Lecture 1: Systems of linear equations, Gaussian elimination.
(book: 1.1, 1.2)

x: chickens.
y: cows.

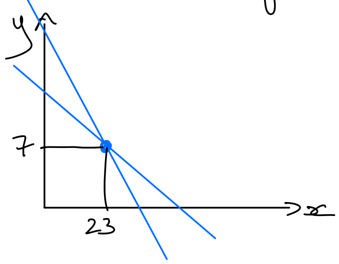
$$\begin{cases} x+y=30 \\ 2x+4y=74 \end{cases}$$

This is a system of linear equations (SLE)

2 equations in 2 variables.
We need a solution (a pair (x,y)) that works for both equations.

We have 2 variables, so we "live" in $\mathbb{R}^2$.
Each equation of this SLE represents a line in $\mathbb{R}^2$.

From a geometric/row point of view: we're looking for an intersection of those two lines.



In general solving an SLE means:
* $\mathbb{R}^2$: find intersection of lines in a plane.
* $\mathbb{R}^3$: " " " planes in a 3d-space.
* $\mathbb{R}^n$: " " " hyperplanes in a hyperspace

} a geometric/row point of view.

$$\begin{cases} x+y=30 \\ 2x+4y=74 \end{cases} \longrightarrow \begin{cases} x+y=30 \\ 2x+4y-2\cdot(x+y)=74-2\cdot 30 \end{cases} \longrightarrow \begin{cases} x+y=30 \\ \tfrac{1}{2}*2y=14*\tfrac{1}{2}. \end{cases}$$

$$\begin{cases} x+y-y=30-7 \\ y=7 \end{cases} \longrightarrow \begin{cases} x=23 \\ y=7 \end{cases}$$

So, the SLE has a solution and solution is unique.

Efficient procedure to solve an SLE: Gaussian elimination/row reduction.

An SLE can be summarized by:
* coefficient matrix A
* vector $\underline{b}$ of RHS numbers.

} augmented matrix $[A \,\vdots\, \underline{b}]$

$$\begin{cases} x+y=30 \\ 2x+4y=74 \end{cases} \qquad A=\begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix} \qquad \underline{b}=\begin{bmatrix} 30 \\ 74 \end{bmatrix} \qquad \left[\begin{array}{cc:c} 1 & 1 & 30 \\ 2 & 4 & 74 \end{array}\right]$$

2 x 2 matrix
rows x # columns.
equations x # variables.
$m \times n$

$$[A \vdots \underline{b}] = \left[\begin{array}{cc:c} 1 & 1 & 30 \\ 2 & 4 & 74 \end{array}\right] \underset{R_2 : R_2 - 2 \cdot R_1}{\sim} \left[\begin{array}{cc:c} 1 & 1 & 30 \\ 0 & 2 & 14 \end{array}\right] \underset{R_2 : R_2 * 1/2}{\sim} \left[\begin{array}{cc:c} 1 & 1 & 30 \\ 0 & 1 & 7 \end{array}\right]$$

$$\underset{R_1 : R_1 - R_2}{\sim} \left[\begin{array}{cc:c} 1 & 0 & 23 \\ 0 & 1 & 7 \end{array}\right] \qquad \begin{cases} 1 \cdot x + 0 \cdot y = 23 \\ 0 \cdot x + 1 \cdot y = 7 \end{cases} \qquad \begin{cases} x = 23 \\ y = 7 \end{cases}$$

So, the unique solution is $\begin{cases} x = 23 \\ y = 7 \end{cases}$

The corresponding SLEs of these augmented matrices are all equivalent ($\sim$) to each other: they share the same solution set.

Two matrices are row equivalent if one matrix can be changed into the other by means of a row operation:

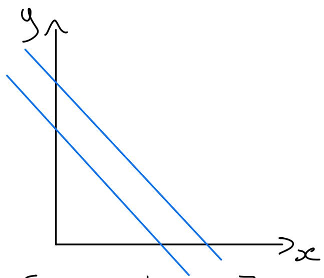
1. Replacement: add a scalar multiple of a row to another row.
2. Scaling: multiply a row by a <u>nonzero</u> scalar.
3. Interchange: swap two rows.

If the augmented matrices of two SLEs are row equivalent, then the two SLEs are equivalent.

An SLE can have:

- * one unique solution } SLE is consistent
- * infinitely many solutions }
- * no solution } SLE is inconsistent.

$$\begin{aligned} x+y&=30 \\ x+y&=49 \end{aligned}$$

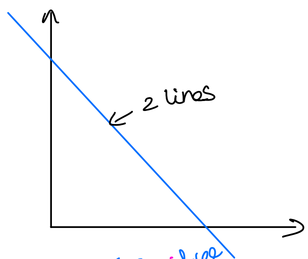


two parallel lines
→ no intersection
→ no solution
→ SLE is inconsistent.

$$\left[\begin{array}{cc:c} 1 & 1 & 30 \\ 1 & 1 & 49 \end{array}\right] \underset{R_2 : R_2 - R_1}{\sim} \left[\begin{array}{cc:c} 1 & 1 & 30 \\ 0 & 0 & 19 \end{array}\right] \qquad 0 = 19 \ \lightning$$

such a row $[0\,0 \cdots 0 \,\vdots\, \alpha]$, $\alpha \neq 0$, is typical for an inconsistent SLE.

$$\begin{cases} x+y=30 \\ 2x+2y=60 \end{cases}$$



they are the same line
→ infinitely many points of intersection
→ infinitely many solutions.

$$\left[\begin{array}{cc:c} 1 & 1 & 30 \\ 2 & 2 & 60 \end{array}\right] \underset{R_2 : R_2 - 2R_1}{\sim} \begin{array}{ccc} \text{basic var} & \text{free var.} & \\ \left[\begin{array}{c} 1 \\ 0 \end{array}\right. & \begin{array}{c} 1 \\ 0 \end{array} & \left.\begin{array}{:c} 30 \\ 0 \end{array}\right] \end{array} \qquad \begin{cases} x+y=30 \\ 0=0 \end{cases} \ ☺$$

Solution set in parametric form $\begin{cases} x = 30 - y \\ y \text{ is free} \end{cases}$.

y is a free variable (we can choose any value we want)
x is a basic variable (we cannot choose it)

Note: having a zero does not necessarily mean that there are infinitely many solutions.
For example:

$$\left[\begin{array}{cc:c} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \qquad \begin{cases} x=1 \\ y=0 \end{cases}$$

So, the SLE has a unique sol.

Let's formalize the row reduction / Gaussian elimination algorithm:

pivot (leading entry): leftmost nonzero element in a row

Row echelon form (REF): (not unique for a matrix)

$$\left[\begin{array}{ccc|c} -2 & 1 & 0 & 3 \\ 0 & 2 & 1 & 0 \\ 0 & 0 & -4 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right] \quad \text{or} \quad \left[\begin{array}{ccc|c} 2 & -3 & 2 & 1 \\ 0 & 0 & 4 & 8 \\ 0 & 0 & 0 & 5 \end{array}\right]$$

one unique sol.

no solution (because there is a pivot in the last column).

① All nonzero rows are above any zero row.
② Every pivot in a row is in a column to the right of the pivot of the row above it.
③ All entries below a pivot are zero.

It tells you:

* whether there is a solution (existence question)
  → is there a row $[0 \cdots 0 \,|\, \alpha]$ with $\alpha \neq 0$? (last column is a pivot column)

* whether the solution is unique (uniqueness question)
  → are there free variables?
  → are all variables basic variables?
  → does every column in the coefficient matrix have a pivot?

Reduced row echelon form (RREF): (unique for a matrix).

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{array}\right] \quad \begin{cases} x_1 = 2 \\ x_2 = 0 \\ x_3 = -1 \end{cases} \quad \text{or} \quad \left[\begin{array}{ccc|c} 1 & 0 & 1 & -1 \\ 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 \end{array}\right] \quad \begin{cases} x_1 = -1 - x_3 \\ x_2 = -2 \\ x_3 \text{ is free} \end{cases}$$

④ All pivots are equal to 1.
⑤ Each pivot is the only nonzero entry in its column.

It tells you:

* the solution.

Example (Gaussian elimination algorithm):

$$\begin{cases} \quad\; 2x_2 - 8x_3 = 8 \\ x_1 - 2x_2 + x_3 = 0 \\ -4x_1 + 5x_2 + 9x_3 = -9 \end{cases}$$

$$\left[\begin{array}{ccc|c} 0 & 2 & -8 & 8 \\ 1 & -2 & 1 & 0 \\ -4 & 5 & 9 & -9 \end{array}\right] \underset{R_1 \leftrightarrow R_2}{\sim}$$

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ -4 & 5 & 9 & -9 \end{array}\right] \underset{R_3 : R_3 + 4\cdot R_1}{\sim}$$

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & -3 & 13 & -9 \end{array}\right] \underset{R_3 : R_3 + 3/2\, R_2}{\sim}$$

$$\left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 0 & 1 & 3 \end{array}\right] \underset{\substack{R_1 : R_1 - R_3 \\ R_2 : R_2 + 8R_3}}{\sim}$$

echelon form!
the SLE is consistent
the solution is unique.

$$\left[\begin{array}{ccc|c} 1 & -2 & 0 & -3 \\ 0 & 2 & 0 & 32 \\ 0 & 0 & 1 & 3 \end{array}\right] \underset{R_2 : R_2 * 1/2}{\sim}$$

$$\left[\begin{array}{ccc|c} 1 & -2 & 0 & -3 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{array}\right] \underset{R_1 : R_1 + 2\cdot R_2}{\sim}$$

$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 16 \\ 0 & 0 & 1 & 3 \end{array}\right]$$

reduced echelon form!

$$\begin{cases} x_1 = 29 \\ x_2 = 16 \\ x_3 = 3 \end{cases}.$$

Lecture 2: Vector equations, Matrix equations.
(book: 1.3, 1.4)

Previous lecture: geometric /row point of view to an SLE
+ Gaussian elimination.

Today: column point of view to an SLE.

$$x + y = 30$$
$$2x + 4y = 74$$

$$A = \begin{bmatrix} 1 & 1 \\ 2 & 4 \end{bmatrix} \qquad \underline{b} = \begin{bmatrix} 30 \\ 74 \end{bmatrix}$$

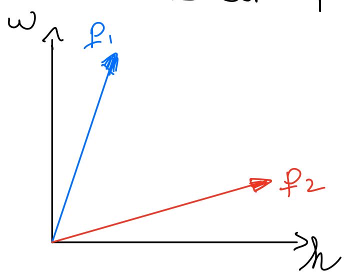
From a machine learning perspective: columns of A are feature vectors (vectors that collect features that we measured for each animal).

For instance, height and weight $\underline{f} = \begin{bmatrix} h \\ w \end{bmatrix}$

Collect data from N animals → N feature vectors:

$$\underline{f}_1 = \begin{bmatrix} h_1 \\ w_1 \end{bmatrix}, \underline{f}_2 = \begin{bmatrix} h_2 \\ w_2 \end{bmatrix}, \dots\dots, \underline{f}_N = \begin{bmatrix} h_N \\ w_N \end{bmatrix}.$$

These vectors are points in the feature space ($\mathbb{R}^2$).

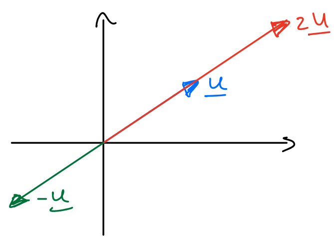


In general: n features → (n×1) vector $\underline{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \in \mathbb{R}^n$

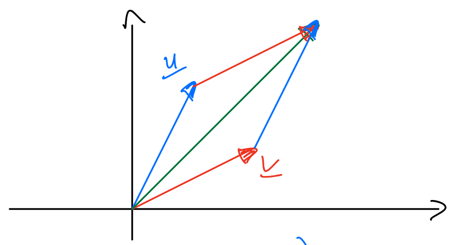
$$\begin{bmatrix} 2 \\ 3 \end{bmatrix} \neq \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

We can perform operations with vectors:

* scaling $\underline{u} = \begin{bmatrix} 3 \\ 2 \end{bmatrix}$ $\qquad 2\underline{u} = \begin{bmatrix} 6 \\ 4 \end{bmatrix}$ $\qquad -\underline{u} = (-1)\cdot\underline{u} = \begin{bmatrix} -3 \\ -2 \end{bmatrix}.$



* addition $\underline{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ $\underline{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ $\underline{u}+\underline{v} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$



(parallellogram rule)

Algebraic properties of $\mathbb{R}^n$: book p53.

Combining these two operations:

Given vectors $\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_p\}$ in $\mathbb{R}^n$ and $c_1, \dots, c_p$ the vector

$$\underline{y} = c_1 \cdot \underline{v}_1 + c_2 \cdot \underline{v}_2 + \cdots + c_p \cdot \underline{v}_p$$

is called a linear combination of $\underline{v}_1, \dots, \underline{v}_p$ with weights $c_1, \dots c_p$.

Example: $\underline{u} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ $\underline{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ $\underline{b} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$

Q: can $\underline{b}$ be written as a linear combination of $\underline{u}$ and $\underline{v}$?
i.e., can we find c_1 and c_2 such that $c_1 \cdot \underline{u} + c_2 \cdot \underline{v} = \underline{b}$?

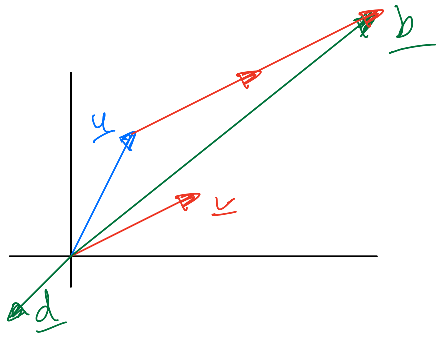
$$c_1 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} + c_2 \cdot \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix} ?$$ ← vector equation

$$\begin{bmatrix} c_1 \\ 2c_1 \end{bmatrix} + \begin{bmatrix} 2c_2 \\ c_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$$

$$\begin{bmatrix} c_1 + 2c_2 \\ 2c_1 + c_2 \end{bmatrix} = \begin{bmatrix} 5 \\ 4 \end{bmatrix}$$

$$\begin{cases} c_1 + 2c_2 = 5 \\ 2c_1 + c_2 = 4 \end{cases}$$ ← hey, that is an SLE!

* Every SLE can be written as a vector equation, and the other way around.
* Solving an SLE means investigating whether $\underline{b}$ can be written as a linear combination of the columns of $\underline{A}$.
(column point of view).



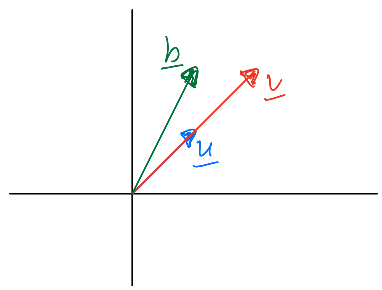
$\underline{u} = \begin{bmatrix}1\\2\end{bmatrix}$ $\underline{v} = \begin{bmatrix}2\\1\end{bmatrix}$

$\underline{b} = \underline{u} + 2\cdot\underline{v}$.

$\underline{d} = \begin{bmatrix}-1\\-1\end{bmatrix}$

$-1/3\ \underline{u} + (-1/3)\ \underline{v}$

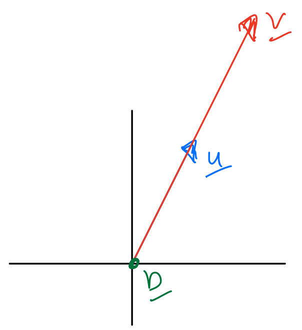
Example: (no sol.) $\underline{u} = \begin{bmatrix}1\\1\end{bmatrix}$ $\underline{v} = \begin{bmatrix}2\\2\end{bmatrix}$ $\underline{b}\ \begin{bmatrix}1\\2\end{bmatrix}$



Hence, there is no way to obtain $\underline{b}$ by taking a linear combination of $\underline{u}$ and $\underline{v}$.

Example (∞ many solutions in $\mathbb{R}^2$): $\underline{u} = \begin{bmatrix}1\\2\end{bmatrix}$ $\underline{v} = \begin{bmatrix}2\\4\end{bmatrix}$ $b = \begin{bmatrix}0\\0\end{bmatrix}$

$$2\cdot\underline{u} + (-1)\cdot\underline{v} = \underline{b}$$
$$0\cdot\underline{u} + 0\cdot\underline{v} = \underline{b}$$
$$u + (-1/2)\cdot\underline{v} = \underline{b}.$$



There are ∞ many ways to linearly combine $\underline{u}$ and $\underline{v}$.

$$\begin{cases} x_1 + 2x_2 = 0 \\ 2x_1 + 4x_2 = 0 \end{cases} \qquad \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 2 & 4 & 0 \end{array}\right] \underset{R_2 : R_2 - 2\cdot R_1}{\sim} \left[\begin{array}{cc|c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

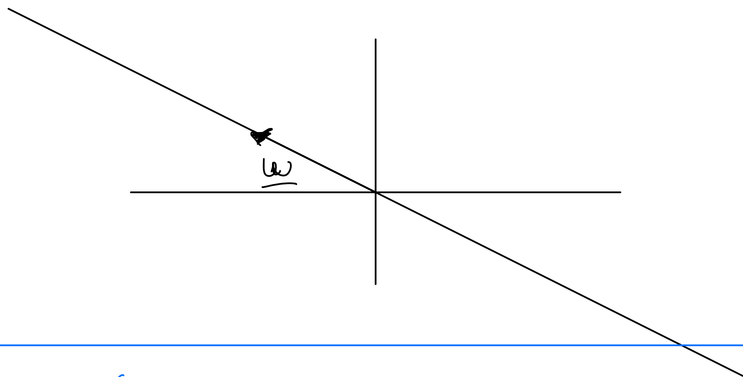
$$\begin{cases} x_1 = -2x_2 \\ x_2 \text{ is free} \end{cases} \quad \text{(parametric form)}$$

$$\underline{x} = \begin{bmatrix}x_1\\x_2\end{bmatrix} = \begin{bmatrix}-2x_2\\x_2\end{bmatrix} = x_2\cdot\begin{bmatrix}-2\\1\end{bmatrix} \quad \text{(parametric vector form)}$$

Hence, the solution set is $x_2 \cdot \underline{w}$, where $\underline{w} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$,

i.e., Span $\{\underline{w}\}$

$\hookrightarrow$ any scalar multiple of $\underline{w}$.



any solution on this line is a solution to the SLE.

