Math 2270-004 Week 4 notes

We will not necessarily finish the material from a given day's notes on that day. We may also add or subtract some material as the week progresses, but these notes represent an in-depth outline of what we plan to cover. These notes cover material in 1.8-1.9, 2.1-2.2.

Mon Jan 29

- 1.8-1.9 Matrix and linear transformations

Announcements: continuing from below, I could compute $T(\vec{x}(t))$ directly:

$$= \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} t \\ 1+2t \end{bmatrix} = \begin{bmatrix} -t \\ 1+2t \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + t\begin{bmatrix} -1 \\ 2 \end{bmatrix} \checkmark$$

$$\left[\begin{array}{l} T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v}) \\ T(c\vec{u}) = cT(\vec{u}) \end{array}\right.$$

Warm-up Exercise: Consider $T: \mathbb{R}^2 \to \mathbb{R}^2$

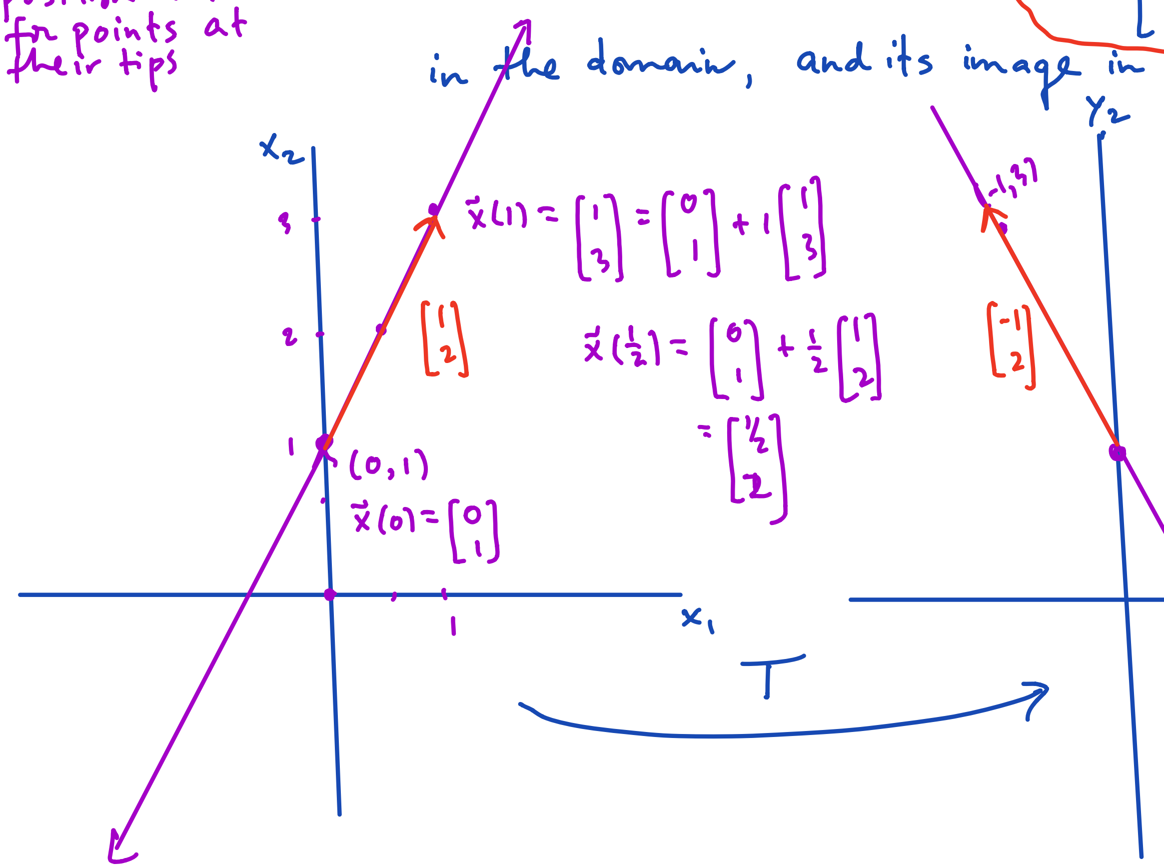
'til 1:00

$$T\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

(reflection across the x_2-axis)

remember, these are position vectors for points at their tips

- sketch the parametric line $\vec{x} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + t\begin{bmatrix} 1 \\ 2 \end{bmatrix}$, $t \in \mathbb{R}$ in the domain, and its image in the codomain



formula for image line

$$T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix} + t\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right)$$
$$= T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) + T\left(t\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right)$$
$$= \begin{bmatrix} 0 \\ 1 \end{bmatrix} + t\,T\begin{bmatrix} 1 \\ 2 \end{bmatrix}$$
$$= \begin{bmatrix} 0 \\ 1 \end{bmatrix} + t\begin{bmatrix} -1 \\ 2 \end{bmatrix} \checkmark$$

Recall from Friday that

Definition: A function T which has domain equal to $\mathbb{R}^n$ and whose range lies in $\mathbb{R}^m$ is called a *linear transformation* if it transforms sums to sums, and scalar multiples to scalar multiples. Precisely, $T : \mathbb{R}^n \to \mathbb{R}^m$ is *linear* if and only if

$$T(\underline{\boldsymbol{u}} + \underline{\boldsymbol{v}}) = T(\underline{\boldsymbol{u}}) + T(\underline{\boldsymbol{v}}) \qquad \forall\ \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}} \in \mathbb{R}^n$$
$$T(c\,\underline{\boldsymbol{u}}) = c\,T(\underline{\boldsymbol{u}}) \qquad \forall\ c \in \mathbb{R}, \underline{\boldsymbol{u}} \in \mathbb{R}^n\ .$$

Notation In this case we call $\mathbb{R}^m$ the *codomain*. We call $T(\underline{\boldsymbol{u}})$ the *image of* $\underline{\boldsymbol{u}}$. The *range* of T is the collection of all images $T(\underline{\boldsymbol{u}})$, for $\underline{\boldsymbol{u}} \in \mathbb{R}^n$.

Important connection to matrices: Each matrix $A_{m \times n}$ gives rise to a linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, namely

$$T(\underline{\boldsymbol{x}}) := A\,\underline{\boldsymbol{x}} \qquad \forall\ \underline{\boldsymbol{x}} \in \mathbb{R}^n.$$

This is because, as we checked last week, matrix transformation satisfies the linearity axioms:

$$A(\underline{\boldsymbol{u}} + \underline{\boldsymbol{v}}) = A\,\underline{\boldsymbol{u}} + A\,\underline{\boldsymbol{v}} \qquad \forall\ \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}} \in \mathbb{R}^n$$
$$A(c\,\underline{\boldsymbol{u}}) = c\,A\,\underline{\boldsymbol{u}} \qquad \forall\ c \in \mathbb{R}, \underline{\boldsymbol{u}} \in \mathbb{R}^n$$

Recall, we verified this by using the column notation $A = \left[\underline{\boldsymbol{a}}_1, \underline{\boldsymbol{a}}_2, \ldots\, \underline{\boldsymbol{a}}_n\right]$, writing

$$\underline{\boldsymbol{u}} = \begin{bmatrix} u_1 \\ u_2 \\ : \\ u_n \end{bmatrix}, \underline{\boldsymbol{v}} = \begin{bmatrix} v_1 \\ v_2 \\ : \\ v_n \end{bmatrix},$$

and computing

$$\begin{aligned} A(\underline{\boldsymbol{u}} + \underline{\boldsymbol{v}}) &= \left[\underline{\boldsymbol{a}}_1, \underline{\boldsymbol{a}}_2, \ldots\, \underline{\boldsymbol{a}}_n\right](\underline{\boldsymbol{u}} + \underline{\boldsymbol{v}}) \\ &= (u_1 + v_1)\underline{\boldsymbol{a}}_1 + (u_2 + v_2)\underline{\boldsymbol{a}}_2 + \ldots + (u_n + v_n)\underline{\boldsymbol{a}}_n \\ &= (u_1\underline{\boldsymbol{a}}_1 + u_2\underline{\boldsymbol{a}}_2 + \ldots + u_n\underline{\boldsymbol{a}}_n) + (v_1\underline{\boldsymbol{a}}_1 + v_2\underline{\boldsymbol{a}}_2 + \ldots + v_n\underline{\boldsymbol{a}}_n) \\ &= A\,\underline{\boldsymbol{u}} + A\,\underline{\boldsymbol{v}}. \end{aligned}$$

And,

$$\begin{aligned} A(c\,\underline{\boldsymbol{u}}) &= \left[\underline{\boldsymbol{a}}_1, \underline{\boldsymbol{a}}_2, \ldots\, \underline{\boldsymbol{a}}_n\right](c\,\underline{\boldsymbol{u}}). \\ &= cu_1\underline{\boldsymbol{a}}_1 + cu_2\underline{\boldsymbol{a}}_2 + \ldots + cu_n\underline{\boldsymbol{a}}_n = c\,A\,\underline{\boldsymbol{u}} \\ &= c\,A\,\underline{\boldsymbol{u}}\ . \end{aligned}$$

Geometry of linear transformations: (We saw this illustrated in a concrete example on Friday)

<u>Theorem:</u> For each matrix $A_{m \times n}$ and corresponding linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$, given by

$$T(\underline{\mathbf{x}}) := A\,\underline{\mathbf{x}} \qquad \forall\, \underline{\mathbf{x}} \in \mathbb{R}^n$$

- (parallel) lines in the domain $\mathbb{R}^n$ are transformed by T into (parallel) lines or points in the codomain $\mathbb{R}^m$.

