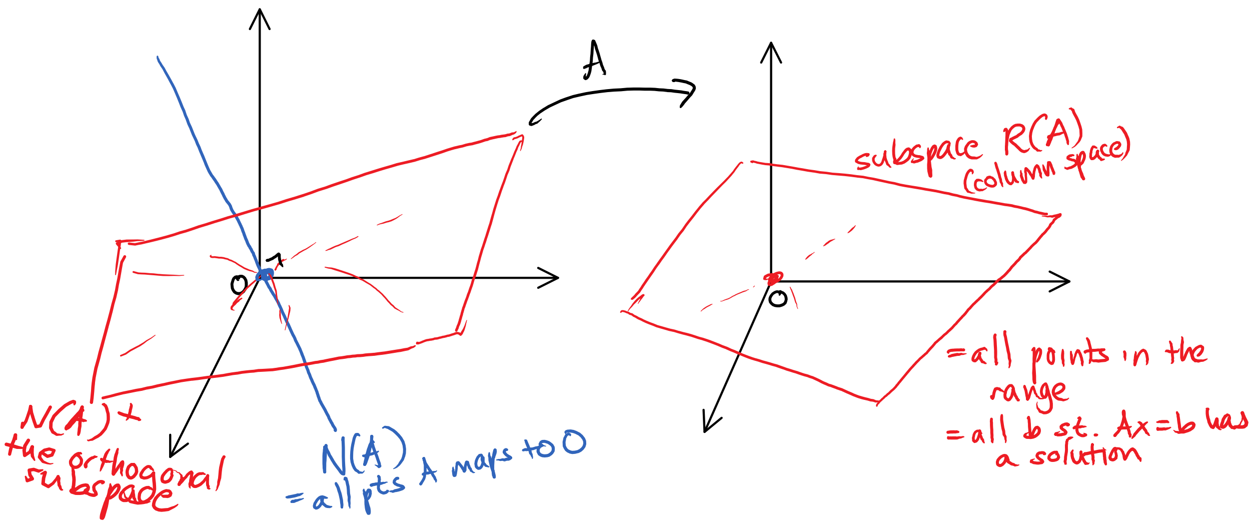
The geometry of matrices

Let A be an $m \times n$ real matrix.



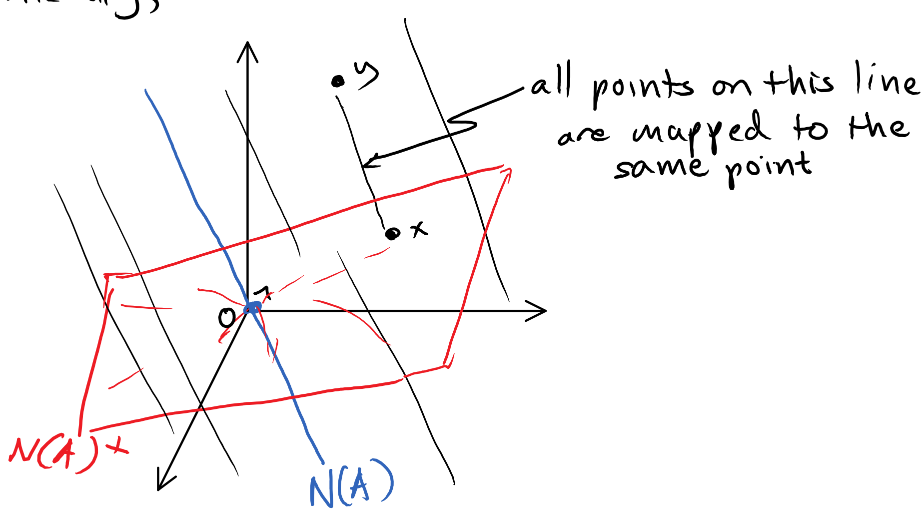
Let's refine this picture...





