MAT 22A - Linear Algebra

Kevin Lamb - MSB 2117 - MWF 1-3pm
kmlamb@math.ucdavis.edu
.../~kmlamb/MAT%2022A%20SS1%2015.html

Systems of Linear Equations

§1.1 - Lay §2.1, 1.3, 2.2 - Strang

Def: A linear equation in the variables $x_1, \ldots, x_n$ is an equation that can be written in the form

$$a_1 x_1 + a_2 x_2 + \cdots + a_n x_n = b$$

where b and the coefficients $a_1, \ldots, a_n$ are real (or complex) numbers.

Ex:

Linear	Non-linear
$4x - 5y + z = x$	$x_1^2 + x_2^2 = 1$
$x_2 = 2(\sqrt{3} - x_1) + x_3$	$\sin x_1 - x_2 + 3 = 0$
$x_1 + x_2 + x_3 - x_4 = 5$	$x_1 x_2 + \sqrt{x_3} - i x_4 = 2$

Def: A <u>system of linear equations</u> is a collection one or more linear equations involving the same variables.

The solution set of the system is the set of all possible ordered n-tuples for $(x_1, \dots, x_n)$ that make all equations true simultaneously. Two systems are equivalent if they have the same solution set.

Ex: ① $\begin{cases} x_1 + 2x_2 = -1 \\ 2x_1 - x_2 = 3 \end{cases}$

$\swarrow$ solve for one var and sub.

$\searrow$

$$\begin{aligned} 2(x_1 + 2x_2) &= 2(-1) \\ -(2x_1 - x_2 &= 3) \\ \hline 5x_2 &= -5 \\ x_2 &= -1 \end{aligned}$$

$\rightarrow x_1 + 2(-1) = -1$

$x_1 = 1$

$(x_1, x_2) = \boxed{(1, -1)}$

Solution set $= \{(1,-1)\}$

② $\begin{cases} x_1 + 2x_2 = -1 \\ 2x_1 + 4x_2 = -2 \end{cases} \rightarrow \begin{cases} x_1 + 2x_2 = -1 \\ 0 = 0 \end{cases} \rightarrow \begin{cases} x_1 = -1 - 2x_2 \\ x_2 \text{ free} \end{cases}$

Solutions $= \{ (1,-1), (-1,0), \dots \}$

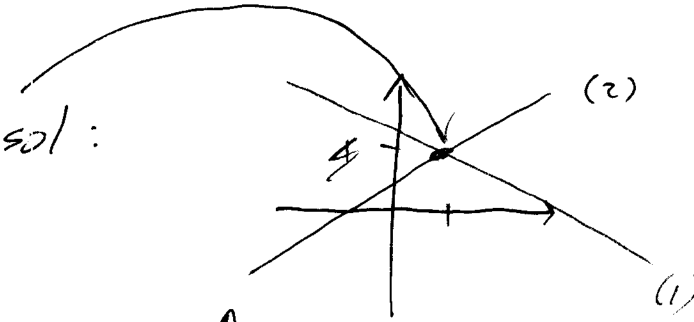
③ $\begin{cases} x_1 + 2x_2 = -1 \\ 2x_1 + 4x_2 = -3 \end{cases} \rightarrow \begin{cases} x_1 + 2x_2 = -1 \\ 0 = 1 \end{cases}$

$\boxed{\text{No solutions!}}$

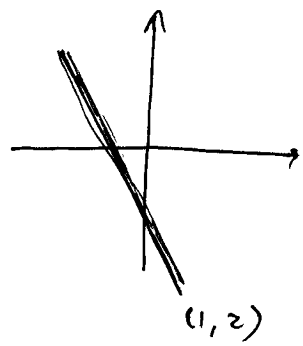
Solution set $= \{\} = \emptyset$

THEME:

① Exactly one sol:

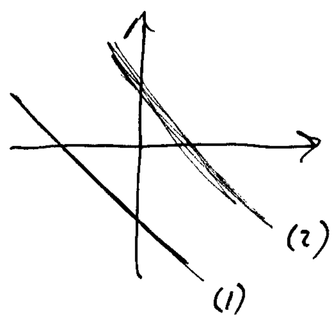


② Inf. many:



Intersect everywhere!

③ No solutions:



Remark: This is a direct consequence of linearity.

Enter: Matrix notation

(left side) coefficient matrix

Ex:

$$\begin{cases} x_1 - 2x_2 + 4x_3 = -1 \\ -x_1 + 5x_2 - x_3 = 0 \\ 2x_1 + x_3 = 2 \end{cases} \longleftrightarrow \left[\begin{array}{ccc|c} 1 & -2 & 4 & -1 \\ -1 & 5 & -1 & 0 \\ 2 & 0 & 1 & 2 \end{array}\right]$$

~~Bar~~ Bar is the equal signs.

GOAL:

Idea -
$$\begin{cases} x_1 \qquad\qquad = a \\ \quad x_2 \qquad = b \\ \qquad\quad x_3 = c \end{cases}$$
Not always possible to get here.

Consolation -
"Back-sub"
$$\begin{cases} x_1 + \cdots \qquad\qquad = a_1 \\ \quad x_2 + \cdots \qquad = a_2 \\ \qquad\quad \ddots \\ \qquad\qquad\quad x_n = a_n \end{cases}$$
Work from bottom up

"Forward-sub"
$$\begin{cases} x_1 \qquad\qquad = a_1 \\ x_2 + \cdots \quad = a_2 \\ \vdots \\ x_n + \cdots \quad = a_n \end{cases}$$
... top down

Ex:
$$\begin{cases} \qquad 2x_2 - 8x_3 = 8 \\ x_1 - 2x_2 + x_3 = 0 \\ -4x_1 + 5x_2 + 9x_3 = -9 \end{cases}$$
"Solve this linear system."

$$\longrightarrow \begin{cases} x_1 - 2x_2 + x_3 = 0 \\ \qquad x_2 - 4x_3 = 4 \\ \qquad\qquad x_3 = 3 \end{cases}$$

$\boxed{(29, 16, 3)}$

As a matrix system:
$$\left[\begin{array}{ccc|c} 0 & 2 & -8 & 8 \\ 1 & -2 & 1 & 0 \\ -4 & 5 & 9 & -9 \end{array}\right] \sim \left[\begin{array}{ccc|c} 1 & -2 & 1 & 0 \\ & 1 & -4 & 4 \\ & & 1 & 3 \end{array}\right]$$

Def: The matrix for a linear system of equations is called the augmented matrix for it. Its size is the ordered pair (m, n) (usually denoted $m \times n$) where m is # of rows and n is # of columns.

Question: How did we solve the systems before?

A: Created an equivalent system where the solution set was easily determined.

Question: How did we find those eq. sys's?

A: ROW REDUCTION

Three Elementary Row Operations

1. Multiply an equation by a (non-zero) number.
2. Interchange two ~~the~~ equations.
3. Replace one equation by adding to it a multiple of another equation.

Question: Why does row reduction work?

A: Homework!

Questions: Given a linear system:

① Does a solution exist? (Consistency)

② If a solution exists, is it the only one? (Uniqueness)

Answer: Tune in next time.

Row Reduction and Echelon Forms

Lay - 1.2 Strang - 2.2, 3.4

Need to find some consistent way to determine whether or not a system has solution (unique)?

~~Def:~~

Def: (Looks scarier than it is)

A rectangular matrix is in ~~ech~~ <u>row echelon form</u> if it has the following properties:

1. All nonzero entries are above any row of all zeros.
2. Each leading entry of a row is in a column to the right of the row ~~above~~. above it.
3. All entries in a column below a leading entry are zeros.

A matrix is in row echelon form (REF) is in <u>reduced row echelon form (RREF)</u> if:

4. The leading entry in each nonzero row is '1'.
5. Each leading 1 is the only nonzero entry in its column.

Ex:

REF	$\left[\begin{array}{ccc\|c} \star & \cdot & \cdot & \cdot \\ 0 & \star & \cdot & \cdot \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$	$\left[\begin{array}{cccccccc\|c} 0 & \star & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \star & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \star & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & \star & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \star \end{array}\right]$
RREF	$\left[\begin{array}{ccc\|c} 1 & 0 & \cdot & \cdot \\ 0 & 1 & \cdot & \cdot \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$	$\left[\begin{array}{cccccccc\|c} 0 & 1 & \cdot & 0 & 0 & 0 & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \cdot & \cdot & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right]$

Theorem: Each matrix is row equivalent to one AND ONLY ONE matrix in RREF.

Remark: "Leading entry" is fluid. The position where it occurs is definite.

Def: A pivot position in a matrix A is a location in A corresponding to a leading 1 of RREF(A).
A pivot column is a column of A containing a pivot.

Ex:
$$\begin{bmatrix} 0 & -3 & -6 & 4 & 9 \\ -1 & -2 & -1 & 3 & 1 \\ -2 & -3 & 0 & 3 & -1 \\ 1 & 4 & 5 & -9 & -7 \end{bmatrix}$$
$\uparrow \quad \uparrow \qquad \quad \uparrow$

Pivots are circled
Pivot columns pointed to

Putting a matrix in RREF (and REF along the way)

Two phases: Forward - Get A into REF
Backward - Use pivots to kill off the rest of their columns

$$A = \begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

Phase I - 1 Start w/ the leftmost (nonzero) column. This is a pivot column w/ pivot on top.

2 Select a nonzero entry in this column and place it at top (2 ele. row. op.?)

pivot

$$\underset{R1 \leftrightarrow R3}{\sim} \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

3 Use el. row. op.'s to zero out below the pivot.

$$\underset{R2 \to R2-R1}{\sim} \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

4 Look at the submatrix that ignores the row containing the pivot just created.

$$\begin{bmatrix} 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

5 Repeat 1 - 4 until no nonzero submatrices remain.

$$\underset{R3 \to 3R2-2R3}{\sim} \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \quad \text{(REF)}$$

Phase II - (1) Scale all rows so that their pivot position's entry is 1.

$$R1 \to \tfrac{1}{3}R1 \qquad R2 \to \tfrac{1}{2}R2 \quad \sim \begin{bmatrix} 1 & -3 & 4 & -3 & 2 & 5 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

(2) Starting with rightmost pivot, create zeros above it.

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

(3) Look at submatrix that ignores the row containing the pivot column just reduced.

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

(4) Repeat (2) & (3) until no submatrices remain.

The matrix A is now in RREF.

Last time:

Explained RREF algorithm with example.

$$A = \left[\begin{array}{ccccc|c} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{array}\right] \longleftrightarrow \begin{cases} 3x_2 - 6x_3 + 6x_4 + 4x_5 = -5 \\ 3x_1 - 7x_2 + 8x_3 - 5x_4 + 8x_5 = 9 \\ 3x_1 - 9x_2 + 12x_3 - 9x_4 + 6x_5 = 15 \end{cases}$$

Phase I $\rightarrow$ REF

Phase II $\rightarrow$ RREF:

$$\left[\begin{array}{ccccc|c} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right]$$

$$\updownarrow$$

$$\begin{cases} x_1 \quad - 2x_3 + 3x_4 = -24 \\ x_2 - 2x_3 + 2x_4 = -7 \\ x_5 = 4 \end{cases}$$

$$\|$$

$$\boxed{\bigstar \quad \begin{cases} x_1 = 2x_3 - 3x_4 - 24 \\ x_2 = 2x_3 - 2x_4 - 7 \\ x_5 = 4. \end{cases} \quad \begin{matrix} x_3 \text{ free} \\ x_4 \text{ free} \end{matrix}}$$

$$\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix} = \begin{bmatrix} -24 + 2x_3 - 3x_4 \\ -7 + 2x_3 - 2x_4 \\ x_3 \\ x_4 \\ 4 \end{bmatrix}$$

$$\vec{x} = \begin{bmatrix} -24 \\ -7 \\ 0 \\ 0 \\ 4 \end{bmatrix} + x_3 \begin{bmatrix} 2 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

x_3, x_4 free

<u>Parametric form</u> of the solution.

<u>Question</u>: When do solutions to lin. sys. exist?

<u>A</u>: <u>Theorem</u>: A lin. sys. is consistent if and only if the rightmost column of its aug. matrix is <u>not</u> a pivot column.

If it is consistent, then the solution set contains either (i) a unique solution (no free var's) or (ii) infinitely many solutions (free var(s))

"free" = not ~~from~~ a pivot column.

Vector Equations

Lay - 1.3   Strang - 1.1

In $\mathbb{R}^2$ (or $\mathbb{C}^2$):

Def: Vectors are 2×1 matrices (also called column vectors)

E.g. $\vec{u} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\vec{w} = \begin{pmatrix} \pi \\ i+1 \end{pmatrix}$, $\vec{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

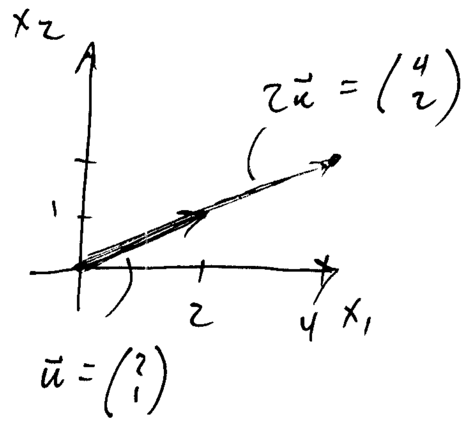
Two vectors are equal when corresponding entries are equal.

E.g. $\vec{u} + \vec{v} = \vec{x}$, but $\vec{u} - \vec{v} \neq \vec{w}$.

The sum of $\vec{u} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$ and $\vec{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ is $\vec{u} + \vec{v} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \end{pmatrix}$.

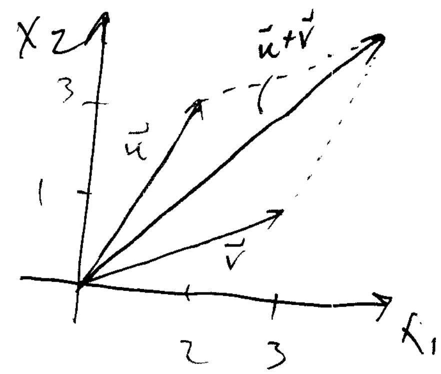
Given a number/scalar c, $c \cdot \vec{u} = \begin{pmatrix} c u_1 \\ c u_2 \end{pmatrix}$.

Geometrically:



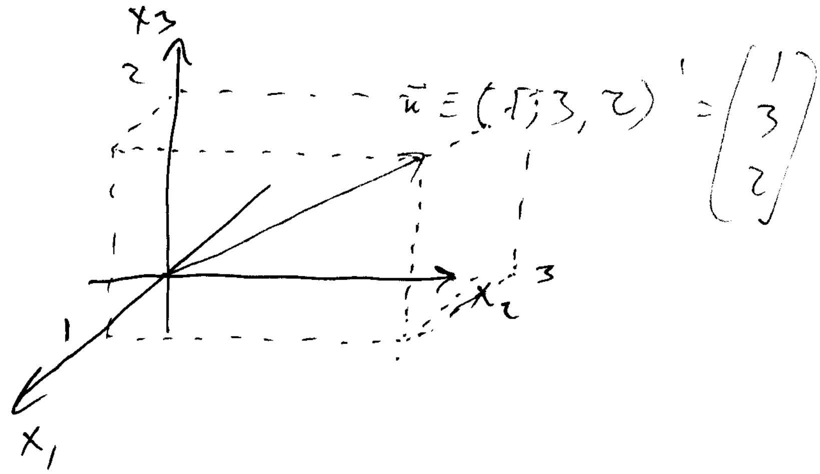
Parallelogram Law:

$\vec{u} + \vec{v}$ is the vector starting at $\vec{0}$ whose endpoint is at the fourth corner of the parallelogram defined by $\vec{0}$, $\vec{u}$, and $\vec{v}$.



In $\mathbb{R}^3$:

Not much different ...



More generally...

In $\mathbb{R}^n$:

Def: A vector is an $n \times 1$ matrix. The zero vector $\vec{0}$ is the vector whose entries are all zero.

Properties of $\mathbb{R}^n$:

For all $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$ and all scalars $c, d \in \mathbb{R}$:

$+$

1. $\vec{u} + \vec{v} = \vec{v} + \vec{u}$
2. $(\vec{u} + \vec{v}) + \vec{w} = \vec{u} + (\vec{v} + \vec{w})$
3. $\vec{u} + \vec{0} = \vec{0} + \vec{u} = \vec{u}$
4. $\vec{u} + (-\vec{u}) = \vec{0}$

$\cdot$

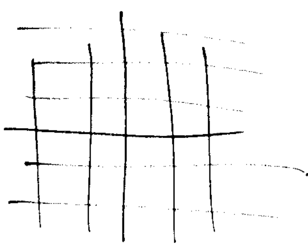
5. $c \cdot (\vec{u} + \vec{v}) = c\vec{u} + c\vec{v}$
6. $(c+d)\vec{u} = c\vec{u} + d\vec{u}$
7. $c(d\vec{u}) = (cd)\vec{u} = cd\vec{u}$
8. $1(\vec{u}) = \vec{u}$.

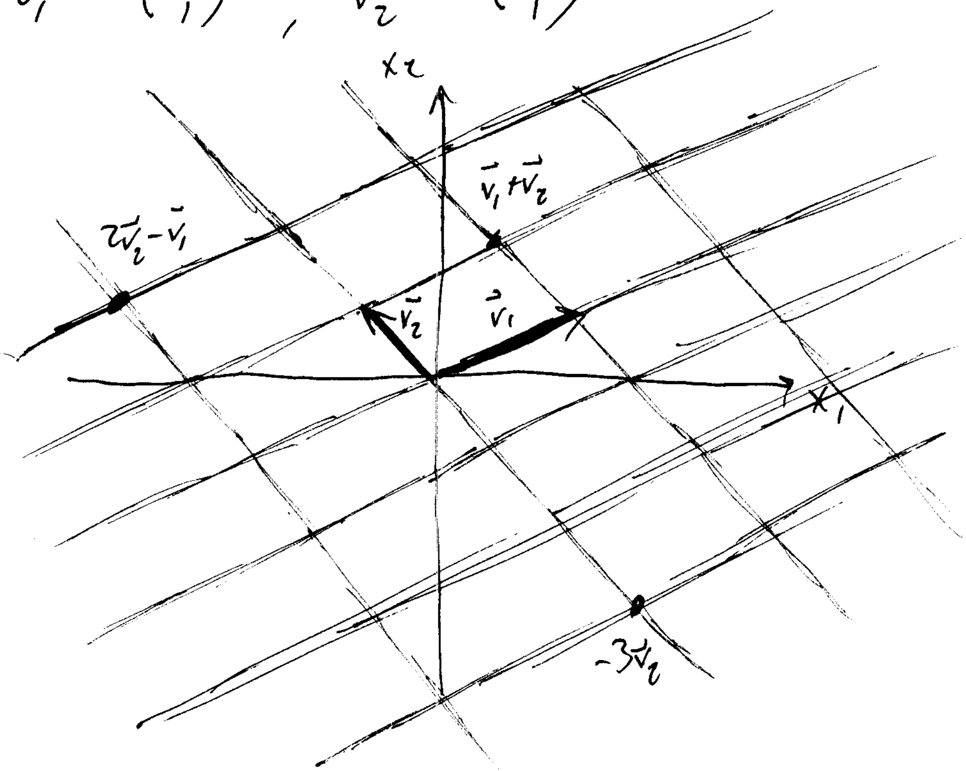
Makes $\mathbb{R}^n$ into a vector space.

Def: A *linear combination* of the vectors $\vec{v}_1, \dots, \vec{v}_k \in \mathbb{R}^n$ with *weights* $c_1, \dots, c_k \in \mathbb{R}$ is the vector

$$\vec{y} = c_1 \vec{v}_1 + c_2 \vec{v}_2 + \cdots + c_k \vec{v}_k = \sum_{i=1}^{k} c_i \vec{v}_i .$$

Ex: $\vec{v}_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix},$





$2\vec{v}_2 - \vec{v}_1$

$-3\vec{v}_2$

Remarks: Linear equations can be view in terms of lin. comb.'s

Geometric (row) picture	Algebraic (column) picture
$\Pi_1 : \sum_{i=1}^{n} a_{1i} x_i = b_1$	Vectors $\vec{v}_1, \dots, \vec{v}_n$ ...
planes $\Pi_2 : \sum_{i=1}^{n} a_{2i} x_i = b_2$	$c_1 \begin{pmatrix} v_{11} \\ \vdots \\ v_{m1} \end{pmatrix} + c_2 \begin{pmatrix} v_{12} \\ \vdots \\ v_{m2} \end{pmatrix} + \cdots + c_n \begin{pmatrix} v_{1n} \\ \vdots \\ v_{mn} \end{pmatrix}$
$\vdots \qquad \vdots$	
$\Pi_m : \sum_{i=1}^{n} a_{mi} x_i = b_m$	$\boxed{x_i\text{'s and } c_i\text{'s play same role!}}$ $\quad = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$

In matrix notation: $A = [\vec{a}_1 \cdots \vec{a}_n]$ w/ $\vec{a}_i \in \mathbb{R}^m$.