- 2-dimensional planes in the domain $\mathbb{R}^n$ are transformed by T into (parallel) planes, (parallel) lines, or points in the codomain $\mathbb{R}^m$.

<u>proof:</u>

- A parametric line through a point with position vector $\underline{\boldsymbol{p}}$ and direction vector $\underline{\boldsymbol{u}}$ in the domain can be expressed as the set of point having position vectors

$$\underline{\boldsymbol{p}} + t\,\underline{\boldsymbol{u}}, \quad t \in \mathbb{R}.$$

The image of this line is the set of points

$$A(\underline{\boldsymbol{p}} + t\,\underline{\boldsymbol{u}}) = A\,\underline{\boldsymbol{p}} + t\,A\,\underline{\boldsymbol{u}}, \quad t \in \mathbb{R}$$

"$A(\vec{p}) + A(t\vec{u})$"

which is either a line through $A\,\underline{\boldsymbol{p}}$ with direction vector $A\,\underline{\boldsymbol{u}}$, when $A\,\underline{u} \neq \underline{\mathbf{0}}$, or the point $A\,\underline{\boldsymbol{p}}$ when $A\,\underline{\boldsymbol{u}} = \underline{\mathbf{0}}$. Since parallel lines have parallel direction vectors, their images also will have parallel direction vectors, and therefore be parallel lines.

parallel line: $\vec{q} + t\vec{u}$

$A(\vec{q} + t\vec{u}) = A(\vec{q}) + t\,A(\vec{u})$

is parallel to 1st line if $A(\vec{u}) \neq \vec{0}$.

- A parametric (2-dimensional) plane through a point with position vector $\underline{\boldsymbol{p}}$ and independent direction vectors $\underline{\boldsymbol{u}}, \underline{\boldsymbol{v}}$ in the domain can be expressed as the set of point having position vectors

$$\underline{\boldsymbol{p}} + t\,\underline{\boldsymbol{u}} + s\,\underline{\boldsymbol{v}}, \quad t, s \in \mathbb{R}.$$

The image of this line is the set of points

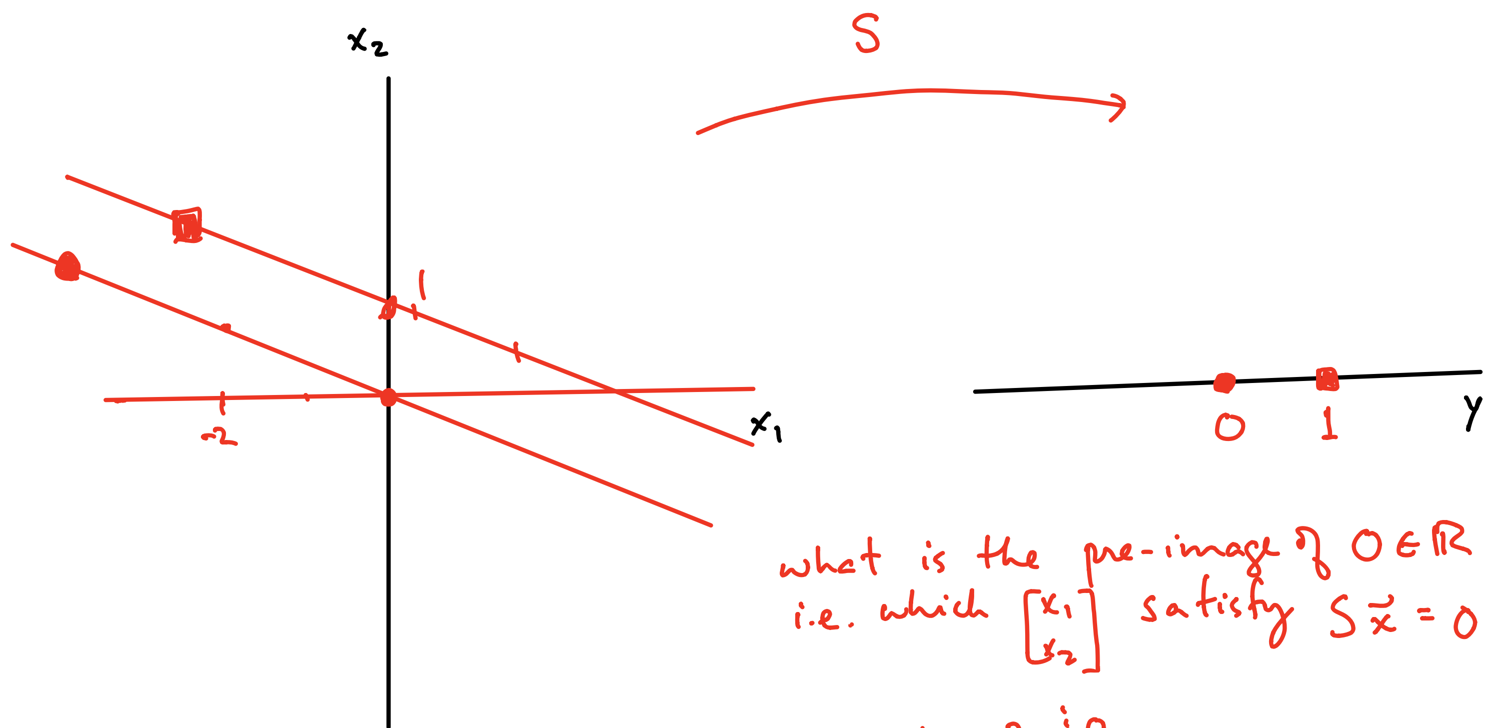
$$A(\underline{\boldsymbol{p}} + t\,\boldsymbol{u} + s\,\underline{\boldsymbol{v}}) = A\,\underline{\boldsymbol{p}} + t\,A\,\underline{\boldsymbol{u}} + s\,A\,\underline{\boldsymbol{v}}, \quad t, s \in \mathbb{R}$$

which is either a plane through $A\,\underline{\boldsymbol{p}}$ with independent direction vectors $A\,\underline{\boldsymbol{u}}, A\,\underline{\boldsymbol{v}}$, or a line through point $A\,\underline{\boldsymbol{p}}$ if $A\,\underline{\boldsymbol{u}}, A\,\underline{\boldsymbol{v}}$ are dependent but not both zero, or the point $A\,\underline{\boldsymbol{p}}$ if $A\,\underline{u} = A\,\underline{\boldsymbol{v}} = \underline{\mathbf{0}}$. Since parallel planes can be expressed with the same direction vectors, their images also will have parallel direction vectors, and therefore be parallel.

Exercise 1) Consider the linear transformation $S : \mathbb{R}^2 \to \mathbb{R}$ given by

$$S\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) := \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = [x_1 + 2x_2].$$

Make a geometric sketch that indicates what the transformation does. In this case the interesting behavior is in the domain.



what is the pre-image of $0 \in \mathbb{R}$

i.e. which $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ satisfy $S\vec{x} = 0$

$$\left[\begin{array}{cc|c} 1 & 2 & 0 \end{array}\right]$$

$x_1 = -2t$
$x_2 = t$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = t \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

preimage of 1

$$\left[\begin{array}{cc|c} 1 & 2 & 1 \end{array}\right]$$

$x_1 = 1 - 2t$
$x_2 = t$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + t \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

For the rest of today we'll consider linear transformations from $\mathbb{R}^2 \rightarrow \mathbb{R}^2$. (See the transformed goldfish in text section 1.9 - it's worth it.) We write the standard basis vectors for $\mathbb{R}^2$ as

$$\underline{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \ \underline{e}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = x_1 \vec{e}_1 + x_2 \vec{e}_2$$

If $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is linear, then

$$\begin{aligned} T\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) &= T(x_1 \underline{e}_1 + x_2 \underline{e}_2) \\ &= T(x_1 \underline{e}_1) + T(x_2 \underline{e}_2) \\ &= x_1 T(\underline{e}_1) + x_2 T(\underline{e}_2) \\ &= \left[T(\underline{e}_1) \ \ T(\underline{e}_1) \right] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} ! \end{aligned}$$

In other words, if $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is a linear transformation, it's actually a matrix transformation. And the columns of the matrix, in order, are $T(\underline{e}_1), T(\underline{e}_2)$. (The same idea shows that every linear $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is actually a matrix transformation....and that the columns of the matrix are T applied to the standard basis vectors in the domain $\mathbb{R}^n$.)