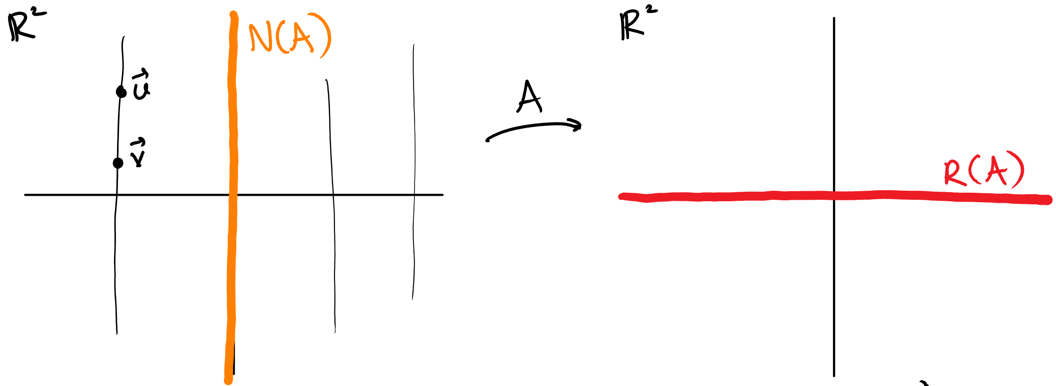
Observations

① If $\vec{x} - \vec{y} \in N(A)$, then $A\vec{x} = A\vec{y}$.
Graphically,



Example: $m = n = 2$
$A = \begin{pmatrix} 1 & 0 \end{pmatrix}$ $\qquad A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ 0 \end{pmatrix}$

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \qquad A(y) = (0)$$



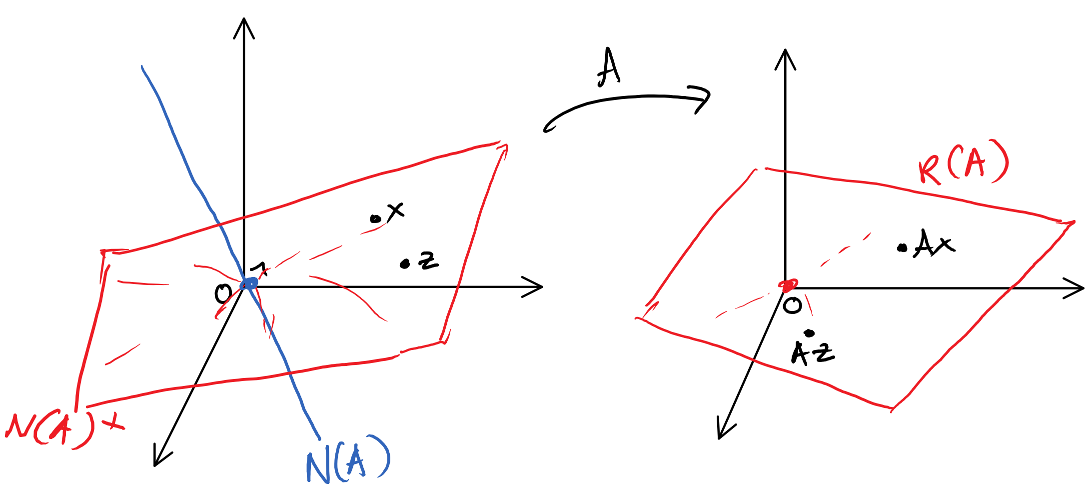
The y-coordinate (direction parallel to N(A)) is irrelevant.
All that matters is the x-coordinate (perpendicular to N(A)).
That's also true in general!

In our heads, we can thus break A into two steps:

1) First map y to x,
   i.e., take a point and move it parallel to N(A) to get to $N(A)^\perp$.
   This is a **projection**; it flattens the space to $N(A)^\perp$.

2) Map $\vec{x}$ to $A\vec{x}$.

② If $\vec{x}, \vec{z}$ are distinct points in $N(A)^\perp$, then $A\vec{x} \neq A\vec{z}$.
(because if $Ax = Az$, then $x - z \in N(A)$.)



Thus A is a 1-to-1 map $N(A)^\perp \longrightarrow R(A)$. It is invertible between these spaces.

By the Rank-Nullity theorem,

$$N(A)^{\perp} = R(A^T) \leftarrow \text{rowspace of } A$$

and $\dim R(A^T) = \dim R(A) \leftarrow$ called the **rank** of A.

SINGULAR-VALUE DECOMPOSITION (SVD)

Informally: Any linear transformation can be split into:
- a rotation, followed by
- scaling vectors in or out

LINEAR TRANSFORMATIONS

Why matrices???

why matrix multiplication?
Why matrix inversion?