Example (∞ many solutions in $\mathbb{R}^3$)

$$\begin{aligned} x_1 - 3x_2 + 2x_3 &= 0 \\ 2x_1 - 6x_2 + 4x_3 &= 0 \end{aligned}$$

$$\left[\begin{array}{ccc|c} 1 & -3 & 2 & 0 \\ 2 & -6 & 4 & 0 \end{array}\right] \sim_{R_2 : R_2 - 2 \cdot R_1} \left[\begin{array}{ccc|c} 1 & -3 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\begin{cases} x_1 = 3x_2 - 2x_3 \\ x_2 \text{ is free} \\ x_3 \text{ is free} \end{cases}$$ parametric form

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 3x_2 - 2x_3 \\ x_2 \\ x_3 \end{bmatrix} = x_2 \cdot \underbrace{\begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}}_{\underline{w}_1} + x_3 \cdot \underbrace{\begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix}}_{\underline{w}_2}$$

Hence, a solution is any linear combination of $\underline{w}_1$ and $\underline{w}_2$.

Span $\{\underline{w}_1, \underline{w}_2\}$

So, the solution set is a plane in $\mathbb{R}^3$.

Given a set of p vectors $\{\underline{v}_1, \ldots, \underline{v}_p\}$ in $\mathbb{R}^n$, the span of this set of vectors is the set of all possible linear combinations of the vectors in this set.

$\longrightarrow$ Span $\{\underline{v}_1, \ldots, \underline{v}_p\}$ contains any vector $\underline{y}$ that can be written as

$$\underline{y} = c_1 \cdot \underline{v}_1 + c_2 \cdot \underline{v}_2 + \cdots + c_p \cdot \underline{v}_p.$$

Hence, solving an SLE boils down to investigating whether $\underline{b}$ belongs to the Span of the columns of A.

We can see a linear combination of vectors as the product of a matrix (A) and a vector ($\underline{x}$).

Definition of $A\underline{x}$: linear combination of the columns of A with the entries of $\underline{x}$ being the weights.

$$A\underline{x} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = 4 \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 3 \cdot \begin{bmatrix} 2 \\ -5 \end{bmatrix} + 7 \cdot \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}.$$

More efficient $A\underline{x} = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix} \begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = \begin{bmatrix} 1\cdot 4 + 2\cdot 3 + (-1)\cdot 7 \\ 0\cdot 4 + (-5)\cdot 3 + 3\cdot 7 \end{bmatrix} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$

We need: # columns of A = # row/entries of $\underline{x}$.

$\boxed{\text{Properties of the matrix-vector product: p. 65}}$

Three things with the same solution set:

* the SLE with augmented matrix $[\underline{a}_1\ \underline{a}_2 \cdots \underline{a}_n \mid \underline{b}]$
* the vector equation $x_1 \cdot \underline{a}_1 + x_2 \cdot \underline{a}_2 + \cdots + x_n \cdot \underline{a}_n = \underline{b}$.
* the matrix equation $[\underline{a}_1\ \underline{a}_2 \cdots \underline{a}_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \underline{b}$. $A\underline{x} = \underline{b}$

Example: $A = \begin{bmatrix} 1 & 5 & -3 \\ 0 & 1 & -2 \\ -2 & -10 & 6 \end{bmatrix}$ Q: Is $A\underline{x} = \underline{b}$ consistent for every $\underline{b} \in \mathbb{R}^3$? No!

$$\left[\begin{array}{ccc|c} 1 & 5 & -3 & b_1 \\ 0 & 1 & -2 & b_2 \\ -2 & -10 & 6 & b_3 \end{array}\right] \underset{R_3:\ R_3 + 2\cdot R_1}{\sim} \left[\begin{array}{ccc|c} 1 & 5 & -3 & b_1 \\ 0 & 1 & -2 & b_2 \\ 0 & 0 & 0 & b_3 + 2b_1 \end{array}\right]$$

The SLE is consistent iff $b_3 + 2b_1 = 0$.

For example, the SLE is inconsistent if $\underline{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

consistent if $\underline{b} = \begin{bmatrix} 1 \\ 2 \\ -2 \end{bmatrix}$

---

4 equivalent statements:

* The columns of A span $\mathbb{R}^m$.
* Each $\underline{b} \in \mathbb{R}^m$ is a linear combination of the columns of A.
* For each $\underline{b} \in \mathbb{R}^m$, $A\underline{x} = \underline{b}$ has a solution.
* A has a pivot position in every row.
  ↳ (coefficient matrix!)

Lecture 3: Solution sets, linear independence
(book: 1.5, 1.7)

Previous lecture: column point of view to an SLE.

Today: homogeneous / nonhomogeneous SLE
+ linear independence.

Homogeneous SLE : $A\underline{x} = \underline{0}$
Is it always consistent? Yes, as there is the trivial solution $\underline{x} = \underline{0}$.

Is there also a nontrivial solution ?
No free variables $\to$ No.
At least one free variable $\to$ Yes.

$$\begin{aligned} 2x_1 + 4x_2 &= 0 \\ x_1 + 2x_2 &= 0 \end{aligned}$$ homogeneous SLE.

$$\left[\begin{array}{cc:c} 2 & 4 & 0 \\ 1 & 2 & 0 \end{array}\right] \underset{R_2 : R_2 - \frac{1}{2}R_1}{\sim} \left[\begin{array}{cc:c} 2 & 4 & 0 \\ 0 & 0 & 0 \end{array}\right] \underset{R_1 * 1/2}{\sim} \left[\begin{array}{cc:c} 1 & 2 & 0 \\ 0 & 0 & 0 \end{array}\right]$$

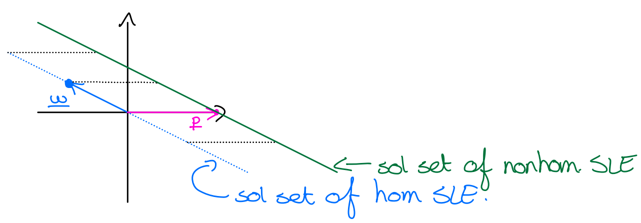
$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -2x_2 \\ x_2 \end{bmatrix} = x_2 \underbrace{\begin{bmatrix} -2 \\ 1 \end{bmatrix}}_{\underline{w}} \qquad \text{Span}\left\{\begin{bmatrix} -2 \\ 1 \end{bmatrix}\right\}.$$

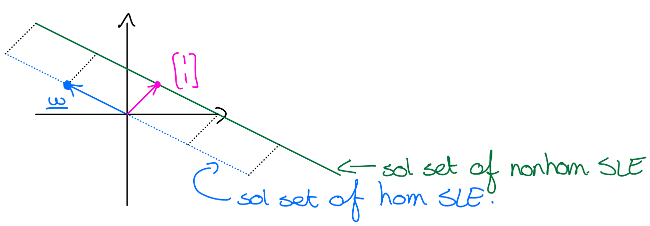
$$\begin{aligned} 2x_1 + 4x_2 &= 6 \\ x_1 + 2x_2 &= 3 \end{aligned}$$ nonhomogeneous SLE.

$$\left[\begin{array}{cc:c} 2 & 4 & 6 \\ 1 & 2 & 3 \end{array}\right] \underset{R_2 : R_2 - \frac{1}{2}R_1}{\sim} \left[\begin{array}{cc:c} 2 & 4 & 6 \\ 0 & 0 & 0 \end{array}\right] \underset{R_1 * 1/2}{\sim} \left[\begin{array}{cc:c} 1 & 2 & 3 \\ 0 & 0 & 0 \end{array}\right]$$

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 - 2x_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} + \begin{bmatrix} -2x_2 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

particular sol of the nonhomogeneous SLE.





Observation: the sol. set of $A\underline{x} = \underline{b}$ (when non-empty) is a **translation** of the sol. set of $A\underline{x} = \underline{0}$ for a special vector $\underline{p}$, where $\underline{p}$ is a particular solution of the nonhom. SLE (take $x_2 = 0$).
Any particular solution works.

Theorem: Assume $A\underline{x} = \underline{b}$ is consistent, and let $\underline{p}$ be a particular solution of $A\underline{x} = \underline{b}$. So, $A\underline{p} = \underline{b}$. Then,

Set of all solutions of $A\underline{x} = \underline{b}$

$=$

Set of vectors that can be written as $\underline{q} + \underline{p}$, where $A\underline{q} = \underline{0}$.

Proof:

"$\subseteq$" Let $\underline{v}$ be a solution of $A\underline{x} = \underline{b}$, so $A\underline{v} = \underline{b}$.
We need to show that we can write $\underline{v} = \underline{q} + \underline{p}$, where $A\underline{q} = \underline{0}$.
So, we need to show that $A\underline{q} = \underline{0}$, where $\underline{q} = \underline{v} - \underline{p}$.
Here we go:
$A\underline{q} = A(\underline{v} - \underline{p}) = A\underline{v} - A\underline{p} = \underline{b} - \underline{b} = \underline{0}$ ✓

"$\supseteq$" Let $\underline{v}$ be a vector such that $\underline{v} = \underline{q} + \underline{p}$, where $A\underline{q} = \underline{0}$.
We need to show that $\underline{v}$ is a solution of $A\underline{x} = \underline{b}$.
Here we go:
$A\underline{v} = A(\underline{q} + \underline{p}) = A\underline{q} + A\underline{p} = \underline{0} + \underline{b} = \underline{b}$ ✓ $\square$

Conclusion:
If we want to solve an SLE $A\underline{x} = \underline{b}$, and we already know the sol. set of the corresponding $A\underline{x} = \underline{0}$, there are three possibilities:

* Row reduce $[A \,\vdots\, \underline{b}]$
* Re-apply the row operations, but now only to $\underline{b}$.
* If we can easily spot a particular solution for $A\underline{x} = \underline{b}$, we add this solution to the sol. set of $A\underline{x} = \underline{0}$.

The set $\{\underline{v}_1, \ldots, \underline{v}_p\}$ is **linearly independent** if

$$c_1 \underline{v}_1 + c_2 \underline{v}_2 + \cdots + c_p \underline{v}_p = \underline{0} \quad \text{implies} \quad c_1 = c_2 = \cdots = c_p = 0.$$

(it has **only** the **trivial solution**).

Otherwise: it's called **linearly dependent**.

Examples: $\begin{bmatrix}1\\1\end{bmatrix}, \begin{bmatrix}1\\-2\end{bmatrix}, \begin{bmatrix}2\\1\end{bmatrix}$ lin indep. ?

For example, $5\cdot\begin{bmatrix}1\\1\end{bmatrix} + 1\cdot\begin{bmatrix}1\\-2\end{bmatrix} + (-3)\cdot\begin{bmatrix}2\\1\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$ So, lin dep.

$\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}$ lin indep?

$c_1\cdot\begin{bmatrix}1\\0\end{bmatrix} + c_2\cdot\begin{bmatrix}0\\1\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$

$\begin{bmatrix}c_1+0\\0+c_2\end{bmatrix} = \begin{bmatrix}0\\0\end{bmatrix}$ $\Rightarrow c_1=0$ $\Rightarrow c_2=0.$

Only the trivial solution. So, lin indep.

How can we answer this question in general?
Consider the corresponding homogeneous SLE and reduce it to REF.
* no free variables $\to$ unique sol (only the trivial sol) $\to$ lin indep.
* some free variables $\to$ infinitely many sols $\to$ lin dep.

If a set contains more vectors than there are entries in each vector.

$$4\Big[\underbrace{\qquad\qquad\qquad\qquad\qquad}_{}\Big]^{10}$$

$\to$ more columns than rows.
$\to$ there must be a column without a pivot.
$\to$ some free variables.
$\to$ lin dep.

What about a set containing only one vector? is $\{\underline{v}\}$ lin indep?

$\begin{bmatrix}0\\0\end{bmatrix}$ lin dep. $\begin{bmatrix}1\\1\end{bmatrix}$ lin indep

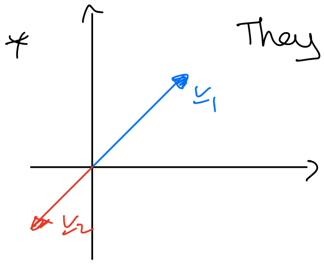
$\begin{bmatrix}1\\0\end{bmatrix}$ lin indep.

* if $\underline{v}\neq\underline{0}$, then we need $c=0$ (only trivial sol). So, $\{\underline{v}\}$ is lin indep.
* if $\underline{v}=\underline{0}$, then c can be anything (also nontrivial sol). So, $\{\underline{v}\}$ is lin dep.

What about a set containing the zero vector?
Is $\{\underline{v}_1,\dots,\underline{v}_p,\underline{0}\}$ lin indep? $c_1\underline{v}_1+c_2\underline{v}_2+\dots+c_p\underline{v}_p+c_{p+1}\underline{0}=\underline{0}.$
$c_1=\dots=c_p=0,\ c_{p+1}=8$ is for example a nontrivial sol.
So, a set containing the zero vector is always lin dep.

What about a set with two vectors? is $\{\underline{v}_1, \underline{v}_2\}$ lin dep?
Assume $\underline{v}_1 \neq \underline{0}$ and $\underline{v}_2 \neq \underline{0}$.

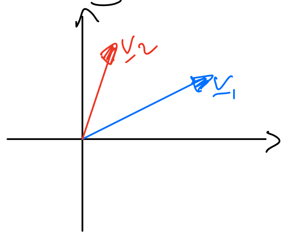
* They lie on the same line.



$$\underline{v}_2 = -\tfrac{2}{3}\,\underline{v}_1$$
$$\underline{v}_2 + \tfrac{2}{3}\,\underline{v}_1 = \underline{0}$$
$$\tfrac{2}{3}\,\underline{v}_1 + \underline{v}_2 = \underline{0}$$

→ we found a nontriv. sol.
→ $\{\underline{v}_1, \underline{v}_2\}$ is lin dep.

* They do not lie on the same line.



→ lin indep.

Proof: by contradiction :)

Suppose $\{\underline{v}_1, \underline{v}_2\}$ is lin dep.

⟹ $c_1 \cdot \underline{v}_1 + c_2 \cdot \underline{v}_2 = \underline{0}$ non triv. sol.

Suppose $c_1 \neq 0$.
→ $c_1 \underline{v}_1 + c_2 \cdot \underline{v}_2 = \underline{0}$
→ $c_1 \underline{v}_1 = -c_2 \cdot \underline{v}_2$
→ $\underline{v}_1 = -\dfrac{c_2}{c_1} \cdot \underline{v}_2$ ↯

So, $c_1 = 0$. So, $c_2 \neq 0$.

$$c_2 \cdot \underline{v}_2 = \underline{0} \quad ↯$$

$c_2 \neq 0$ $\quad$ $\underline{v}_2 \neq \underline{0}$

So, $\{\underline{v}_1, \underline{v}_2\}$ is lin indep.

$\{\underline{v}_1, \underline{v}_2, \underline{v}_3, \underline{v}_4, \underline{v}_5\}$ lin indep?

$\underline{v}_4$ is a lin comb of $\underline{v}_1, \underline{v}_2, \underline{v}_3$.

$$\underline{v}_4 = 2 \cdot \underline{v}_1 + (-8) \cdot \underline{v}_2 + 3.5 \cdot \underline{v}_3.$$

$$c_1 \cdot \underline{v}_1 + c_2 \cdot \underline{v}_2 + c_3 \cdot \underline{v}_3 + c_4 \cdot \underline{v}_4 + c_5 \cdot \underline{v}_5 = \underline{0} \;?$$

$c_1 = 2$, $c_2 = -8$, $c_3 = 3.5$, $c_4 = -1$, $c_5 = 0$

non triv sol
lin dep.

Theorem: $\{\underline{v}_1, \dots, \underline{v}_p\}$ is lin dep $\Longleftrightarrow$ at least one of the vectors is a linear combination of the others.

Proof: "$\Leftarrow$" Assume $\underline{v}_j = c_1 \cdot \underline{v}_1 + \dots + c_{j-1}\underline{v}_{j-1} + c_{j+1}\underline{v}_{j+1} + \dots + c_p \cdot \underline{v}_p$
Then, $c_1 \cdot \underline{v}_1 + \dots + c_{j-1}\underline{v}_{j-1} + (-1)\cdot \underline{v}_j + c_{j+1}\underline{v}_{j+1} + \dots + c_p \underline{v}_p = \underline{0}$.
The weight of $\underline{v}_j$ is nonzero.
So, we found a nontrivial sol.
So, $\{\underline{v}_1, \dots, \underline{v}_p\}$ is lin dep

"$\Rightarrow$" Assume $\{\underline{v}_1, \dots, \underline{v}_p\}$ is lin dep.
Distinguish between two cases:

Case 1: $\underline{v}_1 = \underline{0}$.
Then $\underline{v}_1 = 0 \cdot \underline{v}_2 + \dots + 0 \cdot \underline{v}_p$
So, $\underline{v}_1$ is a lin comb. of the others.