So lin. sys. for $x_1\vec{a}_1 + \cdots + x_n\vec{a}_n = \vec{b}$ is

$$[\vec{a}_1 \cdots \vec{a}_n \mid \vec{b}].$$

<u>Def</u>: If $\vec{v}_1, \ldots, \vec{v}_k \in \mathbb{R}^n$, then the set of all lin. comb.'s of them is $\text{span}\{\vec{v}_1, \ldots, \vec{v}_k\}$, the <u>span</u> of these vectors. If $\vec{b} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k$ for some weights $c_1, \ldots, c_k$, then $\vec{b} \in \text{span}\{\vec{v}_1, \ldots, \vec{v}_k\}$ and say $\vec{b}$ is <u>spanned</u> by $\{\vec{v}_1, \ldots, \vec{v}_k\}$.

<u>What is span visually?</u>

One vectors – $\text{span}\{\vec{v}_1\} = \begin{cases} \text{line} & (1) \\ \text{point} & (0) \end{cases}$

Two vectors – $\text{span}\{\vec{v}_1, \vec{v}_2\} = \begin{cases} \text{plane} & (2) \\ \text{line} & (1) \\ \text{point} & (0) \end{cases}$

Three vectors – $\text{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \begin{cases} \text{hyperplane (copy of } \mathbb{R}^3) & (3) \\ \text{plane} & (2) \\ \text{line} & (1) \\ \text{point} & (0) \end{cases}$

<u>Question</u>: When can I guarantee that the dimension of the span is equal to the number of vectors in $\{\vec{v}_1, \ldots, \vec{v}_k\}$.

To be answered later...

The Matrix Equation $A\vec{x} = \vec{b}$

Lay - 1.4   Strang - 1.3, 3.4

Def: If A is an $m \times n$ matrix ($A \in M_{m\times n}(\mathbb{R})$) w/ columns $\vec{a}_1, \dots, \vec{a}_n$ and if $\vec{x} \in \mathbb{R}^n$, then the <u>product</u> $A\vec{x}$ is the linear combination of the columns of A using the corresponding entries in $\vec{x}$ as weights. That is,

$$A\vec{x} = [\vec{a}_1 \cdots \vec{a}_n]\begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\vec{a}_1 + \cdots + x_n\vec{a}_n = \sum_{i=1}^{n} x_i\vec{a}_i .$$

Ex: ①
$$\begin{bmatrix} 1 & 2 & -1 \\ 0 & -5 & 3 \end{bmatrix}\begin{bmatrix} 4 \\ 3 \\ 7 \end{bmatrix} = 4\begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3\begin{pmatrix} 2 \\ -5 \end{pmatrix} + 7\begin{pmatrix} -1 \\ 3 \end{pmatrix}$$
$$= \begin{pmatrix} 4 \\ 0 \end{pmatrix} + \begin{pmatrix} 6 \\ -15 \end{pmatrix} + \begin{pmatrix} -7 \\ 21 \end{pmatrix}$$
$$= \begin{pmatrix} 3 \\ 6 \end{pmatrix}$$

②
$$\begin{bmatrix} 2 & -3 \\ 8 & 0 \\ -5 & 2 \end{bmatrix}\begin{bmatrix} 4 \\ 7 \end{bmatrix} = 4\begin{pmatrix} 2 \\ 8 \\ -5 \end{pmatrix} + 7\begin{pmatrix} -3 \\ 0 \\ 2 \end{pmatrix}$$
$$= \begin{pmatrix} 8 \\ 32 \\ -20 \end{pmatrix} + \begin{pmatrix} -21 \\ 0 \\ 14 \end{pmatrix} = \begin{pmatrix} -13 \\ 32 \\ -6 \end{pmatrix}$$

Def: Given a system of equations $\vec{a}_i \in \mathbb{R}^m$, $x_1\vec{a}_1 + \cdots + x_n\vec{a}_n = \vec{b}$, its <u>matrix equation</u> is given by $A\vec{x} = \vec{b}$, where

$$A = [\vec{a}_1 \cdots \vec{a}_n] \text{ and } \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \text{ and } \vec{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}.$$

The augmented matrix is $[\vec{a}_1 \cdots \vec{a}_n \mid \vec{b}]$.

Remark (key observation) : Following this definition,

$A\vec{x} = \vec{b}$ has a solution if and only if $\vec{b}$ is some linear combination of the columns of A.

Ex: "Is $\begin{bmatrix} 1 & 3 & 4 \\ -4 & 2 & -6 \\ -3 & -2 & -7 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$ consistent for all $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$? "

Solution - ROW REDUCE!

$$\left[\begin{array}{ccc|c} 1 & 3 & 4 & b_1 \\ -4 & 2 & -6 & b_2 \\ -3 & -2 & -7 & b_3 \end{array}\right] \underset{R3 \to R3 + 3R1}{\overset{R2 \to R2 + 4R1}{\sim}} \left[\begin{array}{ccc|c} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & b_2 + 4b_1 \\ 0 & 7 & 5 & b_3 + 3b_1 \end{array}\right]$$

$$\overset{R3 \to R3 - \frac{1}{2}R2}{\sim} \left[\begin{array}{ccc|c} 1 & 3 & 4 & b_1 \\ 0 & 14 & 10 & b_2 + 4b_1 \\ 0 & 0 & 0 & b_3 + 3b_1 - \frac{1}{2}(b_2 + 4b_1) \end{array}\right]$$

To be consistent, I need $\boxed{b_1 - \frac{1}{2}b_2 + b_3 = 0}$.

$\vec{b}$ must lie in this plane.

Span$\{\vec{a}_1, \vec{a}_2, \vec{a}_3\}$ only really need two of these vectors to by fully described. O.o

Question: Why can we not get to every $\vec{b} \in \mathbb{R}^3$?

A: REF(A) has a free column in it.
RREF(A) has a row of zeros!

Question: What condition do we need on A to guarantee that we can get to every $\vec{b} \in \mathbb{R}^3$?

A: All rows of $A \in M_{m\times n}(\mathbb{R})$ need a pivot!
In this case, the columns of A span $\mathbb{R}^m$.

Theorem: Let $A \in M_{m\times n}(\mathbb{R})$. Then the following are equivalent: (TFAE)

1. For each $\vec{b} \in \mathbb{R}^m$, $A\vec{x} = \vec{b}$ has a solution.
2. Each $\vec{b} \in \mathbb{R}^m$ is a lin. comb. of the columns of A.
3. The columns A span $\mathbb{R}^m$.
4. A has a pivot position in every row

not $[A|\vec{b}]$.

Alternate calculation $A\vec{x}$

row vectors

$$A = \begin{bmatrix} \vec{a}_1 \\ \vec{a}_2 \\ \vec{a}_3 \end{bmatrix}$$

Ex: By our def...

$$\begin{bmatrix} 2 & 3 & 4 \\ -1 & 5 & -3 \\ 6 & -2 & 8 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{pmatrix} 2x_1 + 3x_2 + 4x_3 \\ -x_1 + 5x_2 - 3x_3 \\ 6x_1 - 2x_2 + 8x_3 \end{pmatrix} = \begin{pmatrix} \vec{a}_1^{*} \cdot \vec{x} \\ \vec{a}_2^{*} \cdot \vec{x} \\ \vec{a}_3^{*} \cdot \vec{x} \end{pmatrix}.$$

Solution Sets of Linear Systems

Lay - 1.5 Strang - 3.4

Def: A lin. system $A\vec{x}=\vec{b}$ is <u>homogeneous</u> if $\vec{b}=\vec{0}$.

Remark: Hom. systems always have $\vec{x}=\vec{0}$ (<u>trivial</u>) as a solution. Are there other <u>nontrivial</u> solutions.

Question: When are there other nontrivial solutions to $A\vec{x}=\vec{0}$?

Answer: When there is at least 1 free variable.

Ex: "Find two vectors that span the plane given by
$x-2y+3z=0$."

$$[1\ -2\ 3\ |\ 0] \sim [1\ -2\ 3\ |\ 0]$$

$$\begin{cases} x = 2y-3z \\ y, z \text{ free} \end{cases}$$

$$\vec{x}=\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2y-3z \\ y \\ z \end{bmatrix} = \begin{bmatrix} 2y \\ y \\ 0 \end{bmatrix} + \begin{bmatrix} -3z \\ 0 \\ z \end{bmatrix}$$

$$= y\begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$$

Take these!

Sol's to Nonhom. Equations

We've seen parametric form : $\vec{x} = \vec{x}_p + \vec{x}_h$,
$\vec{x}_p$ solves $A\vec{x} = \vec{b}$ and $\vec{x}_h$ solves $A\vec{x} = \vec{0}$.

__Theorem:__ $A\vec{x} = \vec{b}$ has sol's $\vec{x} = \vec{x}_p + \vec{x}_h$ as

$$A\vec{x} = A(\vec{x}_p + \vec{x}_h) = A\vec{x}_p + A\vec{x}_h = \vec{b} + \vec{0}$$
$$= \vec{b}.$$

$$(a_1 + b_1)\vec{x}_1 + (a_2 + b_2)\vec{x}_2 + \cdots$$
$$= (a_1\vec{x}_1 + a_2\vec{x}_2 + \cdots) + (b_1\vec{x}_1 + b_2\vec{x}_2 + \cdots)$$

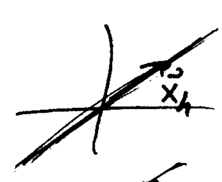
__Remark:__ Homogeneous sol's are closed under lin. comb. :

$$A(a\vec{x}_h + b\vec{y}_h) = A(a\vec{x}_h) + A(b\vec{y}_h)$$
$$= aA\vec{x}_h + bA\vec{y}_h$$
$$= \vec{0} + \vec{0} = \vec{0}.$$

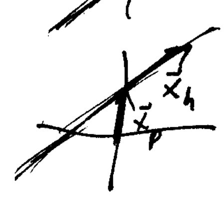
This doesn't work for nonhom. :

$$A(a\vec{x}_p + b\vec{y}_p) = A(a\vec{x}_p) + A(b\vec{y}_p)$$
$$= aA\vec{x}_p + bA\vec{y}_p$$
$$= a\vec{b} + b\vec{b} \neq \vec{b}.$$

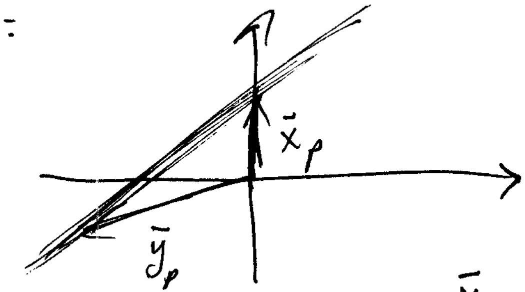
__Visually:__ solutions to $A\vec{x} = \vec{0}$:



Solutions to $A\vec{x} = \vec{b}$:



$\vec{x}_p$ is not unique:



$$\vec{x}_p - \vec{y}_p = \vec{v}_h$$

$$A(\vec{x}_p - \vec{y}_p) = A\vec{x}_p + (-1)A\vec{y}_p$$
$$= \vec{b} - \vec{b} = \vec{0}.$$

Conseq.: Diff. between to particular solutions only ~~diffe~~ a homogeneous solution.

Linear Independence

Lay - 1.7   Strang - 3.5

Def: An indexed set of vectors $\{\vec{v}_1, \dots, \vec{v}_p\} \subseteq \mathbb{R}^n$ is said to be linearly independent if the equation

$$x_1\vec{v}_1 + x_2\vec{v}_2 + \cdots + x_p\vec{v}_p = \vec{0}$$

has only the trivial solution. The set is linearly dependent if there are weights $c_1, \dots, c_p$, not all zero, such that

$$c_1 v_1 + \cdots + c_p\vec{v}_p = \vec{0}. \quad (*)$$

$(*)$ is called a lin. dep. relation.

$$c_p \vec{v}_p = -(c_1 \vec{v}_1 + \cdots + c_{p-1} \vec{v}_{p-1})$$

$$\vec{v}_p = \frac{-c_1}{c_p} \vec{v}_1 - \cdots - \frac{c_{p-1}}{c_p} \vec{v}_{p-1} .$$

Ex: Define $\vec{v}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$, $\vec{v}_3 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}$.

(a) Are these lin. ind.?

(b) If not, find a lin. dep. relation.

Sol.: Solve $c_1 \vec{v}_1 + c_2 \vec{v}_2 + c_3 \vec{v}_3 = \vec{0}$.

$$\swarrow$$

$$\left[\begin{array}{ccc|c} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 & \vec{0} \end{array}\right]$$

$$= \left[\begin{array}{ccc|c} 1 & 4 & 2 & 0 \\ 2 & 5 & 1 & 0 \\ 3 & 6 & 0 & 0 \end{array}\right]$$

$$\sim \left[\begin{array}{ccc|c} 1 & 4 & 2 & 0 \\ 0 & -3 & -3 & 0 \\ 0 & -6 & -6 & 0 \end{array}\right] \sim \left(\begin{array}{ccc|c} 1 & 4 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right)$$

$$\begin{cases} c_1 = -4c_2 - 2c_3 \\ c_2 = -c_3 \\ c_3 \text{ free} \end{cases} \Rightarrow \underline{\text{not}} \text{ lin. ind.}$$

$$\boxed{2\vec{v}_1 - \vec{v}_2 + \vec{v}_3 = \vec{0}}$$

$$\vec{c} = \begin{bmatrix} -4c_2 - 2c_3 \\ -c_3 \\ c_3 \end{bmatrix} = \begin{bmatrix} -4(-c_3) - 2c_3 \\ -c_3 \\ c_3 \end{bmatrix} = c_3 \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix}$$

Vectors often given ~~as~~ columns of matrix $A = [\vec{a}_1 \cdots \vec{a}_n]$.

Then $A\vec{x} = \vec{0} \longleftrightarrow x_1\vec{a}_1 + \cdots + x_n\vec{a}_n = \vec{0}$.

So columns of A are lin. ind.

if and only if

$A\vec{x} = \vec{0}$ has <u>only</u> the trivial solution.

<u>Ex:</u> Then

$$A = \begin{bmatrix} 1 & 4 & 2 \\ 2 & 5 & 1 \\ 3 & 6 & 0 \end{bmatrix}$$ has lin. dep. columns

$$A = \begin{bmatrix} 0 & 1 & 4 \\ 1 & 2 & -1 \\ 5 & 8 & 0 \end{bmatrix}$$ has lin. ind. columns

solve $[A \mid \vec{0}]$

<u>Special Case:</u> Two vectors: $\{\vec{v}_1, \vec{v}_2\}$.

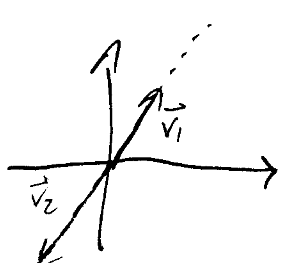
When are these lin. dep.?

$c_1\vec{v}_1 + c_2\vec{v}_2 = \vec{0}$

There exist c_1 and c_2 (not both zero) that solve this.

$c_1 \neq 0 \Rightarrow c_1\vec{v}_1 = -c_2\vec{v}_2$

$\vec{v}_1 = \left(\frac{-c_2}{c_1}\right)\vec{v}_2$.



Theorem: An indexed set $S = \{\vec{v}_1, \dots, \vec{v}_p\}$ of at least two vectors is lin. dep. iff at least one of the vectors is a lin. comb. of the others. In fact, if S is lin. dep., then some $\vec{v}_j$ $(j>1)$ is a lin. comb. of the preceding $\vec{v}_1, \dots, \vec{v}_{j-1}$.

E.g. From before, $2\vec{v}_1 - \vec{v}_2 + \vec{v}_3 = \vec{0}$ $\left(\left\{\begin{pmatrix}1\\2\\3\end{pmatrix}, \begin{pmatrix}4\\5\\6\end{pmatrix}, \begin{pmatrix}2\\1\\0\end{pmatrix}\right\}\right)$

$$\vec{v}_3 = -2\vec{v}_1 + \vec{v}_2$$

$$= -2\begin{pmatrix}1\\2\\3\end{pmatrix} + \begin{pmatrix}4\\5\\6\end{pmatrix} = \begin{pmatrix}2\\1\\0\end{pmatrix} \checkmark$$

Theorem: If a set contains more vectors than there entries in each vector, then the set is lin. dep. That is, any set $\{\vec{v}_1, \dots, \vec{v}_p\} \subseteq \mathbb{R}^n$ is lin. dep. if $p > n$.

Theorem: If a set $S = \{\vec{v}_1, \dots, \vec{v}_p\} \subseteq \mathbb{R}^n$ contains $\vec{0}$, then the set is lin. dep.

Ex: Using these theorems, determine if each set is lin. dep.

(a) $\left\{\begin{pmatrix}1\\7\\6\end{pmatrix}, \begin{pmatrix}2\\0\\9\end{pmatrix}, \begin{pmatrix}3\\1\\5\end{pmatrix}, \begin{pmatrix}4\\1\\6\end{pmatrix}\right\}$ lin. dep.

(b) $\left\{\begin{pmatrix}2\\3\\5\end{pmatrix}, \begin{pmatrix}0\\0\\0\end{pmatrix}, \begin{pmatrix}1\\1\\8\end{pmatrix}\right\}$ lin. dep.

(c) $\left\{\begin{pmatrix}3\\4\end{pmatrix}, \begin{pmatrix}6\\7\end{pmatrix}\right\}$ lin. ind.

Intro to Linear Transformations

Lay - 1.8 Strang - 7.2

Thought - In the vector $A\vec{x}$, A takes $\vec{x}$ to another vector $A\vec{x}$. So A "acts" on vectors (of the appropriate size). In fact, A is a kind of function - not actually a function.

$$T(\vec{x}) = A\vec{x} \quad \text{is the function.}$$

Now, "solve $A\vec{x} = \vec{b}$ for $\vec{x}$" means
"find all $\vec{x}$ that A sends to $\vec{b}$, if any."

Def: A transformation / function / map / mapping T from $\mathbb{R}^n$ to $\mathbb{R}^m$ (denoted $T: \mathbb{R}^n \to \mathbb{R}^m$) is a rule assigning to each vector $\vec{x} \in \mathbb{R}^n$ a vector $T(\vec{x}) \in \mathbb{R}^m$.
The set $\mathbb{R}^n$ is the domain of T (dom T).
The set $\mathbb{R}^m$ is the codomain of T (codom T).
The vector $T(\vec{x}) \in \mathbb{R}^m$ is called the image of $\vec{x}$.
The set of all images is called range of T (ran T).

Ex: $A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$, $\vec{u} = \begin{pmatrix} 2 \\ -1 \end{pmatrix}$, $\vec{b} = \begin{pmatrix} 3 \\ 2 \\ -5 \end{pmatrix}$, $\vec{c} = \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix}$

$$T(\vec{x}) = A\vec{x}.$$

$$\vec{x} \mapsto A\vec{x}$$

$$\vec{u} \mapsto A\vec{u} = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{pmatrix} 5 \\ 1 \\ -9 \end{pmatrix}$$

Solve $A\vec{x} = \vec{b}$ $\Rightarrow$ $\begin{pmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{pmatrix} 3 \\ 2 \\ -5 \end{pmatrix}$

$$\begin{bmatrix} x_1 - 3x_2 \\ 3x_1 + 5x_2 \\ -x_1 + 7x_2 \end{bmatrix} = \begin{pmatrix} 3 \\ 2 \\ -5 \end{pmatrix}$$

$$\vec{x} = \begin{pmatrix} 3/2 \\ -1/2 \end{pmatrix}$$

$A\vec{x} = \vec{b}$ shows $\begin{pmatrix} 3/2 \\ -1/2 \end{pmatrix} \mapsto \begin{pmatrix} 3 \\ 2 \\ -5 \end{pmatrix}$ is the <u>only</u> $\vec{x}$ that maps to $\vec{b}$.

Is there an $\vec{x}$ that maps to $\vec{c}$ via A?

$$\vec{c} = \begin{pmatrix} 3 \\ 2 \\ 5 \end{pmatrix}$$

$$\left[\begin{array}{cc|c} 1 & -3 & 3 \\ 3 & 5 & 2 \\ -1 & 7 & 5 \end{array}\right] \rightarrow \text{inconsistent.}$$

No! No $\vec{x}$ maps to $\vec{c}$.

$\vec{c}$ is <u>not</u> in the range of T.