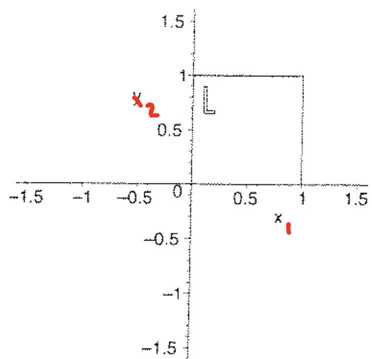
In the following diagrams the unit square in the domain, together with an "L" to keep track of whether the transformation involves a reflection or not, is on the left. On the right is the image of the square and the L under the transformation. Find the matrix for T!!! (Also note, that because parallel lines transform to parallel lines, grids go to transformed grids, so once you know how the unit square transforms, you know everything about the transformation. Or, to put it another way, once you know $T(\underline{e}_1), T(\underline{e}_2)$, you know the matrix for T so you know how every position vector transforms.)

§2.2 Important linear transformations $T: \mathbb{R}^2 \to \mathbb{R}^2$

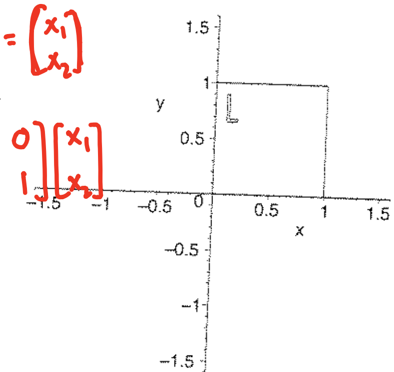
Bretscher likes to use an "L" to represent the geometry of the transformation.
I like to use "L boxes" instead ~ will tie into Maple project I on fractals.

Find the matrix formulas for these transformations!

If T is linear, it's a matrix transformation with matrix $[T(\vec{e}_1)\ T(\vec{e}_2)]$



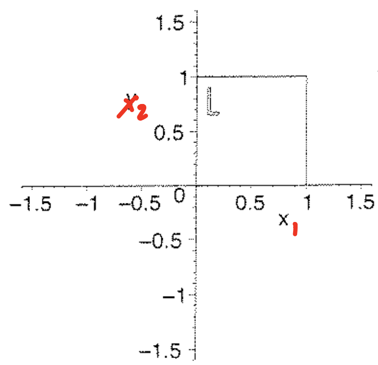
$$T_1 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$



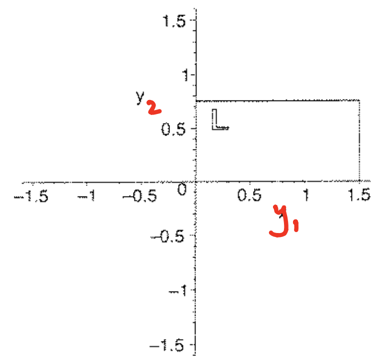
identity transformation

$T(\vec{e}_1) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $T(\vec{e}_2) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ ✓



T_2



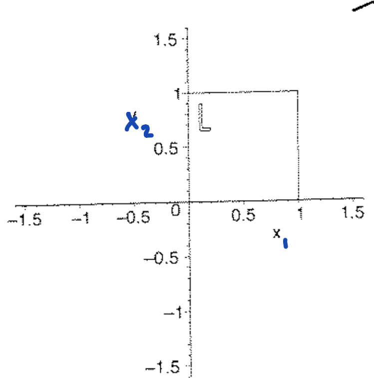
stretch by 1.5 horizontally & .75 vertically (from origin)

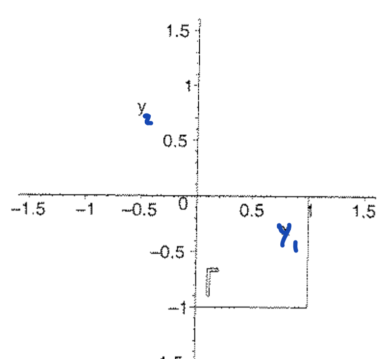
$$T_2 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \left[T_2(\vec{e}_1) \ \vdots\ T_2(\vec{e}_2) \right] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= \begin{bmatrix} 1.5 & 0 \\ 0 & .75 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= \begin{bmatrix} 1.5\,x_1 \\ .75\,x_2 \end{bmatrix}$$

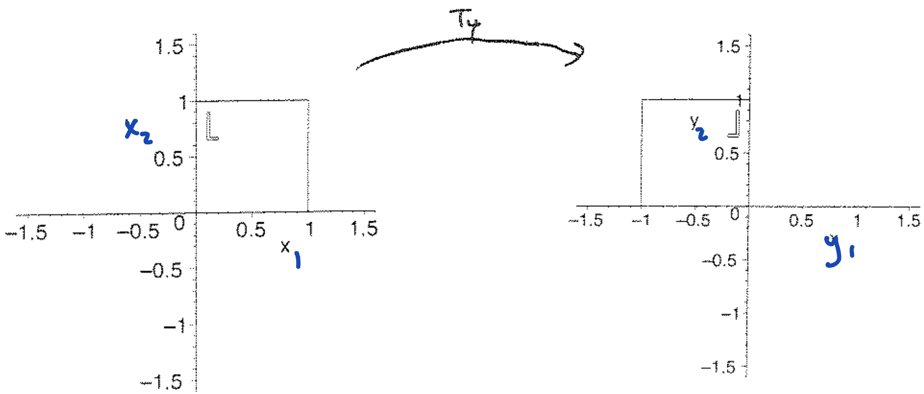
T_3





reflect across x-axis

$$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

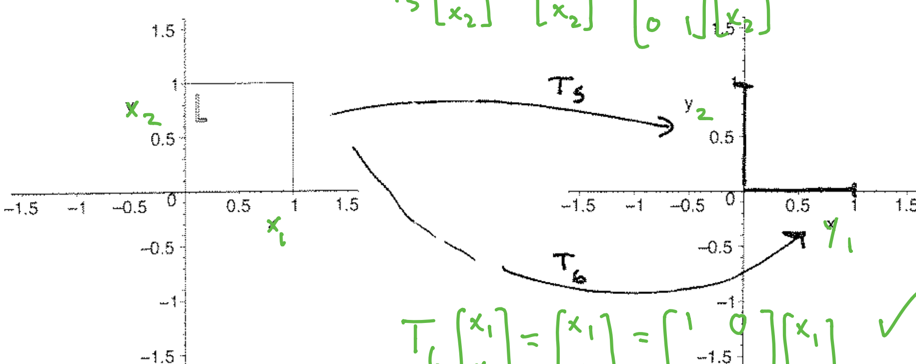


reflect across y-axis

warmup.

$\vec{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ position vector $(1,0)$
project $(1,0)$ onto vertical axis I get $(0,0)$

$$T_5 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$ ✓



project to y_1-axis

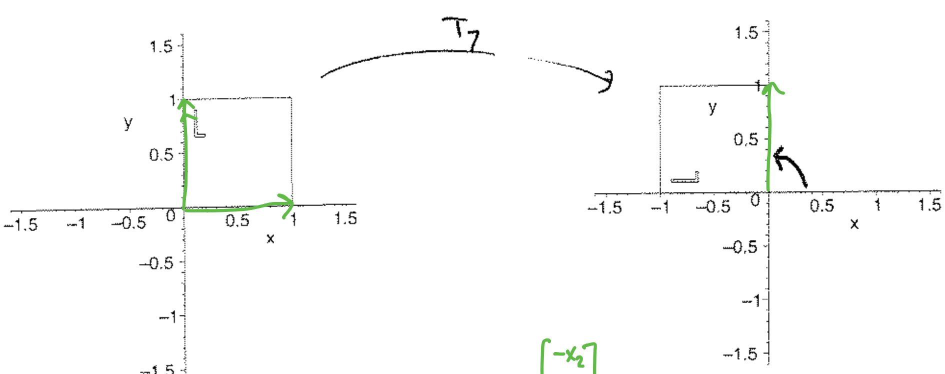
$$A = \left[T_5(\vec{e}_1) \;\vdots\; T_5(\vec{e}_2) \right] = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}.$$

$$T_6 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$ ✓

project to x-axis

$$T_6 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad T_6 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