Case 2: $\underline{v}_1 \neq \underline{0}$.

Since $\{\underline{v}_1, \dots, \underline{v}_p\}$ is lin dep, there is a nontrivial sol $c_1 \cdot \underline{v}_1 + \dots + c_p \cdot \underline{v}_p = \underline{0}$.
Let j be the largest subscript for which $c_j \neq 0$.
Note: this subscript exists because it is a nontrivial sol.
Moreover, note that $j=1$ would imply $c_1 \cdot \underline{v}_1 = \underline{0}$, which is not possible because $c_1 \neq 0$ and $\underline{v}_1 \neq \underline{0}$.
Hence, $j > 1$ and $c_1 \underline{v}_1 + \dots + c_j \underline{v}_j + 0 \cdot \underline{v}_j + \dots + 0 \cdot \underline{v}_p = \underline{0}$

$$\Rightarrow c_j \underline{v}_j = -c_1 \underline{v}_1 - \dots - c_{j-1}\underline{v}_{j-1} + 0 \cdot \underline{v}_{j+1} + \dots + 0 \cdot \underline{v}_p$$

$$\Rightarrow \underline{v}_j = \frac{-c_1}{c_j}\underline{v}_1 + \dots + \frac{-c_{j-1}}{c_j}\underline{v}_{j-1} + 0 \cdot \underline{v}_{j+1} + \dots + 0 \cdot \underline{v}_p$$

So, $\underline{v}_j$ is a lin comb. of the others. $\square$

So, we actually also already proved:
If $\{\underline{v}_1, \dots, \underline{v}_p\}$ is lin dep. and $\underline{v}_1 \neq \underline{0}$, then there is a $j \in \{2, \dots, p\}$ such that $\underline{v}_j$ is a lin comb. of $\{\underline{v}_1, \dots, \underline{v}_{j-1}\}$.

Lecture 4: Linear transformations, Matrix algebra
(book: 1.8, 1.9, 2.1).

Previous lecture: homogeneous / nonhomogeneous SLE
+ linear independence.

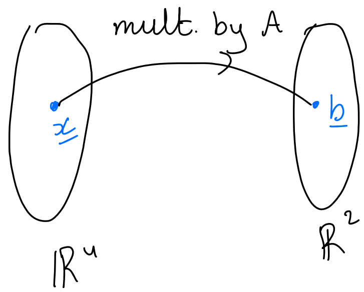
Recall the matrix-vector product:

$$\underbrace{\begin{bmatrix} 4 & -3 & 1 & 3 \\ 2 & 0 & 5 & 1 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}}_{\underline{x}} = \underbrace{\begin{bmatrix} 13 \\ 21 \end{bmatrix}}_{\underline{b}}$$

And recall the properties:
* $A(\underline{u}+\underline{v}) = A\underline{u} + A\underline{v}$
* $A(c\cdot\underline{u}) = c\cdot(A\underline{u})$.

Multiplication by A transforms $\underline{x}$ into $\underline{b}$.

Schematic



Transformation / function / mapping

$T: \mathbb{R}^m \to \mathbb{R}^n$
domain codomain.

$T(\underline{x}) = \underline{y} \Leftarrow$ output
the transformation operator. input

image: $T(\underline{x})$
range: set of all images.

A transformation is linear if:
* $T(\underline{u}+\underline{v}) = T(\underline{u}) + T(\underline{v})$
* $T(c\cdot\underline{u}) = c\cdot T(\underline{u})$.

$\Big\}$ $T(c\cdot\underline{u} + d\cdot\underline{v}) = c\cdot T(\underline{u}) + d\cdot T(\underline{v})$.

If T is linear, then $T(\underline{0}) = \underline{0}$.

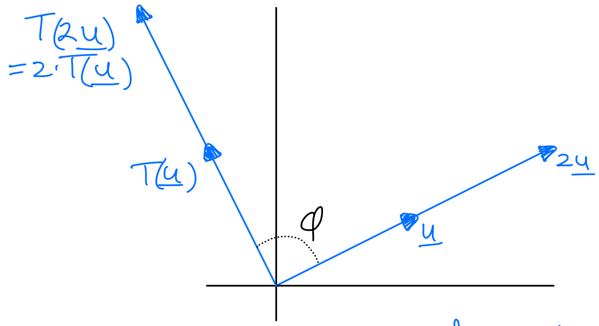
$T: \mathbb{R}^2 \to \mathbb{R}^2$ $\quad T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x^2 \\ y^2 \end{bmatrix}$.

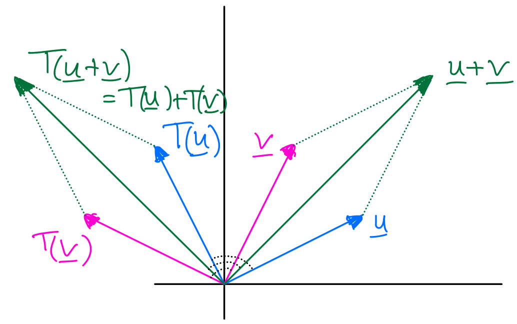
$\underline{u} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$ $\quad \underline{v} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}$ $\quad T(\underline{u}) = \begin{bmatrix} 9 \\ 9 \end{bmatrix}$ $\quad T(\underline{v}) = \begin{bmatrix} 16 \\ 16 \end{bmatrix}$ $\quad T(\underline{u}) + T(\underline{v}) = \begin{bmatrix} 25 \\ 25 \end{bmatrix}$

$T(\underline{u}+\underline{v}) = T\left(\begin{bmatrix} 7 \\ 7 \end{bmatrix}\right) = \begin{bmatrix} 49 \\ 49 \end{bmatrix}$ $\neq$

So, T is not linear.

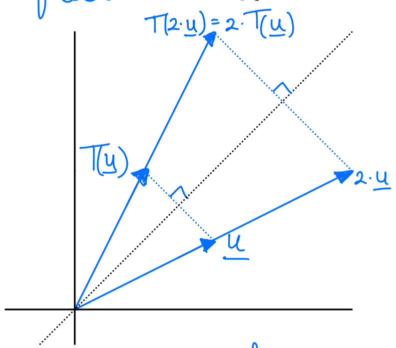
Example: rotation about the origin through an angle φ

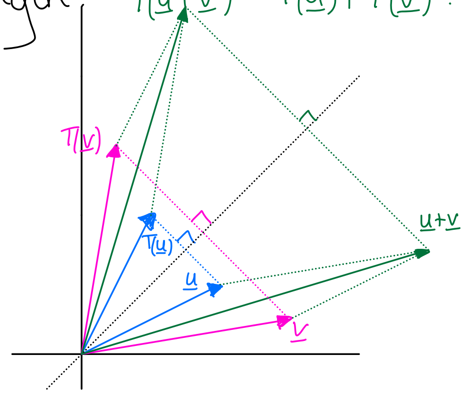




So, it is a linear transformation.

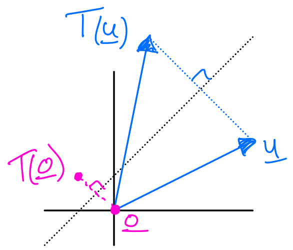
Example: reflection in a line through the origin, $T(\underline{u}+\underline{v}) = T(\underline{u})+T(\underline{v})$.





So, it is a linear transformation.

Example: reflection in a line not through the origin



not a linear transformation because $T(\underline{0}) \neq \underline{0}$.

Let's go back to the transformation of a matrix-vector product.

$$\underset{m\times n}{A}\,\underset{n\times 1}{\underline{u}} = \underset{m\times 1}{\underline{b}} \qquad \mathbb{R}^n \to \mathbb{R}^m \qquad T(\underline{u}) = A\underline{u}.$$

Is a matrix transformation linear?

$* \; T(\underline{u}+\underline{v}) = A(\underline{u}+\underline{v}) = A\underline{u}+A\underline{v} = T(\underline{u})+T(\underline{v})$

$* \; T(c\cdot\underline{u}) = A(c\cdot\underline{u}) = c\cdot(A\underline{u}) = c\cdot T(\underline{u})$.

follow from the properties of a matrix-vector product.

$\Rightarrow$ Every matrix transformation is a linear transformation.
The opposite is also true (at least, in the context: $\mathbb{R}^n \to \mathbb{R}^m$)

Theorem: Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. There is a unique matrix A such that for $\underline{x} \in \mathbb{R}^n$ $T(\underline{x}) = A\underline{x}$.

Proof: Let $\underline{x} \in \mathbb{R}^n$.

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix} = x_1 \underbrace{\begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\underline{e}_1} + x_2 \underbrace{\begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\underline{e}_2} + x_3 \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}}_{\underline{e}_3} + \cdots + x_n \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}}_{\underline{e}_n}.$$

Then, $T(\underline{x}) = T(x_1 \cdot \underline{e}_1 + x_2 \cdot \underline{e}_2 + x_3 \cdot \underline{e}_3 + \cdots + x_n \cdot \underline{e}_n)$

$$\overset{\text{linearity}}{=} T(x_1 \underline{e}_1) + T(x_2 \underline{e}_2) + T(x_3 \underline{e}_3) + \cdots T(x_n \cdot \underline{e}_n)$$

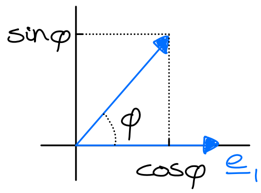
$$\overset{\text{linearity}}{=} x_1 T(\underline{e}_1) + x_2 T(\underline{e}_2) + x_3 T(\underline{e}_3) + \cdots + x_n T(\underline{e}_n).$$

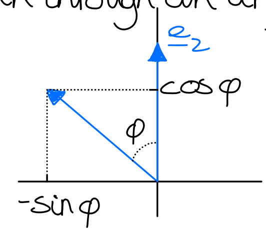
$$= \underbrace{\begin{bmatrix} | & | & | & & | \\ T(\underline{e}_1) & T(\underline{e}_2) & T(\underline{e}_3) & \cdots & T(\underline{e}_n) \\ | & | & | & & | \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{bmatrix}}_{\underline{x}} = A\underline{x} \quad \square$$

Uniqueness of A? DIY (exc. 41 Ch 1.9).

Standard matrix for the linear transformation T: $[T(\underline{e}_1) \ \ldots \ T(\underline{e}_n)]$

Example: rotation about the origin through an angle φ $\quad T: \mathbb{R}^2 \to \mathbb{R}^2$





So, $T(\underline{e}_1) = \begin{bmatrix} \cos\varphi \\ \sin\varphi \end{bmatrix}$ $\qquad$ So, $T(\underline{e}_2) = \begin{bmatrix} -\sin\varphi \\ \cos\varphi \end{bmatrix}$.

So, $A = \begin{bmatrix} T(\underline{e}_1) & T(\underline{e}_2) \end{bmatrix} = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$

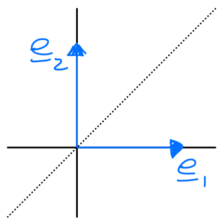
Now it's easy to get the image of $\underline{x} = \begin{bmatrix} 2 \\ -3 \end{bmatrix}$, namely $T\left(\begin{bmatrix} 2 \\ -3 \end{bmatrix}\right) = \begin{bmatrix} 2\cos\varphi + 3\sin\varphi \\ 2\sin\varphi - 3\cos\varphi \end{bmatrix}$

Example: Suppose the standard matrix is $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

What is the geometric interpretation?

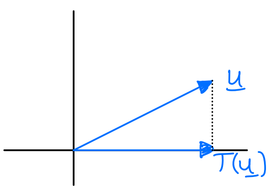
$T(\underline{e}_1) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \underline{e}_2$

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



Hence, the transformation is a reflection in the line $y = x$.

Example: Projection onto the x_1-axis



This is a linear transformation (DIY!) with standard matrix

$$A = \left[T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) \quad T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) \right] = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

So, indeed $T\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$.

Surjectivity: A transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ is surjective / onto if each $\underline{b} \in \mathbb{R}^m$ is the image of at least one $\underline{x} \in \mathbb{R}^n$ (range = codomain).

Injectivity: A transformation $T: \mathbb{R}^n \to \mathbb{R}^m$ is injective / one-to-one if each $\underline{b} \in \mathbb{R}^m$ is the image of at most one $\underline{x} \in \mathbb{R}^n$.

Theorem: Let $\mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation.
T is injective $\iff$ $T(\underline{x}) = \underline{0}$ has only the trivial solution.

Proof:
($\Longrightarrow$) Assume T is injective.
Since T is linear, we have $T(\underline{0}) = \underline{0}$.
Since T is injective, $T(\underline{x}) = \underline{0}$ has at most one solution.
So, $T(\underline{x}) = \underline{0}$ has only the trivial solution.
($\Longleftarrow$) By contrapositive.
Assume T is not injective.
So, there is a $\underline{b} \in \mathbb{R}^m$ that is the image of at least two vectors in $\mathbb{R}^n$.
So, $T(\underline{u}) = \underline{b}$ and $T(\underline{v}) = \underline{b}$ with $\underline{u} \neq \underline{v}$.
So, $T(\underline{u} - \underline{v}) \underset{\text{because } T \text{ is linear}}{=} T(\underline{u}) - T(\underline{v}) = \underline{b} - \underline{b} = \underline{0}$

Note $\underline{u} - \underline{v} \neq \underline{0}$ because $\underline{u} \neq \underline{v}$.
Hence, $T(\underline{x}) = \underline{0}$ has also a nontrivial solution. $\square$

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation with standard matrix A.

T is injective
$\iff$ $T(\underline{x}) = \underline{0}$ has only the trivial solution.
$\iff$ $A\underline{x} = \underline{0}$ has only the trivial solution.
$\iff$ A has a pivot in every column.

T is surjective
$\iff$ for each $\underline{b} \in \mathbb{R}^m$, $T(\underline{x}) = \underline{b}$ has a solution.
$\iff$ for each $\underline{b} \in \mathbb{R}^m$ $A\underline{x} = \underline{b}$ has a solution.
$\iff$ A has a pivot in every row.

Matrix Algebra

* equality: same size and corresponding entries are equal.
* addition: $\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}+\begin{bmatrix}0&1\\1&0\\2&-1\end{bmatrix}=\begin{bmatrix}1&3\\4&4\\7&5\end{bmatrix}$ Note: for A+B to be defined, we need size A = size B.
* scaling: $2\cdot\begin{bmatrix}1&2\\3&4\\5&6\end{bmatrix}=\begin{bmatrix}2&4\\6&8\\10&12\end{bmatrix}$
* matrix-matrix product: it works as a sequence of matrix-vector products.

$$\underbrace{\begin{bmatrix}2&3\\1&-5\end{bmatrix}}_{A}\begin{bmatrix}\underset{\underline{b}_1}{\begin{matrix}4\\1\end{matrix}}&\underset{\underline{b}_2}{\begin{matrix}3\\-2\end{matrix}}&\underset{\underline{b}_3}{\begin{matrix}6\\3\end{matrix}}\end{bmatrix}=\begin{bmatrix}A\underline{b}_1&A\underline{b}_2&A\underline{b}_3\end{bmatrix}=\begin{bmatrix}11&0&21\\-1&13&-9\end{bmatrix}$$

Each column of AB is a linear combination of the columns of A with the entries of the corresponding column of B being the weights.

$$\begin{bmatrix}11\\-1\end{bmatrix}=4\cdot\begin{bmatrix}2\\1\end{bmatrix}+1\cdot\begin{bmatrix}3\\-5\end{bmatrix}$$

Note: we need # columns of A = # rows of B.

$$\left.\begin{matrix}A: m\times n\\ B: n\times p\end{matrix}\right\}\quad C=AB\quad m\times p$$

In general $AB\neq BA$.

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

A matrix is symmetric if $A^T=A$ $\quad\begin{bmatrix}1&1\\1&1\end{bmatrix}\quad\begin{bmatrix}1&4&5\\4&2&6\\5&6&3\end{bmatrix}$

identity matrix $\quad I_2=\begin{bmatrix}1&0\\0&1\end{bmatrix}\quad I_4=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\end{bmatrix}$

$$AI_4=A\qquad I_4A=A$$

Power of a square $(n\times n)$ matrix $\quad A^k=\underbrace{A\cdot A\cdot A\cdot\ldots\cdot A}_{k\text{ times}}$.

$A^0=I_n$.

Composition of linear transformation.

$T_1 : \mathbb{R}^n \to \mathbb{R}^m$ with standard matrix A $m \times n$

$T_2 \quad \mathbb{R}^m \to \mathbb{R}^q$ $\qquad B$ $q \times m$

$$T = T_2 \circ T_1 = T_2(T_1)$$

Then, $T : \mathbb{R}^n \to \mathbb{R}^q$ with standard matrix C $q \times n$ where $C = BA$

Let A, B, and C be matrices of the same size, and let r and s be scalars.

a. $A + B = B + A$
b. $(A + B) + C = A + (B + C)$
c. $A + 0 = A$
d. $r(A + B) = rA + rB$
e. $(r + s)A = rA + sA$
f. $r(sA) = (rs)A$

Let A be an $m \times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined.

a. $A(BC) = (AB)C$ (associative law of multiplication)
b. $A(B + C) = AB + AC$ (left distributive law)
c. $(B + C)A = BA + CA$ (right distributive law)
d. $r(AB) = (rA)B = A(rB)$ for any scalar r
e. $I_m A = A = AI_n$ (identity for matrix multiplication)

WARNINGS:

1. In general, $AB \neq BA$.

2. The cancellation laws do *not* hold for matrix multiplication. That is, if $AB = AC$, then it is *not* true in general that $B = C$. (See Exercise 10.)