More generally now!

<u>Def</u>: A transformation is <u>linear</u> provided:

① $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$ for all $\vec{u}, \vec{v} \in \operatorname{dom} T$.

② $T(c\vec{u}) = c\,T(\vec{u})$ for all $\vec{u} \in \operatorname{dom} T$ and all scalars c.

<u>Remark</u>: ① Not all ~~trans~~ transformation are lin. trans.

$f(x) = x^2$, $\ln x$, $\sin x$ <u>not lin.</u>

② Matrices are <u>not</u> transformations. (and vice-versa)

They represent transformations.

Wake-up Call!

Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a function. What does it mean for T to be linear?

---

① For all $\vec{u}, \vec{v} \in \mathbb{R}^n$, $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$.

② For all $\vec{u} \in \mathbb{R}^n$, $c \in \mathbb{R}$, $T(c\vec{u}) = cT(\vec{u})$.

Why is this important?

Let $\vec{v} = \sum_{i=1}^{n} c_i \vec{v}_i$ and T be linear.

What is $T(\vec{v})$?

$$= T\left(\sum_{i=1}^{n} c_i \vec{v}_i\right) = \sum_{i=1}^{n} \underbrace{T(c_i \vec{v}_i)}$$

$$= \sum_{i=1}^{n} c_i \cdot T(\vec{v}_i)$$

The Matrix of a Linear Transformation

Lay - 1.9 Strang - 7.2

Thought: Are matrix transformations linear? $\vec{x} \mapsto A\vec{x}$ linear?

$$A(\vec{x}+\vec{y}) = (x_1+y_1)\vec{a}_1 + \cdots + (x_n+y_n)\vec{a}_n$$

$$= (x_1\vec{a}_1 + \cdots + x_n\vec{a}_n) + (y_1\vec{a}_1 + \cdots + y_n\vec{a}_n)$$

$$= A\vec{x} + A\vec{y}.$$

Similarly, $A(c\vec{x}) = \cdots = c(A\vec{x})$

Yes, they are! Are all lin. trans. matrix trans.?

Def: The identity matrix $I_n = [\vec{e}_1 \cdots \vec{e}_n] = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}$.

The vectors $\vec{e}_1, \dots, \vec{e}_n$ are the standard coordinate vectors of $\mathbb{R}^n$.

Ex: ~~Eg~~ Suppose I have a lin. map $T: \mathbb{R}^2 \to \mathbb{R}^3$. Assume I know that $T(\vec{e}_1) = \begin{pmatrix} 5 \\ -7 \\ 2 \end{pmatrix}$ and $T(\vec{e}_2) = \begin{pmatrix} -3 \\ 8 \\ 0 \end{pmatrix}$.

What is $T(x,y)$? Think of (x,y) as $x\vec{e}_1 + y\vec{e}_2$.

So $T(x,y) = T(x\vec{e}_1 + y\vec{e}_2) = x\,T(\vec{e}_1) + y\,T(\vec{e}_2)$

$$= \begin{pmatrix} 5x - 3y \\ -7x + 8y \\ 2x + 0y \end{pmatrix} = \begin{bmatrix} 5 & -3 \\ -7 & 8 \\ 2 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$= [T(\vec{e}_1)\ T(\vec{e}_2)] \begin{pmatrix} x \\ y \end{pmatrix}.$$

Theorem: Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be a lin. trans. Then there exists a unique matrix $A \in M_{m \times n}(\mathbb{R})$ such that $T(\vec{x}) = A\vec{x}$ for all $\vec{x} \in \mathbb{R}^n$. In fact, A is the matrix whose j^{th} column is the vector $T(\vec{e}_j)$, where $\vec{e}_j$ is the j^{th} standard coordinate of $\mathbb{R}^n$.

So $A = [T(\vec{e}_1) \cdots T(\vec{e}_n)]$.

Proof: $T(\vec{x}) = \sum_{j=1}^{n} x_j T(\vec{e}_j) = [T(\vec{e}_1) \cdots T(\vec{e}_n)]\,\vec{x} = A\vec{x}$. ∎

Def: A above is the standard matrix for the lin. trans. T.

Question: How can we reframe our questions about lin. systems in terms of lin. trans.?

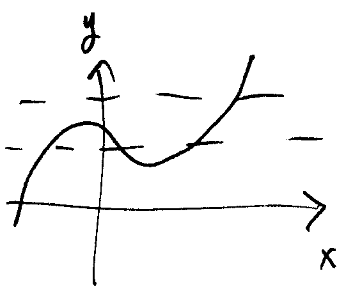
① Does the system have a solution? (consistent)

Reframe: $T(\vec{x}) = \vec{b} \Leftrightarrow A\vec{x} = \vec{b}$

Is there an $\vec{x}$ where ~~$T(\vec{x})$~~ $\vec{x}$ is "sent to $\vec{b}$"?

② If there is a solution, is it unique? (unique)

Reframe: Is T one-to-one?



T 1-1 means: ~~Does~~ For every ~~$\vec{b}$~~ $\vec{b}$ ~~$\vec{y}$~~ where $\vec{b} \in$ range T, $T(\vec{x}) = \vec{b} = T(\vec{y})$

$\Rightarrow \vec{x} = \vec{y}$.

Def: A mapping $T: \mathbb{R}^n \to \mathbb{R}^m$ is onto $\mathbb{R}^m$ (surjective) if each $\vec{b} \in \mathbb{R}^m$ is the image of at least one $\vec{x} \in \mathbb{R}^n$.

T is one-to-one if ~~for~~ each $\vec{b} \in \mathbb{R}^m$, ~~there is~~ is the image of at most one $\vec{x} \in \mathbb{R}^n$.

Ex: T given by $T(\vec{x}) = A\vec{x}$ where $A = \begin{bmatrix} 1 & -4 & 8 & 1 \\ 0 & 2 & -1 & 3 \\ 0 & 0 & 0 & 5 \end{bmatrix}$.

Every row has a pivot, so $[A \mid \vec{b}]$ is consistent.

Then T is onto $\mathbb{R}^3$.

Is T one-to-one? No ... why? $T(\vec{x}) = \vec{0}$ ...

Solve $\vec{b} = \vec{0}$ in $A\vec{x} = \vec{b}$.

I know that $\vec{0} \in \mathbb{R}^4 \mapsto \vec{0} \in \mathbb{R}^3$.

Also ... $\begin{pmatrix} 12 \\ 1 \\ 2 \\ 0 \end{pmatrix} \mapsto \vec{0}$. $\Rightarrow$ T is not one-to-one (has a free column in A).

Theorem: Let $T: \mathbb{R}^n \to \mathbb{R}^m$ be linear. T is 1-1 iff $T(\vec{x}) = \vec{0}$ has only the trivial solution.

Theorem: Let A be the std. mat. for a lin. trans. $T: \mathbb{R}^n \to \mathbb{R}^m$

(a) T is onto iff the columns of A span $\mathbb{R}^m$.

(b) T is 1-1 iff the columns of A ~~are~~ are lin. ind.

Chapter 2 - Matrix Algebra

Matrix Operations

Lay - 2.1 Strong - 2.4, 2.7, 1.3

Shorthand - $A = (a_{ij})_{m\times n} \in M_{m\times n}(\mathbb{F}) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$

Def: Two matrices are equal if they have have the same dimension and corresponding entries are equal.
For two matrices of the same size, their sum is the matrix whose entries are the sums of corresponding entries. That is, for $A = (a_{ij})_{m\times n}$ and $B = (b_{ij})_{m\times n}$,

define $A+B = (a_{ij} + b_{ij})_{m\times n}$.

If r is a scalar and $A = (a_{ij})_{m\times n}$, then we define the scalar multiple of A by r as $rA = (ra_{ij})_{m\times n}$.

Ex: $r=2$, $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, $B = \begin{bmatrix} -1 & 0 & 3 \\ 2 & -4 & 5 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 2 \\ -2 & 5 \end{bmatrix}$.

$A+B = \begin{bmatrix} 0 & 2 & 6 \\ 6 & 1 & 11 \end{bmatrix}$, $A+C$ is undefined

$rB = \begin{bmatrix} -2 & 0 & 6 \\ 4 & -8 & 10 \end{bmatrix}$, $rC = \begin{bmatrix} 2 & 4 \\ -4 & 10 \end{bmatrix}$

<u>Theorem</u>: Let $A, B, C \in M_{m \times n}(\mathbb{F})$ with $r, s \in \mathbb{F}$.

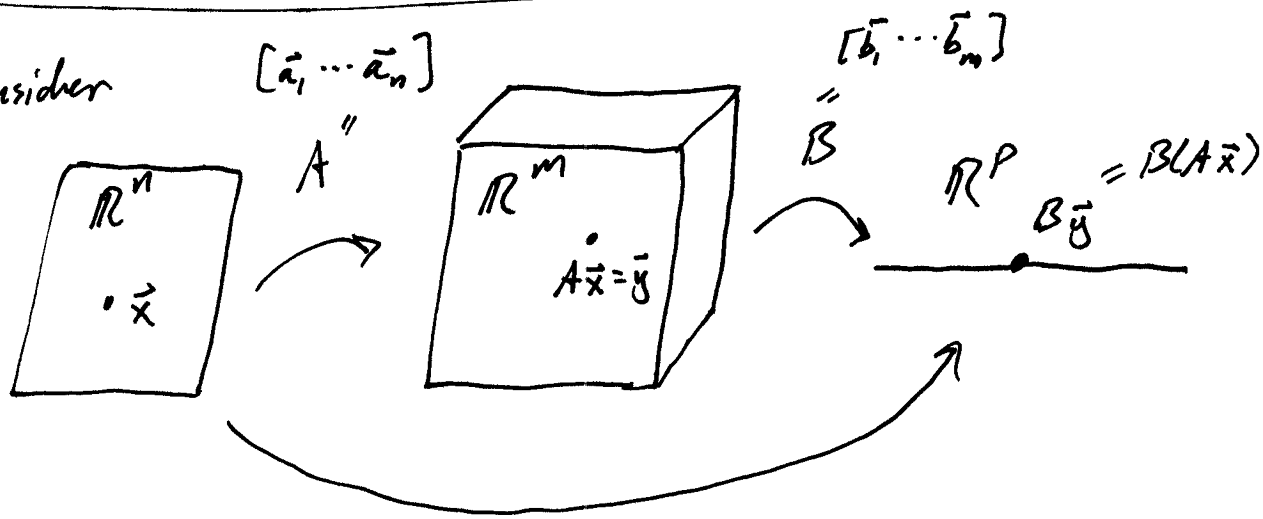
mat. $+$

$$\left\{\begin{array}{lll} (a) & A + B = B + A & \text{(commutativity)} \\ (b) & (A+B)+C = A + (B+C) & \text{(associativity)} \\ (c) & A + O_{m \times n} = A & \text{(existence of identity)} \end{array}\right.$$

scal. $\cdot$

$$\left\{\begin{array}{lll} (d) & r(A+B) = rA + rB & \text{(matrix distributivity)} \\ (e) & (r+s)A = rA + sA & \text{(scalar dist.)} \\ (f) & r(sA) = (rs)A & \end{array}\right.$$

Consider



? $B \cdot A$

So $\left.\begin{array}{l} A \in M_{m \times n}(\mathbb{F}) \\ B \in M_{p \times m}(\mathbb{F}) \end{array}\right\} BA \in M_{p \times n}(\mathbb{F})$

<u>Moral</u>: $(p \times m) \cdot (m \times n) = p \times n$

Why? Consider $A\vec{x} = x_1 \vec{a}_1 + \cdots + x_n \vec{a}_n$

$$\text{so } B(A\vec{x}) = B(x_1 \vec{a}_1 + \cdots + x_n \vec{a}_n)$$

$$= x_1 B\vec{a}_1 + \cdots + x_n B\vec{a}_n$$

$$= [B\vec{a}_1 \ \cdots \ B\vec{a}_n]\,\vec{x}.$$

$B(A\vec{x}) = (BA)\vec{x}$

Define: $BA = [B\vec{a}_1 \; \cdots \; B\vec{a}_n]$.

Ex: $A = \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix}$, $B = \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix}$.

$$AB = \left[A\begin{bmatrix} 4 \\ 1 \end{bmatrix} \; A\begin{bmatrix} 3 \\ -2 \end{bmatrix} \; A\begin{bmatrix} 6 \\ 3 \end{bmatrix} \right] = \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix}.$$

check: $(2\times 2)\cdot(2\times 3) = 2\times 3$

BA undefined?: ~~$(2\times 3)\cdot(2\times 2)$~~

$(2\times 3)\cdot(2\times 2)$

not equal $\Rightarrow$ undefined.

Observation: Each column of BA is a linear combination of the columns of B using weights from the corresponding column of A.

Alternatively - $(BA)_{ij} = b_{i1}a_{1j} + \cdots + b_{in}a_{nj} = \sum_{k=1}^{n} b_{ik}a_{kj}$.

$$B = \begin{bmatrix} \vec{b}_1 \\ \vdots \\ \vec{b}_m \end{bmatrix}, \quad A = [\vec{a}_1 \cdots \vec{a}_n]$$

$$(BA)_{ij} = (\vec{b}_i^{\,T} \cdot \vec{a}_j).$$

Theorem: Let $A \in M_{m\times n}(F)$ and B, C have "appropriate size": $(rA)B = A(rB)$ $(r \in F)$

(a) $A(BC) = (AB)C$

(b) $A(B+C) = AB + AC$

(c) $(A+B)C = AC + BC$

(d) $r(AB) = (rI_m)(AB)$ ~~$= (rA)B$~~

(e) $I_m A = A = A I_n$

Observe: $A = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix}$, $B = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}$.

$AB = \begin{bmatrix} 14 & 3 \\ -2 & -6 \end{bmatrix}$, $BA = \begin{bmatrix} 10 & 2 \\ 24 & -2 \end{bmatrix}$.

So $AB \neq BA$ in general

Also noteworthy: (i) $AB = AC \not\Rightarrow B = C$ in general.
(ii) $AB = 0 \not\Rightarrow A$ or $B = 0$.

Ex: (i) $A = 0$ matrix ...
(ii) $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \left[\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right]$
$= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0 \ldots$

More operation: $A^k = \underbrace{A \cdot A \cdot \cdots \cdot A}_{k \text{ times}}$ // $A^0 = I$.

Ex: Transpose of A: $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$, $A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$.

In general, $A = (a_{ij})_{m \times n}$, $A^T = (a_{ji})_{n \times m}$

Theorem: A, B of "appropriate sizes":
(a) $(A^T)^T = A$
(b) $(A+B)^T = A^T + B^T$
(c) $(rA)^T = r \cdot A^T$
(d) $(AB)^T = B^T A^T$ ← this reverses the order of the product.

<u>Warm-up</u>: Give examples $A, B \in M_{2\times 2}(\mathbb{R})$ where $AB \neq BA$.

" " $C, D, E \in M_{2\times 2}(\mathbb{R})$ where $(C \neq 0)$

$A \neq \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} = B$

$CD = CE$ but $D \neq E$.

$\begin{cases} A = \begin{bmatrix} 5 & 1 \\ 3 & -2 \end{bmatrix} \\ B = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix} \end{cases}$ Check: $AB = \begin{bmatrix} 14 & 3 \\ -2 & 6 \end{bmatrix}$, $BA = \begin{bmatrix} 10 & 2 \\ 24 & -2 \end{bmatrix}$.

$\begin{cases} C = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ D = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \\ E = \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} \end{cases}$ $\neq$ Check: $CD = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0.$

$CE = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0.$

<u>The Inverse of a Matrix</u> $rr^{-1} = r^{-1}r$ $\left(\begin{array}{l} ra = rb \\ \Rightarrow a = b. \end{array} \right)$ and $ar = br \Rightarrow a = b$

Lay - 2.2 Strang - 2.5, 2.3, (1.3)

$2^{-1} = 1/2$, $20^{-1} = 1/20$, 0^{-1} = Undefined.

What ~~dom~~ does it mean for A to have an inverse?

<u>Def</u>: An $n\times n$ matrix A is <u>invertible</u> if there is an $n\times n$ matrix C where $AC = CA = I_n = [\vec{e}_1 \cdots \vec{e}_n] = \text{eye}(n)$
$= \text{diag}(\underbrace{1, \dots, 1}_{n \text{ times}})$

<u>Ex</u>: Let $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$, $C = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}$.

$AC = \left[\begin{pmatrix} 2 & 5 \\ -3 & -7 \end{pmatrix}\begin{pmatrix} -7 \\ 3 \end{pmatrix}, \begin{pmatrix} 2 & 5 \\ -3 & -7 \end{pmatrix}\begin{pmatrix} -5 \\ 2 \end{pmatrix} \right] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$

$CA = \left[\begin{pmatrix} -7 & -5 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 \\ -3 \end{pmatrix}, \begin{pmatrix} -7 & -5 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 5 \\ -7 \end{pmatrix} \right] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$ ✓

<u>Def</u>: A matrix C such that $AC = CA = I_n$ is called <u>the</u> <u>inverse</u> of A (denoted $C = A^{-1}$).

<u>Theorem</u>: If A^{-1} exists, it is unique.

<u>Pf</u>: Suppose $BA = AB = I_n$ and $CA = AC = I_n$.

Consider
$$\begin{aligned} B = BI_n &= B(AC) \\ &= (BA)C \\ &= (I_n)C = C. \end{aligned}$$
$\blacksquare$

<u>Theorem</u>: Let $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in M_{2\times 2}(\mathbb{R})$. If $ad - bc \neq 0$, then A is invertible. Moreover,
$$A^{-1} = \frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

If $ad-bc = 0$, A^{-1} doesn't exist (A is not invertible).

<u>Ex</u>: $A = \begin{bmatrix} 2 & 5 \\ -3 & -7 \end{bmatrix}$, $ad-bc = -14 - (-15) = 1$.

$$A^{-1} = \frac{1}{1}\begin{bmatrix} -7 & -5 \\ -(-3) & 2 \end{bmatrix} = \begin{bmatrix} -7 & -5 \\ 3 & 2 \end{bmatrix}. \ (= C).$$

$B = \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}$, $ad-bc = 18 - 20 = -2$.

$$B^{-1} = \frac{1}{-2}\begin{bmatrix} 6 & -4 \\ -5 & 3 \end{bmatrix} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix}.$$

Theorem: If A is invertible, then $A\vec{x} = \vec{b}$ has a unique solution for every $\vec{b}$. Moreover, $\vec{x} = A^{-1}\vec{b}$.

Pf: Start w/ $A\vec{x} = \vec{b}$.

$$A^{-1}(A\vec{x}) = A^{-1}(\vec{b})$$
$$(A^{-1}A)\vec{x} = A^{-1}\vec{b}$$
$$I_n \vec{x} = A^{-1}\vec{b}$$
$$\vec{x} = A^{-1}\vec{b}. \blacksquare$$

Ex: "Solve the system $\begin{cases} 3x + 4y = 3 \\ 5x + 6y = 7 \end{cases}$."

Equiv. to $\left[\begin{array}{cc|c} 3 & 4 & 3 \\ 5 & 6 & 7 \end{array}\right] \leftrightarrow \begin{bmatrix} 3 & 4 \\ 5 & 6 \end{bmatrix}\vec{x} = \begin{bmatrix} 3 \\ 7 \end{bmatrix}$.

$ad - bc = -2 \Rightarrow A^{-1}$ exists.

$$A^{-1} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix}.$$

So $\vec{x} = A^{-1}\vec{b} = \begin{bmatrix} -3 & 2 \\ 5/2 & -3/2 \end{bmatrix}\begin{bmatrix} 3 \\ 7 \end{bmatrix} = \begin{pmatrix} 5 \\ -3 \end{pmatrix}$.

Theorem: (a) If A^{-1} exists, $(A^{-1})^{-1} = A$.

(b) A, B are invertible $\Rightarrow$ AB and BA are inv.

$$(AB)^{-1} = B^{-1}A^{-1}$$
$$(BA)^{-1} = A^{-1}B^{-1}.$$

$(AB)\cdot C$
$= C\cdot(AB)$
$= I_n$

(c) A inv. $\Rightarrow$ A^T invertible and $(A^T)^{-1} = (A^{-1})^T = A^{-T}$.

Elementary Row Operations as Matrix Actions

Since lin. trans. $\leftrightarrow$ matr. trans., let's characterize the E.R.O.'s matrices since they underlie all of our row-reduction procedures.