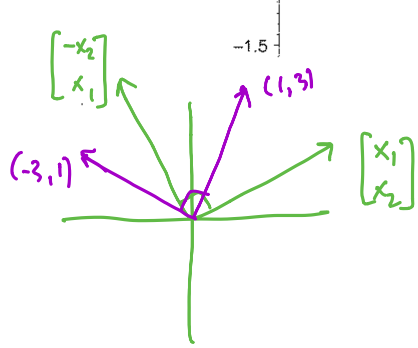


rotate by $\pi/2$ radians counterclockwise

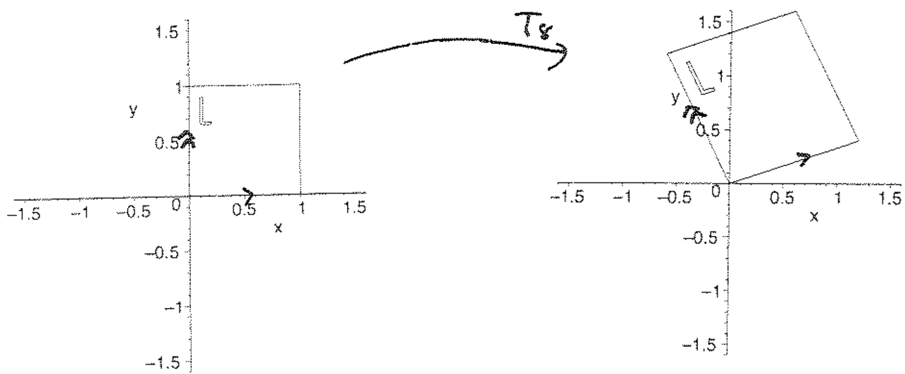
$$T_7 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad T_7 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$$

$$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

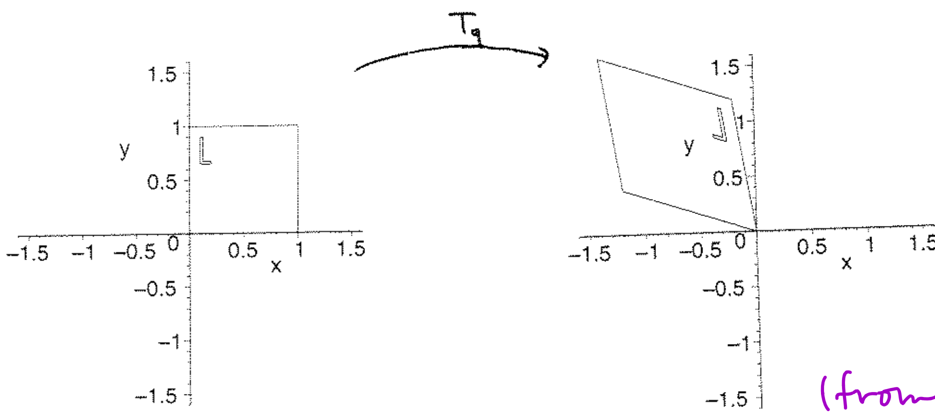
$$T_7 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix}$$



$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \cdot \begin{bmatrix} -x_2 \\ x_1 \end{bmatrix} = -x_1x_2 + x_2x_1 = 0$$



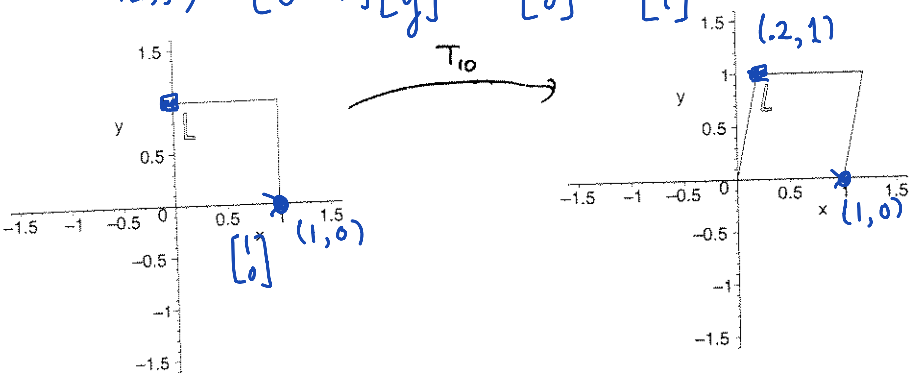
mystery linear trans.



another mystery!

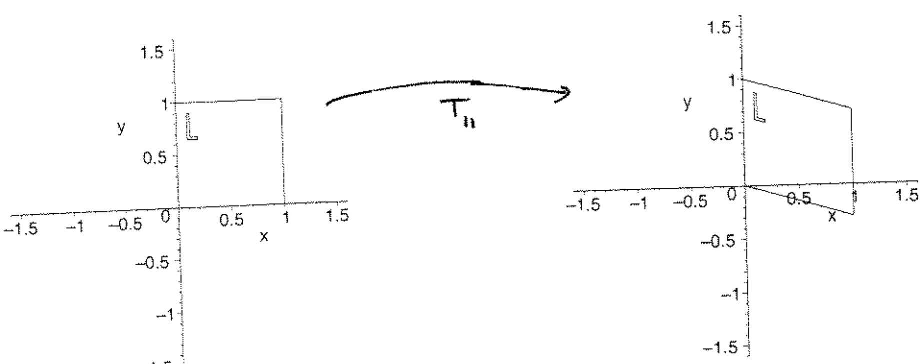
(from "Monday" notes)

so $T_{10}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 1 & .2 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = x\begin{bmatrix} 1 \\ 0 \end{bmatrix} + y\begin{bmatrix} .2 \\ 1 \end{bmatrix}$ Try on Wed.

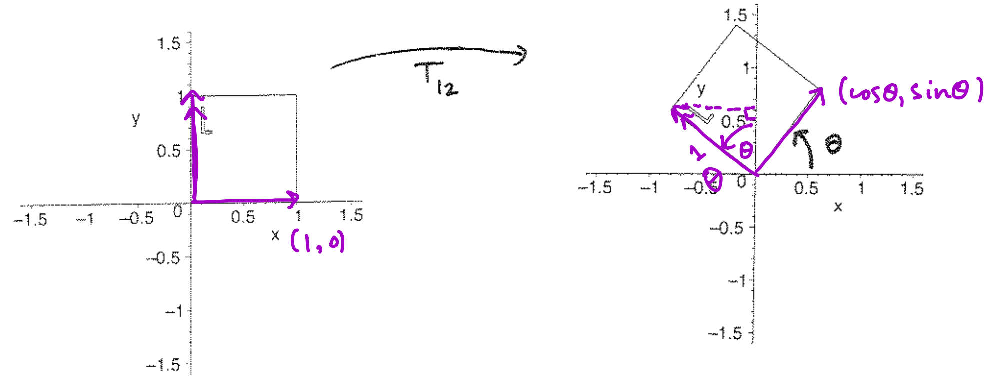


horizontal shear with strength .2

$$A = \left[T(\vec{e}_1) \mid T(\vec{e}_2) \right] = \begin{bmatrix} 1 & .2 \\ 0 & 1 \end{bmatrix}$$



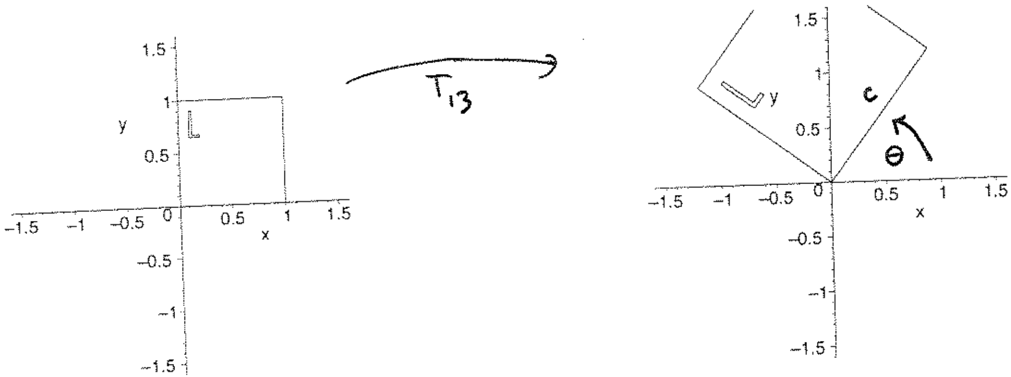
vertical shear with strength -.3



rotation (c.c.) by angle θ

$$T_{12}\begin{bmatrix}1\\0\end{bmatrix} = \begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix} \qquad T_{12}\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}-\sin\theta\\ \cos\theta\end{bmatrix}$$

$$A = \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}$$



rotate by θ and scale uniformly by factor of c ("rotation dilation")