3. If a product AB is the zero matrix, you *cannot* conclude in general that either $A = 0$ or $B = 0$. (See Exercise 12.)

Let A and B denote matrices whose sizes are appropriate for the following sums and products.

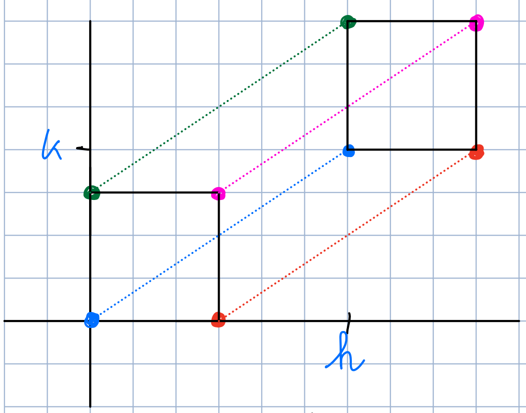
a. $(A^T)^T = A$
b. $(A + B)^T = A^T + B^T$
c. For any scalar r, $(rA)^T = rA^T$
d. $(AB)^T = B^T A^T$

The transpose of a product of matrices equals the product of their transposes in the *reverse* order.

Lecture 5: Perspective projections + The Inverse of a Matrix
(book: 2.2, 2.3, 2.7).

Previous lecture: Linear transformations, Matrix algebra

Applications to Computer Graphics.



This is a transformation, but not linear because $T(\underline{0}) \neq \underline{0}$.

We cannot find a 2x2 matrix A such that $A\underline{x} = T(\underline{x}) = \begin{bmatrix} x_1 + h \\ x_2 + k \end{bmatrix}$

$$\begin{bmatrix} 1 & \cdot \\ \cdot & \cdot \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \overset{\text{NOT possible}}{=} \begin{bmatrix} x_1 + h \\ x_2 + k \end{bmatrix}$$

Homogeneous coordinates: work on a specific plane in $\mathbb{R}^3$.
For example the plane with $x_3 = 1$.

(Note: a plane in $\mathbb{R}^3$ is not $\mathbb{R}^2$).

$$\begin{bmatrix} 1 & 0 & h \\ 0 & 1 & k \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ 1 \end{bmatrix} = \begin{bmatrix} x_1 + h \\ x_2 + k \\ 1 \end{bmatrix}$$

LAB 2: Perspective projections
↳ project a 3-dimensional object on a 2-dimensional computer screen.

(not on the exam).

Recall the three row operations,
* replacement
* scaling
* interchange.

$A \sim A_2 \sim A_3 \sim A_4 \sim A_5$.

Can we find an elementary matrix E such that $EA = A_5$.

Example: $\underset{A}{\begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}} \underset{1}{\sim} \underset{A_2}{\begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}} \underset{2}{\sim} \underset{A_3}{\begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}} \underset{3}{\sim} \underset{A_4}{\begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}} \underset{4}{\sim} \underset{A_5}{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}$

1 $R_1 \leftrightarrow R_2$
2 $R_2 : R_2 - 2 \cdot R_1$
3 $R_2 : R_2 * -1$
4 $R_1 : R_1 - 2 \cdot R_2$

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \underset{1}{\sim} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = E_1$ $\quad E_1 A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = A_2$

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \underset{2}{\sim} \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} = E_2$ $\quad E_2 A_2 = \begin{bmatrix} 1 & 0 \\ -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} = A_3$

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \underset{3}{\sim} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} = E_3$ $\quad E_3 A_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = A_4$

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \underset{4}{\sim} \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} = E_4$ $\quad E_4 A_4 = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = A_5$

Hence, $A_5 = E_4 A_4 = E_4 (E_3 A_3) = E_4 E_3 (E_2 A_2) = E_4 E_3 E_2 (E_1 A)$
$= \underbrace{E_4 E_3 E_2 E_1}_{E} A = EA$.

Check: $E = \ldots = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix}$ $\quad EA = \begin{bmatrix} 2 & -3 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

So, E takes us to the reduced echelon form of A.

So, E is obtained by applying the same row operations on the identity matrix.

An $n \times n$ matrix A is invertible if there exists an $n \times n$ matrix C such that $CA = I_n$ <u>and</u> $AC = I_n$.

This C is the inverse of A.
The inverse of a matrix is unique.
Notation: A^{-1}

For a matrix A to be invertible:
* A must be square ($n \times n$)
* the RREF of A is the identity matrix: $A \sim \dots \sim I_n$ (Thm 7)
  (A has a pivot position in every row/column: A has n pivot positions).

How to find the inverse of an invertible matrix A?
We need to find E such that $EA = I_n$.

$$n\begin{array}{c} 2n \\ \left[A \,\vdots\, I_n\right] \end{array} \sim E\left[A \,\vdots\, I_n\right] = \left[EA \,\vdots\, EI_n\right] = \left[I_n \,\vdots\, E\right]$$ hey there is A^{-1} :)

An algorithm for finding A^{-1}:
* Row reduce $[A : I_n]$
* If this leads to $[I_n : E]$, then A is invertible and $A^{-1} = E$.
* Otherwise, A is not invertible.

Example: $A = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$ inverse of A?

$$\begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix} \underset{R_2 : R_2 - 2\cdot R_1}{\sim} \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$$ A is not invertible because we cannot reduce it to I_2.

Example: $A = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$ inverse of A?

$$\left[\begin{array}{cc:cc} 1 & 2 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{array}\right] \underset{R_1 : R_1 - 2R_2}{\sim} \left[\begin{array}{cc:cc} 1 & 0 & 1 & -2 \\ 0 & 1 & 0 & 1 \end{array}\right]$$ So, $A^{-1} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \end{bmatrix}$.

$AA^{-1} = I_n$ ✓

Thm: The inverse of a matrix is unique.
Proof: Let A be an $n \times n$ matrix.
Suppose B and C are both an inverse of A and $B \neq C$.
$B = BI_n = B(AC) = (BA)C = I_nC = C$ ↯ □

Thm: Let A be an $n\times n$ invertible matrix. Then, for each $\underline{b} \in \mathbb{R}^n$, the equation $A\underline{x} = \underline{b}$ has the unique solution $\underline{x} = A^{-1}\underline{b}$.

Proof:
* existence: A is invertible $\Rightarrow$ A has a pivot in every row $\Rightarrow$ $A\underline{x} = \underline{b}$ is consistent for every $\underline{b} \in \mathbb{R}^n$.
* uniqueness: A is invertible $\Rightarrow$ A has a pivot in every col. $\Rightarrow$ no free variables $\Rightarrow$ solution must be unique.
* Show that $\underline{x} = A^{-1}\underline{b}$ is the unique solution:

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

$A^{-1}(A\underline{x}) = A^{-1}\underline{b}$

$A\underline{x} = A(A^{-1}\underline{b}) = (AA^{-1})\underline{b} = I_n\underline{b} = \underline{b}$ ✓

$$(A^{-1}A)\underline{x} = A^{-1}\underline{b}$$
$$I_n \underline{x} = A^{-1}\underline{b}$$
$$\underline{x} = A^{-1}\underline{b}. \checkmark \qquad \square$$

Thm: Let A be an $n \times n$ matrix.
If $CA = I_n$, then also $AC = I_n$.

Proof: Assume $CA = I_n$.

Claim/Lemma: For each $\underline{b} \in \mathbb{R}^n$, $A\underline{x} = \underline{b}$ has a solution.
Proof: Suppose $\exists \underline{b} \in \mathbb{R}^n$ st $A\underline{x} = \underline{b}$ has no solution.
Then, A does not have a pivot in every row.
So, A also does not have a pivot in every column.
So, there is a nontrivial solution $\underline{y}$: $A\underline{y} = \underline{0}$ with $\underline{y} \neq \underline{0}$.
Then, $\underline{y} = I_n\underline{y} = (CA)\underline{y} = C(A\underline{y}) = C\underline{0} = \underline{0}$ ↯
So, $A\underline{x} = \underline{b}$ has a solution for every $\underline{b} \in \mathbb{R}^n$ $\square$

Let $\underline{b} \in \mathbb{R}^n$. $A\underline{x} = \underline{b}$ has a sol.
$\Rightarrow C(A\underline{x}) = C\underline{b}$
$\Rightarrow (CA)\underline{x} = C\underline{b}$
$\Rightarrow I_n\underline{x} = C\underline{b}$
$\Rightarrow \underline{x} = C\underline{b}$ So, this is the sol of $A\underline{x} = \underline{b}$.

Hence, $\underline{b} = A\underline{x} = A(C\underline{b}) = (AC)\underline{b}$ So, we need $AC = I_n$ $\square$

Properties of invertible matrices:
* $(A^{-1})^{-1} = A$
* If A and B are invertible, then $(AB)^{-1} = B^{-1}A^{-1}$ (verify yourself)
* $(A^T)^{-1} = (A^{-1})^T$.

Summary (so far):

Let A be an $m \times n$ matrix with columns $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n$.

$m \geq n$: $\begin{bmatrix} A \\ \\ \\ \end{bmatrix}$

The following statements are equivalent:

① A has a pivot in every column

② A has n pivot positions

③ There are no free variables

④ $A\underline{x} = \underline{0}$ has only the trivial sol.

⑤ $\{\underline{a}_1, \underline{a}_2, \dots \underline{a}_n\}$ is linearly indep.

⑥ $T: \underline{x} \mapsto A\underline{x}$ is one-to-one/injective

$m \leq n$: $\begin{bmatrix} & & A & & \end{bmatrix}$

The following statements are equivalent:

ⓐ A has a pivot in every row.

ⓑ A has m pivot positions.

ⓒ The echelon form of A does not contain a row of all zeros.

ⓓ $A\underline{x} = \underline{b}$ is consistent for every $\underline{b}$ in $\mathbb{R}^m$.

ⓔ $\text{Span}\{\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n\} = \mathbb{R}^m$.

ⓕ $T: \underline{x} \mapsto A\underline{x}$ is onto/surjective.

The Invertible Matrix Theorem:

If A is square ($n = m$), then statements ② and ⓑ are equivalent. Hence, the following statements are equivalent for square matrices

* ① – ⑥, ⓐ – ⓕ
* A is invertible
* There is a matrix C such that $CA = I_n$ and $AC = I_n$
* A is row equivalent to I_n.
* A^T is invertible.
* $\det A \neq 0$

Lecture 6 : Determinants
(book: 3.1, 3.2).

Previous lecture: the inverse of a matrix.

Application of the inverse matrix: Cryptography.

Imitation game. A is used to encrypt the message.

Hill algorithm.

$A = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}$ ATTACK-NOW. $\begin{bmatrix} A \\ T \end{bmatrix} = \begin{bmatrix} 1 \\ 20 \end{bmatrix}$ $\begin{bmatrix} T \\ A \end{bmatrix} = \begin{bmatrix} 20 \\ 1 \end{bmatrix}$

$\begin{bmatrix} C \\ K \end{bmatrix} = \begin{bmatrix} 3 \\ 11 \end{bmatrix}$ $\begin{bmatrix} - \\ N \end{bmatrix} = \begin{bmatrix} 27 \\ 14 \end{bmatrix} \dots$

$\begin{bmatrix} A \end{bmatrix}\begin{bmatrix} A \\ T \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 20 \end{bmatrix} = \begin{bmatrix} 41 \\ 61 \end{bmatrix}$

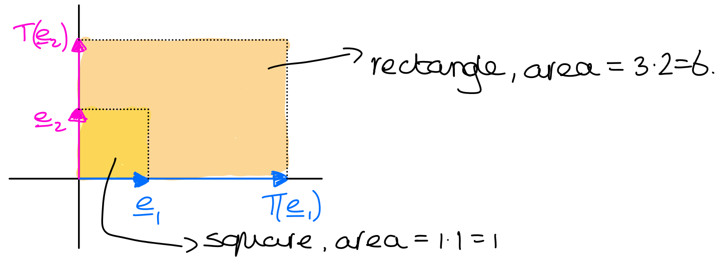
$\begin{bmatrix} A^{-1} \end{bmatrix}\begin{bmatrix} 41 \\ 61 \end{bmatrix} = \begin{bmatrix} 1 \\ 20 \end{bmatrix} = \begin{bmatrix} A \\ T \end{bmatrix}$.

Inverse of a 2×2 matrix $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

* If $ad - bc \neq 0$, then A is invertible and $A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$.
  ↳ determinant.
* If $ad - bc = 0$, then A is not invertible
  ↳ singular.

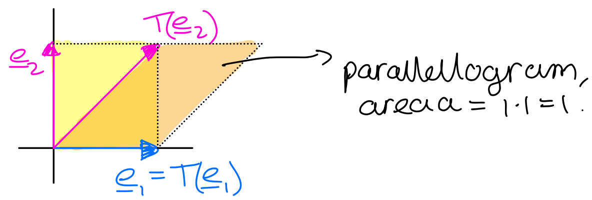
Example: $A = \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix}$ $\mathbb{R}^2 \to \mathbb{R}^2$

$\uparrow T(\underline{e}_1)$ $\uparrow T(\underline{e}_2)$



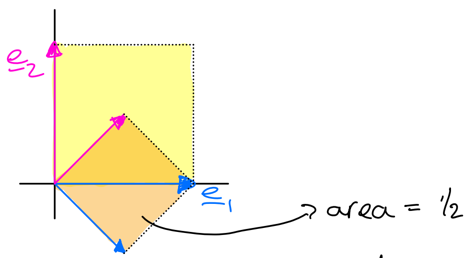
So, A is stretching objects in $\mathbb{R}^2$.
The stretching/scaling factor is $6 = \det(A)$.
$\hookrightarrow > 1$ because the area increases.

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



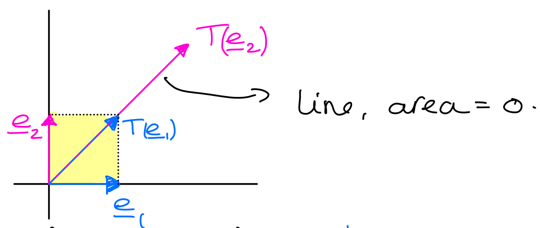
So, $\det(A) = 1$ (because the area stays the same).

Example: $A = \begin{bmatrix} 1/2 & 1/2 \\ -1/2 & 1/2 \end{bmatrix}$



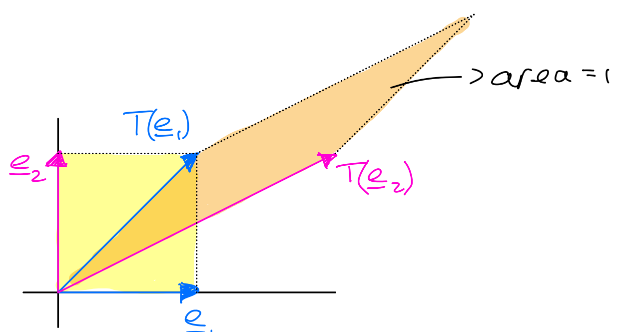
So, $\det(A) = 1/2$ (because the area squishes with a factor $1/2$)

Example: $A = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$



So, $\det(A) = 0$ (because the unit square is crushed in a line).

Example: $A = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}$



The orientation of space has been "inverted"
So, $\det(A) = -1$.

The determinant of a square (n×n) matrix is a scalar associated with the matrix.
Notation: $\det(A)$ $|A|$.

It measures how the transformation $T: \underline{x} \to A\underline{x}$ "scales" space:
* in $\mathbb{R}^2$ it measures the change in areas of objects by T.
* in $\mathbb{R}^3$ it measures the change in volumes of objects by T.

$\det(A)=0$ $\Rightarrow$ spaces are "flattened" / we are losing one dimension
$\Rightarrow$ range $\neq$ codomain
$\Rightarrow$ transformation is not surjective (onto).
$\Rightarrow$ A is not invertible.

How to compute the determinant?
* Gaussian elimination.
* cofactor expansion.

Recall $\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc$.

Cofactor expansion for an n×n matrix:
* Focus on a specific row i or column j.
* For example, for row i:

$$\det(A) = \sum_{j=1}^{n} a_{ij} \cdot C_{ij}.$$

* a_{ij}: entry of A at location (i,j)
* C_{ij}: (i,j)-cofactor $= (-1)^{i+j} \cdot \det(A_{ij})$
* A_{ij}: submatrix obtained by removing row i and column j.

Example $A = \begin{bmatrix} 3 & 5 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 1 \end{bmatrix}$

Cofactor expansion across the first row

$$\det(A) = 3 \cdot (-1)^{1+1} \cdot \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} + 5 \cdot (-1)^{1+2} \cdot \begin{vmatrix} 0 & 1 \\ 0 & 1 \end{vmatrix} + 1 \cdot (-1)^{1+3} \begin{vmatrix} 0 & 2 \\ 0 & 0 \end{vmatrix}$$
$$= 3 \cdot 2 - 5 \cdot 0 + 1 \cdot 0 = 6.$$

Cofactor expansion across the first column.

$$\det(A) = 3 \cdot (-1)^{1+1} \begin{vmatrix} 2 & 1 \\ 0 & 1 \end{vmatrix} + 0 + 0 = 3 \cdot 2 = 6.$$

So, be smart: choose a row/column with many 0s.