Scalar mult. by $r \neq 0$ in i^{th} row:

$$E_{S_i}(r) = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & r & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}.$$

Interchange rows i and j $(i<j)$:

$$E_I = \begin{array}{c} \left[\vec{e}_1 \cdots \underset{\uparrow \, i^{th} \text{ col.}}{\vec{e}_j} \cdots \underset{\uparrow \, j^{th} \text{ column}}{\vec{e}_i} \cdots \vec{e}_n\right] \end{array}.$$

Add $r \times (\text{row } j)$ to row i:

$$E_{SS} = I_n + E_S$$

$$= \begin{array}{cc} & \\ i^{th} \text{ row} \rightarrow & \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \ddots & & \\ \cdots 0 & r & 0 & \cdots & 1 & \\ & & & & \ddots & \\ & & & & & 1 \end{bmatrix} \\ & \quad\quad \underset{j^{th} \text{ col.}}{\uparrow} \end{array}$$

What are their inverses?

$$(E_{S_i}(r))^{-1} = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1/r & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix}$$

$$(E_{I_{i,j}})^{-1} = E_{I_{i,j}} \quad \text{(its own inverse)}$$

$$(E_{SS_{i,j}}(r))^{-1} = (I_n + E_{S_j}(r)\vec{e}_i)^{-1} = I_n - E_{S_j}(r)\vec{e}_i.$$

Ex: $A = \begin{bmatrix} 2 & 5 & -1 \\ 0 & 4 & 2 \\ 1 & 3 & 6 \end{bmatrix}.$

$R1 \leftrightarrow R3 \sim \begin{bmatrix} 1 & 3 & 6 \\ 0 & 4 & 2 \\ 2 & 5 & -1 \end{bmatrix}$ $\quad E_1 \left(E_{I_{1,3}} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \right)$

$R3 \to R3 - 2R1 \sim \begin{bmatrix} 1 & 3 & 6 \\ 0 & 4 & 2 \\ 0 & -1 & -13 \end{bmatrix}$ $\quad E_2 \left(E_{SS_{1,3}}(-2) \right) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix}$

$R2 \leftrightarrow R3 \sim \begin{bmatrix} 1 & 3 & 6 \\ 0 & -1 & -13 \\ 0 & 4 & 2 \end{bmatrix}$ $\quad E_3 \left(E_{I_{2,3}} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \right).$

$R2 \to -R2 \sim \begin{bmatrix} 1 & 3 & 6 \\ 0 & 1 & 13 \\ 0 & 4 & 2 \end{bmatrix}$ $\quad E_4 \left(E_{S_2}(-1) = \begin{bmatrix} 1 & & \\ & -1 & \\ & & 1 \end{bmatrix} \right).$

$R3 \to R3 - 4R2 \sim \begin{bmatrix} 1 & 3 & 6 \\ 0 & 1 & 13 \\ 0 & 0 & -50 \end{bmatrix} = B$ $\quad E_5 \left(E_{SS_{2,3}}(-4) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{bmatrix} \right).$

Then ~~A~~ $E_5 E_4 E_3 E_2 E_1 A = B$.

<u>Theorem:</u> An $n \times n$ matrix A is invertible iff $A \sim I_n$.

In this case, any sequence of row operations that reduces A to I_n also reduces I_n to A^{-1}.

IDEA: $\overbrace{(E_k \cdots E_1)}^{C} A = I_n$

$$A \sim I_n$$

$$\Downarrow$$

$$E_k \cdots E_1 A \sim E_k \cdots E_1 I_n$$

$$I_n \sim (A^{-1}) I_n = A^{-1}.$$

<u>Algorithm to find A^{-1}:</u>

Row reduce $[A \mid I_n] \sim [I_n \mid A^{-1}]$.

<u>Verify 2×2 formula:</u> $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ w/ $\underbrace{ad-bc}_{\Delta} \neq 0$.

Then assume $a \neq 0$ and $c \neq 0$.

So $\left[\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right] \overset{R1 \to \frac{1}{a}R1}{\sim} \left[\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ c & d & 0 & 1 \end{array}\right]$

$\overset{R2 \to R2 - cR1}{\sim} \left[\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & d-\frac{bc}{a} & -c/a & 1 \end{array}\right] = \left[\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & \frac{ad-bc}{a} & -c/a & 1 \end{array}\right]$

$\overset{R2 \to \frac{a}{ad-bc}R2}{\sim} \left[\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right] \overset{R1 \to R1 - \frac{b}{a}R2}{\sim} \left[\begin{array}{cc|cc} 1 & 0 & \frac{1}{a}+\frac{bc}{a(\Delta)} & \frac{-b}{\Delta} \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right]$

$$\frac{1}{a} + \frac{bc}{a\Delta} = \frac{\Delta}{a\Delta} + \frac{bc}{a\Delta} = \frac{ad-bc+bc}{a\Delta} = \frac{ad}{a\Delta}$$
$$= \frac{d}{\Delta}.$$

Finally,
$$\sim \left[\begin{array}{cc|cc} 1 & 0 & \frac{d}{ad-bc} & \frac{-b}{ad-bc} \\ 0 & 1 & \frac{-c}{ad-bc} & \frac{a}{ad-bc} \end{array}\right]$$

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

For HW ... ($a \neq 0$ and $c = 0$)
and ($a = 0$ and $c \neq 0$).

Characterization of Invertible Matrices

Lay - 2.3 Strang - 2.5

the following are equiv.

Theorem: Let $A \in M_{n\times n}(\mathbb{R})$. Then TFAE (all T or all F).

(a) A^{-1} exists.

(b) $A \sim I_n$.

(c) A ~~[illegible]~~ has n pivots

(d) $A\vec{x} = \vec{0}$ has only trivial sol.

(e) Columns of A are lin. ind.

(f) $\vec{x} \mapsto A\vec{x}$ is one-to-one.

(g) $A\vec{x} = \vec{b}$ has at least one solution for each $\vec{b} \in \mathbb{R}^n$ ($\vec{x} \mapsto A\vec{x}$ is onto)

(h) Columns of A span $\mathbb{R}^n$.

(i) $\vec{x} \mapsto A\vec{x}$ is onto.

(j) There is a left inverse - $CA = I_n$.

(k) There is a right inverse - $AD = I_n$.

(l) A^T is invertible.

Remark: These are all ways to detect if A^{-1} exists.
Any one of them being F $\Rightarrow$ A^{-1} DNE.

FACT: If $AB = I_n$ and $A, B \in M_{n\times n}(\mathbb{R})$,
then A^{-1}, B^{-1} both exist. $A^{-1} = B$, $B^{-1} = A$.

Ex: "Is $A = \begin{bmatrix} 1 & 0 & -2 \\ 3 & 1 & -2 \\ -5 & -1 & 9 \end{bmatrix}$ invertible?"

$$A \sim \begin{bmatrix} 1 & 0 & -2 \\ & 1 & 4 \\ & & 3 \end{bmatrix} \Rightarrow \text{invertible (by (c))}.$$

Recall: Linear transformations vs. Matrix Transformations

$\curvearrowright$ represented by.

Def: A lin. trans. $T: \mathbb{R}^n \to \mathbb{R}^n$ is invertible if there exists a function $S: \mathbb{R}^n \to \mathbb{R}^n$ such that

$$T(S(\vec{x})) = S(T(\vec{x})) = \vec{x}.$$

S is then called the inverse of T ($S = T^{-1}$).

Questions How many inverses can I have for a lin. trans.?

A: 1! Because... $T \leftrightarrow A$ invertible $\Rightarrow$ A^{-1} exists
$A^{-1} \leftrightarrow S$.

<u>Theorem</u>: Let $T: \mathbb{R}^n \to \mathbb{R}^n$ be linear and let A be the std. matrix for T. Then T is invertible iff A^{-1} exists. In that case, $S(\vec{x}) = A^{-1}\vec{x}$ is the <u>unique</u> function in the definition above.

<u>Ex</u>: "What can be said about a 1-1 lin. trans. $T: \mathbb{R}^n \to \mathbb{R}^n$?"

1-1 $\Rightarrow$ columns of A in $T(\vec{x}) = A\vec{x}$ are lin. ind.

$\Rightarrow A^{-1}$ exists $\Rightarrow A\vec{x} = \vec{b}$ has $\vec{x} = A^{-1}\vec{b}$ always (for each $\vec{b} \in \mathbb{R}^n$).

$\Rightarrow T(x) = A\vec{x}$ is onto.

function is 1-1 and onto $\Leftrightarrow$ invertible.

lin. trans. 1-1 and onto $\Leftrightarrow$ isomorphism.

Note: Prob 3 on HW 2 will come back in a few lectures. If you don't understand it, come to office hours (or make an appointment) to talk about it.

Matrix Factorizations

Lay - 2.5 Strang - 2.6

Motivation: Given some diff.'l eq., use $A\vec{x} = \vec{b}$ to estimate a solution at a single point. To estimate the entire solution, we have solve $A\vec{x} = \vec{b}$ for many different $\vec{b}$ but with the same A.

Problem... $A\vec{x} = \vec{b}$ takes $O(n^3)$ (or $O(n^4)$) flops each time. (by row reduction)
This means a long time to output.

Solution: Break up $A\vec{x} = \vec{b}$ into smaller, more efficient computations.

Questions: Any suggestions?

~~Ans~~ Answer: Foward / Backward sub are easier: Take ~~on~~ $O(n^2)$ instead.

Question: Can we turn $A\vec{x}=\vec{b}$ into smaller parts using upper/lower triangular systems?

Answer: Yes! Yes, we can.

Ex: Notice that

$$A = \begin{bmatrix} 3 & -7 & -2 & 2 \\ -3 & 5 & 1 & 0 \\ 6 & -4 & 0 & -5 \\ -9 & 5 & -5 & 12 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 2 & -5 & 1 & 0 \\ -3 & 8 & 3 & 1 \end{bmatrix}}_{L} \underbrace{\begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix}}_{U}$$

Consider $\vec{b} = \begin{bmatrix} -4 \\ 3 \\ -3 \\ 3 \end{bmatrix}$ in $A\vec{x}=\vec{b}$.

$$A\vec{x} = b$$
$$(LU)\vec{x} = \vec{b}$$
$$L(U\vec{x}) = \vec{b}$$
$$L\vec{y} = \vec{b}$$

Instead of solving $A\vec{x}=b$,
I will solve:
① $L\vec{y}=\vec{b}$
then ② $U\vec{x}=\vec{y}$.

So ... $L\vec{y} = \begin{bmatrix} -4 \\ 3 \\ -3 \\ 3 \end{bmatrix}$

That is,

$$\begin{bmatrix} 1 & & & \\ -1 & 1 & & \\ 2 & -5 & 1 & \\ -3 & 8 & 3 & 1 \end{bmatrix} \bar{y} = \begin{bmatrix} -4 \\ 3 \\ -3 \\ 3 \end{bmatrix}$$

$$y_1 = -4$$

$$-y_1 + y_2 = 3$$

$$y_2 = -1$$

$$2y_1 - 5y_2 + y_3 = -3$$

$$-8 + 5 + y_3 = -3$$

$$y_3 = 0$$

$$-3y_1 + 8y_2 + 3y_3 + y_4 = 3$$

$$12 - 8 + 0 + y_4 = 3$$

$$y_4 = -1$$

So $\vec{y} = \begin{bmatrix} -4 \\ -1 \\ 0 \\ -1 \end{bmatrix}$.

Now solve $U\vec{x} = \vec{y}$:

$$\begin{bmatrix} 3 & -7 & -2 & 2 \\ 0 & -2 & -1 & 2 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{bmatrix} \vec{x} = \begin{bmatrix} -4 \\ -1 \\ 0 \\ -1 \end{bmatrix}$$

$$-x_4 = -1$$

$$-x_3 + x_4 = 0$$

$$x_3 = 1$$

$$\cdots \quad x_2 = 1$$

$$x_1 = 1.$$

$$\vec{x} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

Cost: Factoring $A = LU$ ~~is~~ is a row-reduction algorithm, so it is $O(n^3)$. But we only need to do this once! Then after factoring A, $A\vec{x} = \vec{b}$ takes only $O(n^2)$ thereafter.

Question: How do we get $A = LU$? (LU-factorization of A)

Method by example:

Notice, $REF(A)$ above is $\underline{\underline{U}}$.

Recall elem. row. op. matrices:

$$E_S = diag(1, 1, \dots, r, \dots, 1) \quad (r \neq 0).$$

inv. is diag.

$$E_I = [\vec{e}_1 \cdots \vec{e}_j \cdots \vec{e}_i \cdots \vec{e}_n]$$

i^{th} $\quad$ j^{th} $\quad$ $(i<j)$

inv. is itself

$$E_{SS} = I_n + r\tilde{e}_j$$

Assume j is subdiagonal. Then E_{SS} always lower tri. and has 1's on its diagonal.

Inverse is also low. tri.

Then $A \sim U \iff \underbrace{E_k E_{k-1} \cdots E_1}_{L^{-1}} A = U.$

$$A = LU$$

<u>Ex:</u> $A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ -6 & 0 & 7 & -3 & 1 \end{bmatrix}$

$(R1 \to R1)$
$R2 \to R2 + 2R1$
$R3 \to R3 - R1$
$R4 \to R4 + 3R1$

$$\sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & -9 & -3 & -4 & 10 \\ 0 & 12 & 4 & 12 & -5 \end{bmatrix}$$

$$\parallel$$

$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \end{bmatrix} \to E_1 A$

$\begin{pmatrix} R1 \to R1 \\ R2 \to R2 \end{pmatrix}$
$R3 \to R3 + 3R2$
$R4 \to R4 - 4R2$

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

$$\parallel$$

$E_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 3 & 1 & 0 \\ 0 & -4 & 0 & 1 \end{bmatrix} \to E_2(E_1 A)$

$$\begin{pmatrix} R1 \to R1 \\ R2 \to R2 \\ R3 \to R3 \end{pmatrix} \sim \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ 0 & 3 & 1 & 2 & -3 \\ 0 & 0 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}$$

$$R4 \to R4 - 2R3$$

$$= U$$

$$E_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{bmatrix} \to \underbrace{E_3(E_2 E_1 A)}_{=U} = \underbrace{(E_3 E_2 E_1)}_{=L^{-1}} A$$

$$\text{So} \quad A = (E_3 E_2 E_1)^{-1} U$$

$$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 3 & 0 & 0 & 1 \end{bmatrix} \to E_1^{-1} = \begin{bmatrix} 1 & & & \\ -2 & 1 & & \\ +1 & 0 & 1 & \\ -3 & 0 & 0 & 1 \end{bmatrix}$$

$$E_2^{-1} = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & -3 & 1 & \\ 0 & 4 & 0 & 1 \end{bmatrix}$$

$$E_3^{-1} = \begin{bmatrix} 1 & & & \\ 0 & 1 & & \\ 0 & 0 & 1 & \\ 0 & 0 & 2 & 1 \end{bmatrix}$$

$$(E_3 E_2 E_1)^{-1} = E_1^{-1} E_2^{-1} E_3^{-1} = \begin{bmatrix} 1 & & & \\ -2 & 1 & & \\ 1 & -3 & 1 & \\ -3 & 4 & 2 & 1 \end{bmatrix} = L$$

So, $A = \begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ -4 & -5 & 3 & -8 & 1 \\ 2 & -5 & -4 & 1 & 8 \\ 6 & 0 & 7 & -3 & 1 \end{bmatrix}$

$$= \underbrace{\begin{bmatrix} 1 & & & \\ -2 & 1 & & \\ 1 & -3 & 1 & \\ -3 & 4 & 2 & 1 \end{bmatrix}}_{"L} \underbrace{\begin{bmatrix} 2 & 4 & -1 & 5 & -2 \\ & 3 & 1 & 2 & -3 \\ & & & 2 & 1 \\ & & & & 5 \end{bmatrix}}_{"U}$$

Ex: "Solve $A\vec{x} = \begin{bmatrix} 8 \\ -13 \\ 2 \\ 1 \end{bmatrix}$ using the LU-fact. from above."

$LU\vec{x} = \vec{b}$

$\rightarrow$ (1) $L\vec{y} = \vec{b}$

(2) $U\vec{x} = \vec{y}$.

$L\vec{y} = \vec{b}$:

$$\left[\begin{array}{cccc|c} 1 & & & & 8 \\ -2 & 1 & & & -13 \\ 1 & -3 & 1 & & 2 \\ 3 & 4 & 2 & 1 & 1 \end{array}\right] \xrightarrow[\text{Sub.}]{\text{Fwd.}} \vec{y} = \begin{bmatrix} 8 \\ 3 \\ 3 \\ 5 \end{bmatrix}$$

$U\vec{x} = \vec{y}$:

$$\left[\begin{array}{ccccc|c} 2 & 4 & -1 & 5 & -2 & 8 \\ & 3 & 1 & 2 & -3 & 3 \\ & & & 2 & 1 & 3 \\ & & & & 5 & 5 \end{array}\right] \rightarrow \vec{x} = \begin{bmatrix} \frac{1}{2}(5-4x_2+x_3) \\ \frac{1}{3}(4-x_3) \\ x_3 \\ 1 \\ 1 \end{bmatrix}$$

$$\vec{x} = \begin{pmatrix} -1/6 \\ 4/3 \\ 0 \\ 1 \\ 1 \end{pmatrix} + x_3 \begin{pmatrix} 7/6 \\ -1/3 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad x_3 \text{ free}$$

NEW CHAPTER! NEW UNIT!

Vector Spaces and Subspaces

Lay - 4.1 Strang - 3.1

Thought: We've had such success with $\mathbb{R}^n$. Are there any other sets that behave like it so that we can apply our techniques and results from $\mathbb{R}^n$ to these new sets as well?

Def: A vector space V (over a field $\mathbb{F}$ of scalars) is a nonempty set of objects (called vectors) on which two operations are defined subject to the following axioms:

vector $+$

1. $\vec{u}+\vec{v} \in V$ for all ~~every~~ $\vec{u},\vec{v} \in V$. (closed under add.)
2. $\vec{u}+\vec{v} = \vec{v}+\vec{u}$
3. $(\vec{u}+\vec{v})+\vec{w} = \vec{u}+(\vec{v}+\vec{w})$
4. There is a $\vec{0} \in V$ where $\vec{u}+\vec{0} = \vec{u}$.
5. For each $\vec{u} \in V$, there is a $-\vec{u} \in V$ where $\vec{u}+(-\vec{u}) = \vec{0}$.

scalar $\cdot$

6. $c\vec{u} \in V$ for all $\vec{u} \in V$ and $c \in \mathbb{F}$.
7. $c(\vec{u}+\vec{v}) = c\vec{u}+c\vec{v}$
8. $(c+d)\vec{u} = c\vec{u}+d\vec{u}$
9. $c(d\vec{u}) = (cd)\vec{u}$
10. $1\vec{u} = \vec{u}$.

Scratchwork: Prove to yourselves using (1) - (10) only that

(i) $0 \cdot \vec{u} = \vec{0}$

(ii) $c \cdot \vec{0} = \vec{0}$

(iii) $-\vec{u} = (-1)\vec{u}$.

Ex: As they are the archetypical spaces, $\mathbb{R}^n$ $(n \geq 1)$ are all vector spaces over $\mathbb{R}$ (also over $\mathbb{Q}$).

Ex: Let $C(\mathbb{R})$ be the set of all real-valued functions on $\mathbb{R}$ ($f : \mathbb{R} \to \mathbb{R}$ continuous). Define two op's as

$$(\vec{f} + \vec{g})(t) \overset{\text{def.}}{:=} \vec{f}(t) + \vec{g}(t)$$

$$(c\vec{f})(t) := c \cdot \vec{f}(t).$$

This is a vector space over $\mathbb{R}$ (and $\mathbb{Q}$).

Ex: Let $\mathbb{P}_n^f$ be the set of all (formal) polynomials of degree n (or less) with $\mathbb{R}$-coefficients:

$$\vec{p}(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n.$$
$$\vec{q}(t) = b_0 + b_1 t + b_2 t^2 + \cdots + b_n t^n.$$

~~The~~ This is a v.s. over $\mathbb{F}$.

Check: $(\vec{p} + \vec{q})(t) := \vec{p}(t) + \vec{q}(t)$

$$= (a_0 + b_0) + (a_1 + b_1)t + \cdots + (a_n + b_n)t^n$$

$(\ldots, 1, 2, 3, \ldots)$
$+ (\ldots, -4, 6, 2, \ldots)$

$$(c\vec{p})(t) := (c a_0) + (c a_1)t + \cdots + (c a_n)t^n.$$

Ex: Let $\$(\mathbb{R})$ be the set of bi-infinite sequences w/ $\mathbb{R}$ entries.

$$\vec{x} = \{x_k\}_{k=-\infty}^{\infty} = (\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots)$$

Define $(\vec{x} + \vec{y}) = \underbrace{\{x_k\}_{k=-\infty}^{\infty} + \{y_k\}_{k=-\infty}^{\infty}}_{\vec{x} + \vec{y}} := \{x_k + y_k\}_{k=-\infty}^{\infty}$

$$(c\vec{x}) = c\{\vec{x}_k\}_{k=-\infty}^{\infty} := \{c x_k\}_{k=-\infty}^{\infty} = c \cdot \vec{x}.$$

Def: A subspace H of a vector space V is a subset $H \subseteq V$ that is also a vector space (over the same $\mathbb{F}$).