Definition: A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is **linear** if

$$f(\alpha \vec{u}) = \alpha \cdot f(\vec{u}) \quad \text{for all vectors } \vec{u} \in \mathbb{R}^n \text{ and scalars } \alpha \in \mathbb{R}$$

$$f(\vec{u}+\vec{v}) = f(\vec{u}) + f(\vec{v}) \quad \text{for all } \vec{u}, \vec{v} \in \mathbb{R}^n$$

Examples: For $n=m=2$,
- $f(x,y) = (0,0)$ ✓
- f a rotation by θ: $f(x,y) = \begin{pmatrix} (\cos\theta)x - (\sin\theta)y, \\ (\sin\theta)x + (\cos\theta)y \end{pmatrix}$ ✓
- $f(x,y) = (x^2, \sin y)$ ✗ not linear
- $f(x,y) = (1+x, y)$ ✗ not linear

One more example: For a polynomial p,

eg., $5x^2 + 3x + 2$,

let

$$f(p) = (2+3x)\cdot p.$$

f maps polynomials in x to polynomials in x

$\leftarrow$ a vector space

and it is a linear transformation!

We'll see lots more examples later (e.g., differentiation, ...).

$$\boxed{\begin{array}{c}\text{LINEAR TRANSFORMATIONS}\\ \updownarrow \\ \text{MATRICES}\end{array}}$$

We'll see the correspondence today for maps $\mathbb{R}^n \to \mathbb{R}^m$, and generalize it to arbitrary vector spaces next week.

Theorem 1: Let A be an $m \times n$ matrix. Define $f: \mathbb{R}^n \to \mathbb{R}^m$ by

$$f(\vec{v}) = A\vec{v}.$$

Then f is a linear transformation.

Proof: $f(\alpha u) = A \cdot (\alpha u) = \alpha \cdot (Au) = \alpha f(u)$ ✓

$f(u+v) = A \cdot (u+v) = Au + Av = f(u) + f(v)$. ✓ □

Theorem 2: Let $f: \mathbb{R}^n \to \mathbb{R}^m$ be any linear transformation. Then there exists an $m \times n$ matrix A such that

$$f(\vec{v}) = A\vec{v}.$$

Proof: Let $\vec{e}_j = (0, 0, \ldots, \underset{\substack{\uparrow \\ j\text{th coord}}}{1}, 0, \ldots, 0) \in \mathbb{R}^n$.

Let

$$A = \begin{pmatrix} | & | & & | \\ f(\vec{e}_1) & f(\vec{e}_2) & \cdots & f(\vec{e}_n) \\ | & | & & | \end{pmatrix}$$; its jth column is $f(\vec{e}_j)$.

We claim that $f(\vec{u}) = A\vec{u}$ for any $\vec{u} \in \mathbb{R}^n$. Indeed,

$$\vec{u} = (u_1, u_2, u_3, \ldots, u_n) = u_1\vec{e}_1 + u_2\vec{e}_2 + \cdots + u_n\vec{e}_n$$

Then

$$A\vec{u} = u_1 f(\vec{e}_1) + \cdots + u_n f(\vec{e}_n),$$

while, by applying the linearity property repeatedly

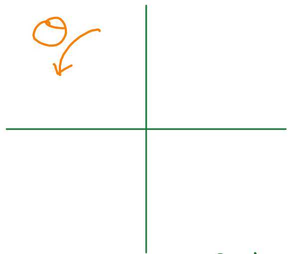
$$f(\vec{u}) = u_1 f(\vec{e}_1) + \cdots + u_n f(\vec{e}_n).$$

Hence indeed $f(\vec{u}) = A\vec{u}$. ✓ $\square$

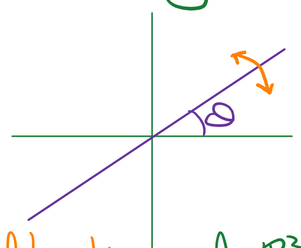
Note: This proof is important! It constructs the matrix A.

Examples: Give the matrices for these linear transformations:

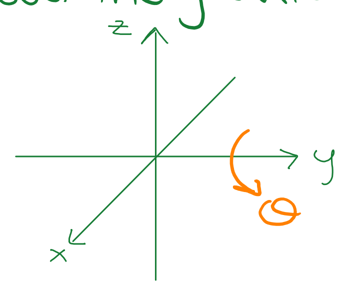
a) Rotation of $\mathbb{R}^2$ by θ:



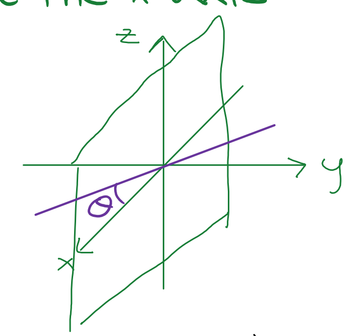
b) Reflection of $\mathbb{R}^2$ about the line at angle θ:



c) Rotation in $\mathbb{R}^3$ by θ about the y-axis:



d) Reflection of $\mathbb{R}^3$ about the line in the xz-plane at angle θ above the x-axis:



These theorems are why matrix-vector multiplication is defined the way it is.