Triangular matrix: the entries below / above the main diagonal are all 0s.

upper triangular / lower triangular.

$$\begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{bmatrix} \qquad \begin{bmatrix} * & 0 & 0 & 0 \\ * & * & 0 & 0 \\ * & * & * & 0 \\ * & * & * & * \end{bmatrix}$$

Diagonal matrix: a square matrix whose nondiagonal entries are all 0s.

$$\begin{bmatrix} * & 0 & 0 & 0 \\ 0 & * & 0 & 0 \\ 0 & 0 & * & 0 \\ 0 & 0 & 0 & * \end{bmatrix}.$$

For triangular or diagonal matrices, the determinant equals the product of the entries on the main diagonal.

$$\begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \\ 0 & 0 & 0 & * \end{bmatrix} \qquad \det(A) = a_{11} \cdot a_{22} \cdot a_{33} \cdot a_{44} \cdot \ldots \cdot a_{nn}.$$

REF is upper triangular.
So, maybe we can use Gaussian elimination to compute the determinant?

How do row operations change the determinant?

* two rows of A are interchanged to produce B: $\det(B) = -\det(A)$.
* one row of A is multiplied by k to produce B: $\det(B) = k \cdot \det(A)$
* a multiple of one row of A is added to another row to produce B: $\det(B) = \det(A)$.

Example:

$$\begin{vmatrix} 0 & 5 & 1 \\ 4 & -3 & 0 \\ 2 & 4 & 1 \end{vmatrix} \underset{R_1 \leftrightarrow R_3}{=} -1 \cdot \begin{vmatrix} 2 & 4 & 1 \\ 4 & -3 & 0 \\ 0 & 5 & 1 \end{vmatrix} \underset{R_2 : R_2 - 2 \cdot R_1}{=}$$

$$-1 \cdot \begin{vmatrix} 2 & 4 & 1 \\ 0 & -11 & -2 \\ 0 & 5 & 1 \end{vmatrix} \underset{R_3 : R_3 + 5/11 \cdot R_2}{=} -1 \cdot \begin{vmatrix} 2 & 4 & 1 \\ 0 & -11 & -2 \\ 0 & 0 & 1/11 \end{vmatrix} = (-1) \cdot 2 \cdot (-11) \cdot 1/11 = 2.$$

---

square matrix A not invertible

$\Rightarrow$ A is not row equivalent to I_n.

$\Rightarrow$ a pivot is missing.

$\Rightarrow$ $\det(\text{REF of } A) = 0$.

$\Rightarrow$ $\det(A) = (-1)^{\#\text{swaps}} \cdot$ $\det(\text{REF of } A) = (-1)^{\#\text{swaps}} \cdot 0 = 0$

$\Rightarrow$ $\det(A) = 0$.

Conclusion: square matrix A is not invertible $\Leftrightarrow$ $\det(A) = 0$.

Properties of determinants:

* $\det(A^T) = \det(A)$
* $\det(AB) = \det(A) \cdot \det(B)$ (Thm 6).
* but $\det(A+B) \neq \det(A) + \det(B)$ in general.
* $\det(c \cdot A) = c^n \cdot \det(A)$

Theorem: $\det(A^{-1}) = \dfrac{1}{\det(A)}$ (for all invertible matrices).

Proof:

$$I_n = A \cdot A^{-1}$$

$$\Rightarrow \det(I_n) = \det(A \cdot A^{-1})$$

$$\Rightarrow 1 = \det(A) \cdot \det(A^{-1})$$

$$\Rightarrow \frac{1}{\det(A)} = \det(A^{-1}) \qquad \square$$

Summary (so far):

Let A be an $m\times n$ matrix with columns $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n$.

$m \geq n$: $\begin{bmatrix} A \\ \\ \\ \end{bmatrix}$	$m \leq n$: $\begin{bmatrix} & A & \end{bmatrix}$
The following statements are equivalent:	The following statements are equivalent:
① A has a pivot in every column	ⓐ A has a pivot in every row.
② A has n pivot positions	ⓑ A has m pivot positions.
③ There are no free variables	ⓒ The echelon form of A does not contain a row of all zeros.
④ $A\underline{x}=\underline{0}$ has only the trivial sol.	ⓓ $A\underline{x}=\underline{b}$ is consistent for every $\underline{b}$ in $\mathbb{R}^m$.
⑤ $\{\underline{a}_1, \underline{a}_2, \cdots \underline{a}_n\}$ is linearly indep.	ⓔ Span$\{\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n\} = \mathbb{R}^m$.
⑥ $T: \underline{x} \mapsto A\underline{x}$ is one-to-one/injective	ⓕ $T: \underline{x} \mapsto A\underline{x}$ is onto/surjective.

The Invertible Matrix Theorem:

If A is square $(n=m)$, then statements ② and ⓑ are equivalent. Hence, the following statements are equivalent for square matrices

* ① - ⑥, ⓐ - ⓕ
* A is invertible
* There is a matrix C such that $CA = I_n$ <u>and</u> $AC = I_n$
* A is row equivalent to I_n.
* A^T is invertible.
* $\det A \neq 0$

Lecture 7: Vector Spaces (part 1)
(book: 4.1, 4.2).

Previous lecture: Determinants

Applications of Vector Spaces:
* Vector Space Search Engines. (see the YouTube video on Canvas)
* Digital Signal Processing (Section 4.7)
* Fourier Series (follow-up course Numerical Mathematics)

Intro to Vector Space: 3Blue1Brown - Abstract vector spaces (see Canvas)

So far: $\mathbb{R}^1, \mathbb{R}^2, \mathbb{R}^3, \ldots, \mathbb{R}^n, \ldots$ were our vector spaces.

* adding vectors in $\mathbb{R}^n \to$ another vector in $\mathbb{R}^n$.
* scaling vectors in $\mathbb{R}^n \to$ another vector in $\mathbb{R}^n$.

Definition of a vector space V: (axioms)

A non-empty set of objects (vectors) with the following 10 rules:

1. $\underline{u}, \underline{v} \in V \Rightarrow \underline{u} + \underline{v}$ is another vector in V (closed under addition)
2. $\underline{u} \in V, c \in \mathbb{R} \Rightarrow c \cdot \underline{u}$ is another vector in V (closed under scalar mult.)
3. $\underline{u} + \underline{v} = \underline{v} + \underline{u}$ (commutativity)
4. $\underline{u} + (\underline{v} + \underline{w}) = (\underline{u} + \underline{v}) + \underline{w}$ (associativity)

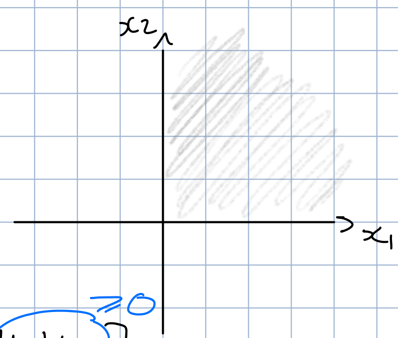
5. + 6. $\exists\, \underline{0} \in V$: $\underline{u} + \underline{0} = \underline{u}$ and $\underline{u} + (-\underline{u}) = \underline{0}$ (zero vector)

7. $1 \cdot \underline{u} = \underline{u}$

8. - 10. see the book.

Example: Is $\mathbb{R}^2_+$ a vector space? No.

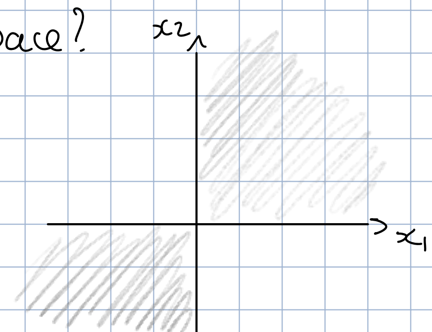
$\hookrightarrow \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0 \right\}$



② ✗ e.g. $\begin{bmatrix} 2 \\ 3 \end{bmatrix} \in \mathbb{R}^2_+$, but $-1 \cdot \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} -2 \\ -3 \end{bmatrix} \notin \mathbb{R}^2_+$

① ✓ Let $\underline{u}, \underline{v} \in \mathbb{R}^2_+$. Then, $\underline{u} + \underline{v} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} \underbrace{u_1 + v_1}_{\geq 0} \\ \underbrace{u_2 + v_2}_{\geq 0} \end{bmatrix}$

Example: Is $\left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \mathbb{R}^2 : x_1 \cdot x_2 \geq 0 \right\}$ a vector space?



① ✗ e.g. $\underline{u} = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ and $\underline{v} = \begin{bmatrix} -2 \\ -2 \end{bmatrix}$.

Then $\underline{u} + \underline{v} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ $\quad 1 \cdot (-1) < 0$.

Example: Is $\mathbb{P}_n$ a vector space?
↳ set of polynomials of degree at most n.
i.e., all polynomials of the form
$p(x) = a_0 + a_1 x + a_2 \cdot x^2 + a_3 \cdot x^3 + \dots + a_n \cdot x^n$.

$n=3$ $\mathbb{P}_3$

$p(x) = x^3 \in \mathbb{P}_3$
$8.5x^4 \notin \mathbb{P}_3$
$9x^2 + x^3 \in \mathbb{P}_3$.
$7.5x^3 + x^4 \notin \mathbb{P}_3$

① Let $p, q \in \mathbb{P}_n$ $p+q \in \mathbb{P}_n$?
$p(x) = a_0 + a_1 x + \dots + a_n \cdot x^n$
$q(x) = b_0 + b_1 x + \dots + b_n \cdot x^n$.

$(p+q)(x) = p(x) + q(x)$
$= (a_0 + b_0) + (a_1 + b_1) x + \dots + (a_n + b_n) x^n$

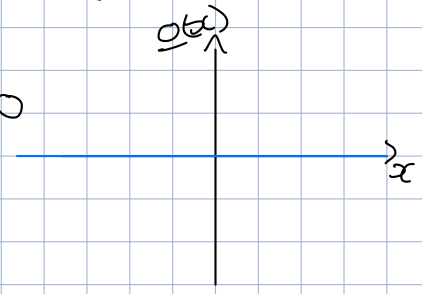
So, $p + q \in \mathbb{P}_n$. ✓

② Let $p \in \mathbb{P}_n$ and $c \in \mathbb{R}$. $p(x) = a_0 + a_1 x + \dots + a_n \cdot x^n$

$(c \cdot p)(x) = c \cdot p(x) = c \cdot (a_0 + a_1 x + \dots + a_n \cdot x^n)$
$= c \cdot a_0 + (c \cdot a_1) \cdot x + \dots + (c \cdot a_n) \cdot x^n$.

So, $c \cdot p \in \mathbb{P}_n$ ✓

③-⑩ DIY ✓ But what is $\underline{0}$? $\underline{0}(x) = 0$



Similarly: $\mathbb{P}$ is also a vector space.
↳ set of all polynomials

Example: Set of polynomials of the form
$p(x) = 8 + a_1 x + a_2 x^2 + \dots + a_n x^n$

① ✗ → So, not a vector space.
② ✗
③ ✗

Let V be a vector space and let $W \subseteq V$.
When is W also a vector space?
① $\underline{w}_1, \underline{w}_2 \in W \Rightarrow \underline{w}_1 + \underline{w}_2 \in W$. (closed under addition)
② $\underline{w} \in W$, $c \in \mathbb{R} \Rightarrow c \cdot \underline{w} \in W$ (closed under scalar mult.)
③ $\underline{0} \in W$ (if $W \neq \emptyset$, then ③ follows from ②)

(All the other axioms are fulfilled because $W \subseteq V$ and V is a vector space).

W is called a subspace of V.

Example: Vector space V. $W = \emptyset$ $W \subseteq V$.

Is W a vector space?
Is W a subspace of V?

① ✓
② ✓
③ ✗ (because $W = \emptyset$ and thus $\underline{0} \notin W$) $\rightarrow$ So, W is not a subspace of V.

Example: $\mathbb{R}^3$ is a vector space.

$W = \left\{ \begin{bmatrix} a \\ b \\ 1 \end{bmatrix} : a, b \in \mathbb{R} \right\}$ $W \subseteq \mathbb{R}^3$.

① ✗
② ✗
③ ✗ $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \notin W$. $\rightarrow$ So, W is not a subspace of $\mathbb{R}^3$.

Example: $W = \left\{ \begin{bmatrix} a \\ b \\ 0 \end{bmatrix} : a, b \in \mathbb{R} \right\}$ is a subspace of $\mathbb{R}^3$

Example: $W = \text{Span}\left\{ \begin{bmatrix} 2 \\ 4 \\ 7 \end{bmatrix} \right\}$ Subspace of $\mathbb{R}^3$? Yes.

(W is a line in $\mathbb{R}^3$). ✓

Theorem: Let V be a vector space.
If $\underline{v}_1, \dots, \underline{v}_p \in V$, then $W = \text{Span}\{\underline{v}_1, \dots, \underline{v}_p\}$ is a subspace of V.

Proof:

⓪ We need to show $W \subseteq V$.
Since V is a vector space:
So, $c_1 \underline{v}_1 \in V, \dots, c_p \underline{v}_p \in V$. (because of property ② of V).
So, $c_1 \underline{v}_1 + \dots + c_p \underline{v}_p \in V$ (because of prop ① of V)
So, $W \subseteq V$. ✓

①. If $\underline{x}, \underline{y} \in W$, then $\underline{x} + \underline{y} = c_1 \cdot \underline{v}_1 + \dots + c_p \cdot \underline{v}_p + d_1 \underline{v}_1 + \dots + d_p \underline{v}_p$
$= (c_1 + d_1) \underline{v}_1 + \dots + (c_p + d_p) \underline{v}_p$.
So, $\underline{x} + \underline{y} \in W$ ✓

②. If $\underline{x} \in W$, $a \in \mathbb{R}$, then $a \cdot \underline{x} = a \cdot (c_1 \cdot \underline{v}_1 + \dots + c_p \cdot \underline{v}_p)$
$= (a \cdot c_1) \cdot \underline{v}_1 + \dots + (a \cdot c_p) \underline{v}_p$
So, $a \cdot \underline{x} \in W$. ✓

③. $\underline{0} = 0 \cdot \underline{v}_1 + \dots + 0 \cdot \underline{v}_p$ So, $\underline{0} \in W$. ✓ □

Theorem: Let A be an $m \times n$ matrix.
$\underbrace{\text{The set of all solutions of } A\underline{x}=\underline{0}}_{= Nul(A).\quad W.}$ is a subspace of $\mathbb{R}^n$.

Proof:

⓪. A is $m \times n$, so $\underline{x}$ needs to have n elements. So, $W \subseteq \mathbb{R}^n$. ✓

①. If $\underline{u}, \underline{v} \in W$, then $A(\underline{u}+\underline{v}) = A\underline{u} + A\underline{v} = \underline{0} + \underline{0} = \underline{0}$ ✓

$\underline{u} \in W \Rightarrow A\underline{u} = \underline{0}$
$\underline{v} \in W \Rightarrow A\underline{v} = \underline{0}$

②. If $\underline{u} \in W$ and $c \in \mathbb{R}$, then $A(c \cdot \underline{u}) = c \cdot A\underline{u} = c \cdot \underline{0} = \underline{0}$ ✓

$\underline{u} \in W \Rightarrow A\underline{u} = \underline{0}$

③. $A\underline{0} = \underline{0}$. So, $\underline{0} \in W$. ✓

$\square$

Lecture 8: Vector Spaces (part 2)
(book: 4.2, 4.3, 4.4, 4.5)

Today: basis, dimension, Nul, Col, Row.

A basis of a vector space V is a set $\{\underline{v}_1, \dots, \underline{v}_p\}$ of vectors in V such that:
* $\{\underline{v}_1, \dots, \underline{v}_p\}$ is lin. indep.
* it spans the whole vector space.

Example: A basis for $\mathbb{R}^3$, standard basis $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$.

But also $\left\{ \begin{bmatrix} 3 \\ 0 \\ -6 \end{bmatrix}, \begin{bmatrix} -4 \\ 1 \\ 7 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix} \right\}$. Why?

$$\begin{bmatrix} 3 & -4 & -2 \\ 0 & 1 & 1 \\ -6 & 7 & 5 \end{bmatrix} \sim \quad \dots \quad \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

* pivot in every column $\Rightarrow$ lin. indep.
* pivot in every row $\Rightarrow$ spans $\mathbb{R}^3$.

Example: The standard basis for $\mathbb{P}_2$: $\{1, x, x^2\}$

But a basis is also: $\{1+x^2,\ x-3x^2,\ 1+x-3x^2\}$ (later).

Two views of a basis $\mathcal{B}$:
* it is maximal: adding any vector $\underline{v} \in V$ to $\mathcal{B}$, makes it lin dep.
* it is minimal: removing any vector $\underline{v}$ from $\mathcal{B}$, causes $\mathcal{B}$ to no longer span V.

dim (V): # vectors in a basis of V.

$\dim(\mathbb{R}^3) = 3 \qquad \dim(\mathbb{R}^n) = n$

$\dim(\mathbb{P}_n) = n+1$.

Let $\mathcal{B} = \{\underline{b}_1, \dots, \underline{b}_n\}$ be a basis for V.
For every $\underline{x} \in V$ we can find weights $\underbrace{c_1, \dots, c_n}$ such that
$\underline{x} = c_1 \cdot \underline{b}_1 + \dots + c_n \cdot \underline{b}_n$

coordinates of $\underline{x}$ wrt basis $\mathcal{B}$.

$$[\underline{x}]_{\mathcal{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

Note: this set of coordinates is unique.

Example: $\mathcal{B} = \left\{ \begin{bmatrix} 3 \\ 0 \\ -6 \end{bmatrix}, \begin{bmatrix} -4 \\ 1 \\ 7 \end{bmatrix}, \begin{bmatrix} -2 \\ 1 \\ 5 \end{bmatrix} \right\}$ basis for $\mathbb{R}^3$.