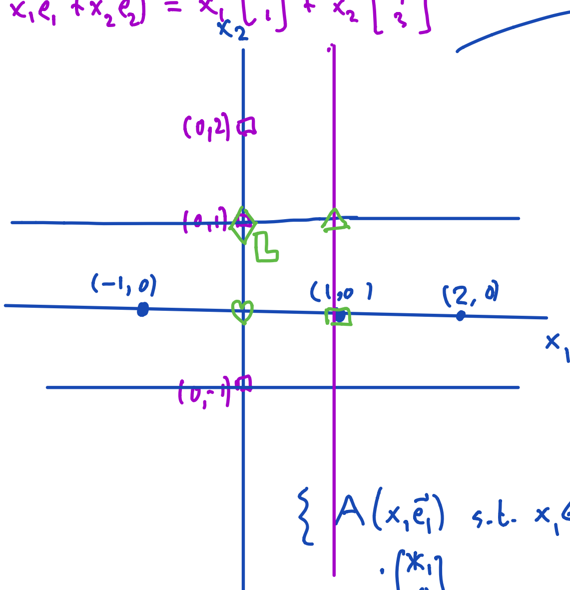
$T: \mathbb{R}^2 \to \mathbb{R}^2$ linear. $\boxed{T(x_1\vec{e}_1 + x_2\vec{e}_2) = x_1 T(\vec{e}_1) + x_2 T(\vec{e}_2)} = \begin{bmatrix}T(\vec{e}_1) \,\vdots\, T(\vec{e}_2)\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}$

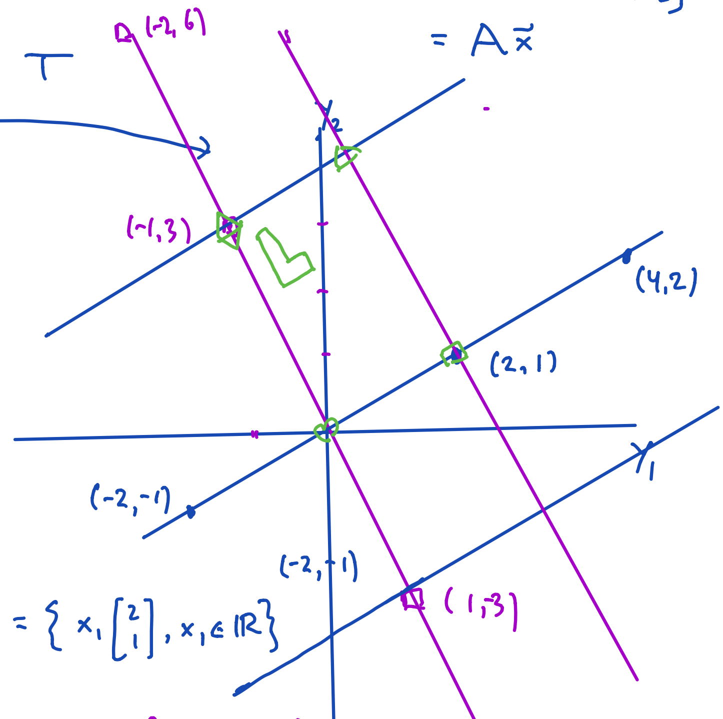
$= A\vec{x}$

- $T\begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}2 & -1\\ 1 & 3\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix} = x_1\begin{bmatrix}2\\1\end{bmatrix} + x_2\begin{bmatrix}-1\\3\end{bmatrix}$

$T(x_1\vec{e}_1 + x_2\vec{e}_2) = x_1\begin{bmatrix}2\\1\end{bmatrix} + x_2\begin{bmatrix}-1\\3\end{bmatrix}$

T





$\left\{ A(x_1\vec{e}_1) \text{ s.t. } x_1 \in \mathbb{R} \right\} = \left\{ x_1\begin{bmatrix}2\\1\end{bmatrix}, x_1 \in \mathbb{R} \right\}$

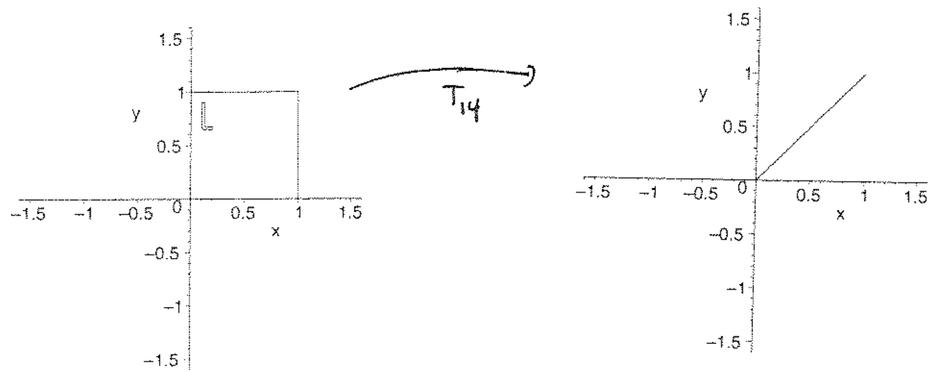
$\begin{bmatrix}x_1\\0\end{bmatrix}$

$\left\{ A(x_2\vec{e}_2) \text{ s.t. } x_2 \in \mathbb{R} \right\} = \left\{ x_2\begin{bmatrix}-1\\3\end{bmatrix}, x_2 \in \mathbb{R} \right\}$

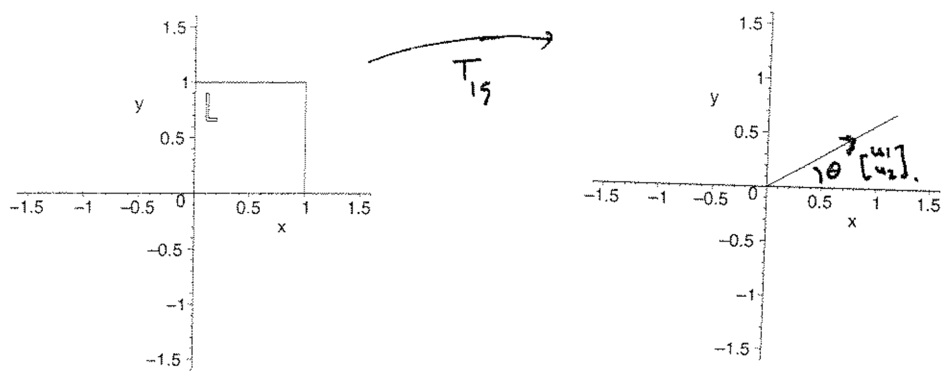
$\begin{bmatrix}0\\x_2\end{bmatrix}$

$$\left\{ A(x_1\vec{e}_1 + 1\cdot\vec{e}_2) = x_1 A(\vec{e}_1) + A(\vec{e}_2) = x_1\begin{bmatrix}2\\1\end{bmatrix} + \begin{bmatrix}-1\\3\end{bmatrix} \right\}$$ 6

$$\left\{ A(\vec{e}_1 + x_2\vec{e}_2) = A\vec{e}_1 + x_2 A\vec{e}_2 = \begin{bmatrix}2\\1\end{bmatrix} + x_2\begin{bmatrix}-1\\3\end{bmatrix} \right\}$$



project onto the line $y=x$



project onto line thru origin at angle θ, with unit direction

$$\begin{bmatrix}u_1\\u_2\end{bmatrix} = \begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix}.$$

Tues Jan 30

- 1.9 the matrix of a linear transformation
- 2.1 Matrix operations, especially matrix multiplication

postpone all 2.1 HW until next assignment

Announcements:

- 2.1 HW: do 1, 3, 9, 11. postpone 23, 25, 27 until next week
- discuss "one to one" and "onto" today.

function concepts : $f : X \longrightarrow Y$
domain codomain

f is "1-1" means the equation $f(x) = b$ $x \in X, b \in Y$ has at most one solution $x \in X$.
(if $f(x) = f(z)$, then actually $x = z$)

HW questions?

~~Warm-up Exercise~~:

f is "onto" Y means
for each $b \in Y$ there is (at least one) $x \in X$ with $f(x) = b$

If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, given by
$T(\vec{x}) = A\vec{x}$ $(A_{m \times n})$

- then, T being 1-1 means $A\vec{x} = \vec{b}$ has at most one solution. $\Longleftrightarrow$ rref(A) has pivot in each column
(i.e. no free parameters).
- then T being onto means $A\vec{x} = \vec{b}$ has (at least) a solution $\vec{x}$ for each $\vec{b} \in \mathbb{R}^m$
$\Longleftrightarrow$ rref(A) has no zero rows, i.e. a pivot in each row.

- continue understanding the geometry of linear transformations, using Monday's notes

1.9 Theorem: Every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is actually a matrix transformation!

Some general defintions first, which we've already alluded to:

- In $\mathbb{R}^n$ we write

$$\underline{\boldsymbol{e}}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ \underline{\boldsymbol{e}}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ \ldots \ \underline{\boldsymbol{e}}_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}.$$

- We call

$$\{\underline{\boldsymbol{e}}_1, \underline{\boldsymbol{e}}_2, \ldots \underline{\boldsymbol{e}}_n\}$$

the *standard basis* for $\mathbb{R}^n$ because each

$$\underline{\boldsymbol{x}} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n$$

is easily and uniquely expressed as a linear combination of the standard basis vectors,

$$\underline{\boldsymbol{x}} = x_1 \underline{\boldsymbol{e}}_1 + x_2 \underline{\boldsymbol{e}}_2 + \ \ldots \ + x_n \underline{\boldsymbol{e}}_n$$

Example 1:

$$\begin{bmatrix} 3 \\ 7 \\ -6 \end{bmatrix} = 3\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 7\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} - 6\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 3\underline{\boldsymbol{e}}_1 + 7\underline{\boldsymbol{e}}_2 - 6\underline{\boldsymbol{e}}_3 .$$

you used to call these basis vectors $\hat{i}, \hat{j}, \hat{k}$.