Ex: $\{\vec{0}\} \subseteq V$ is always a subspace of any v.s.

It's called the (trivial) zero subspace.

Similarly, $V \subseteq V$ is a subspace of itself.

Ex: $\mathbb{P}_n^f(\mathbb{R})$ ... $\mathbb{P}_1^f(\mathbb{R}), \dots, \mathbb{P}_{n-1}^f(\mathbb{R}) \subseteq \mathbb{P}_n^f(\mathbb{R})$

Theorem: (Vector ~~space~~ subspace check) A set $H \subseteq V$ of a v.s. is a subspace of V if:

1. $\vec{0}_H = \vec{0}_V \in H$.
2. $+_H = +_V$ and $\vec{u}, \vec{v} \in H \Rightarrow \vec{u} + \vec{v} \in H$.
3. $\cdot_H = \cdot_V$ and $\vec{u} \in H, c \in \mathbb{F}_H = \mathbb{F}_V \Rightarrow c\vec{u} \in H$.

Wake-up Call!

Show that $W = \{ \vec{v} \in \mathbb{R}^3 \mid x - z = 0 \text{ and } y + z = 0 \}$, where $\vec{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$,

is a subspace of $\mathbb{R}^3$. (Think: What do you need to check?)

To show W is a subspace:

1. $\vec{0} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \in W$
2. $\vec{v}_1, \vec{v}_2 \in W$, show $\vec{v}_1 + \vec{v}_2 \in W$.
3. $c \in \mathbb{R}$, $\vec{v} \in W$, show $c \cdot \vec{v} \in W$.

Proof:

① $\vec{0} = \begin{pmatrix} 0 = x \\ 0 = y \\ 0 = z \end{pmatrix}$

$x - z = 0 - 0 = 0$ ✓

$y + z = 0 + 0 = 0$ ✓

$\vec{0} \in W$.

② $\vec{v}_1 = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} \in W$.

$\vec{v}_1 + \vec{v}_2 = \begin{pmatrix} x_1 + x_2 \\ y_1 + y_2 \\ z_1 + z_2 \end{pmatrix}$.

$(x_1 + x_2) - (z_1 + z_2)$

$= (x_1 - z_1) + (x_2 - z_2)$

$= 0 + 0 = 0$ ✓

$(y_1 + y_2) + (z_1 + z_2)$

$= (y_1 + z_1) + (y_2 + z_2)$

$= 0 + 0 = 0$ ✓.

③ $c \in \mathbb{R}$, $\vec{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in W$.

Check $c\vec{v} = \begin{pmatrix} cx \\ cy \\ cz \end{pmatrix}$

$(cx) - (cz)$

$= c(x - z)$

$= c \cdot 0 = 0$ ✓

$(cy) + (cz)$

$= c(y + z)$

$= c \cdot 0 = 0$. ✓

By ①, ②, and ③, W is a subspace of $\mathbb{R}^3$. ∎

Important Distinction:

Is $\mathbb{R}^2 \underset{\text{subspace}}{\subseteq} \mathbb{R}^3$? $\mathbb{R}^2$ not even subset of $\mathbb{R}^3$.

In general, $\mathbb{R}^n \not\subseteq \mathbb{R}^{n+k}$ $(k \geq 1)$.

Ex: Consider $W = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\} \subseteq \mathbb{R}^3$.

x-axis y-axis.

So W = xy-plane in $\mathbb{R}^3$.

$W \neq \mathbb{R}^2$ ... but W "looks like" $\mathbb{R}^2$.

We say $W \cong \mathbb{R}^2$ (isomorphic).

($\mathbb{R}^2$ embeds in $\mathbb{R}^3$)

Ex: Solutions to $A\vec{x} = \vec{b}$ form a vector space

iff

$\vec{b} = \vec{0}$. otherwise, $\vec{x} = \vec{0} \notin$ Sol. Set.

$\Rightarrow$ Sol. set not a v.s.

Def: Lin. comb. of $\{\vec{v}_1, \ldots, \vec{v}_p\} \subseteq V$ w/ weights $c_1, \ldots, c_p \in \mathbb{F}$ is $\vec{v} = c_1 \vec{v}_1 + \cdots + c_p \vec{v}_p = \sum_{i=1}^{p} c_i \vec{v}_i$.

Theorem: If $\vec{v}_1, \ldots, \vec{v}_p \in V$, then $\text{span}\{\vec{v}_1, \ldots, \vec{v}_p\} \underset{}{\overset{\text{subspace}}{\subseteq}} V$.

Proof: $S = \text{span}\{\vec{v}_1, \dots, \vec{v}_p\} = \{\text{all. lin. comb. of } \vec{v}_1, \dots, \vec{v}_p\}$.

$\Rightarrow \vec{v} \in S, \quad \vec{v} = \sum_{i=1}^{p} c_i \vec{v}_i$ for some c_i's.

① $\vec{0} \in V$? : $\quad c_1 = c_2 = \cdots = c_p = 0 \Rightarrow \vec{v} = \sum_{i=1}^{p} c_i \vec{v}_i = \vec{0}$. ✓

So $\vec{0} \in S$.

② For any $\vec{v}, \vec{w} \in S \Rightarrow \vec{v} + \vec{w} \in S$:

Let $\vec{v} = \sum c_i \vec{v}_i, \quad \vec{w} = \sum d_i \vec{v}_i$.

Then $\vec{v} + \vec{w} = \sum c_i \vec{v}_i + \sum d_i \vec{v}_i = \sum (c_i + d_i) \vec{v}_i$. ✓

$\vec{v} + \vec{w} \in S$

③ For any $r \in \mathbb{F}, \vec{v} \in S \Rightarrow r\vec{v} \in S$:

Let $\vec{v} = \sum c_i \vec{v}_i$. Then $r\vec{v} = r(\sum c_i \vec{v}_i)$

$= \sum r(c_i \vec{v}_i)$

$= \sum (r c_i) v_i$. ✓

$r\vec{v} \in S$. ✓

By ①, ②, and ③, $S \subseteq V$ is a subspace. ∎

Null Spaces, Column Spaces, and Linear Transformations

Lay - 4.2 Strang - 3.2

Observation: $\mathbb{R}^n$ is so commonly used as it arises in two ways:

① set of all solutions to a homogeneous system.

② set of lin. comb.'s of certain vectors. (spans)

We've seen ② creates subspaces. Let's look at ①.

Def: The <u>null space</u> of an $m \times n$ matrix A ($Null(A)$) is the set of all solutions to the homogeneous eq. $A\vec{x} = \vec{0}$.

$$Null(A) = \{ \vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0} \}.$$

Theorem: The null space of $A \in M_{m \times n}(\mathbb{R})$ is a subspace of $\mathbb{R}^n$. Equivalently, the solutions to $A\vec{x} = \vec{0}$ form a subspace of $\mathbb{R}^n$.

Proof: By def, $Null(A) \overset{\text{subset}}{\subseteq} \mathbb{R}^n$.

① <u>$\vec{0} \in Null(A)$?</u> : $A\vec{0}_n = \vec{0}_m$ ✓. $\vec{0} \in Null(A)$.

② Let $\vec{v}, \vec{w} \in Null(A)$. $A(\vec{v} + \vec{w}) = A\vec{v} + A\vec{w} = \vec{0} + \vec{0} = \vec{0}$. ✓

$\vec{v} + \vec{w} \in Null(A)$.

③ Let $c \in \mathbb{R}$, $\vec{v} \in Null(A)$. $A(c\vec{v}) = c(A\vec{v}) = c \cdot \vec{0} = \vec{0}$. ✓

$c\vec{v} \in Null(A)$. ✓

∎

Ex: "Find a description of Null(A) where

$$A = \begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}.$$"

$(A\vec{x} = \vec{0})$.

First, put $[A \mid \vec{0}]$ in RREF:

$$[A \mid \vec{0}] \sim \left[\begin{array}{ccccc|c} 1 & -2 & 0 & -1 & 3 & 0 \\ 0 & 0 & 1 & 2 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

$$\begin{cases} x_1 = 2x_2 + x_4 - 3x_5 \\ x_3 = -2x_4 + 2x_5 \\ x_2, x_4, x_5 \text{ free} \end{cases} \longrightarrow \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = x_2 \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} + x_5 \begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix}.$$

x_2, x_4, x_5 free

Solutions to $A\vec{x} = \vec{0}$.

$$\boxed{\mathrm{Null}(A) = \mathrm{span}\left\{ \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix} \right\}.}$$

$$A\begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \vec{0}, \quad A\begin{pmatrix} 1 \\ 0 \\ -2 \\ 1 \\ 0 \end{pmatrix} = \vec{0}, \quad A\begin{pmatrix} -3 \\ 0 \\ 2 \\ 0 \\ 1 \end{pmatrix} = \vec{0}.$$

$\Rightarrow$ all lin. comb.'s of them map to $\vec{0}$.

Def: The <u>column space</u> of $A \in M_{m\times n}(\mathbb{R})$ $(\mathrm{Col}(A))$ is the set of all lin. comb.'s of the columns of A. If

$$A = [\vec{a}_1 \cdots \vec{a}_n], \text{ then } \mathrm{Col}(A) = \mathrm{span}\{\vec{a}_1, \dots, \vec{a}_n\}.$$
$$= \{\vec{b} \in \mathbb{R}^m \mid A\vec{x} = \vec{b} \text{ has a sol.}\}.$$

Theorem: $\mathrm{Col}(A) \subseteq \mathbb{R}^m$ is a subspace.

~~A~~ Moreover, $\mathrm{Col}(A) = \mathbb{R}^m$ iff $A\vec{x} = \vec{b}$ has a solution for all $\vec{b} \in \mathbb{R}^m$.

iff $T(\vec{x}) = A\vec{x}$ is onto $(\mathbb{R}^m)$.

Ex: $A = \begin{bmatrix} 2 & 4 & -2 & 1 \\ -2 & -5 & 7 & 3 \\ 3 & 7 & -8 & 6 \end{bmatrix}$

① $\mathrm{Col}(A) \subseteq \mathbb{R}^{\underline{3}}$

② $\mathrm{Null}(A) \subseteq \mathbb{R}^{\underline{4}}$.

Notice that these two spaces "live" in entirely <u>different</u> vector spaces.

In particular, $\mathrm{Col}(A) = \mathrm{span}\left\{ \begin{pmatrix} 2 \\ -2 \\ 3 \end{pmatrix}, \begin{pmatrix} 4 \\ -5 \\ 7 \end{pmatrix}, \begin{pmatrix} -2 \\ 7 \\ -8 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 6 \end{pmatrix} \right\}$.

$= \mathrm{span}\{ \underset{\dots}{<4} \}$.

From before: $\mathrm{Col}(A)$ for $A \sim \begin{bmatrix} 1 & -2 & 0 & -1 & 3 \\ 0 & 0 & 1 & 2 & -2 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

$$\begin{bmatrix} -3 & 6 & -1 & 1 & -7 \\ 1 & -2 & 2 & 3 & -1 \\ 2 & -4 & 5 & 8 & -4 \end{bmatrix}$$

$\mathrm{Col}(A) = \mathrm{span}\left\{ \begin{pmatrix} -3 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 5 \end{pmatrix} \right\}$.

Look for pivot columns in A (not $\mathrm{REF}(A)$).

Def: A lin. trans. $T: V \to W$ is a rule assigning $\vec{x} \in V$ to a $T(\vec{x}) \in W$ such that

① $T(\vec{u}+\vec{v}) = T(\vec{u}) + T(\vec{v})$ for all $\vec{u}, \vec{v} \in V$.

② $T(c\vec{u}) = c\,T(\vec{u})$ for all $c \in \mathbb{F}_V = \mathbb{F}_W$ and $\vec{u} \in V$.

The kernel of T ($\ker T$) is the set of all $\vec{u} \in W$ such that $T(\vec{u}) = \vec{0}$. The range of T ~~is~~ (range T) is the set of all $\vec{b} \in W$ such that $T(\vec{x}) = \vec{b}$ for some $\vec{x} \in V$.

Linearly Independent Sets; Bases

Lay - 4.3 Strang - 3.5

Same def's and thm's as before, just in a diff. context.

Def: $\{\vec{v}_1, \dots, \vec{v}_p\} \subseteq V_{\mathbb{F}}$ is lin. ind. if $c_1\vec{v}_1 + \cdots + c_p\vec{v}_p = \vec{0}$ (*) has only the trivial solution ($c_1 = c_2 = \cdots = c_p = 0$).

~~Df~~ It's lin. dep. if (*) has nontrivial solutions.

Theorem: $\{\vec{v}_1, \dots, \vec{v}_p\}$ w/ $\vec{v}_1 \neq \vec{0}$ is lin. dep. iff for some $\vec{v}_j$ $(j>1)$ is lin. comb. of the vectors $\vec{v}_1, \dots, \vec{v}_{j-1}$.

Ex: $\{1, t, 4-t\} \subseteq \mathbb{P}_2^t$. Set is lin. dep. since

$$4-t = 4\cdot 1 + (-1)\cdot t.$$

Ex: $\{\sin t, \cos t\} \subseteq C^{\infty}(\mathbb{R})$ is lin. ind.

Notice that $\sin t \neq c \cdot \cos t$ for <u>any</u> $c \in \mathbb{R}$.

But $\{\sin t \cos t, \sin 2t\}$ is lin. dep. as

$$\sin(2t) = 2 \cdot (\sin t \cos t).$$

Def: Let $H \subseteq V$ be a subspace of the v.s. V (over $\mathbb{F}$). An indexed set of vectors $\beta = \{\vec{b}_1, \ldots, \vec{b}_p\} \subseteq V$ is a <u>basis</u> for H if

☆☆

① β is lin. ind. (gives uniqueness)

② $\operatorname{span}(\beta) = H$ (gives weights)

Ex: $\{\vec{e}_1, \ldots, \vec{e}_n\} \subseteq \mathbb{R}^n$ is the <u>standard basis</u> for $\mathbb{R}^n$.

$\uparrow$ columns of I_n.

E.g. in $\mathbb{R}^3$
$$\left[\begin{array}{ccc|c} 1 & 0 & 0 & b_1 \\ 0 & 1 & 0 & b_2 \\ 0 & 0 & 1 & b_3 \end{array}\right].$$

Ex: $A \in M_{n \times n}(\mathbb{R})$ invertible with $A = [\vec{a}_1 \cdots \vec{a}_n]$.

Then $\{\vec{a}_1, \ldots, \vec{a}_n\}$ form a basis for $\mathbb{R}^n$ as well.

since $A \sim I_n$.

Def: The columns of

E.g. $A = \begin{bmatrix} 3 & -4 & -2 \\ 0 & 1 & 1 \\ -6 & 7 & 5 \end{bmatrix}$. Is $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$ a basis for $\mathbb{R}^3$?

$$\begin{array}{ccc} \uparrow & \uparrow & \uparrow \\ \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{array}$$

$A \sim \begin{bmatrix} 3 & -4 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix} \Rightarrow A^{-1}$ exists $\Rightarrow$ ~~the~~ columns are a basis for $\mathbb{R}^3$.

Ex: $\{1, t, t^2, \dots, t^n\} \subseteq \mathbb{P}_n$ is a basis ~~for~~

Question: Do I need all vectors in $\{\vec{v}_1, \dots, \vec{v}_p\}$ to describe their span?

Ex: $\vec{v}_1 = \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}$, $\vec{v}_3 = \begin{pmatrix} 6 \\ 16 \\ -5 \end{pmatrix}$, $H = \text{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$.

Notice $\vec{v}_3 = 5\vec{v}_1 + 3\vec{v}_2$. So $\vec{v}_3$ is "extra".

So $H = \text{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \text{span}\{\vec{v}_1, \vec{v}_2\}$.

Generally,

Theorem: (Spanning set / Replacement Theorem)

on next page.

Let $S = \{\vec{v}_1, \dots, \vec{v}_p\} \subseteq V^{(v.s)}$ be a set with $H = \text{span}\{S\} \subseteq V$.

(a) If one of the vectors (say $\vec{v}_k$) is a lin. comb. of the remaining vectors in S, then the set $S - \{\vec{v}_k\}$ still spans H. That is,

$$\text{span}\{\vec{v}_1, \dots, \vec{v}_p\} = \text{span}\{\vec{v}_1, \dots, \vec{v}_{k-1}, \vec{v}_{k+1}, \dots, \vec{v}_p\}.$$

(b) If $H \neq \{\vec{0}\}$, then some subset of S is a basis for H.

~~Ex~~. <u>Warm-up</u>: $H = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ -4 \end{pmatrix}, \begin{pmatrix} 2 \\ 4 \\ -3 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$

$= \text{span}\{ \_\_ , \_\_ \}$?

<u>Theorem</u>: (Spanning Set/Replacement Theorem)

Let $S = \{\vec{v}_1, \dots, \vec{v}_p\} \subseteq V$ (a v.s.) be a set with $H = \text{span}(S)$.

(a) If one of the vectors (say $\vec{v}_k$) is a linear combination of the remaining vectors in S, then the set $S - \{\vec{v}_k\}$ still spans H. That is,

$$\text{span}\{\vec{v}_1, \dots, \vec{v}_k, \dots, \vec{v}_p\} = \text{span}\{\vec{v}_1, \dots, \vec{v}_{k-1}, \vec{v}_{k+1}, \dots, \vec{v}_p\}.$$

(b) If $H \neq \{\vec{0}\}$, then some subset of S is a basis for H.

$$A = \begin{bmatrix} 1 & 3 & 2 & 1 \\ 0 & 4 & 4 & 1 \\ -1 & -4 & -3 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & \cdot & \cdot & \cdot \\ 0 & 1 & \cdot & \cdot \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

spans H

pivots tell you which columns <u>of A</u> are lin. ind.

$$\Rightarrow H = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \\ -4 \end{pmatrix} \right\} = \text{Col}(A)$$

<u>Question</u>: Can we find bases for ($A \in M_{m\times n}(\mathbb{R})$)

$$\text{Null}(A) = \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\}. \quad \text{(kernel of trans.)}$$

and $$\text{Col}(A) = \{\vec{b} \in \mathbb{R}^m \mid A\vec{x} = \vec{b} \text{ is consistent}\}.$$

(image of trans.)

Col(A) is easy:

How? Row reduce A to find pivot columns.

Pivot columns <u>of A</u> correspond to a basis of Col(A).

<u>Ex:</u>

$$A = \begin{pmatrix} 1 & 4 & 0 & 2 & -1 \\ 3 & 12 & 1 & 5 & 5 \\ 2 & 8 & 1 & 3 & 2 \\ 5 & 20 & 2 & 8 & 8 \end{pmatrix} \sim \begin{bmatrix} 1 & 4 & 0 & 2 & 0 \\ 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} = \text{rref}(A) = B$$

$$\text{Col}(A) = \text{span}\left\{ \begin{pmatrix} 1 \\ 3 \\ 2 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 5 \\ 2 \\ 8 \end{pmatrix} \right\}.$$

Notice: Col 2 = $4 \cdot$ Col 1

Col 4 = $2 \cdot$ Col 1 $-$ Col 3.

<u>Justify:</u> Since $A\vec{x} = \vec{0}$ and $B\vec{x} = \vec{0}$ have same solutions, their columns have the same dependence relations.

To find a basis for Null(A):

① Solve $[A \mid \vec{0}]$.

② Vectors corresponding to each free variable form a basis for Null(A).

E.g. $$\vec{x} = x_2 \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} 0 \\ 2 \\ 1 \\ 1 \end{pmatrix}$$

$$\text{Null}(A) = \text{span}\left\{ \begin{pmatrix} 3 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

Ex:

"$A = \begin{bmatrix} 1 & 4 & 0 & 2 & -1 \\ 3 & 12 & 1 & 5 & 5 \\ 2 & 8 & 1 & 3 & 2 \\ 5 & 20 & 2 & 8 & 8 \end{bmatrix}$. Describe a basis for Null(A)."

$$[A \mid \vec{0}] \sim \left[\begin{array}{ccccc|c} 1 & 4 & 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

$$\updownarrow$$

$$\begin{cases} x_1 = -4x_2 - 2x_4 \\ x_2 \text{ free } (x_2 = x_2) \\ x_3 = x_4 \\ x_4 = x_4 \\ x_5 = 0 \end{cases}$$

So $\vec{x} = x_2 \begin{pmatrix} -4 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -2 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \text{span}\left\{ \begin{pmatrix} -4 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}$.