Consider $\underline{x} = \begin{bmatrix} 5 \\ 2 \\ 4 \end{bmatrix}$.

What are the coordinates of $\underline{x}$ wrt the basis $\mathcal{B}$?

$$\left[\begin{array}{ccc|c} 3 & -4 & -2 & 5 \\ 0 & 1 & 1 & 2 \\ -6 & 7 & 5 & 4 \end{array}\right] \sim \dots \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 8 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & -2 \end{array}\right].$$

Matlab: rref

$$[\underline{x}]_{\mathcal{B}} = \begin{bmatrix} 8 \\ 5 \\ -2 \end{bmatrix}$$

Example: $\mathcal{B} = \{1+x^2, x-3x^2, 1+x-3x^2\}$ is a basis for $\mathbb{P}_2$. (later).

Consider $p(x) = 6+3x-x^2$.
Find $[p]_{\mathcal{B}}$.

$$6 + 3x - x^2 = c_1 \cdot (1+x^2) + c_2 \cdot (x - 3x^2) + c_3 \cdot (1 + x - 3x^2).$$

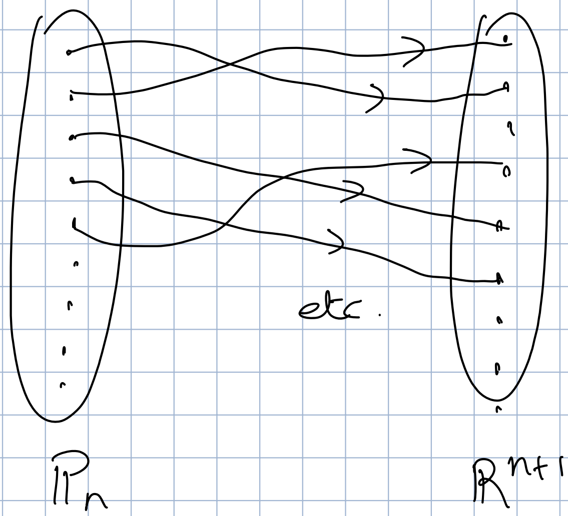
$$\text{LS} \begin{cases} c_1 + c_3 = 6 \\ c_2 + c_3 = 3 \\ c_1 - 3c_2 - 3c_3 = -1 \end{cases} \quad \left[\begin{array}{ccc|c} 1 & 0 & 1 & 6 \\ 0 & 1 & 1 & 3 \\ 1 & -3 & -3 & -1 \end{array}\right] \sim \dots \sim \left[\begin{array}{ccc|c} 1 & 0 & 0 & 8 \\ 0 & 1 & 0 & 5 \\ 0 & 0 & 1 & -2 \end{array}\right].$$

So, $[p]_{\mathcal{B}} = \begin{bmatrix} 8 \\ 5 \\ -2 \end{bmatrix}$. Check: $p(x) = 8(1+x^2) + 5 \cdot (x - 3x^2) + (-2) \cdot (1 + x - 3x^2)$. ✓

Theorem: Let $\mathcal{B} = \{\underline{b}_1, \ldots, \underline{b}_n\}$ be a basis for V. Then, the linear transformation $T: V \rightarrow \mathbb{R}^n$ defined by $T(\underline{x}) = [\underline{x}]_{\mathcal{B}}$ is a bijection.

Proof: See the book (Theorem 9).

Hence, the vectors can be abstract objects, but using the coordinates we can translate them to real numbers.



bijection.

Example: Show that $\{1+x^2, x-3x^2, 1+x-3x^2\}$ forms a basis for $\mathbb{P}_2$.

The coordinate mapping of the standard basis $\{1, x, x^2\}$ produces the coordinate vectors $\begin{bmatrix}1\\0\\1\end{bmatrix}, \begin{bmatrix}0\\1\\-3\end{bmatrix}, \begin{bmatrix}1\\1\\-3\end{bmatrix}$.

Do these vectors form a basis in $\mathbb{R}^3$? Yes.

$$\begin{bmatrix}1 & 0 & 1\\0 & 1 & 1\\1 & -3 & -3\end{bmatrix} \sim \cdots \sim \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}.$$

So, $\{1+x^2, x-3x^2, 1+x-3x^2\}$ forms a basis for $\mathbb{P}_2$.

Three important subspaces associated with an $m \times n$ matrix A.

* $\text{Nul}(A)$
* $\text{Col}(A)$
* $\text{Row}(A)$

Recall from Monday: $\text{Nul}(A) = \{\underline{x} \in \mathbb{R}^n : A\underline{x} = \underline{0}\}$ is a subspace of $\mathbb{R}^n$.

$$A = \begin{bmatrix} 1 & -3 & 2 \\ 2 & -6 & 4 \end{bmatrix} \qquad \left[\begin{array}{ccc|c} 1 & -3 & 2 & 0 \\ 2 & -6 & 4 & 0 \end{array}\right] \sim \cdots \sim \left[\begin{array}{ccc|c} 1 & -3 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\underline{x} = x_2 \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \qquad \text{Nul}(A) = \text{Span}\left\{ \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}$$

Note: * # vectors in the spanning set = # free variables.
* the vectors in the spanning set are always lin indep.

Basis for $\text{Nul}(A)$: $\left\{ \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \end{bmatrix} \right\}$. $\quad \dim(\text{Nul}(A)) = 2$

So, in general:
* $\dim(\text{Nul}(A))$ = # free vars.
* the vectors from parametric vector form form a basis for $\text{Nul}(A)$.

$\text{Col}(A) = \text{Span}\{\underline{a}_1, \ldots, \underline{a}_n\}$ is a subspace of $\mathbb{R}^m$

How to find a basis for $\text{Col}\,A$?

Intermezzo.

The Spanning Set Theorem:

Let $\underline{v}_1, \dots, \underline{v}_p \in V$ and let $H = \text{Span}\{\underline{v}_1, \dots, \underline{v}_p\}$.

* If $\underline{v}_k$ is a lin comb. of the other vectors, then $\{\underline{v}_1, \dots, \underline{v}_p\} \setminus \{\underline{v}_k\}$ still spans H.

* If $H \neq \{\underline{0}\}$, some subset of $\{\underline{v}_1, \dots, \underline{v}_p\}$ spans H.

---

* 1st strategy: keep deleting columns if they are dependent on others.

  The columns that you end up with form a basis for $\text{Col}(A)$.

  Easy when in reduced echelon form.

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

$$\left.\begin{aligned} \underline{b}_2 &= 3 \cdot \underline{b}_1 \\ \underline{b}_4 &= -1 \cdot \underline{b}_1 - 2 \cdot \underline{b}_3 \end{aligned}\right\} \text{Remove } \underline{b}_2 \text{ and } \underline{b}_4.$$

So, $\underbrace{\{\underline{b}_1, \underline{b}_3\}}_{\hookrightarrow \text{ pivot columns.}}$ forms a basis and $\dim(\text{Col}(B)) = 2$. True in general.

Intermezzo.

Linear dependence relations between colums do not change under row operations.

If $A \sim B$, then $A\underline{x} = \underline{0}$ and $B\underline{x} = \underline{0}$ have the same set of solutions.

$$A = \begin{bmatrix} 1 & 3 & 0 & -1 \\ 0 & 0 & 1 & -2 \\ 2 & 6 & -1 & 0 \\ 1 & 3 & 0 & -1 \end{bmatrix} \sim \dots \sim B$$

So, we also have
$$\underline{a}_2 = 3 \cdot \underline{a}_1$$
$$\underline{a}_4 = -1 \cdot \underline{a}_1 - 2 \cdot \underline{a}_3$$

So, $\{\underline{a}_1, \underline{a}_3\}$ forms a basis for $\text{Col}(A)$. and $\dim(\text{Col}(A)) = 2$.

be careful! of A! Not of REF!

* 2nd strategy: the pivot columns of A form a basis for $\text{Col}(A)$.

  Hence, $\dim(\text{Col}(A)) = \#\text{pivot cols}$.

$Row(A) = Span\{\underline{r}_1, \dots, \underline{r}_m\}$ is a subspace of $\mathbb{R}^n$.

$Row(A) = Col(A^T)$.

Note: $\boxed{A \sim B \Rightarrow Row(A) = Row(B)}$.

(Warning $Col(A) \neq Col(B)$).

idea: $A \sim B$

$\Rightarrow$ rows of B are a lin comb. of the rows of A.

$\Rightarrow$ any lin comb of the rows of B are also a lin comb of the rows of A.

$\Rightarrow Row(B) \subseteq Row(A)$.

$A \sim B \Rightarrow B \sim A \Rightarrow \dots \Rightarrow Row(A) \subseteq Row B$.

So, $Row(A) = Row(B)$.

If B is RREF, then the nonzero rows of B are lin indep and thus form a basis for both $Row(B)$ and $Row(A)$.

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

Hence, $\dim(Row(A)) = \#$ pivot rows.

$\boxed{Rank(A)} = \boxed{\dim(Col(A))} = \#$pivot cols $= \#$pivots $= \#$ pivot rows $= \boxed{\dim(Row(A))}$

$\hookrightarrow = n - \#$ nonpivot cols.

$= n - \#$ free variables.

$\boxed{= n - \dim(Nul(A))}$

Recall a linear transformation $T: \underline{x} \longrightarrow A\underline{x}$

$A: m \times n$
domain: $\mathbb{R}^n$
codomain: $\mathbb{R}^m$.

* Range $= \{T(\underline{x}) : \underline{x} \in \mathbb{R}^n\} = \{A\underline{x} : \underline{x} \in \mathbb{R}^n\} = Col(A)$

So, T is surjective $\Leftrightarrow$ range = codomain $\Leftrightarrow Col(A) = \mathbb{R}^m$

$\Leftrightarrow \dim(Col(A)) = m \Leftrightarrow Rank(A) = m$

* Kernel $= \{\underline{x} \in \mathbb{R}^n : T(\underline{x}) = \underline{0}\} = \{\underline{x} \in \mathbb{R}^n : A\underline{x} = \underline{0}\} = Nul(A)$.

So T is injective $\Leftrightarrow T(\underline{x}) = \underline{0}$ has only the trivial sol.

$\Leftrightarrow$ Kernel $= \{\underline{0}\} \Leftrightarrow Nul(A) = \{\underline{0}\} \Leftrightarrow \dim(Nul(A)) = 0$.

$\Leftrightarrow Rank(A) = n$

Summary (so far):

Let A be an $m \times n$ matrix with columns $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n$.

$m \geq n$: $\begin{bmatrix} A \end{bmatrix}$

The following statements are equivalent:

① A has a pivot in every column

② A has n pivot positions

③ There are no free variables

④ $A\underline{x} = \underline{0}$ has only the trivial sol.

⑤ $\{\underline{a}_1, \underline{a}_2, \dots \underline{a}_n\}$ is linearly indep.

⑥ $T: \underline{x} \mapsto A\underline{x}$ is one-to-one/injective

⑦ $\text{Nul } A = \{\underline{0}\}$

⑧ $\dim \text{Nul } A = 0$

⑨ $\text{rank } A = n$

$m \leq n$: $\begin{bmatrix} & A & \end{bmatrix}$

The following statements are equivalent:

ⓐ A has a pivot in every row.

ⓑ A has m pivot positions.

ⓒ The echelon form of A does not contain a row of all zeros.

ⓓ $A\underline{x} = \underline{b}$ is consistent for every $\underline{b}$ in $\mathbb{R}^m$.

ⓔ $\text{Span}\{\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n\} = \mathbb{R}^m$.

ⓕ $T: \underline{x} \mapsto A\underline{x}$ is onto/surjective.

ⓖ $\text{Col } A = \mathbb{R}^m$

ⓗ $\dim \text{Col } A = m$

ⓘ $\text{rank } A = m$

If A is square ($n = m$), then statements ② and ⓑ are equivalent.
Hence, the following statements are equivalent for square matrices.

* ① – ⑨, ⓐ – ⓘ
* A is invertible
* There is a matrix C such that $CA = I_n$ and $AC = I_n$
* A is row equivalent to I_n.
* A^T is invertible.
* $\det A \neq 0$
* The columns of A form a basis for $\mathbb{R}^n$.

A: $m \times n$ matrix.

	definition	subspace of	equal to $\mathbb{R}^n/\mathbb{R}^m/\mathbb{R}^n$ if	equal to $\{\underline{0}\}$ if
Nul A	$\{\underline{x} : A\underline{x} = \underline{0}\}$	$\mathbb{R}^n$	A is the zero matrix	A has a pivot in every column
Col A	Span$\{\underline{a}_1, \dots, \underline{a}_n\}$	$\mathbb{R}^m$	A has a pivot in every row	A is the zero matrix
Row A	Span$\{\underline{r}_1, \dots, \underline{r}_m\}$	$\mathbb{R}^n$	A^T has a pivot in every row	A is the zero matrix

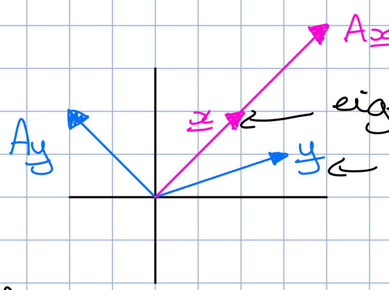
	dimension
Nul A	# free var's in the equation $A\underline{x} = \underline{0}$ = # nonpivot columns in A = n - rank A
Col A	# pivot columns in A. = # pivot columns in A = rank A
Row A	# non-zero rows in the echelon form of A = # pivot columns in A = rank A

	finding a basis
Nul A	Find the general sol. of $A\underline{x} = \underline{0}$. Write the solution in parametric vector form where the weights are the free var's. The corresponding vectors form a basis for Nul A.
Col A	The pivot columns of A (so, of A itself, and thus not the pivot columns of a reduced form of A)
Row A	The nonzero rows of an echelon form of A.

Lecture 9: Eigenvalues and Eigenvectors.
(book: 5.1, 5.2)

Previous episode: Vector Spaces.
Next episode: Diagonalization.

$n \times n$ matrix A $\qquad T: \underline{x} \mapsto A\underline{x}$



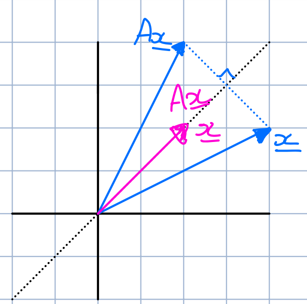
eigenvector of A.

not an eigenvector of A.

A non-zero $\underline{x} \in \mathbb{R}^n$ is an eigenvector of A when $A\underline{x} = \lambda \underline{x}$ for some scalar $\lambda \to$ eigenvalue.

in words: A produces a scalar multiple of $\underline{x}$ (the direction does not change).
(apart from a minus sign) ⚠

Example: $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ $\qquad A\underline{x} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}$

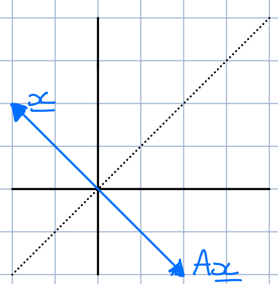


A doesn't change the direction of the vectors on the $x_2 = x_1$

$A\begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = 1 \cdot \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

So, $\begin{bmatrix} 2 \\ 2 \end{bmatrix}$ is an eigenvector with eigenvalue 1.
$\lambda_1 = 1$.

↳ any vector of the form $\begin{bmatrix} t \\ t \end{bmatrix}$ with $t \neq 0$.



Each vector perpendicular to the line $x_2 = x_1$ is also an eigenvector.

$A\begin{bmatrix} -3 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \end{bmatrix} = -1 \cdot \begin{bmatrix} -3 \\ 3 \end{bmatrix}$.

So, $\begin{bmatrix} -3 \\ 3 \end{bmatrix}$ is an eigenvector with eigenvalue -1.
$\lambda_2 = -1$.

↳ any vector of the form $\begin{bmatrix} -t \\ t \end{bmatrix}$ with $t \neq 0$.

Suppose $A\underline{x} = \underline{0}$ has a non-trivial solution.
$\Rightarrow \exists \underline{x} \neq \underline{0} : A\underline{x} = 0 \cdot \underline{x}$

So, each non-trivial solution is an eigenvector with eigenvalue 0.

Recall: A is invertible $\Leftrightarrow A\underline{x} = \underline{0}$ has only the trivial solution.
$\Leftrightarrow$ 0 is not an eigenvalue of A.

Example: Is $\underline{u} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ an eigenvector of $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$? No.

$A\underline{u} = \lambda \underline{u}$?

$$A\underline{u} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \neq \lambda \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Is a scalar p an eigenvalue of A?

$A\underline{x} = p \cdot \underline{x}$ for some vector $\underline{x} \neq \underline{0}$?

$\Leftrightarrow A\underline{x} - p\underline{x} = \underline{0} \Leftrightarrow (A - pI)\underline{x} = \underline{0}$ has a nontrivial sol?

$\Leftrightarrow A - pI$ has a free variable?

Example: $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ Is 1 an eigenvalue?

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

x_2 is a free var
So, 1 is an eigenvalue.

What are the corresponding eigenvectors?

$$\left[\begin{array}{cc|c} 1 & -1 & 0 \\ 0 & 0 & 0 \end{array}\right] \qquad \underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = x_2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

The eigenspace of λ is $\mathrm{Nul}(A - \lambda I) = \mathrm{Span}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}$

$\hookrightarrow$ a vector space that consists of $\underline{0}$ and all eigenvectors corresponding to the eigenvalue λ.