Example 2: For A written in terms of its columns, $A = [\underline{\boldsymbol{a}}_1, \underline{\boldsymbol{a}}_2, \ldots \underline{\boldsymbol{a}}_n]$,

$$A\,\underline{\boldsymbol{e}}_j = \underline{\boldsymbol{a}}_j \ ,$$

the j^{th} column of A.

$$\left[a_1 \,\vdots\, a_2 \,\vdots\, \cdots \ \vec{a}_n \right] \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \vec{a}_2 .$$

Theorem: Every linear transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ is actually a matrix transformation,

$$T(\underline{\boldsymbol{x}}) = A\,\underline{\boldsymbol{x}},$$

where the j^{th} column of $A_{m \times n}$, is $T(\underline{\boldsymbol{e}}_j), j = 1, 2., \ldots n$. In other words the matrix of T is

$$A = [\,T(\underline{\boldsymbol{e}}_1)\;\; T(\underline{\boldsymbol{e}}_1)\;\; \ldots.\;\; T(\underline{\boldsymbol{e}}_1)\,]$$

proof: T is linear, which means it satisfies

(i) $\quad T(\underline{\boldsymbol{u}} + \underline{\boldsymbol{v}}) = T(\underline{\boldsymbol{u}}) + T(\underline{\boldsymbol{v}}) \qquad \forall\, \underline{\boldsymbol{u}}, \underline{\boldsymbol{v}} \in \mathbb{R}^n$

(ii) $\quad T(c\,\underline{\boldsymbol{u}}) = c\,T(\underline{\boldsymbol{u}}) \qquad \forall\, c \in \mathbb{R}, \underline{\boldsymbol{u}} \in \mathbb{R}^n$.

$$T((\vec{u} + \vec{v}) + \vec{w}) \overset{(i)}{=} T(\vec{u} + \vec{v}) + T(\vec{w})$$
$$\overset{(i)}{=} (T(\vec{u}) + T(\vec{v})) + T(\vec{w})$$
$$= T(\vec{u}) + T(\vec{v}) + T(\vec{w})$$

Let's compute $T(\underline{\boldsymbol{x}})$ for $\underline{\boldsymbol{x}} \in \mathbb{R}^n$:

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) = T(x_1\underline{\boldsymbol{e}}_1 + x_2\underline{\boldsymbol{e}}_2 + \;\ldots.\; + x_n\underline{\boldsymbol{e}}_n)$$

$$= T(x_1\underline{\boldsymbol{e}}_1) + T(x_2\underline{\boldsymbol{e}}_2) + \ldots T(x_n\underline{\boldsymbol{e}}_n)$$

by repeated applications of the sum property (i).

$$= x_1 T(\underline{\boldsymbol{e}}_1) + x_2 T(\underline{\boldsymbol{e}}_2) + \ldots + x_n T(\underline{\boldsymbol{e}}_n)$$

by applications fo the scalar multiple property (ii).

$$= A\,\underline{\boldsymbol{x}}$$

for the matrix A given in column form as

$$A = [\,T(\underline{\boldsymbol{e}}_1)\;\; T(\underline{\boldsymbol{e}}_2)\;\; \ldots.\;\; T(\underline{\boldsymbol{e}}_n)\,].$$

Q.E.D.

Exercise 1 Illustrate the linear transformation theorem with the projection function $T : \mathbb{R}^3 \to \mathbb{R}^2$,

$$T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

we did this. warmup Wed.

2.1 Matrix multiplication

Suppose we take a composition of linear transformations:

$$T_1 : \mathbb{R}^n \to \mathbb{R}^m, \quad T_1(\underline{\boldsymbol{x}}) = A\,\underline{\boldsymbol{x}}, \qquad (A_{m \times n}).$$

$$T_2 : \mathbb{R}^m \to \mathbb{R}^p, \quad T_2(\underline{\boldsymbol{y}}) = B\,\underline{\boldsymbol{y}}. \qquad (B_{p \times m})$$

- Then the composition $T_2 \circ T_1 : \mathbb{R}^n \to \mathbb{R}^p$ is linear:

$$\begin{aligned}(i) \quad (T_2 \circ T_1)(\underline{\boldsymbol{u}} + \underline{\boldsymbol{v}}) &:= (T_2(T_1(\underline{\boldsymbol{u}} + \underline{\boldsymbol{v}}))) \\ &= T_2(T_1(\underline{\boldsymbol{u}}) + T_1(\underline{\boldsymbol{v}})) && T_1 \text{ linear} \\ &= T_2(T_1(\underline{\boldsymbol{u}})) + T_2(T_1(\underline{\boldsymbol{v}})) && T_2 \text{ linear} \\ &:= (T_2 \circ T_1)(\underline{\boldsymbol{u}}) + (T_2 \circ T_1)(\underline{\boldsymbol{v}})\end{aligned}$$

$$\begin{aligned}(ii) \quad (T_2 \circ T_1)(c\,\underline{\boldsymbol{u}}) &:= (T_2(T_1(c\,\underline{\boldsymbol{u}}))) \\ &= T_2(c\,T_1(\underline{\boldsymbol{u}})) && T_1 \text{ linear} \\ &= c\,T_2(T_1(\underline{\boldsymbol{u}})) && T_2 \text{ linear} \\ &:= c\,(T_2 \circ T_1)(\underline{\boldsymbol{u}})\end{aligned}$$

- So $T_2 \circ T_1 : \mathbb{R}^n \to \mathbb{R}^p$ is a matrix transformation, by the theorem on the previous page.

- Its j^{th} column is

$$\begin{aligned}T_2 \circ T_1(\vec{e}_j) = T_2(T_1(\underline{\boldsymbol{e}}_j)) &= T_2(A(\underline{\boldsymbol{e}}_j)) \\ &= T_2(col_j(A)) \\ &= B(col_j(A)).\end{aligned}$$

- i.e. the matrix of $T_2 \circ T_1 : \mathbb{R}^n \to \mathbb{R}^p$ is

$$[B\,\underline{\boldsymbol{a}}_1 \; B\,\underline{\boldsymbol{a}}_2 \; \ldots \; B\,\underline{\boldsymbol{a}}_n] := B\,A.$$

where $col_j(A) = \underline{\boldsymbol{a}}_j$.

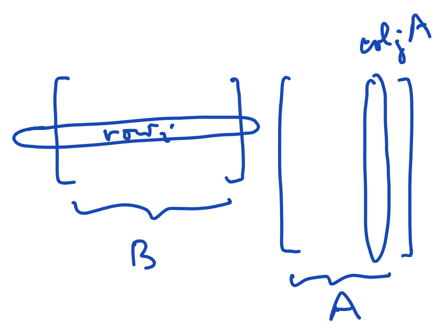
<u>Summary</u>: For $B_{p \times m}$ and $A_{m \times n}$

- the matrix product $(BA)_{p \times n}$ is defined by

$$col_j(B\,A) = B\,col_j(A) \qquad j = 1\,..\,n$$

- or equivalently

$$entry_{ij}(B\,A) = row_i(B) \cdot col_j(A) \quad i = 1\,..p, j = 1\,..n$$



- And,

$$(BA)\underline{\boldsymbol{x}} = B\,(A\,\underline{\boldsymbol{x}})$$

because BA is the matrix of $T_2 \circ T_1$.

$$entry_{ij}(BA) = row_i(B) \cdot col_j(A)$$

Exercise 2 Compute

$$\underset{B}{\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}}\underset{A}{\begin{bmatrix} 0 & 1 & -2 & 3 \\ 1 & 0 & 4 & 1 \end{bmatrix}} = \begin{bmatrix} 2 & 1 & 6 & 5 \\ 3 & -1 & 14 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0\begin{bmatrix} 1 \\ -1 \end{bmatrix} + 1\begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

or $\begin{bmatrix} 0\cdot 1 + 1\cdot 2 \\ -1\cdot 0 + 3\cdot 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

$\to$ 3rd col. $\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} -2 \\ 4 \end{bmatrix} = \begin{bmatrix} -2+8 \\ 2+12 \end{bmatrix}$