What about matrix-matrix multiplication?

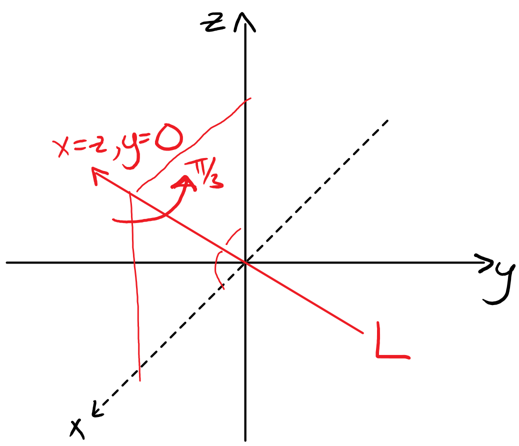
$$\boxed{\text{MATRIX MULTIPLICATION} \longleftrightarrow \text{LINEAR FUNCTION COMPOSITION}}$$

$$\mathbb{R}^m \xrightarrow{f} \mathbb{R}^n \xrightarrow{g} \mathbb{R}^p$$

$f(u) = Au \qquad g(v) = Bv$

$$g(f(u)) = B \cdot A \cdot u$$

Exercise: Let $f: \mathbb{R}^3 \to \mathbb{R}^3$ be the linear transformation that rotates by $\pi/3$ radians about the line $\{x=z, y=0\}$. What is the 3×3 matrix representing f?



Answer: In three steps:
1. Rotate L to the x-axis
2. Rotate by $\pi/3$ about x-axis
3. Go back.

Exercise: Constructing linear transformations.

a) Give a matrix mapping $(1,2,3) \mapsto (4,5,6)$.

Answer:

There are many possible answers, eg.,

$$\begin{pmatrix} 4 & 0 & 0 \\ 5 & 0 & 0 \\ 6 & 0 & 0 \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} 4 & 0 & 0 \\ 3 & 1 & 0 \\ 3 & 0 & 1 \end{pmatrix}.$$

We can find all solutions: $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$.

$$a_{11} + 2a_{12} + 3a_{13} = 4$$
$$a_{21} + 2a_{22} + 3a_{23} = 5$$
$$a_{31} + \underbrace{2a_{32} + 3a_{33}}_{\text{these are all free variables.}} = 6$$

b) Give a matrix mapping

$$(1,2,3) \mapsto (4,5,6)$$
$$(1,1,1) \mapsto (1,0,0)$$
$$(2,3,4) \mapsto (0,1,0)$$

Answer:

This is impossible!

$$\begin{pmatrix}2\\3\\4\end{pmatrix}=\begin{pmatrix}1\\2\\3\end{pmatrix}+\begin{pmatrix}1\\1\\1\end{pmatrix} \quad \text{but} \quad \begin{pmatrix}1\\0\end{pmatrix}\neq\begin{pmatrix}5\\6\end{pmatrix}+\begin{pmatrix}0\\0\end{pmatrix}.$$

c) Give a matrix mapping

$$(1,2,3)\longmapsto(4,5,6)$$
$$(1,1,1)\longmapsto(1,0,0)$$
$$(1,3,2)\longmapsto(1,2,3)$$

Answer 1:

This can be solved like in ⓐ, but now with 9 equations for the 9 unknowns.