How to find the eigenvalues?

λ is an eigenvalue $\Leftrightarrow (A - \lambda I)\underline{x} = \underline{0}$ has nontrivial sols.
$\Leftrightarrow A - \lambda I$ is not invertible.
$\Leftrightarrow \det(A - \lambda I) = 0$.

So, solve $\underline{\det(A-\lambda I)=0}$ for λ.

$\Leftrightarrow$ polynomial of degree n. (characteristic equation / polynomial).

Example: Find the eigenvalues of $A=\begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix}$.

$$A-\lambda I=\begin{bmatrix}2 & 1\\ 1 & 2\end{bmatrix}-\lambda\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}=\begin{bmatrix}2-\lambda & 1\\ 1 & 2-\lambda\end{bmatrix}$$

$$\det(A-\lambda I)=(2-\lambda)^2-1=\lambda^2-4\lambda+3.$$

$$\det(A-\lambda I)=0 \iff \lambda^2-4\lambda+3=0 \iff (\lambda-1)(\lambda-3)=0$$
$$\iff \lambda_1=1,\ \lambda_2=3.$$

And find the corresponding eigenvectors.

$$\lambda_1=1:\quad A-\lambda_1 I=\left[\begin{array}{cc|c}1 & 1 & 0\\ 1 & 1 & 0\end{array}\right]\sim\left[\begin{array}{cc|c}1 & 1 & 0\\ 0 & 0 & 0\end{array}\right]\qquad \underline{x}=x_2\begin{bmatrix}-1\\ 1\end{bmatrix}\text{ with } x_2\neq 0.$$

$$\lambda_2=3:\quad A-\lambda_2 I=\left[\begin{array}{cc|c}-1 & 1 & 0\\ 1 & -1 & 0\end{array}\right]\sim\left[\begin{array}{cc|c}1 & -1 & 0\\ 0 & 0 & 0\end{array}\right]\qquad \underline{x}=x_2\begin{bmatrix}1\\ 1\end{bmatrix}\text{ with } x_2\neq 0.$$

Example: $A=\begin{bmatrix}5 & 5\\ 5 & 5\end{bmatrix}$

$$\underline{v}=\begin{bmatrix}1\\ 1\end{bmatrix}\qquad A\underline{v}=\begin{bmatrix}5 & 5\\ 5 & 5\end{bmatrix}\begin{bmatrix}1\\ 1\end{bmatrix}=\begin{bmatrix}10\\ 10\end{bmatrix}=10\cdot\begin{bmatrix}1\\ 1\end{bmatrix}\qquad \lambda_1=10.$$

$$\underline{u}=\begin{bmatrix}1\\ -1\end{bmatrix}\qquad A\underline{u}=\begin{bmatrix}5 & 5\\ 5 & 5\end{bmatrix}\begin{bmatrix}1\\ -1\end{bmatrix}=\begin{bmatrix}0\\ 0\end{bmatrix}=0\cdot\begin{bmatrix}1\\ -1\end{bmatrix}\qquad \lambda_2=0.$$

Example: $A=\begin{bmatrix}1 & 2 & 5\\ 4 & 2 & 2\\ 6 & 1 & 1\end{bmatrix}$

$$\begin{bmatrix}1 & 2 & 5\\ 4 & 2 & 2\\ 6 & 1 & 1\end{bmatrix}\begin{bmatrix}1\\ 1\\ 1\end{bmatrix}=\begin{bmatrix}8\\ 8\\ 8\end{bmatrix}=8\cdot\begin{bmatrix}1\\ 1\\ 1\end{bmatrix}$$

Example: $A=\begin{bmatrix}5 & 1 & 0\\ 0 & -3 & 1\\ 0 & 0 & -3\end{bmatrix}$

$$A - \lambda I = \begin{bmatrix} 5-\lambda & 1 & 0 \\ 0 & -3-\lambda & 1 \\ 0 & 0 & -3-\lambda \end{bmatrix}$$

$$|A - \lambda I| = (5-\lambda)(-3-\lambda)(-3-\lambda) = 0$$

So, $\lambda_1 = 5$ and $\lambda_2 = -3$ (with multiplicity 2).

So, for triangular or diagonal matrices, the eigenvalues are the entries on the main diagonal !

Example $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ $\left(= \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix} \text{ with } \varphi = \pi/2 \right)$

Give me the eigenvalues.

$\mathbb{R}^2 \to \mathbb{R}^2$.

$\mathbb{C}^2 \to \mathbb{C}^2$

$$|A - \lambda I| = \begin{vmatrix} -\lambda & -1 \\ 1 & -\lambda \end{vmatrix} = \lambda^2 + 1$$

$\lambda^2 + 1 = 0 \iff \lambda^2 = -1$

So, $\lambda_1 = i$ and $\lambda_2 = -i$

So, eigenvalues can also be complex numbers.

$\lambda_1 = i$: $A - \lambda_1 I = \begin{bmatrix} -i & -1 \\ -1 & -i \end{bmatrix} \sim \begin{bmatrix} -i & -1 \\ 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix}$

$$\underline{x} = x_2 \begin{bmatrix} i \\ 1 \end{bmatrix}$$

$\lambda_2 = -i$: . . . $\underline{x} = x_2 \begin{bmatrix} -i \\ 1 \end{bmatrix}$.

DIY : $A = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$ Determine the eigenvalues and eigenvectors.

Properties:

* A is invertible $\iff$ 0 is not an eigenvalue of A.

* Exc 19, Ch 5.2: $\det(A) = \lambda_1 \cdot \lambda_2 \cdot \ldots \cdot \lambda_n$.

* $\underbrace{\sum_{i=1}^{n} a_{ii}}_{\to \text{ trace}(A).} = \lambda_1 + \lambda_2 + \ldots + \lambda_n$.

* Thm 2: If $\underline{v}_1, \ldots, \underline{v}_r$ are eigenvectors that correspond to distinct eigenvalues $\lambda_1, \ldots, \lambda_r$ of a matrix A, then

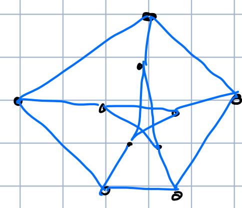
$\{\underline{v}_1, \ldots, \underline{v}_r\}$ is linearly independent.

Applications to Graph Theory.

Given adjacency matrix A ($n \times n$, symmetric)

$\lambda_1 \geq \lambda_2 \geq \lambda_3 \geq \cdots \geq \lambda_n$.

$\chi(G)$ chromatic number.



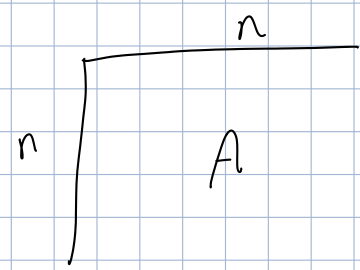
$\chi(G) = 3$.

Petersen graph.

* Hoffmann lower bound

$$\chi(G) \geq 1 + \frac{\lambda_1}{-\lambda_n}$$

* Wilf upper bound $\chi(G) \leq \lambda_1 + 1$.



More applications? Read the Google Page Rank algorithm paper.

Summary (so far):

Let A be an $m \times n$ matrix with columns $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n$.

$m \geq n$: $\begin{bmatrix} A \end{bmatrix}$

The following statements are equivalent:

1. A has a pivot in every column
2. A has n pivot positions
3. There are no free variables
4. $A\underline{x} = \underline{0}$ has only the trivial sol.
5. $\{\underline{a}_1, \underline{a}_2, \dots \underline{a}_n\}$ is linearly indep.
6. $T: \underline{x} \mapsto A\underline{x}$ is one-to-one/injective
7. $\operatorname{Nul} A = \{\underline{0}\}$
8. $\dim \operatorname{Nul} A = 0$
9. $\operatorname{rank} A = n$

$m \leq n$: $\begin{bmatrix} \quad A \quad \end{bmatrix}$

The following statements are equivalent:

a. A has a pivot in every row.
b. A has m pivot positions.
c. The echelon form of A does not contain a row of all zeros.
d. $A\underline{x} = \underline{b}$ is consistent for every $\underline{b}$ in $\mathbb{R}^m$.
e. $\operatorname{Span}\{\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n\} = \mathbb{R}^m$.
f. $T: \underline{x} \mapsto A\underline{x}$ is onto/surjective.
g. $\operatorname{Col} A = \mathbb{R}^m$
h. $\dim \operatorname{Col} A = m$
i. $\operatorname{rank} A = m$

If A is square ($n = m$), then statements 2 and b are equivalent.
Hence, the following statements are equivalent for square matrices.

* 1 - 9, a - i
* A is invertible
* There is a matrix C such that $CA = I_n$ and $AC = I_n$
* A is row equivalent to I_n.
* A^T is invertible.
* $\det A \neq 0$
* The columns of A form a basis for $\mathbb{R}^n$.
* 0 is not an eigenvalue of A.

Lecture 10: Diagonalization
(book: 5.2, 5.3)

Previous episode: Eigenvalues and Eigenvectors.
Next episode: Orthogonality and Symmetric Matrices.

Let A and B be two $n \times n$ matrices.
A and B are similar $\iff$ $\exists$ invertible matrix P st
$A = PBP^{-1}$ or $B = P^{-1}AP$.

Theorem: If A and B are similar, then they have the same eigenvalues.

Proof: $|B - \lambda I| = |P^{-1}AP - \lambda I| = |P^{-1}AP - \lambda P^{-1}P| = |P^{-1}(AP - \lambda P)|$

$= |P^{-1}(AP - \lambda IP)| = |P^{-1}(A - \lambda I)P| = |P^{-1}| \cdot |A - \lambda I| \cdot |P|$

$= \frac{1}{\cancel{|P|}} \cdot |A - \lambda I| \cdot \cancel{|P|} = |A - \lambda I|$

So, A and B have the same characteristic equation
$\Rightarrow$ same eigenvalues $\square$

$A^k = \underbrace{A \cdot A \cdot A \cdot \ldots \cdot A \cdot A}_{k \text{ times}}$

For a diagonal matrix this is easy.

$D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ $\quad D^2 = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2^2 & 0 \\ 0 & 3^2 \end{bmatrix}$

$D^3 = D^2 \cdot D = \begin{bmatrix} 2^2 & 0 \\ 0 & 3^2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 2^3 & 0 \\ 0 & 3^3 \end{bmatrix}$ $\quad D^k = \begin{bmatrix} 2^k & 0 \\ 0 & 3^k \end{bmatrix}$

If A is similar to a diagonal matrix, then it's also easy.
$\hookrightarrow A = PDP^{-1}$.

$A^k = \underbrace{A \cdot A \cdot A \cdot \ldots \cdot A \cdot A}_{k \text{ times}} = \underbrace{PD\cancel{P^{-1}} \cdot \cancel{P}D\cancel{P^{-1}} \cdot \cancel{P}D\cancel{P^{-1}} \cdot \ldots \cdot \cancel{P}D\cancel{P^{-1}} \cdot \cancel{P}DP^{-1}}_{k \text{ times}}$.

$= PD^kP^{-1}$.

A is called diagonalizable if A is similar to a diagonal matrix.

* How to build D?

$\begin{cases} A \text{ is similar to } D \Rightarrow A \text{ and } D \text{ have the same eigenvalues } \lambda_1, \dots, \lambda_n. \\ D \text{ is a diagonal matrix} \Rightarrow \text{the eigenvalues are on the diagonal.} \end{cases}$

$$D = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}.$$

* How to build P?

$P = [\underline{v}_1 \ \cdots \ \underline{v}_n]$ where $\{\underline{v}_1, \dots, \underline{v}_n\}$ are lin indep vectors in $\mathbb{R}^n$.
↳ because P is invertible.

$A = PDP^{-1} \Rightarrow AP = PD$, where

$AP = A[\underline{v}_1 \ \cdots \ \underline{v}_n] = [A\underline{v}_1 \ A\underline{v}_2 \ \cdots \ A\underline{v}_n]$

and $PD = [\underline{v}_1 \ \cdots \ \underline{v}_n] \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} = [\lambda_1 \underline{v}_1 \ \lambda_2 \underline{v}_2 \ \cdots \ \lambda_n \underline{v}_n].$

So, for $AP = PD$ we need $A\underline{v}_i = \lambda_i \underline{v}_i \quad \forall i \in \{1, \dots, n\}$.

$\Rightarrow \underline{v}_1, \dots, \underline{v}_n$ are eigenvectors corresponding to the eigenvalues $\lambda_1, \dots, \lambda_n$.

So, A is diagonalizable $\Leftrightarrow$ A has n lin indep eigenvectors.

If there are n distinct eigenvalues $\Rightarrow$ n lin indep. eigenvectors.

Monday: eigenvectors from different eigenspaces are lin indep.

$\Rightarrow A$ is diagonalizable.

Example: $A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}$ $\quad \lambda_1 = 1 \quad \lambda_2 = 2 \quad \lambda_1 \neq \lambda_2$:)

So, A is diagonalizable. and $D = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ What is P?

eigenspace of $\lambda_1 = 1$: $[A - 1 \cdot I \mid 0] = \left[\begin{array}{cc|c} 0 & 2 & 0 \\ 0 & 1 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 0 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \quad \underline{x} = \begin{bmatrix} x_1 \\ 0 \end{bmatrix} = x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$

So, take for example $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$.

eigenspace of $\lambda_2 = 2$: $[A-2I \,\vdots\, 0] = \begin{bmatrix} -1 & 2 & \vdots & 0 \\ 0 & 0 & \vdots & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & \vdots & 0 \\ 0 & 0 & \vdots & 0 \end{bmatrix}$ $\underline{x} = x_2 \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

So, take for example $\begin{bmatrix} 2 \\ 1 \end{bmatrix}$.

Then, $P = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}$

Check: $A = PDP^{-1}$. $AP = PD$.

$$AP = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 0 & 2 \end{bmatrix}$$

$$PD = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 0 & 2 \end{bmatrix} \quad \checkmark$$

---

* What if the eigenvalues of A are not all distinct?
(some have mult. >1)

$\lambda_1 = 1,\ \lambda_2 = 2,\ \lambda_3 = 2.$
$(\lambda - 1)(\lambda - 2)^2 = 0.$

Is A diagonalizable? It depends....

# of lin indep eigenvectors corresponding to λ = dim of the eigenspace of λ = $\dim(\text{Nul}(A-\lambda I)) \leq$ mult. of λ.

if strictly $<$ for some λ, then A is not diagonalizable.

Theorem: an $n \times n$ matrix A is diagonalizable $\iff$ the sum of the dimensions of the eigenspaces equals n.

Example: $A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ Is A diagonalizable?

$\lambda = 0$ with mult 2.

$A - 0I = A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$ 1 free var. $x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

So, $\dim(\text{Nul}(A-\lambda I)) = 1 < 2 =$ multiplicity of λ.
and thus A is not diagonalizable.

Example $A = \begin{bmatrix} -4 & 1 & 1 \\ 2 & -3 & 2 \\ 3 & 3 & -2 \end{bmatrix}$

* Is -5 an eigenvalue?
* Is A diagonalizable?

$A - (-5)I = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{bmatrix} \sim \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

There are free variables, so -5 is indeed an eigenvalue.
$\dim(\text{Nul}(A - (-5)I)) = 2$.

So, the multiplicity of $\lambda=-5$ is either 2 or 3.
yes ← 2 ; 3 → no.

$\text{trace}(A) = (-4)+(-3)+(-2) = -9$
$\text{trace}(A) = \lambda_1+\lambda_2+\lambda_3 = (-5)+(-5)+\lambda_3 = -10+\lambda_3$

$\left.\begin{matrix} \\ \\ \end{matrix}\right\} -9 = -10+\lambda_3 \Rightarrow \lambda_3 = 1.$

So, A is diagonalizable because the multiplicity of $\lambda=-5$ is 2.

$$D = \begin{bmatrix} -5 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

eigenspace of -5: $\underline{x} = x_2 \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ $\quad \underline{v}_1 = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$ $\quad \underline{v}_2 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$

eigenspace of 1: $A - 1\cdot I = \begin{bmatrix} -5 & 1 & 1 \\ 2 & -4 & 2 \\ 3 & 3 & -3 \end{bmatrix} \sim \dots \sim \begin{bmatrix} 1 & 0 & -1/3 \\ 0 & 1 & -2/3 \\ 0 & 0 & 0 \end{bmatrix}$

$\underline{x} = x_3 \begin{bmatrix} 1/3 \\ 2/3 \\ 1 \end{bmatrix}$ $\qquad$ So, $\underline{v}_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$

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

DIY: Verify that $AP = PD$. ✓

Applications of diagonalization:
* Markov Processes.
* Dynamical Systems.
* Difference Equations.

Summary. Is A diag?

Does A have n distinct eigenvalues?

Yes

Yes, $A = PDP^{-1}$

No.

Is the sum of the dimensions of the eigenspaces equal to n?

Yes

Yes, $A = PDP^{-1}$

No

No, A is not diag.

$\mathbb{C}^2 \to \mathbb{C}^2$ $\quad T(\underline{x}) = A\underline{x}$ $\quad A = \begin{bmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{bmatrix}$

$\lambda_1 = \cos\varphi + i\cdot\sin\varphi$ $\quad \underline{x} = x_2 \begin{bmatrix} i \\ 1 \end{bmatrix}$

$\lambda_2 = \cos\varphi - i\cdot\sin\varphi$ $\quad \underline{x} = x_2 \begin{bmatrix} -i \\ 1 \end{bmatrix}$ DIY.

$V = \mathbb{R}^+$.

* addition of vectors u and : $u * v$.
* multiplication of a vector u by a scalar c : u^c.