Exercise 3 For $T_1 : \mathbb{R}^4 \to \mathbb{R}^2$ $\qquad$ $T_2 : \mathbb{R}^2 \to \mathbb{R}^2$

$$T_1(\underline{\boldsymbol{x}}) = \underset{A}{\begin{bmatrix} 0 & 1 & -2 & 3 \\ 1 & 0 & 4 & 1 \end{bmatrix}}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} \qquad T_2(\underline{\boldsymbol{y}}) = \underset{B}{\begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

compute $T_2\left(T_1(\underline{\boldsymbol{x}})\right)$. How does this computation relate to Exercise 2?

$$T_2(T_1(\vec{x})) = T_2(A\vec{x}) = T_2\left(\begin{bmatrix} x_2 - 2x_3 + 3x_4 \\ x_1 + 4x_3 + x_4 \end{bmatrix}\right)$$

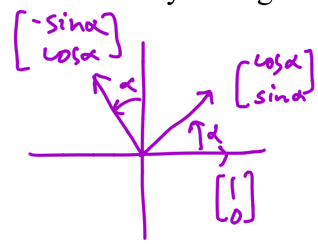
$$= \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}\begin{bmatrix} x_2 - 2x_3 + 3x_4 \\ x_1 + 4x_3 + x_4 \end{bmatrix}$$

$$= \begin{bmatrix} 1\cdot(x_2 - 2x_3 + 3x_4) + 2(x_1 + 4x_3 + x_4) \\ -1(x_2 - 2x_3 + 3x_4) + 3(x_1 + 4x_3 + x_4) \end{bmatrix}$$

$$= \begin{bmatrix} 2x_1 + x_2 + 6x_3 + 5x_4 \\ 3x_1 - x_2 + 14x_3 + 0x_4 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 1 & 6 & 5 \\ 3 & -1 & 14 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$$

Exercise 4 Recall from Monday that the linear transformation which rotates counterclockwise by an angle α has matrix



$$[Rot_\alpha] = \begin{bmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{bmatrix}$$

and

$$[Rot_\beta] = \begin{bmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{bmatrix}$$

use this in Friday FFT

Compute the product

$$[Rot_\beta][Rot_\alpha].$$

What do you see?

Wed Jan 31

- 2.1 Matrix operations

Announcements: quiz ☺

§2.1 Tuesday's notes: matrix multiplication

today's: more matrix algebra

Warm-up Exercise: Let $T\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) := \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ be projection from $\mathbb{R}^3$ to the $x_1 - x_2$ plane

a) Find the matrix A so that
$T(\vec{x}) = A\vec{x}$.

b) is T one to one? : NO (not true that $T(\vec{x}) = \vec{b}$ has unique solt'ns (when it has solt'ns)) (there were columns without pivots)

c) is T onto $\mathbb{R}^2$? : YES every row rref(A) has pivot i.e. $T(\vec{x}) = \vec{b}$ can always solved for $\vec{x}$

a) $\begin{bmatrix} 1 \cdot x_1 + 0x_2 + 0x_3 \\ 0x_1 + x_2 + 0x_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$

OR

$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3$

standard basis of $\mathbb{R}^3$
$\{\vec{e}_1, \vec{e}_2, \vec{e}_3\} = \left\{\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right\}$

T linear means

$$T(x_1\vec{e}_1 + x_2\vec{e}_2 + x_3\vec{e}_3)$$
$$= T(x_1\vec{e}_1) + T(x_2\vec{e}_2) + T(x_3\vec{e}_3)$$
$$= x_1 T(\vec{e}_1) + x_2 T(\vec{e}_2) + x_3 T(\vec{e}_3)$$
$$= \left[\, T(\vec{e}_1) \,\vdots\, T(\vec{e}_2) \,\vdots\, T(\vec{e}_3) \,\right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$
$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \quad \text{(same)}$$

$\boxed{\begin{array}{l} T \text{ linear} \\ T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v}) \\ T(c\vec{u}) = cT(\vec{u}) \end{array}}$

$T\left(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$T\left(\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

$T\left(\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

2.1 matrix algebra....we already talked about matrix multiplication. It interacts with matrix addition in interesting ways. We can also add and scalar multiply matrices of the same size, just treating them as oddly-shaped vectors:

Matrix algebra:

- addition and scalar multiplication: Let $A_{m \times n}$, $B_{m \times n}$ be two matrices of the same dimensions (m rows and n columns). Let $entry_{i\,j}(A) = a_{i\,j}$, $entry_{i\,j}(B) = b_{i\,j}$. (In this case we write $A = \left[a_{i\,j}\right]$, $B = \left[b_{i\,j}\right]$.) Let c be a scalar. Then

$$entry_{i\,j}(A + B) := a_{i\,j} + b_{i\,j}\,.$$
$$entry_{i\,j}(c\,A) := c\,a_{i\,j}\,.$$

In other words, addition and scalar multiplication are defined analogously as for vectors. In fact, for these two operations you can just think of matrices as vectors written in a rectangular rather than row or column format.

Exercise 1) Let $A := \begin{bmatrix} 1 & -2 \\ 3 & -1 \\ 0 & 3 \end{bmatrix}$ and $B := \begin{bmatrix} 0 & 27 \\ 5 & -1 \\ -1 & 1 \end{bmatrix}$. Compute $4\,A - B$.

$$4A - B = 4\begin{bmatrix} 1 & -2 \\ 3 & -1 \\ 0 & 3 \end{bmatrix} - \begin{bmatrix} 0 & 27 \\ 5 & -1 \\ -1 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 4 & -8 \\ 12 & -4 \\ 0 & 12 \end{bmatrix} - \begin{bmatrix} 0 & 27 \\ 5 & -1 \\ -1 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 4 & -35 \\ 7 & -3 \\ 1 & 11 \end{bmatrix}$$

vector properties of matrix addition and scalar multiplication

But other properties you're used to do hold:

- $+$ is commutative $\qquad A + B = B + A$

$$entry_{ij}(A + B) = a_{ij} + b_{ij} = b_{ij} + a_{ij} = entry_{ij}(B + A)$$

- $+$ is associative $\qquad (A + B) + C = A + (B + C)$

the ij entry of each side is $a_{ij} + b_{ij} + c_{ij}$

- scalar multiplication distributes over $+$ $\qquad c(A + B) = cA + cB$.

ij entry of LHS is $c\left(a_{ij} + b_{ij}\right) = c\left(a_{ij} + b_{ij}\right) = ij$ *entry of RHS*

More interesting are how matrix multiplication and addition interact:

Check some of the following. Let I_n be the $n \times n$ identity matrix, with $I_n \underline{\mathbf{x}} = \underline{\mathbf{x}}$ for all $\underline{\mathbf{x}} \in \mathbb{R}^n$. Let A, B, C have compatible dimensions so that the indicated expressions make sense. Then

a $A(BC) = (AB)C$ (associative property of multiplication)

$$A_{m\times n}(B_{n\times p}\, C_{pq}) \qquad (A_{m\times n} B_{n\times p}) C_{pq}$$

$$\underbrace{A_{m\times n}(BC)_{n\times q}}_{[A(BC)]_{m\times q}} \qquad (AB)_{m\times p} C_{p\times q}$$

$$[(AB)C]_{m\times q}$$

look at j^{th} columns:

$$\begin{aligned} \mathrm{col}_j(A(BC)) &= A\, \mathrm{col}_j(BC) \\ &= A(B\, \mathrm{col}_j C) \\ &= (AB)\, \mathrm{col}_j C \\ &= \mathrm{col}_j[(AB)C] \end{aligned}$$

because of composition rule!!

b $A(B + C) = AB + AC$ (left distributive law)

c $(A + B)C = AC + BC$ (right distributive law)

d $rAB = (rA)B = A(rB)$ for any scalar r.

$$A_{m\times n} I_{n\times n}$$

e) If $A_{m \times n}$ then $I_m A = A$ and $A I_n = A$.

Warning: $AB \neq BA$ in general. In fact, the sizes won't even match up if you don't use square matrices.