$= \text{Null}(A)$

$$\boxed{\text{Basis} = \left\{ \begin{pmatrix} -4 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} \right\}}$$

Check:

$$A \begin{pmatrix} -4 \\ 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} = -4 \begin{pmatrix} 1 \\ 3 \\ 2 \\ 5 \end{pmatrix} + 1 \begin{pmatrix} 4 \\ 12 \\ 8 \\ 20 \end{pmatrix} = \vec{0}.$$

$$A \begin{pmatrix} -2 \\ 0 \\ 1 \\ 1 \\ 0 \end{pmatrix} = -2 \begin{pmatrix} 1 \\ 3 \\ 2 \\ 5 \end{pmatrix} + 1 \begin{pmatrix} 0 \\ 1 \\ 1 \\ 2 \end{pmatrix} + 1 \begin{pmatrix} 2 \\ 5 \\ 3 \\ 8 \end{pmatrix} = \vec{0}.$$

Subtle Observation : 3 pivots + 2 free var's = 5 (# of columns).

~~Why are these independent? First vector has x_2 leading term, second has x_4 l.t. $\Rightarrow$~~

Observation: Bases are exactly the "right size."

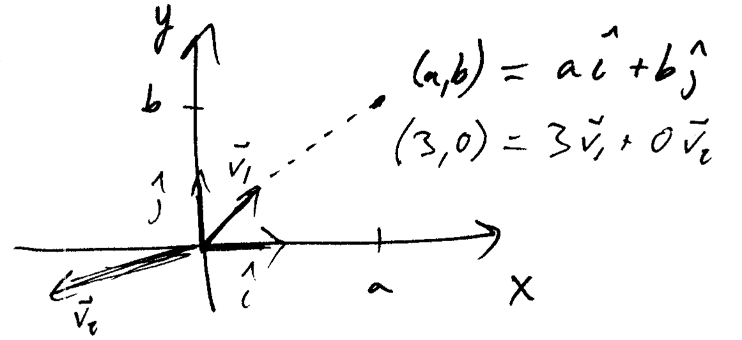
① Too many $\bar{v}_i$ $\Rightarrow$ lin. dep. (just small enough)

② Too few $\vec{v}_i$ $\Rightarrow$ won't span (just big enough)

Coordinate Systems

Lay - 4.4 Strang - 3.5

Motivation : In $\mathbb{R}^2$



Axes already define a coordinate system.

Can we make this idea rigorous? How many vectors do I need to describe this point? Uniquely? At all?

onto 1-1

Theorem: (Unique Representation) Let $\beta = \{\vec{b}_1, \dots, \vec{b}_n\}$ be a basis for some vector space V (over a field $\mathbb{F}$). The for each $\vec{x} \in V$, there exist unique scalar weights $c_1, \dots, c_n \in \mathbb{F}$ such that
$$\vec{x} = c_1 \vec{b}_1 + \cdots + c_n \vec{b}_n = \sum_{i=1}^{n} c_i \vec{b}_i .$$

(Def: The vector $[\vec{x}]_\beta = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{F}^n$ is called the β-coordinates of $\vec{x}$.)

Proof: Because β spans V (i.e. $\text{span}(\beta) = V$), there must exist weights $c_1, \dots, c_n$ where $\vec{x} = \sum c_i \vec{b}_i$.

To show these c_i's are unique, suppose there are some other d_i's $\in \mathbb{F}$ such that
$$\vec{x} = \sum c_i \vec{b}_i = \sum d_i \vec{b}_i .$$

Then $\vec{0} = \vec{x} - \vec{x} = \sum c_i \vec{b}_i - \sum d_i \vec{b}_i = \sum (c_i - d_i) \vec{b}_i$.

Lin. ind. of β $\Rightarrow$ $c_i - d_i = 0$ for all $1 \le i \le n$.

$c_i = d_i \Rightarrow c_i$'s are unique. ∎

Ex: "Let $S = \{\vec{e}_1, \vec{e}_2\}$, $\beta = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$, $\gamma = \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 5 \end{pmatrix} \right\}$.

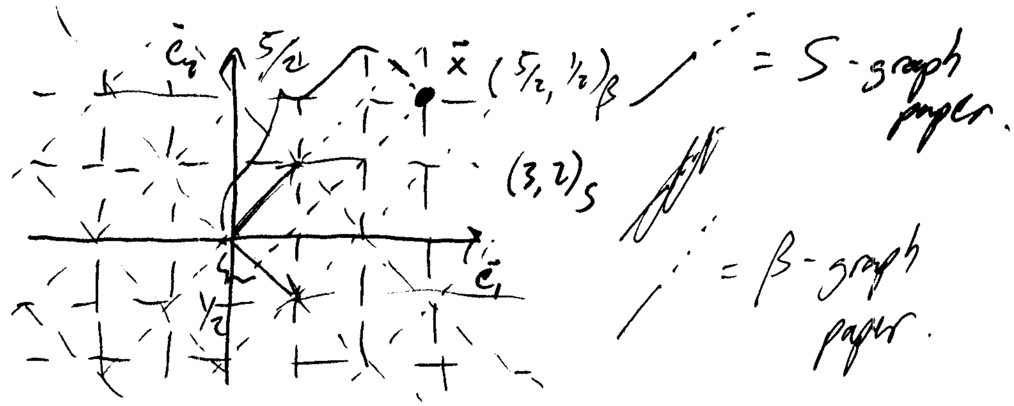
Find the β- and γ-coordinates of $\vec{x}$ ($[\vec{x}]_S = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$)"

I.e. Find c_1, c_2 such that $c_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$.

So solve $\left[\begin{array}{cc|c} 1 & 1 & 3 \\ 1 & -1 & 2 \end{array}\right] \Rightarrow c_1 = 5/2,\ c_2 = 1/2.$

Check: $\frac{5}{2}\begin{pmatrix}1\\1\end{pmatrix} + \frac{1}{2}\begin{pmatrix}1\\-1\end{pmatrix} = \begin{pmatrix}3\\2\end{pmatrix}.$

Pictorially:



For γ-coords: $c_1\begin{pmatrix}2\\1\end{pmatrix} + c_2\begin{pmatrix}0\\5\end{pmatrix} = \begin{pmatrix}3\\2\end{pmatrix}$

Solve $\left[\begin{array}{cc|c} 2 & 0 & 3 \\ 1 & 5 & 2 \end{array}\right] \Rightarrow \begin{array}{l} c_1 = 3/2 \\ c_2 = 1/10 \end{array}$

So $\begin{pmatrix}3\\2\end{pmatrix}_S = \begin{pmatrix}3/2\\1/10\end{pmatrix}_\gamma.$

Coordinates in $\mathbb{R}^n$: $S = \{\vec{e}_1, \vec{e}_2\}$

~~One~~ One more ex: $\beta = \left\{\begin{pmatrix}2\\1\end{pmatrix}, \begin{pmatrix}-1\\1\end{pmatrix}\right\}.\quad [\vec{x}]_S = \begin{pmatrix}4\\5\end{pmatrix}$

$[\vec{x}]_\beta$: $c_1\begin{pmatrix}2\\1\end{pmatrix} + c_2\begin{pmatrix}-1\\1\end{pmatrix} = \begin{pmatrix}4\\5\end{pmatrix}$

$$\begin{pmatrix}2 & -1\\1 & 1\end{pmatrix}\begin{pmatrix}c_1\\c_2\end{pmatrix} = \begin{pmatrix}4\\5\end{pmatrix}$$

$$\begin{pmatrix}2 & -1\\1 & 1\end{pmatrix}[\vec{x}]_\beta = [\vec{x}]_S.$$

Uniqueness theorem $\Rightarrow \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}^{-1}$ exists.

So $$[\vec{x}]_\beta = \begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}^{-1} [\vec{x}]_s.$$
$$= \frac{1}{3} \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix} [\vec{x}]_s.$$

$$[\vec{x}]_\beta = \frac{1}{3} \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 4 \\ 5 \end{pmatrix}$$
$$= \frac{1}{3} \begin{pmatrix} 9 \\ 6 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

$$\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \end{pmatrix} = 3 \begin{pmatrix} 2 \\ 1 \end{pmatrix} + 2 \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix} + \begin{pmatrix} -2 \\ 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \end{pmatrix}_S \checkmark$$

<u>From HW:</u> $\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix} = A_\beta^S$, $\begin{pmatrix} 1/3 & 1/3 \\ -1/3 & 2/3 \end{pmatrix} = A_S^\beta$

takes β-coords to S-coords.

<u>In Fact:</u> $T([\vec{x}]_\beta) = A_\beta^S [\vec{x}]_\beta = [\vec{x}]_S$.

$A_\beta^S = [\vec{b}_1 \cdots \vec{b}_n]$ for $\beta = \{\vec{b}_1, \dots, \vec{b}_n\}$.

<u>MORAL:</u> Changing coordinates is a lin. trans.!!!

Note: $(A_\beta^S)^{-1} = A_S^\beta$ so
$$\begin{tikzcd} \beta\text{-coords} \arrow[r, bend left, "A_\beta^S"] & S\text{-coords} \arrow[l, bend left, "A_S^\beta"] \end{tikzcd}$$

Def: (Standard notation) $A_\beta^S = \underset{S \leftarrow \beta}{P}$ is the change-of-coordinates matrix from β-coords to S-coords.

Theorem: For any basis $\beta = \{\vec{b}_1, \dots, \vec{b}_n\}$ of a v.s. V, the coordinate map $T(\vec{x}) = [\vec{x}]_\beta$ $(T: V \to \mathbb{R}^n)$ is a one-to-one, onto lin. trans.

(**)

Def: (1-1, onto lin. trans. = isomorphisms of v.s.).

A 1-1, lin. trans. of one v.s. V that is onto W is called a v.s. isomorphism. ($V \cong W$, V is isomorphic to W.)

Ex: $\mathbb{P}_3^f$ w/ $S = \{1, t, t^2, t^3\}$ and

$$\beta = \{1 + 2t,\ 2t + 3t^2,\ 3t^2 + 4t^3,\ 1 + 4t^3\}.$$

(check that β is a basis for $\mathbb{P}_3^f$).

Theorem says $\mathbb{P}_3^f \cong \mathbb{R}^4$: $A_S^S = \underset{S \leftarrow S}{P} = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} = I_4$.

$$T(p(t) = a_0 + a_1 t + a_2 t^2 + a_3 t^3) = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix}. \quad = \underset{\beta \leftarrow \beta}{P}$$

So $A_\beta^S = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 2 & 2 & 0 & 0 \\ 0 & 3 & 3 & 0 \\ 0 & 0 & 4 & 4 \end{bmatrix} = \underset{S \leftarrow \beta}{P}$ is an isomorphism for $T([\vec{x}]_\beta) = [\vec{x}]_S$.

Observe: Since matrices represent lin. trans., notice that A_β^s represents $T(\vec{x}) = \vec{x}$, the identity transformation.

The Dimension of a Vector Space

Lay - 4.5 Strang - 3.5

Observation: Just had a theorem saying that (most) vector spaces are isomorphic to $\mathbb{R}^n$ for some integer $n \geq 0$. This means that all vector spaces "look identical" to some ~~some~~ $\mathbb{R}^n$, ~~differing~~ differing only in the number n. ~~The~~ Is this number unique to the vector space? That is, $T(\vec{x}) = [\vec{x}]_\beta$ can only have one codomain?.

Theorem: If $\beta = \{\vec{b}_1, \dots, \vec{b}_n\} \subseteq V$ is a basis for V, then every set of more than n vectors of V must be lin. dep.

Pf: Let $\{\vec{u}_1, \dots, \vec{u}_p\} \subseteq V$ with $p > n$, then $\{[\vec{u}_1]_\beta, \dots, [\vec{u}_p]_\beta\} \subseteq \mathbb{R}^n$.

Theorem from "matrix land" $\Rightarrow \left[[\vec{u}_1]_\beta \cdots [\vec{u}_p]_\beta \right]_{n \times p}$ has lin. dep. columns.

So there are c_i's such that $\sum_{i=1}^{p} c_i [\vec{u}_i]_\beta = \vec{0}$ w/ ~~not~~ not all $c_i = 0$.

Recall, $T(\vec{x}) = [\vec{x}]_\beta$ is a lin. trans.

Notice $\sum_{i=1}^{p} c_i [\vec{u}_i]_\beta = \sum_{i=1}^{p} c_i T(\vec{u}_i)$.

$= \sum_{i=1}^{p} T(c_i \vec{u}_i)$

$\vec{0} = T\left(\sum_{i=1}^{p} c_i \vec{u}_i\right)$.

T 1-1 $\Rightarrow \left(T(\vec{x}) = \vec{0} \Rightarrow \vec{x} = \vec{0}\right)$.

So $\sum_{i=1}^{p} c_i \vec{u}_i = \vec{0}$ for c_i not all 0.

$\Rightarrow \{\vec{u}_1, \dots, \vec{u}_p\}$ lin. dep. ∎

Warm-up: $S = \{\vec{e}_1, \vec{e}_2\}$, $\alpha = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$, $\beta = \left\{ \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right\}$.

Find α- and β-coordinates for $\begin{pmatrix} -5 \\ 3 \end{pmatrix}_S$.

Hint: $\underset{S \leftarrow \alpha}{P} = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$, $\underset{S \leftarrow \beta}{P} = \begin{pmatrix} 2 & 3 \\ 3 & 2 \end{pmatrix}$.

---

take $[\vec{x}]_S \mapsto [\vec{x}]_\alpha$, $[\vec{x}]_\beta$.

Same as $\underset{S \leftarrow \alpha}{P} [\vec{x}]_\alpha = [\vec{x}]_S$.

$$\underset{\alpha \leftarrow S}{P} = \left(\underset{S \leftarrow \alpha}{P} \right)^{-1} = \frac{1}{-2} \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix}$$

$$\underset{\beta \leftarrow S}{P} = \left(\underset{S \leftarrow \beta}{P} \right)^{-1} = \frac{1}{-5} \begin{pmatrix} 2 & -3 \\ -3 & 2 \end{pmatrix}$$

So
$$\left[\begin{pmatrix} -5 \\ 3 \end{pmatrix}_S \right]_\alpha = \frac{1}{-2} \begin{pmatrix} -1 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} -5 \\ 3 \end{pmatrix} = \frac{1}{-2} \begin{pmatrix} 2 \\ 8 \end{pmatrix} = \begin{pmatrix} -1 \\ -4 \end{pmatrix}.$$

$$\left[\begin{pmatrix} -5 \\ 3 \end{pmatrix}_S \right]_\beta = \frac{1}{-5} \begin{pmatrix} 2 & -3 \\ -3 & 2 \end{pmatrix} \begin{pmatrix} -5 \\ 3 \end{pmatrix} = \frac{1}{-5} \begin{pmatrix} -19 \\ 21 \end{pmatrix} = \begin{pmatrix} +19/5 \\ -21/5 \end{pmatrix}.$$

Theorem: If $\beta = \{\vec{b}_1, \dots, \vec{b}_n\} \subseteq V$ is a basis for V, then every set of more than n vectors of V must be lin. dep.

Theorem: If a vector space V has a basis of n vectors, then any basis of V must consist of exactly n vectors.

Pf: β Basis spans $V \Rightarrow$ # vectors in $\beta = |\beta| \geq n$.

Theorem above $\Rightarrow |\beta| \leq n$.

$n \leq |\beta| \leq n \Rightarrow |\beta| = n$. ∎

Def: n is called the dimension of V (denoted by $\dim(V)$).

Def: If V is spanned by a finite set, V is finite-dimensional. Otherwise, V is infinite-dimensional.

Note: We use the convention that $\dim(\{\vec{0}\}) = 0$.

(Strange fact: The only basis for $\{\vec{0}\}$ is $\emptyset$.)

Ex: $\mathbb{P}_n^f = \operatorname{span}\{1, t, t^2, \dots, t^n\}$ $\quad \dim(\mathbb{P}_n^f) = n+1$

$\mathbb{P}^f = \operatorname{span}\{1, t, t^2, t^3, \dots\}$ $\quad \dim(\mathbb{P}^f) \text{ “}=\text{'' } \infty$.

Ex: Subspaces of $\mathbb{R}^n$: $\{\vec{0}\} \subseteq \mathbb{R}^n$ is the 0-dim'l subspace. It is a point.

$\operatorname{span}\{\vec{v}_1\} \subseteq \mathbb{R}^n$: 1-dim'l line through origin.
$\vec{v}_1 \neq \vec{0}$

$\operatorname{span}\{\vec{v}_1, \vec{v}_2\} \subseteq \mathbb{R}^n$: 2-dim'l plane through origin.
$\{\vec{v}_1, \vec{v}_2\}$ lin. ind.

$\operatorname{span}\{\vec{v}_1, \dots, \vec{v}_p\} \subseteq \mathbb{R}^n$: p-dim'l hyperplanes through origin.
$(3 \leq p \leq n)$ lin. ind.

Remark: The dimension of a fin-dim'l v. s. classifies it entirely. Dimension is called a v. s. invariant.

Theorem: Let $H \subseteq V$ be a subspace of a finite-dim'l V. Any lin. ind. subset of H can be extended to a basis of H. Moreover, H is fin-dim'l with $\dim H \leq \dim V$.

Ex: Consider $\{\vec{e}_1, \vec{e}_2\} \subseteq \mathbb{R}^3$. $H = \operatorname{span}\{\vec{e}_1, \vec{e}_2\}$.
Add $\vec{e}_3$ to $\uparrow$, to make $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ which is a basis for $\mathbb{R}^3$.

Theorem: Let V be a p-dim'l v.s. $(p \geq 1)$. Any lin. ind. set of exactly p elements in V is a basis for V. Any set of exactly p elements that span V is a basis for V.

Remark: Any two of the following determines the third:

1. set has ~~$\geq$~~ p elements and is lin. ind.
2. set has p elements and spans V.
3. $\dim V = p \geq 1$.

Rank:

Lay - 4.6 Strang - 3.3

Def: For a matrix $A \in M_{m \times n}(\mathbb{R})$, the set of all lin. comb.'s of the rows of A is called the <u>row space</u> of A (denoted by $\text{Row}(A)$).

Ex:

$$A = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix}$$

$$= [\vec{v}_1 \ \vec{v}_2 \ \vec{v}_3 \ \vec{v}_4 \ \vec{v}_5]$$

col vect's

$$= \begin{bmatrix} \vec{w}_1^T \\ \vec{w}_2^T \\ \vec{w}_3^T \\ \vec{w}_4^T \end{bmatrix}$$

row vect's

$\text{Col}(A) = \text{span}\{\vec{v}_1, \ldots, \vec{v}_5\} \subseteq \mathbb{R}^4$

$\text{Row}(A) = \text{span}\{\vec{w}_1, \ldots, \vec{w}_4\} \subseteq \mathbb{R}^5$.

$\text{Col}(A) \subseteq \mathbb{R}^m$

$\text{Row}(A) \subseteq \mathbb{R}^n$

Theorem: If $A, B \in M_{m\times n}(\mathbb{R})$ are row eq., then their row spaces are the same ($\text{Row}(A) = \text{Row}(B)$ if $A \sim B$). If B is in REF, then the nonzero rows of B form a basis for $\text{Row}(A) = \text{Row}(B) \leq \mathbb{R}^n$.

Col (A)	versus	Row (A)
$\text{Col}(A) \neq \text{Col}(\text{REF}(A))$		$\text{Row}(A) = \text{Row}(\text{REF}(A))$
Pivot columns of A form a basis for $\text{Col}(A)$.		"Pivot rows" of either A or $\text{REF}(A)$ form a basis of $\text{Row}(A)$.

Ex: $A = \begin{bmatrix} -2 & -5 & 8 & 0 & -17 \\ 1 & 3 & -5 & 1 & 5 \\ 3 & 11 & -19 & 7 & 1 \\ 1 & 7 & -13 & 5 & -3 \end{bmatrix} \sim B = \begin{bmatrix} 1 & 3 & -5 & 1 & 5 \\ 0 & 1 & -2 & 2 & -7 \\ 0 & 0 & 0 & -4 & 20 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$

So $\text{Row}(A) = \text{span}\left\{ \begin{pmatrix} -2 \\ -5 \\ 8 \\ 0 \\ -17 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ -5 \\ 1 \\ 5 \end{pmatrix}, \begin{pmatrix} 3 \\ 11 \\ -19 \\ 7 \\ 1 \end{pmatrix} \right\}$.

$= \text{span}\left\{ \begin{pmatrix} 1 \\ 3 \\ -5 \\ 1 \\ 5 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -2 \\ 2 \\ -7 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \\ -4 \\ 20 \end{pmatrix} \right\}$.

Def: The dimension of the column space (of a matrix A) is called the rank of A. $\text{rank}(A) = \dim(\text{Col}(A)) = \dim(\text{Row}(A))$.
The dimension of the null space is called the nullity of A. $\text{nul}(A) = \dim(\text{Null}(A))$ $[= \dim(\text{LNull}(A))]$.