Answer 2:

Let

$$B=\begin{pmatrix}1&1&1\\2&1&3\\3&1&2\end{pmatrix} \qquad C=\begin{pmatrix}4&1&1\\5&0&2\\6&0&3\end{pmatrix}.$$

Then B maps

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

C maps

$$\begin{pmatrix}1\\0\\0\end{pmatrix}\longmapsto\begin{pmatrix}4\\5\\6\end{pmatrix}, \quad \begin{pmatrix}0\\1\\0\end{pmatrix}\longmapsto\begin{pmatrix}1\\0\\0\end{pmatrix}, \quad \begin{pmatrix}0\\0\\1\end{pmatrix}\longmapsto\begin{pmatrix}1\\2\\3\end{pmatrix}$$

So $C\cdot B^{-1}$ does what we want:

```
>> B = [1 1 1; 2 1 3; 3 1 2]; C = [4 1 1; 5 0 2; 6 0 3];
>> A = C * inv(B)

A =

    0.0000   -1.0000    2.0000
   -2.3333   -0.3333    2.6667
   -3.0000   -0.0000    3.0000
```

```
>> A * [1 2 3]'

ans =

     4
     5
     6

>> A * [1 1 1]'

ans =

    1.0000
    0.0000
    0.0000

>> A * [1 3 2]'

ans =

     1
     2
     3
```

Let $f:\mathbb{R}^3\longrightarrow\mathbb{R}^3$ be the linear transformation that

- First rotates the xz-plane by $\pi/3$ radians counterclockwise about the y-axis,

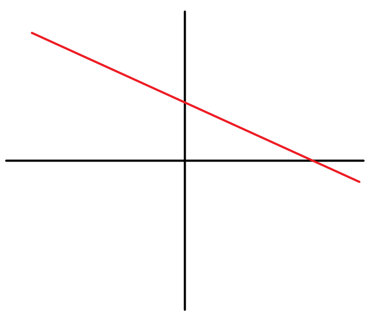
- Then reflects everything about the yz-plane (i.e., switching the sign of the x coordinate),
- Then rotates the xy-plane by $\pi/6$ radians counterclockwise about the z-axis.

What is the 3×3 matrix representing f?
What is the determinant of this matrix?

Trivia:

How to Work with Affine Spaces & Affine Transformations

Q: Is this a subspace?



No. Subspaces have to go through $\vec{0}$, to be closed under multiplication (by 0).

Q: Is this a linear transformation?

$$(x,y,z) \longmapsto (x+1,y,z)$$

No. A linear transformation has to take $\vec{0}$ to $\vec{0}$, in order to satisfy linearity under multiplication:

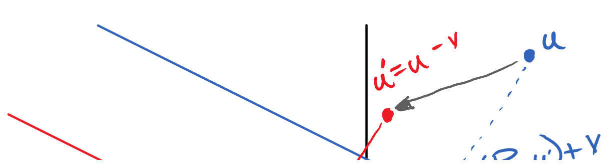
$$f(\alpha\vec{v}) = \alpha \cdot f(\vec{v})$$

$$\Rightarrow f(\vec{0}) = \vec{0}, \text{ setting } \alpha = 0.$$

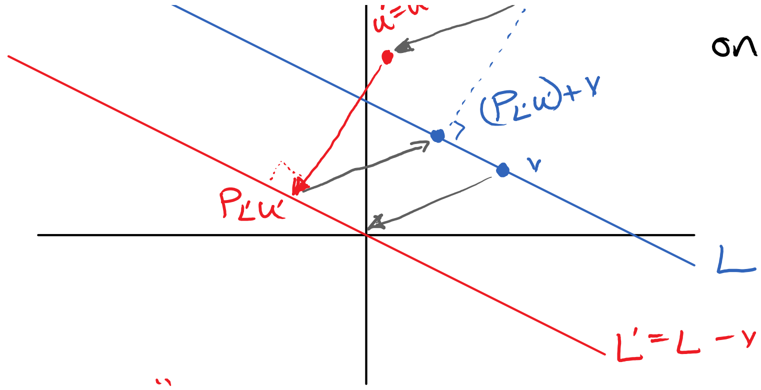
Thus translations are not linear.

Of course, affine subspaces and affine transformations, that shift the origin. How can we manipulate them?

(1) Shift coordinates so the space goes through $\vec{0}$,
work there,
then shift/translate back.



← Example: How to project u onto the affine subspace L.



onto the affine subspace L.

② Use "homogeneous coordinates"

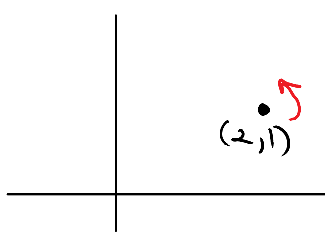
Add a new coordinate, $w=1$. Then,

$$T:(x,y,z)\longmapsto(x+1,y,z)$$

corresponds to

$$\begin{pmatrix}1&0&0&1\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{pmatrix}:(x,y,z,w)\longmapsto(x+w,y,z,w)$$

Example: Using homogeneous coordinates, give a 3×3 matrix for the 2D affine transformation



rotate 30 degrees about the point $(2,1)$.