What is the zero vector in this context? 1

Verify the 10 axioms. ✓

Lecture 11: Orthogonality and Symmetric Matrices
(book: 6.1, 6.2, 7.1)

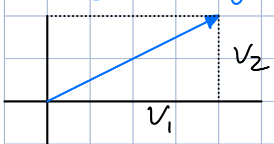
Previous episode: Diagonalization
Next episode: Old Exam (Resit 2022-2023)

---

inner/dot product $\underline{u}, \underline{v} \in \mathbb{R}^n$: $\underline{u} \bullet \underline{v} = \underline{u}^T \cdot \underline{v} = [u_1 \ \dots \ u_n] \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = u_1v_1 + \dots + u_nv_n$

Properties:
- $\underline{u} \cdot \underline{v} = \underline{v} \cdot \underline{u}$ (commutativity)
- $(\underline{u} + \underline{v}) \cdot \underline{w} = \underline{u} \cdot \underline{w} + \underline{v} \cdot \underline{w}$ (distributivity)
- $\underline{0} \cdot \underline{u} = 0$
- $(c \cdot \underline{u}) \cdot \underline{v} = c \cdot (\underline{u} \cdot \underline{v}) = u \cdot (c \cdot \underline{v})$
- $\underline{u} \cdot \underline{u} \geq 0$
- $\underline{u} \cdot \underline{u} = 0 \iff \underline{u} = \underline{0}$.

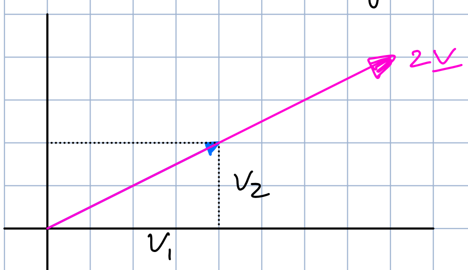
Length of a vector:



$\underline{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ Length $= \sqrt{v_1^2 + v_2^2} = \sqrt{\underline{v} \bullet \underline{v}}$

Similar in $\mathbb{R}^3$: $\underline{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$ length $= \sqrt{v_1^2 + v_2^2 + v_3^2} = \sqrt{\underline{v} \cdot \underline{v}}$

length/norm of a vector $\underline{v} \in \mathbb{R}^n$: $\|\underline{v}\| = \sqrt{\underline{v} \cdot \underline{v}}$



$c\underline{v} = \begin{bmatrix} c \cdot v_1 \\ c \cdot v_2 \end{bmatrix}$ $\|c \underline{v}\| = |c| \cdot \|\underline{v}\|$ DIY.

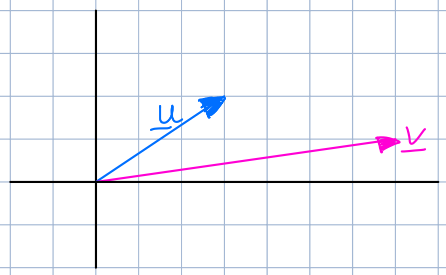
unit vector: Vector of length 1. For example, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$

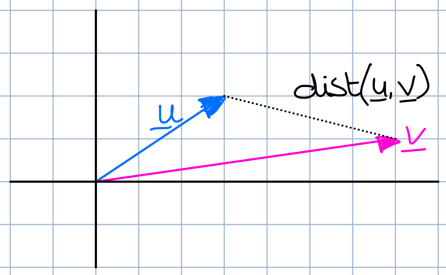
$\underline{v} = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$ $\|\underline{v}\| = \sqrt{2^2 + 1^2} = \sqrt{5}$.

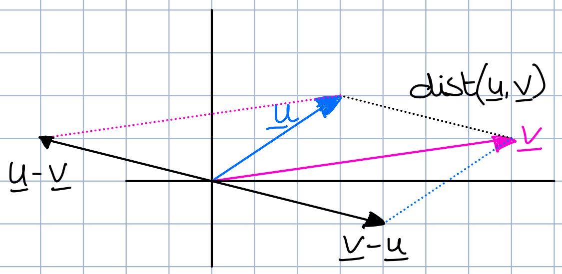
$\underline{u} = \frac{1}{\|\underline{v}\|} \underline{v} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$ $\underline{w} = \begin{bmatrix} -2/\sqrt{5} \\ -1/\sqrt{5} \end{bmatrix}$

normalizing vector $\underline{v}$.

Distance between two vectors:



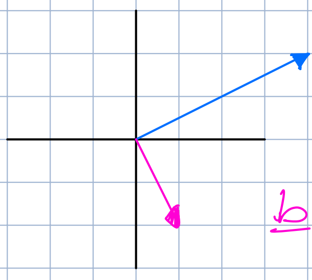




Hence, $\text{dist}(\underline{u}, \underline{v}) = \|\underline{v} - \underline{u}\| = \|\underline{u} - \underline{v}\|$.

Two vectors are orthogonal ($\underline{u} \perp \underline{v}$) if $\underline{u} \cdot \underline{v} = 0$.
"perpendicular"

$\iff \|\underline{u}+\underline{v}\|^2 = \|\underline{u}\|^2 + \|\underline{v}\|^2$.



$\underline{a} = \begin{bmatrix} 4 \\ 2 \end{bmatrix}$

$\underline{b} = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$

$\underline{a} \cdot \underline{b} = 4 \cdot 1 + 2 \cdot (-2) = 4 - 4 = 0$

$\underline{a}$ is perpendicular to $\underline{b}$.

Applications:

* it relates to the correlation coefficient in statistics
* it is important for matching & feature detection in signal and image processing.

Recall the Null space of a matrix A.

$\mathrm{Nul}(A) = \{\underline{x} \in \mathbb{R}^n : A\underline{x} = \underline{0}\}$

$$\begin{bmatrix} \text{---} \; \underline{r}_1 \; \text{---} \\ \text{---} \; \underline{r}_2 \; \text{---} \\ \vdots \\ \text{---} \; \underline{r}_m \; \text{---} \end{bmatrix} \begin{bmatrix} \\ \underline{x} \\ \\ \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$\begin{cases} \underline{r}_1^T \cdot \underline{x} = 0 \\ \underline{r}_2^T \cdot \underline{x} = 0 \\ \vdots \\ \underline{r}_m^T \cdot \underline{x} = 0 \end{cases}$$

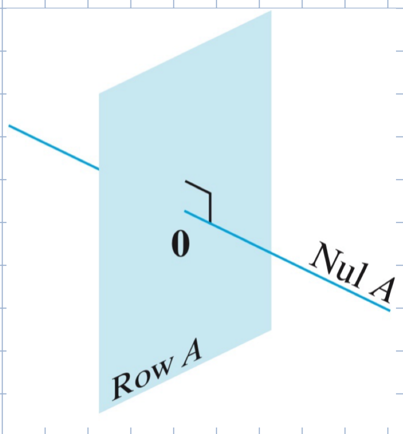
So, every $\underline{x} \in \mathrm{Nul}(A)$ is orthogonal to each of the rows of A.

$$(c_1 \cdot \underline{r}_1^T + c_2 \cdot \underline{r}_2^T + \dots + c_m \cdot \underline{r}_m^T) \cdot \underline{x} = c_1(\underline{r}_1^T \cdot \underline{x}) + c_2 (\underline{r}_2^T \cdot \underline{x}) + \dots + c_m(\underline{r}_m^T \cdot \underline{x})$$
$$= c_1 \cdot 0 + c_2 \cdot 0 + \dots + c_m \cdot 0 = 0.$$

So, every vector in $\mathrm{Nul}(A)$ is orthogonal to every vector in $\mathrm{Row}(A)$

$\Rightarrow \underset{\substack{\downarrow \\ \text{in } \mathbb{R}^n \\ \dim = \#\text{ free vars}}}{\mathrm{Nul}(A)} \perp \underset{\substack{\downarrow \\ \text{in } \mathbb{R}^n \\ \dim = \#\text{ pivot rows}}}{\mathrm{Row}(A)}.$

Recall: $\sum \text{dims} = n.$



And, since $\mathrm{Col}(A) = \mathrm{Row}(A^T)$, we also have $\mathrm{Col}(A) \perp \mathrm{Nul}(A^T)$.

W: subspace of $\mathbb{R}^n$.

$W^\perp$ ("W perpendicular"): orthogonal complement of W.
$\hookrightarrow$ all vectors in $\mathbb{R}^n$ that are orthogonal to W.

$(\mathrm{Row}(A))^\perp = \mathrm{Nul}(A)$. $\quad ((\mathrm{Col}(A))^\perp = \mathrm{Nul}(A^T)$.

$W^\perp$ is also a subspace of $\mathbb{R}^n$ (exc. 38 Ch 6.1).
$(W^\perp)^\perp = W$.

In general, $\dim(W) + \dim(W^\perp) = n$.

$\{\underline{v}_1, \dots, \underline{v}_k\}$ is an orthogonal set if $\underline{v}_i \cdot \underline{v}_j = 0$ for all $i \neq j$.

Theorem: If $S = \{\underline{v}_1, \dots, \underline{v}_k\}$ is an orthogonal set and $\underline{0} \notin S$, then S is linearly independent and thus S forms a basis for $\mathrm{Span}\{\underline{v}_1, \dots, \underline{v}_k\}$.

Proof: Book Thm 4.

$\{\underline{v}_1, \dots, \underline{v}_k\}$ is an orthonormal set if it is an orthogonal set of unit vectors.
$\hookrightarrow$ vectors of length 1.

How to test whether $\{\underline{v}_1, \dots, \underline{v}_k\}$ is orthogonal/orthonormal?

Create $A = \begin{bmatrix} | & & | \\ \underline{v}_1 & \cdots & \underline{v}_k \\ | & & | \end{bmatrix}$

Compute $A^TA = \begin{bmatrix} - & v_1^T & - \\ & \vdots & \\ - & v_k^T & - \end{bmatrix} \begin{bmatrix} | & & | \\ v_1 & \cdots & v_k \\ | & & | \end{bmatrix} = \begin{array}{cc} & j \\ i & \begin{bmatrix} & & \\ & \underline{v}_i \cdot \underline{v}_j & \\ & & \end{bmatrix} \end{array}$

$\{\underline{v}_1, \dots, \underline{v}_k\}$ is orthogonal $\iff A^TA$ is diagonal.

$\{\underline{v}_1, \dots, \underline{v}_k\}$ is orthonormal $\iff A^TA$ is identity matrix.

A square matrix A is an orthogonal matrix $\iff A^TA = I_n \iff A^{-1} = A^T$.

Watch out the terminology: an orthgonal matrix has orthonormal columns

---

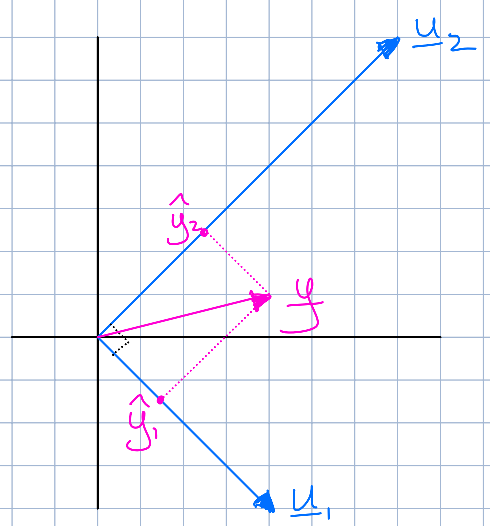
Orthogonal basis for a subspace W of $\mathbb{R}^n$: it is a basis of W, where the vectors form an orthogonal set.

Let $\{\underline{u}_1, \ldots, \underline{u}_k\}$ be an orthogonal basis of W.
Let $\underline{y} \in W$.
Then, $\underline{y} = c_1 \cdot \underline{u}_1 + \ldots + c_k \cdot \underline{u}_k$.
What are the **weights** $c_1, \ldots, c_k$?

$$\underline{y} \bullet \underline{u}_1 = (c_1 \cdot \underline{u}_1 + \ldots + c_k \cdot \underline{u}_k) \bullet \underline{u}_1 = c_1\, \underline{u}_1 \bullet \underline{u}_1 + \cancel{c_2\, \underline{u}_2 \bullet \underline{u}_1} + \cancel{\ldots} + \cancel{c_k\, \underline{u}_k \bullet \underline{u}_1}$$
$$= c_1\, \underline{u}_1 \bullet \underline{u}_1$$

$$\Rightarrow c_1 = \frac{\underline{y} \bullet \underline{u}_1}{\underline{u}_1 \bullet \underline{u}_1} \quad \ldots \ldots \quad c_k = \frac{\underline{y} \bullet \underline{u}_k}{\underline{u}_k \bullet \underline{u}_k}$$

So, it's easy to find the weights ☺
(non-orthogonal basis: solving an SLE).



$$\underline{u}_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad \underline{y} = \begin{bmatrix} 3 \\ 5 \end{bmatrix}.$$

$$\frac{3 \cdot 2 + 5 \cdot 1}{2 \cdot 2 + 1 \cdot 1}$$

$$\underline{y} = \hat{y}_1 + \hat{y}_2 = \underbrace{\frac{\underline{y} \bullet \underline{u}_1}{\underline{u}_1 \bullet \underline{u}_1}\, \underline{u}_1}_{\substack{\text{orthogonal projection} \\ \text{of } \underline{y} \text{ onto } \underline{u}_1.}} + \underbrace{\frac{\underline{y} \bullet \underline{u}_2}{\underline{u}_2 \bullet \underline{u}_2}\, \underline{u}_2}_{\substack{\text{orthogonal projection} \\ \text{of } \underline{y} \text{ onto } \underline{u}_2.}}$$

Recall from a previous episode:
A is diagonalizable $\Longleftrightarrow$ the sum of the dimensions of the eigenspaces equals n.

Symmetric matrix: $A = A^T$.

$$\begin{bmatrix} -1 & 6 & -4 \\ 6 & 2 & 0 \\ -4 & 0 & 3 \end{bmatrix}$$

For an $n \times n$ symmetric matrix:

* All eigenvalues are real numbers.
* Eigenvectors from different eigenspaces are orthogonal
* A is diagonalizable. !

Proof: DIY..

A is called orthogonally diagonalizable if there is orthogonal matrix P and a diagonal matrix D such that $A = PDP^{-1}$ or $A = PDP^T$

A is orthogonally diagonalizable $\Longleftrightarrow$ A is symmetric.

Example: $A = \begin{bmatrix} 3 & -2 & 4 \\ -2 & 6 & 2 \\ 4 & 2 & 3 \end{bmatrix}$ with eigenvalues -2 and 7.

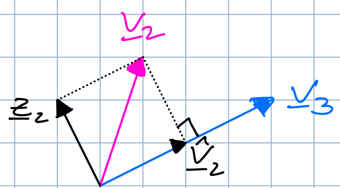
Orthogonally diagonalize A.

$$A-(-2)I = \begin{bmatrix} 5 & -2 & 4 \\ -2 & 8 & 2 \\ 4 & 2 & 5 \end{bmatrix} \sim \dots \sim \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1/2 \\ 0 & 0 & 0 \end{bmatrix} \quad \underline{x} = x_3 \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix} \quad \underline{v}_1 = \begin{bmatrix} -1 \\ -1/2 \\ 1 \end{bmatrix}$$

$$A - 7 \cdot I = \begin{bmatrix} -4 & -2 & 4 \\ -2 & -1 & 2 \\ 4 & 2 & -4 \end{bmatrix} \sim \dots \sim \begin{bmatrix} 1 & 1/2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \underline{x} = x_2 \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} + x_3 \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

$$\underline{v}_2 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} \qquad \underline{v}_3 = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

We need to make $\underline{v}_2$ and $\underline{v}_3$ orthogonal.



Projection of $\underline{v}_2$ onto $\underline{v}_3$: $\hat{\underline{v}}_2 = \dfrac{\underline{v}_2 \cdot \underline{v}_3}{\underline{v}_3 \cdot \underline{v}_3} \underline{v}_3 = \dfrac{-1/2}{2} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} -1/4 \\ 0 \\ -1/4 \end{bmatrix}$

Component of $\underline{v}_2$ orthogonal to $\underline{v}_3$:

$$\underline{z}_2 = \underline{v}_2 - \hat{\underline{v}}_2 = \begin{bmatrix} -1/2 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} -1/4 \\ 0 \\ -1/4 \end{bmatrix} = \begin{bmatrix} -1/4 \\ 1 \\ 1/4 \end{bmatrix}$$

Note: $\underline{z}_2$ is also an eigenvector because it's a linear combination of $\underline{v}_2$ and $\underline{v}_3$.

Moreover $\underline{z}_2 \perp \underline{v}_3$

So $\{\underline{z}_2, \underline{v}_3\} = \left\{ \begin{bmatrix} -1/4 \\ 1 \\ 1/4 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ forms an orthogonal basis for the eigenspace.

Normalize $\underline{v}_1, \underline{z}_2, \underline{v}_3$: $\underline{u}_1 = \begin{bmatrix} -2/3 \\ -1/3 \\ 2/3 \end{bmatrix}$, $\underline{u}_2 = \begin{bmatrix} -1/\sqrt{18} \\ 4/\sqrt{18} \\ 1/\sqrt{18} \end{bmatrix}$, $\underline{u}_3 = \begin{bmatrix} 1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$

↳ make the length equal to 1.

Then, $P = [\underline{u}_1 \; \underline{u}_2 \; \underline{u}_3]$ and $D = \begin{bmatrix} -2 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 7 \end{bmatrix}$. $A = PDP^{-1}$. $A = PDP^T$

↳ orthogonal matrix

$P^TP = I_n$.