Theorem: (Rank/Dimension)

(☆☆☆☆☆)

Let $A \in M_{m \times n}(\mathbb{R})$. The dimensions of $\mathrm{Col}(A)$ and $\mathrm{Row}(A)$ are equal, $\mathrm{rank}(A) = \#$ of pivots in A, and

$$\boxed{\mathrm{rank}(A) + \mathrm{nul}(A) = n.}$$

(Talks about decomposing $\mathbb{R}^n$ ...)

Remark: Recall that a column is either a pivot or free. Rank counts pivots, nullity counts free columns.

Theorem says pivot col's + free col's = all col's.

Ex: (a) If $A \in M_{7 \times 9}(\mathbb{R})$ w/ 2-dim'l $\mathrm{Null}(A)$, what is $\mathrm{rank}(A)$?

(b) Can $A \in M_{6 \times 9}(\mathbb{R})$ have a 2-dim'l $\mathrm{Null}(A)$?

Sol: (a)

$$\mathrm{rank}(A) + \mathrm{nul}(A) = 9$$

$$\mathrm{rank}(A) + 2 = 9$$

$$\mathrm{rank}(A) = 7. \Rightarrow 7 \text{ pivot rows!}$$

(b)

$$\mathrm{ran}(A) + \mathrm{nul}(A) = 9$$

$$\mathrm{rank}(A) = 7 \quad \text{only have } \underline{6 \text{ rows}}!$$

Subspaces can't have dimension exceeding that of the space in which they sit.

So... No! $\underline{\dim(\mathrm{Null}(A)) \geq 3}$.

Theorem: (Inv. mat. thm. part II)

Let $A \in M_{n\times n}(\mathbb{R})$. Then TFAE to the previous statements:

(m) columns of A are a basis for $\mathbb{R}^n$.

(n) $\mathrm{Col}(A) = \mathbb{R}^n$.

(p) $\mathrm{rank}(A) = n$.

(q) $\mathrm{Null}(A) = \{\vec{0}\}$.

(r) $\dim(\mathrm{Null}(A)) = \mathrm{nul}(A) = 0$.

Change of Basis (Coordinates)

Lay - 4.7 Strang - 7.2

Notice: $\mathbb{R}^2$ is comprised of points/locations. I can assign to each point a coordinate/address.

There are many different ways to describe the same points.

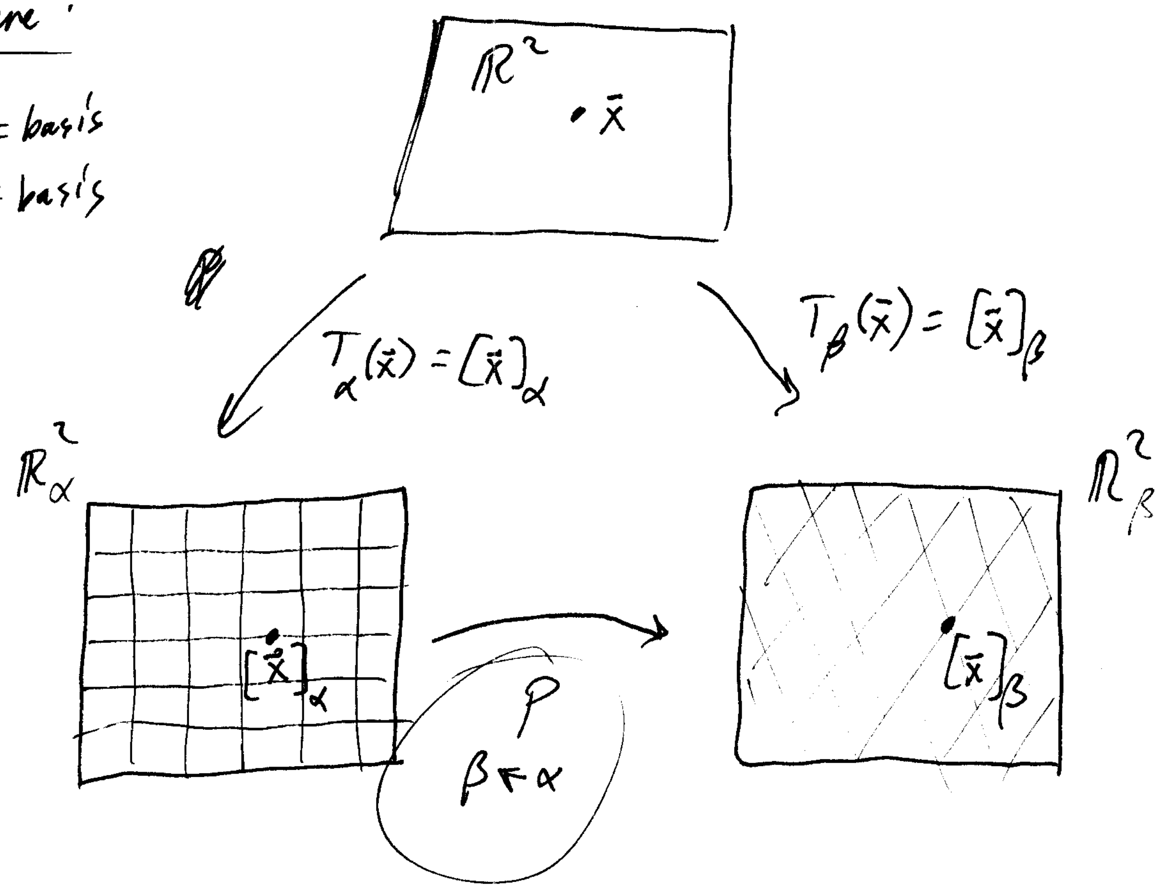
Question: Are these addresses related to one another?

Answer: Yes! Change-of-coordinates matrices.

Picture:

α = basis

β = basis



Ex: $V = \mathbb{P}_2^f$ // $\alpha = \{1, t, t^2\}$, $\beta = \{1-t^2, t-1, t^2+t\}$.

Let $p(t) = 2t$

$$[p(t)]_\alpha = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, \quad [p(t)]_\beta = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Question: How do we go from α-coords to β-coords?

IDEA: $p(t) \in \mathbb{P}_2^f$ has α-coords $\alpha = \{\vec{p}_1, \vec{p}_2, \vec{p}_3\}$

$\beta = \{\vec{q}_1, \vec{q}_2, \vec{q}_3\}$.

Find out what each $\vec{p}_i$ looks like in β.

~~Using the example:~~ α-coords are the "standard vectors"

$$\rho = c_1\rho_1 + c_2\rho_2 + c_3\rho_3.$$

If I know what happens to each ρ_i,
then I know what happens to ρ.

In particular: Idea is To send α, coordinate-by-coordinate, to β-coords.

Hence, $$\underset{\beta \leftarrow \alpha}{P} = \left[[\rho_1]_\beta \; [\rho_2]_\beta \; [\rho_3]_\beta \right].$$

Wake-up call!

Let $p(t) = 1$, $\alpha = \{1+t, t+t^2, 1+t^2\}$.

What is $[p(t)]_\alpha$?

Let $A = \begin{pmatrix} 2 & 1 \\ 0 & 4 \end{pmatrix}$, $\beta = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right\}$.

What is $[A]_\beta$?

So ... $|\alpha| = \dim \mathbb{P}_2^f \Rightarrow \mathbb{P}_2^f \cong \mathbb{R}^3$.

$$c_1(1+t) + c_2(t+t^2) + c_3(1+t^2) = 1 + 0t + 0t^2$$

Then $[p(t)]_\alpha = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix}$.

$$\begin{array}{l} 1 \\ t \\ t^2 \end{array} \left\{ \begin{array}{rcl} c_1 \qquad\quad + c_3 &=& 1 \\ c_1 + c_2 \qquad\quad &=& 0 \\ c_2 + c_3 &=& 0 \end{array} \right. \Leftrightarrow \left[\begin{array}{ccc|c} 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{array} \right]$$

$$[p(t)]_\alpha = \begin{pmatrix} 1/2 \\ -1/2 \\ 1/2 \end{pmatrix}$$

So ... $|\beta| = 4 = \dim M_{2\times 2}(\mathbb{R}) \Rightarrow M_{2\times 2}(\mathbb{R}) \cong \mathbb{R}^4$.

$$c_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + c_2 \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + c_3 \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} + c_4 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 4 \end{pmatrix}$$

$$a_{11}: \quad c_1 + c_2 + c_3 + c_4 = 2$$
$$a_{12}: \quad c_2 + c_3 + c_4 = 1$$
$$a_{21}: \quad c_3 + c_4 = 0$$
$$a_{22}: \quad c_4 = 4$$

$c_4 = 4$
$c_3 = -4$
$c_2 = 1$
$c_1 = 1$

$$\left[\begin{pmatrix} 2 & 1 \\ 0 & 4 \end{pmatrix}\right]_\beta = \begin{pmatrix} 1 \\ 1 \\ -4 \\ 4 \end{pmatrix}.$$

Question: How do we send α-coords to β-coords given $\alpha = \{\vec{a}_1, \dots, \vec{a}_n\}$ and $\beta = \{\vec{b}_1, \dots, \vec{b}_n\}$?

A: Theorem: The change-of-coordinates matrix $\underset{\beta \leftarrow \alpha}{P}$ defines a matrix transformation taking α-coords to β-coords. It is defined by $\underset{\beta \leftarrow \alpha}{P} [\vec{x}]_\alpha = [\vec{x}]_\beta$, and

$$\underset{\beta \leftarrow \alpha}{P} = \left[\, [\vec{a}_1]_\beta \;\; [\vec{a}_2]_\beta \;\; \cdots \;\; [\vec{a}_n]_\beta \,\right].$$

Ex: From before, $\alpha = \{1+t,\ t+t^2,\ 1+t^2\}$, $S = \{1, t, t^2\}$.

$$\underset{\alpha \leftarrow S}{P} = \left[\, [1]_\alpha \;\; [t]_\alpha \;\; [t^2]_\alpha \,\right] = \begin{bmatrix} 1/2 & 1/2 & -1/2 \\ -1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 \end{bmatrix}.$$

So, verify: $\underset{\alpha \leftarrow S}{P} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1/2 \\ -1/2 \\ 1/2 \end{pmatrix}$. ✓

Let $g(t) = 4 + 2t - t^2$.

$$[g(t)]_S = \begin{pmatrix} 4 \\ 2 \\ -1 \end{pmatrix}. \quad \text{So} \ldots \; [g(t)]_\alpha = ?$$

$$\underset{\alpha \leftarrow S}{P} [g(t)]_S = \begin{bmatrix} 1/2 & 1/2 & -1/2 \\ -1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 \end{bmatrix} \begin{pmatrix} 4 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 7/2 \\ -3/2 \\ 1/2 \end{pmatrix} = [g(t)]_\alpha .$$

$$\tfrac{7}{2}(1+t) - \tfrac{3}{2}(t+t^2) + \tfrac{1}{2}(1+t^2)$$
$$= 4 + 2t - t^2 . \quad \checkmark$$

Remarks: ① Matrices describe lin. trans.'s. P describes the identity trans. $T(\vec{x}) = \vec{x}$.
The bases just describe the point $\vec{x}$ in two different ways.

② If $\underset{\beta \leftarrow \alpha}{P}$ is a C.o.C. matrix, how do we take $\beta \to \alpha$ coords?

$$\underset{\alpha \leftarrow \beta}{P} = \left(\underset{\beta \leftarrow \alpha}{P} \right)^{-1}$$

Ex:

$$S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\} \subseteq M_{2\times 2}(\mathbb{R})$$

$$\alpha = \left\{ \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \right\}.$$

Bases for $M_{2\times 2}(\mathbb{R})$.

$$\underset{\alpha \leftarrow S}{P} = \left[\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}_\alpha \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}_\alpha \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}_\alpha \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}_\alpha \right].$$

For $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}_\alpha$: $c_1 \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + c_2 \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix} + c_3 \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} + c_4 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}_\alpha = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

For $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}_\alpha$: $c_1 \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} + c_2 \begin{pmatrix} 1 & 2 \\ 3 & 0 \end{pmatrix} + c_3 \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix} + c_4 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}_\alpha = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = ?$$

$$\begin{array}{ll} a_{11}: & \\ a_{12}: & \\ a_{21}: & \\ a_{22}: & \end{array} \left\{ \begin{array}{ll} c_1 + c_2 + c_3 + c_4 & = 0 \\ 2c_1 + 2c_2 + 2c_3 & = 1 \\ 3c_1 + 3c_2 & = 0 \\ 4c_1 & = 0 \end{array} \right. \leftrightarrow \begin{pmatrix} 0 \\ 0 \\ 1/2 \\ -1/2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}_\alpha$$

Similarly, $\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}_\alpha = \begin{pmatrix} 0 \\ 1/3 \\ -1/3 \\ 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}_\alpha = \begin{pmatrix} 1/4 \\ -1/4 \\ 0 \\ 0 \end{pmatrix}$.

So... $P_{\alpha FS} = \begin{pmatrix} 0 & 0 & 0 & 1/4 \\ 0 & 0 & 1/3 & -1/4 \\ 0 & 1/2 & -1/3 & 0 \\ 1 & -1/2 & 0 & 0 \end{pmatrix}$.

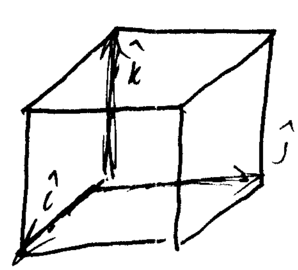
Simply, $P_{SF\alpha} = \left(P_{\alpha FS} \right)^{-1} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 0 \\ 3 & 3 & 0 & 0 \\ 4 & 0 & 0 & 0 \end{pmatrix}$

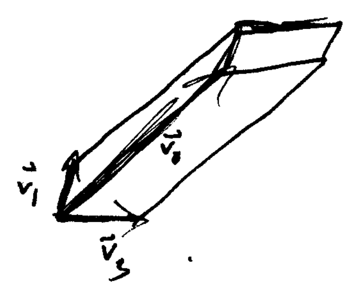
Introduction to Determinants

Lay - 3.1 Strang - 5.2

Historical Note: Analytic Geometry uses vectors as tools to determine geometric features of objects. Things like: distances, angles, areas/volumes, etc. In the subject's early developments, it was of particular interest to find volumes of objects whose edges were described by a collection of vectors.

E.g.





IDEA: Want "determinant of A" to be the volume of the object whose edges are determined by the vectors that are the columns of A. E.g. $I_3 = [\hat{i} \; \hat{j} \; \hat{k}]$.

want $\det(I_3) = 1 \cdot 1 \cdot 1 = 1$

Want $D = [\, d_1 \vec{e}_1 \;\; d_2 \vec{e}_2 \;\; \cdots \;\; d_n \vec{e}_n \,]$

to have $\det(D) = d_1 d_2 \cdots d_n$.

Question: How does this extend to "non-rectangular" objects?

Thought: Recall from 2×2 inverse formula, $ad-bc$ showed up.

How?

For a 3×3 $A = (a_{ij})$, row reduction gives

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \sim \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{11}a_{22} - a_{12}a_{21} & a_{11}a_{23} - a_{13}a_{21} \\ 0 & a_{11}a_{32} - a_{12}a_{31} & a_{11}a_{33} - a_{13}a_{31} \end{pmatrix}$$

$$\sim \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ 0 & a_{11}a_{22} - a_{12}a_{21} & a_{11}a_{23} - a_{13}a_{21} \\ 0 & 0 & a_{11}\triangle \end{pmatrix}$$

$$\triangle = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{31}a_{22})$$

$$= \det(A).$$

$$= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}$$

Def: For $n \geq 2$, the <u>determinant</u> of an $n \times n$ matrix $A = (a_{ij})_{n \times n}$ is the sum of n terms of the form $\pm a_{1j} |A_{1j}|$ with alternating sign where $a_{11}, \ldots, a_{1n}$ are the entries from the first row of A, and A_{1j} is created by removing the first row and j^{th} column of A. That is,

$$|A| = \det(A) = a_{11} |A_{11}| - a_{12} |A_{12}| + a_{13} |A_{13}| - \cdots + (-1)^{n+1} a_{1n} |A_{1n}|.$$

$$= \sum_{j=1}^{n} (-1)^{1+j} a_{1j} |A_{1j}|.$$

This is a <u>cofactor expansion of A</u> (along the first row).

<u>Definition by example:</u> "Let $A = \begin{pmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{pmatrix}$. Find $|A| = \det(A)$."

$$A_{11} = \begin{pmatrix} \cancel{1} & \cancel{5} & \cancel{0} \\ \cancel{2} & 4 & -1 \\ \cancel{0} & -2 & 0 \end{pmatrix} = \begin{pmatrix} 4 & -1 \\ -2 & 0 \end{pmatrix} \to |A_{11}| = -2$$

$$A_{12} = \begin{pmatrix} \cancel{1} & \cancel{5} & \cancel{0} \\ 2 & \cancel{4} & -1 \\ 0 & \cancel{-2} & 0 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ 0 & 0 \end{pmatrix} \to |A_{12}| = 0$$

$$A_{13} = \begin{pmatrix} \cancel{1} & \cancel{5} & \cancel{0} \\ 2 & 4 & \cancel{-1} \\ 0 & -2 & \cancel{0} \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 0 & -2 \end{pmatrix} \to |A_{13}| = -4$$

Then $\det(A) = 1 \cdot (-2) - 5(0) + 0(-4) = \boxed{-2}$.

Theorem: The determinant of $A = (a_{ij})_{n\times n}$ can be computed by a cofactor expansion across any row or column. That is,

$$|A| = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} |A_{ij}| \quad \text{(across the } i^{th} \text{ row)}$$

$$= \sum_{i=1}^{n} (-1)^{i+j} a_{ij} |A_{ij}| \quad \text{(across the } j^{th} \text{ column)}.$$

Remark: signs are determined by $Cb = ((-1)^{i+j})_{n\times n}$:

$$Cb = \begin{bmatrix} 1 & -1 & 1 & \cdots \\ -1 & 1 & -1 & \cdots \\ 1 & -1 & 1 & \\ \vdots & & \vdots & \ddots \end{bmatrix}_{n\times n}.$$

Ex: Same $A = \begin{pmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{pmatrix}$.

(a) Expand across 2^{nd} row:

$$|A| = -2\begin{vmatrix} 5 & 0 \\ -2 & 0 \end{vmatrix} + 4\begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix} - (-1)\begin{vmatrix} 1 & 5 \\ 0 & -2 \end{vmatrix}$$

$$= -2\cdot 0 + 4 \cdot 0 + 1\cdot(-2) = -2. \checkmark$$

(b) Expand dow 3rd col:

$$|A| = +0\begin{vmatrix} 2 & 4 \\ 0 & -2 \end{vmatrix} - (-1)\begin{vmatrix} 1 & 5 \\ 0 & -2 \end{vmatrix} + 0\begin{vmatrix} 1 & 5 \\ 2 & 4 \end{vmatrix}$$

$$= -2. \checkmark$$

Ex: ~~det~~ "Compute $|A|$ for

$$A = \begin{bmatrix} 3 & -7 & 8 & 9 & -6 \\ 0 & 2 & -5 & 7 & 3 \\ 0 & 0 & 1 & 5 & 0 \\ 0 & 0 & 2 & 4 & -1 \\ 0 & 0 & 0 & -2 & 0 \end{bmatrix}."$$

which col/row to expand:

Bad choices: 1^{st} row, 4^{th} col.

Good choices: 1^{st} col, 5^{th} row.

First col: $\cancel{|A| = 0|A_{15}| + 0|A_{25}|} +$

$$|A| = 3|A_{11}| - 0|A_{21}| + 0|A_{31}| - 0|A_{41}| + 0|A_{51}|.$$

$$= 3\begin{vmatrix} 2 & -5 & 7 & 3 \\ 0 & 1 & 5 & 0 \\ 0 & 2 & 4 & -1 \\ 0 & 0 & -2 & 0 \end{vmatrix} \overset{(1^{st}\text{ col})}{=} 3\cdot\left(2\begin{vmatrix} 1 & 5 & 0 \\ 2 & 4 & -1 \\ 0 & -2 & 0 \end{vmatrix}\right)$$

$$\overset{(3^{rd}\text{ row})}{=} 6\cdot\left(0|A_{31}| - (-2)\begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} + 0|A_{33}|\right)$$

$$= 6\cdot 2\begin{vmatrix} 1 & 0 \\ 2 & -1 \end{vmatrix} = \boxed{-12}$$

<u>Theorem</u>: If A is a triangular matrix, then $|A|$ is the product of the digonal entries of A.

<u>Proof</u>: For $n=1$ ... this is a # ... $\det(1\times1) = $ entry.