Answer: Shift to the origin $(0,0)$, rotate there, and shift back.

$$\begin{pmatrix}1&0&2\\0&1&1\\0&0&1\end{pmatrix}\begin{array}{c}\begin{array}{ccc}x&y&w\end{array}\\\begin{pmatrix}c&-s&0\\s&c&0\\0&0&1\end{pmatrix}\end{array}\begin{array}{c}\begin{array}{ccc}x&y&w\end{array}\\\begin{pmatrix}1&0&-2\\0&1&-1\\0&0&1\end{pmatrix}\end{array}$$

$c=\cos 30^\circ=\frac{\sqrt{3}}{2}$

$s=\sin 30^\circ=\frac{1}{2}$

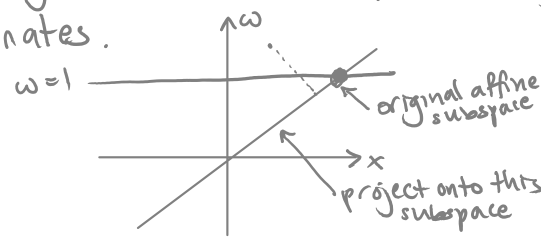
$$=\begin{array}{c}\begin{array}{ccc}x&y&w\end{array}\\\begin{pmatrix}c&-s&2-2c+s\\s&c&1-c-2s\\0&0&1\end{pmatrix}\end{array}=M \quad ✓$$

Sanity check: $M\begin{pmatrix}2\\1\end{pmatrix}=\begin{pmatrix}2\\1\end{pmatrix}$ ✓ $\quad M\begin{pmatrix}3\\1\end{pmatrix}$

Sanity check: $M\begin{pmatrix}2\\1\\1\end{pmatrix} = \begin{pmatrix}2\\1\\1\end{pmatrix}$ ✓ $\quad M\begin{pmatrix}3\\1\\1\end{pmatrix}$

$$M\begin{pmatrix}3\\1\\1\end{pmatrix} = M\begin{pmatrix}2\\1\\1\end{pmatrix} + M\begin{pmatrix}1\\0\\0\end{pmatrix} = \begin{pmatrix}2+c\\1+s\\1\end{pmatrix} ✓$$

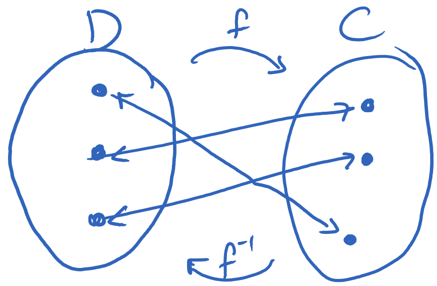
(**Note**: You can also project to affine subspaces using homogeneous coordinates.



)

Matrix and function inverses

Definition: A function $f : \underset{\text{domain}}{D} \to \underset{\text{codomain}}{C}$ is **invertible** if every point in C is the image of exactly one point in D.

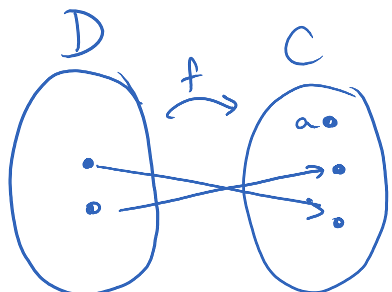


The inverse function $f^{-1} : C \to D$ takes each point in C to its unique preimage.

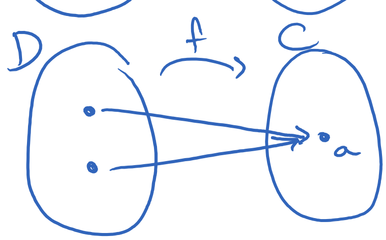
Equivalently, f^{-1} satisfies

$f^{-1} \circ f$ = identity on D

$f \circ f^{-1}$ = identity on C



a isn't the image of anything
$\Rightarrow f$ not invertible



a has two preimages
$\Rightarrow f$ not invertible

Exercise: Prove that the inverse of a linear function, if it exists, is also linear.

Definition: The inverse of a matrix A is a matrix B that satisfies

BA = the identity matrix and AB = the identity matrix.