The transpose operation:

<u>Definition:</u> Let $B_{m \times n} = [b_{ij}]$. Then the <u>transpose</u> of B, denoted by B^T is an $n \times m$ matrix defined by

$$entry_{ij}(B^T) := entry_{ji}(B) = b_{ji}.$$

The effect of this definition is to turn the columns of B into the rows of B^T:

$$entry_i(col_j(B)) = b_{ij}.$$

$$entry_i(row_j(B^T)) = entry_{ji}(B^T) = b_{ij}.$$

And to turn the rows of B into the columns of B^T:

$$entry_j(row_i(B)) = b_{ij}$$

$$entry_j(col_i(B^T)) = entry_{ji}(B^T) = b_{ij}.$$

<u>Exercise 2)</u> explore these properties with the identity

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}.$$

Algebra of transpose:

<u>a</u> $(A^T)^T = A$

<u>b</u> $(A + B)^T = A^T + B^T$

<u>c</u> for every scalar r $(rA)^T = r A^T$

<u>d</u> (the only surprising property) $(A B)^T = B^T A^T$

Fri Feb 2

- 2.2 matrix inverses

Announcements:

- T/W notes.
  (Monday - Tuesday 2.2-2.3)
- all HW & quizzes should be graded & in "Return" folder.
- goal: FFT's returned Monday
- exam 1 next Friday. Feb 9

Warm-up Exercise: Compute

'til 12:58

$$\underbrace{\begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \end{bmatrix}}_{B}\underbrace{\begin{bmatrix} 1 & 2 \\ 0 & -1 \\ 1 & 0 \end{bmatrix}}_{A} = \begin{bmatrix} 4 & 0 \\ 3 & \boxed{-2} \end{bmatrix}$$

$1 \cdot 2 - 2 \cdot 1 + 3 \cdot 0 = 0$

$-1 \cdot 2 + 0 \cdot (-1) + 4 \cdot 0 = \boxed{-2}$

and

$$\begin{bmatrix} 1 & 2 \\ 0 & -1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \end{bmatrix} = \begin{bmatrix} -1 & 2 & 11 \\ \boxed{1} & 0 & -4 \\ \boxed{1} & 2 & 3 \end{bmatrix}$$

$$B_{m\times n} A_{n\times p} = [BA]_{mp}$$

- $col_j(BA) = B\, col_j(A)$

$BA \neq AB$

- $entry_{ij}(BA) = row_i(B) \cdot col_j(A)$

2.2 matrix inverses.

We've been talking about matrix algebra: addition, scalar multiplication, multiplication, and how these operations combine. But I haven't told you what the algebra on the previous page is good for. Today we'll start to find out. By way of comparison, think of a scalar linear equation with known numbers a, b, c, d and an unknown number x,

$$a\,x + b = c\,x + d$$

We know how to solve it by collecting terms and doing scalar algebra:

$$a\,x - c\,x = d - b$$

$$(a - c)\,x = d - b \qquad *$$

$$x = \frac{d - b}{a - c}\,.$$

How would you solve such an equation if A, B, C, D were square matrices, and X was a vector (or matrix) ? Well, you could use the matrix algebra properties we've been discussing to get to the $*$ step. And then if X was a vector you could solve the system * with Gaussian elimination. In fact, if X was a matrix, you could solve for each column of X (and do it all at once) with Gaussian elimination.

But you couldn't proceed as with scalars and do the final step after the * because it is not possible to divide by a matrix. Today we'll talk about a potential shortcut for that last step that is an analog of of dividing, in order to solve for X. It involves the concept of *inverse matrices.*

Matrix inverses: A square matrix $A_{n \times n}$ is invertible if there is a matrix $B_{n \times n}$ so that

$$AB = BA = I\,.$$

In this case we call B the inverse of A, and write $B = A^{-1}$.

Remark: A matrix A can have at most one inverse, because if we have two candidates B, C with

$$AB = BA = I \quad \text{and also} \quad AC = CA = I$$

then

$$(BA)C = IC = C$$
$$B(AC) = BI = B$$

so since the associative property $(BA)C = B(AC)$ is true, it must be that

$$B = C.$$

Exercise 1a) Verify that for $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ the inverse matrix is $B = \begin{bmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{bmatrix}$.

Inverse matrices are very useful in solving algebra problems. For example

Theorem: If A^{-1} exists then the only solution to $A\underline{\boldsymbol{x}} = \underline{\boldsymbol{b}}$ is $\underline{\boldsymbol{x}} = A^{-1}\underline{\boldsymbol{b}}$.

Exercise 1b) Use the theorem and A^{-1} in 2a, to write down the solution to the system

$$x + 2\,y = 5$$
$$3\,x + 4\,y = 6$$

Exercise 2a) Use matrix algebra to verify why the Theorem on the previous page is true. Notice that the correct formula is $\underline{\boldsymbol{x}} = A^{-1}\underline{\boldsymbol{b}}$ and not $\underline{\boldsymbol{x}} = \underline{\boldsymbol{b}}\,A^{-1}$ (this second product can't even be computed because the dimensions don't match up!).

Corollary: If A^{-1} exists, then the reduced row echelon form of A is the identity matrix.
proof: For a square matrix, solutions to $A\,\underline{\boldsymbol{x}} = \underline{\boldsymbol{b}}$ are unique only when A reduces to the identity, and when A^{-1} exists, the solutions to $A\,\underline{\boldsymbol{x}} = \underline{\boldsymbol{b}}$ are unique.

2b) Assuming A is a square matrix with an inverse A^{-1}, and that the matrices in the equation below have dimensions which make for meaningful equation, solve for X in terms of the other matrices:

$$XA + C = B$$

But where did that formula for A^{-1} come from?

<u>Answer:</u> Consider A^{-1} as an unknown matrix, $A^{-1} = X$. We want
$$A\,X = I\,.$$
We can break this matrix equation down by the columns of X. In the two by two case we get:

$$A\left[\begin{array}{c|c} col_1(X) & col_2(X) \end{array}\right] = \left[\begin{array}{c|c} 1 & 0 \\ 0 & 1 \end{array}\right].$$
In other words, the two columns of the inverse matrix X should satisfy
$$A\left(col_1(X)\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix},\quad A\left(col_2(X)\right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$
We can solve for both of these mystery columns at once, as we've done before when we had different right hand sides:

<u>Exercise 3:</u> Reduce the double augmented matrix
$$\left[\begin{array}{cc|c|c} 1 & 2 & 1 & 0 \\ 3 & 4 & 0 & 1 \end{array}\right]$$
to find the two columns of A^{-1} for the previous example.

For 2×2 matrices there's also a cool formula for inverse matrices:

Theorem: $\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1}$ exists if and only if the determinant $D = ad - bc$ of $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is non-zero. And in this case,

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

(Notice that the diagonal entries have been swapped, and minus signs have been placed in front of the off-diagonal terms. This formula should be memorized.)

Exercise 4: Check that the magic formula reproduces the answer you got in Exercise 3 for

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}^{-1}$$

Exercise 4: Will this always work? Can you find A^{-1} for

$$A := \begin{bmatrix} 1 & 5 & 1 \\ 2 & 5 & 0 \\ 2 & 7 & 1 \end{bmatrix}?$$

Exercise 5) Will this always work? Try to find B^{-1} for $B := \begin{bmatrix} 1 & 5 & 5 \\ 2 & 5 & 0 \\ 2 & 7 & 4 \end{bmatrix}$.

Hint: We'll discover that it's impossible for B to have an inverse.

Theorem: Let $A_{n \times n}$ be a square matrix. Then A has an inverse matrix if and only if its reduced row echelon form is the identity. In this case the algorithm illustrated on the previous page will always yield the inverse matrix.

explanation: By the previous theorem, when A^{-1} exists, the solutions to linear systems

$$A\,\underline{\boldsymbol{x}} = \underline{\boldsymbol{b}}$$

are unique ($\underline{\boldsymbol{x}} = A^{-1}\underline{\boldsymbol{b}}$). From our previous discussions about reduced row echelon form, we know that for square matrices, solutions to such linear systems exist and are unique only if the reduced row echelon form of A is the identity matrix. (Do you remember why?) Thus by logic, whenever A^{-1} exists, A reduces to the identity.

In this case that A does reduce to I, we search for A^{-1} as the solution matrix X to the matrix equation

$$A\,X = I$$

i.e.

$$A \left[\begin{array}{c|c|c|c} col_1(X) & col_2(X) & .... & col_n(X) \end{array} \right] = \left[\begin{array}{c|c|c|c} 1 & 0 & & 0 \\ 0 & 1 & & 0 \\ 0 & 0 & .... & 0 \\ 0 & 0 & & 1 \end{array} \right]$$

Because A reduces to the identity matrix, we may solve for X column by column as in the examples we just worked, by using a chain of elementary row operations:

$$[\,A\,|\,I\,] \rightarrow \rightarrow \rightarrow \rightarrow \rightarrow [\,I\,|\,B\,],$$

and deduce that the columns of X are exactly the columns of B, i.e. $X = B$. Thus we know that

$$A\,B = I\,.$$

To realize that $B\,A = I$ as well, we would try to solve $B\,Y = I$ for Y, and hope $Y = A$. But we can actually verify this fact by reordering the columns of $[\,I\,|\,B\,]$ to read $[\,B\,|\,I\,]$ and then reversing each of the elementary row operations in the first computation, i.e. create the chain

$$[\,B\,|\,I\,] \rightarrow \rightarrow \rightarrow \rightarrow \rightarrow [\,I\,|\,A\,]\,.$$

so $B\,A = I$ also holds. (This is one of those rare times when matrix multiplication actually is commuative.)

To summarize: If A^{-1} exists, then solutions $\underline{\boldsymbol{x}}$ to $A\,\underline{\boldsymbol{x}} = \underline{\boldsymbol{b}}$ always exist and are unique, so the reduced row echelon form of A is the identity. If the reduced row echelon form of A is the identity, then A^{-1} exists, because we can find it using the algorithm above. That's exactly what the Theorem claims.