Observe:
- Not every matrix is invertible, e.g., $A = (0)$ is not invertible.
  A matrix that is not invertible is called **singular**.
- If an inverse exists, then it is unique.

Proof: Assume B and C are both inverses of A.

Consider BAC.

$$\begin{array}{ccc} & BAC & \\ \rotatebox{45}{$=$} & & \rotatebox{-45}{$=$} \\ C=(BA)C & & B(AC)=B \end{array} \quad \Rightarrow B=C \quad \square$$

Example: The 2×2 matrix $A=\begin{pmatrix} a & b \\ c & d\end{pmatrix}$ is invertible/nonsingular if and only if $ad-bc\neq 0$. The inverse is given by

$$A^{-1}=\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a\end{pmatrix}.$$

(Check this!)

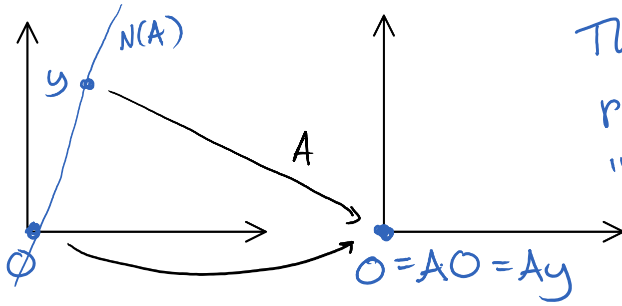
How to compute the inverse of a matrix?

Gaussian elimination, of course ...

When is a matrix invertible?

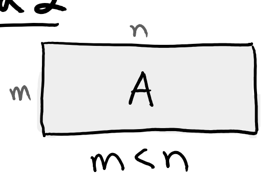
Lemma 1: If $N(A)\neq\{\vec{0}\}$, then A is not invertible.

Proof: Take $y\in N(A)$, $y\neq 0$.



Then the point 0 has two preimages, so A is not invertible. $\square$

Lemma 2:



$\Rightarrow N(A)\neq\{\vec{0}\}$

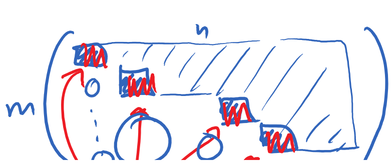
ie., it is **more**!

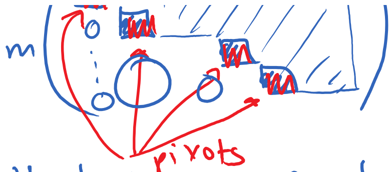
Intuition:

A "squashes" n dimensions down to $m<n$ dimensions.

$\therefore$ Some directions must be collapsed to $\vec{0}$.

Proof: To find $N(A)$, apply Gaussian elimination:

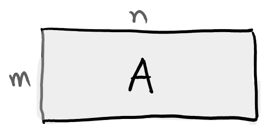




The point is that there are at most m pivots. Since $m<n$, there are necessarily at least $n-m \geq 1$ free variables so the nullspace is infinite. ✓ $\square$

Corollary: No non-square matrix can be inverted.

Proof: Either

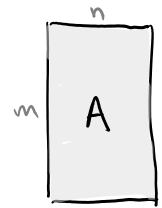


$m<n$

$\Rightarrow N(A) \neq \{\vec{0}\}$ (Lemma 2)

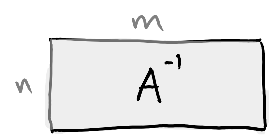
$\Rightarrow A$ not invertible (Lemma 1) ✓

or



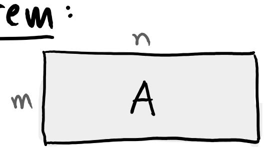
$m>n$

$\Rightarrow A^{-1}$ would be $n \times m$



and this can't be inverted ✓ $\square$

Theorem:



is invertible $\iff$ $m=n$, $N(A)=\{\vec{0}\}$

Remark: This doesn't work in infinite dimensions, eg.,

$$A = \begin{pmatrix} 0 & 0 & 0 & \cdots \\ 1 & 0 & 0 & \\ 0 & 1 & 0 & \\ \vdots & 0 & 1 & \\ & \vdots & 0 & \ddots \\ & & \vdots & \end{pmatrix}$$

$N(A)=\{\vec{0}\}$, but A is not invertible.

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ & & \vdots & & \ddots \end{pmatrix}$$

$N(A) = \mathrm{Span}(\vec{e}_1)$, but $R(A)$ = everything.