(Induction/ Domino effect) Assume the theorem in the n^{th} case.

Consider the $n+1^{st}$ case:

$$\det(A) = \begin{vmatrix} a_{11} & * & \cdots & * \\ 0 & a_{22} & & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{(n+1)(n+1)} \end{vmatrix}$$

(1st row/col)

$$= a_{11} \begin{vmatrix} a_{22} & * & \cdots & * \\ 0 & a_{33} & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{(n+1)(n+1)} \end{vmatrix}$$

$$= a_{11} \left(a_{22}\, a_{33} \cdots a_{(n+1)(n+1)} \right)$$

∎

Properties of Determinants

Lay - 3.2 Strang - 5.1

Theorem: Let $A \in M_{n\times n}(\mathbb{R})$

(a) If a multiple of one row of A is added to another row of A to produce a matrix B, then $|A| = |B|$.

(b) If two rows of A are interchanged to produce B, then $|A| = -|B|$.

(c) If one row of A is multiplied by k to produce B, then $|B| = k|A|$.

Ex: "Compute $|A|$ for $A = \begin{pmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{pmatrix}$."

$$\begin{vmatrix} 1 & -4 & 2 \\ -2 & 8 & -9 \\ -1 & 7 & 0 \end{vmatrix} = \begin{vmatrix} 1 & -4 & 2 \\ 0 & 0 & -5 \\ 0 & 3 & 2 \end{vmatrix} = - \begin{vmatrix} 1 & -4 & 2 \\ 0 & 3 & 2 \\ 0 & 0 & -5 \end{vmatrix}$$

$$= -1\left(1 \cdot 3 \cdot (-5)\right) = \boxed{15}.$$

Warm-up:

① $\begin{vmatrix} 3 & -1 & 5 \\ 0 & -4 & 3 \\ 0 & 0 & 2 \end{vmatrix} = ?$ ② $\begin{vmatrix} 3 & 0 & 0 \\ 2 & -4 & 0 \\ 9 & 0 & 2 \end{vmatrix} = ?$

③ ... $\begin{vmatrix} 0 & 0 & 2 \\ 0 & -4 & 9 \\ 3 & 2 & 6 \end{vmatrix} = ?$

---

①, ② -24

③ $R1 \leftrightarrow R3 \Rightarrow \begin{vmatrix} 3 & 2 & 6 \\ 0 & -4 & 9 \\ 0 & 0 & 2 \end{vmatrix} = -\begin{vmatrix} 0 & 0 & 2 \\ 0 & -4 & 9 \\ 3 & 2 & 6 \end{vmatrix} = -(-24)$
$= \boxed{24}$

From last time:

① Det(up./low. tria.) = product of diagonal entries.

② $|A| = \_\_\_ \, |REF(A)|$ ← hope this is triangular

2a) Row Replacement: Does nothing to $|A|$.

2b) Interchange two rows: Multiplies $|A|$ by -1.

2c) Scale by scalar k: Then $|B| = k|A|$.

$A \underset{R_i \to kR_i}{\sim} B$

Ex: "Compute $|A|$ ($=\det(A)$) for $A = \begin{pmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{pmatrix}$."

$$\begin{vmatrix} 2 & -8 & 6 & 8 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{vmatrix} \xrightarrow{R1 \leftrightarrow R4} - \begin{vmatrix} 1 & -4 & 0 & 6 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 2 & -8 & 6 & 8 \end{vmatrix} \xrightarrow{RR's} - \begin{vmatrix} 1 & -4 & 0 & 6 \\ 0 & 3 & 5 & -8 \\ 0 & -12 & 1 & 16 \\ 0 & 0 & 6 & -4 \end{vmatrix}$$

RR's

$$- \begin{vmatrix} 1 & -4 & 0 & 6 \\ 0 & 3 & 5 & -8 \\ 0 & 0 & 21 & -8 \\ 0 & 0 & 6 & -4 \end{vmatrix} \xrightarrow{R3 \to \frac{6}{21} R3} -\frac{21}{6} \begin{vmatrix} 1 & -4 & 0 & 6 \\ 0 & 3 & 5 & -8 \\ 0 & 0 & 6 & -48/21 \\ 0 & 0 & 6 & -4 \end{vmatrix} \to \begin{vmatrix} 1 & -4 & 0 & 6 \\ 0 & 3 & 5 & -8 \\ 0 & 0 & 6 & -48/21 \\ 0 & 0 & 0 & -36/21 \end{vmatrix}$$

$$-4 + 48/21 = \frac{-84+48}{21} = \frac{-36}{21}$$

$R1 \to \frac{1}{2} R1$

$$= 2 \begin{vmatrix} 1 & -4 & 3 & 4 \\ 3 & -9 & 5 & 10 \\ -3 & 0 & 1 & -2 \\ 1 & -4 & 0 & 6 \end{vmatrix} = 2 \begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & -12 & 10 & 10 \\ 0 & 0 & -3 & 2 \end{vmatrix} = 2 \begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & -6 & 2 \\ 0 & 0 & -3 & 2 \end{vmatrix}$$

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

$$= 2 \cdot (1 \cdot 3 \cdot ([illegible]) \cdot -2)$$

$$= 2 \cdot (-2) \begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & -3 & 2 \end{vmatrix}$$

$$= 2 \cdot (-2) \begin{vmatrix} 1 & -4 & 3 & 4 \\ 0 & 3 & -4 & -2 \\ 0 & 0 & 3 & -1 \\ 0 & 0 & 0 & 1 \end{vmatrix} = \boxed{-36}$$

Remark: RREF is $O(n^3)$, cofactors are $O(n!)$

Theorem: For $A \in M_{n \times n}(\mathbb{R})$ such that $A \sim U \in M_{n \times n}(\mathbb{R})$ (up. tri.),

then $$|A| = \begin{cases} 0 & , \underline{A^{-1} \text{ DNE}} \\ (-1)^r |U| & , A^{-1} \text{ exists} \end{cases}$$

where r is # of row exchange done to take A to U <u>without</u> scaling any rows!

Corollary: A is invertible iff $|A| \neq 0$.

Corollary: If one of the columns of A is free, then A^{-1} DNE.

Corollary: If the columns of A are lin. dep., then A^{-1} DNE.

$\vdots$

Get a corollary for each condition in the Inv. Mat. Theorem.

Theorem: If $A \in M_{n\times n}(\mathbb{R})$, $|A^T| = |A|$.

Why? Bc. this is the same as expanding down a column/row rather than a row/column.

Theorem: If $A, B \in M_{n\times n}(\mathbb{R})$, then $|AB| = |A||B|$.

(**)

Why? A can be thought to act as a product of elem. row op's on B. So $A = E_n E_{n-1} \cdots E_1 I_n$.

Then A row-reduces B as

$$
\begin{aligned}
|E_n E_{n-1} \cdots E_1 B| &= |E_n||E_{n-1} \cdots E_1 B| \\
&= |E_n| \cdots |E_1|\,|B| \\
&= |E_n E_{n-1} \cdots E_1|\,|B| \\
&= |A||B|.
\end{aligned}
$$

Check: $A = \begin{bmatrix} 6 & 1 \\ 3 & 2 \end{bmatrix}$, $B = \begin{bmatrix} 4 & 3 \\ 1 & 2 \end{bmatrix}$.

$|A| = 9$, $|B| = 5$

$AB = \begin{bmatrix} 25 & 20 \\ 14 & 13 \end{bmatrix} \rightarrow |AB| = 325 - 280 = 45 = 9 \cdot 5$.

Chapter 5 - Eigenvalues and Eigenvectors

Eigenvectors and Eigenvalues

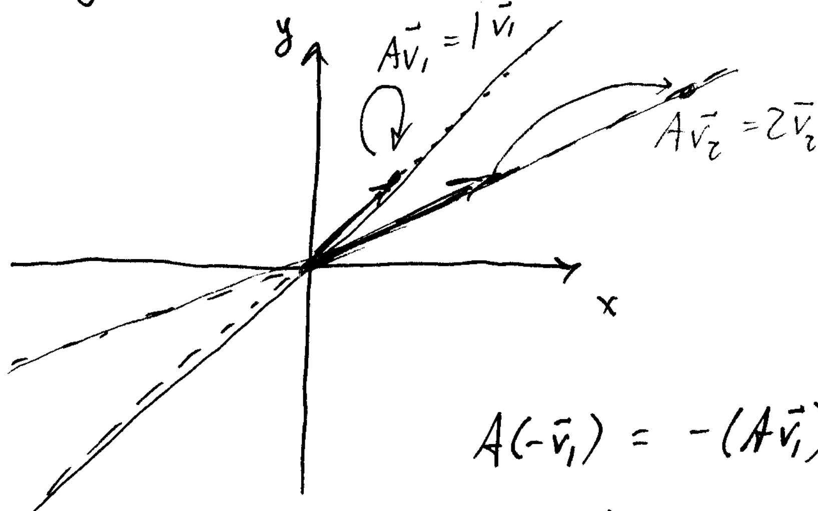
Lay - 5.1 Strang - 6.1

Motivating Examples: Let $A = \begin{bmatrix} 3 & -2 \\ 1 & 0 \end{bmatrix}$, $\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

Consider that $A\vec{v}_1 = \begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 1 \cdot \vec{v}_1$.

$A\vec{v}_2 = \begin{pmatrix} 3 & -2 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix} = 2 \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} = 2 \cdot \vec{v}_2$.

So mult. by A just stretches $\vec{v}_1$ and $\vec{v}_2$.



$A(-\vec{v}_1) = -(A\vec{v}_1) = -(\vec{v}_1)$

$A(-5\vec{v}_2) = -5(A\vec{v}_2) = -5(2\vec{v}_2)$
$= -10\,\vec{v}_2$.

> **Eigenvector** - A direction where mat. mult. is really just scalar mult.

Def: An *eigenvector* of $A \in M_{n\times n}(\mathbb{R})$ is a nonzero vector $\vec{x}$ such that $A\vec{x} = \lambda \vec{x}$ for some scalar λ $(\in \mathbb{C})$.

A scalar λ is called an *eigenvalue* of A if there is a nontrivial solution $\vec{x}$ of $A\vec{x} = \lambda\vec{x}$. If we have $A\vec{x} = \lambda\vec{x}$, then denote $\vec{x} = \vec{v}_\lambda$ to be an *eigenvector associated to λ*.

Ex: $A = \begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix}$, $\vec{u} = \begin{pmatrix} 6 \\ -5 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} 3 \\ -2 \end{pmatrix}$

$$A\vec{u} = \begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix}\begin{pmatrix} 6 \\ -5 \end{pmatrix} = \begin{pmatrix} -24 \\ 20 \end{pmatrix} = -4\begin{pmatrix} 6 \\ -5 \end{pmatrix} = -4\vec{u}. \checkmark$$

$$A\vec{v} = \begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix}\begin{pmatrix} 3 \\ -2 \end{pmatrix} = \begin{pmatrix} -9 \\ 11 \end{pmatrix} \neq \lambda\begin{pmatrix} 3 \\ -2 \end{pmatrix} \text{ for any } \lambda.$$

Aside: $A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
$\lambda = \pm i$

$\vec{u}$ is an evec for A w/ eval $\lambda = -4$.

$\vec{v}$ is *not* an evec for A.

Ex: "Show that 7 is another e.val. for $A = \begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix}$."

Show $A\vec{x} = 7\vec{x}$ has a nontrivial solution $\vec{x}$.

$$A\vec{x} - 7\vec{x} = \vec{0} \rightarrow A\vec{x} - 7I_2\vec{x} = \vec{0}$$

$$\cancel{(A-7)\vec{x} = \vec{0}} \qquad (A - 7I_2)\vec{x} = 0$$

not matrix algebra!

I.e. $\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} - 7$ ... wuh??

Need $\mathrm{Null}(A - 7I) \neq \{\vec{0}\}$.

To get an eigenvector for 7, take any nonzero vec. from

Compute $\begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix} - 7\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1-7 & 6 \\ 5 & 2-7 \end{pmatrix}$.

Solve $\left[\begin{array}{cc|c} -6 & 6 & 0 \\ 5 & -5 & 0 \end{array}\right] \leftrightarrow (A-7I)\vec{x} = \vec{0}$.

$$\sim \left[\begin{array}{cc|c} -1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] \rightarrow \begin{cases} x_1 = x_2 \\ x_2 \text{ free} \end{cases}$$

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_2 \\ x_2 \end{pmatrix} = x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

x_2 free

eigenvector

Check: $\begin{pmatrix} 1 & 6 \\ 5 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 7 \end{pmatrix} = 7\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. ✓

There can be more than one direction/line of evec's.

Makes sense – null spaces can have many dimensions.

Ex: "Let $A = \begin{bmatrix} 4 & -1 & 6 \\ 2 & 1 & 6 \\ 2 & -1 & 8 \end{bmatrix}$. $\lambda = 2$ is an eval. Find a basis of eigenvectors for $\lambda = 2$."

$$\mathrm{Null}(A-2I) = \mathrm{Null}\begin{pmatrix} 2 & -1 & 6 \\ 2 & -1 & 6 \\ 2 & -1 & 6 \end{pmatrix} \rightarrow A - 2I \sim \left[\begin{array}{ccc|c} 2 & -1 & 6 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

So $\begin{cases} 2x_1 = +x_2 - 6x_3 \\ x_2, x_3 \text{ free} \end{cases} \rightarrow \vec{x} = x_2\begin{pmatrix} 1/2 \\ 1 \\ 0 \end{pmatrix} + x_3\begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}$

Could also take $\left\{\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} -3 \\ 0 \\ 1 \end{pmatrix}\right\}$. evec basis.

Def: $\mathrm{Null}(A-\lambda I)$ is called the eigenspace of the eval λ.

Theorem: The evals of a triangular matrix are the entries along its diagonal.

Pf: $A = \begin{bmatrix} a_{11} & * & \cdots & * \\ & a_{22} & \cdots & * \\ & & \ddots & \vdots \\ & & & a_{nn} \end{bmatrix}$, So $A - \lambda I = \begin{pmatrix} a_{11}-\lambda & * & \cdots & * \\ & a_{22}-\lambda & \cdots & * \\ & & \ddots & \vdots \\ & & & a_{nn}-\lambda \end{pmatrix}$.

Setting $\lambda = a_{ii}$ for any $1 \le i \le n$ creates a free column for that i^{th} column.

$\Rightarrow \text{Null}(A - a_{ii} I) \neq \{\vec{0}\}$.

Ex: $A = \begin{pmatrix} 3 & 6 & -8 \\ 0 & 0 & 6 \\ 0 & 0 & 2 \end{pmatrix}$ $\quad \lambda(A) = \{3, 0, 2\}$. eval's of A

$B = \begin{pmatrix} 4 & 0 & 0 \\ -2 & 1 & 0 \\ 5 & 3 & 4 \end{pmatrix}$ $\quad \lambda(B) = \{4, 1, 4\} = \{4, 1\}$. eval's of B.

repeated

Question: What if $\lambda = 0$ is an eval for A? What is $\text{Null}(A)$?

$(A - 0I)\vec{x} = \vec{0}$ ... Eigenspace for 0 is $\text{Null}(A)$...

$\Rightarrow A^{-1}$ DNE. since $\text{Null}(A) \neq \{\vec{0}\}$.

Theorem: If $\vec{v}_1, \ldots, \vec{v}_r$ are evec's that correspond to <u>distinct</u> eval's $\lambda_1, \lambda_2, \ldots, \lambda_r$ of $A \in M_{n \times n}(\mathbb{R})$, then the set $S = \{\vec{v}_1, \ldots, \vec{v}_r\}$ is lin. ind.

Pf: (contradiction) Suppose S is lin. dep. Then because $\vec{v}_1 \neq \vec{0}$, Theorem from §1.7 tells us that

$$(*) \quad \vec{v}_{p+1} = c_1 \vec{v}_1 + \cdots + c_p \vec{v}_p$$

for some $p \geq 1$. Let it be the smallest index p that achieves this property.

(nontrivial $\Rightarrow$ at least one $c_I \neq 0$).

Consider $A(\vec{v}_{p+1}) = A(c_1 \vec{v}_1 + \cdots + c_p \vec{v}_p)$

$$= c_1 (A\vec{v}_1) + \cdots + c_p (A\vec{v}_p)$$

$$(\#) \quad \lambda_{p+1} \vec{v}_{p+1} = c_1 (\lambda_1 \vec{v}_1) + \cdots + c_p (\lambda_p A\vec{v}_p).$$

But from $(*)$, $\lambda_{p+1} \vec{v}_{p+1} = \lambda_{p+1} (c_1 \vec{v}_1 + \cdots + c_p \vec{v}_p)$.

So ... $\vec{0} = \lambda_{p+1} \vec{v}_{p+1} - \lambda_{p+1} \vec{v}_{p+1}$

$$= (*) - (\#)$$

$$\vec{0} = c_1 (\lambda_{p+1} - \lambda_1) \vec{v}_1 + \cdots + c_p (\lambda_{p+1} - \lambda_p) \vec{v}_p.$$

nonzero $\lambda_{p+1} - \lambda_i$.

$\Rightarrow c_1 = c_2 = \cdots = c_p = 0.$

But $c_I \neq 0$ ... Contradiction! ∎

The Characteristic Equation

Lay - 5.2 Strang - 6.2

Remark: Once we have eval's, we can do a lot w/ their evecs... But we haven't shown how to get the eval's in the first place...

Motivating Ex: Let $A = \begin{pmatrix} 2 & 3 \\ 3 & -6 \end{pmatrix}$.

Want all λ such that $A - \lambda I$ has a nontrivial null space.

When does this occur? When $\det(A - \lambda I) = 0$... since $(A-\lambda I)^{-1}$ doesn't exist here.

> **For evals:**
> Compute $|A - \lambda I|$ and set it to zero!

E.g. $|A - \lambda I| = \begin{vmatrix} 2-\lambda & 3 \\ 3 & -6-\lambda \end{vmatrix} = (2-\lambda)(-6-\lambda) - 9 = 0.$

$$\lambda^2 + 4\lambda - 21 = 0$$

$$(\lambda+7)(\lambda-3) = 0$$

$$\boxed{\lambda = 3, -7.}$$

evals for A.

Def: The polynomial $|A-\lambda I| = p(\lambda)$ is called the ~~char~~ ~~characteristic~~ characteristic polynomial of A. The zeroes of $p(\lambda)$ are the eval's of A, and $p(\lambda)=0$ is called the characteristic equation of A.

Ex: "Find the characteristic polynomial $p(\lambda)$ for $A = \begin{pmatrix} 5 & -2 & 6 & -1 \\ 0 & 3 & -8 & 0 \\ 0 & 0 & 5 & 4 \\ 0 & 0 & 0 & 1 \end{pmatrix}$."

Notice, $A-\lambda I$ is up. tri.

$$|A-\lambda I| = \left|\begin{pmatrix} 5-\lambda & * & * & * \\ & 3-\lambda & * & * \\ & & 5-\lambda & * \\ & & & 1-\lambda \end{pmatrix}\right| = \boxed{(5-\lambda)^2(3-\lambda)(1-\lambda)}$$

Observe: $p(\lambda)$ can always be factored (in $\mathbb{C}$) as

$$p(\lambda) = \prod_{i=1}^{k} (\lambda_i - \lambda)^{m_i} \ \ldots \text{ what is } m_i?$$

Def: In the above factorization, m_i is a positive integer called the algebraic multiplicity of the eval λ_i.

Ex: $p(\lambda) = (5-\lambda)^2(3-\lambda)(1-\lambda)\ldots$

eval	Alg. Mult.
$3=\lambda_1$	1
$1=\lambda_2$	1
$5=\lambda_3$	2

NOTICE!! HW 5 due 9/9 (Wed. of week 6).

Warm-up: ① "Define eigenvalue and eigenvector for ~~a~~ $A \in M_{n\times n}(\mathbb{R})$."

② "What are the ~~ ~~ evals for

$$A = \begin{bmatrix} 1 & 3 \\ 0 & 2 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 3 & 2 \end{bmatrix}, \; \& \; C = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} ?"$$

---

① Evec: $\bar{x}$ is an evec for A if there exists a scalar λ such that $A\vec{x} = \lambda \vec{x}$.

λ is an eval for A if there exists a ~~ ~~ nonzero vector $\vec{x}$ such that $A\vec{x} = \lambda \bar{x}$.

② $\boxed{\lambda(A) = \{1, 2\} = \lambda(B)}$

$$|C - \lambda I| = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = 0$$

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

$$\lambda = \pm i = \boxed{\pm\sqrt{-1}}.$$

Similar Matrices

Clash of Notation: $A \sim B$ can be meant as "row equivalent" or "similar."

Def: Given $A, B \in M_{n \times n}(\mathbb{C})$, if there is an invertible P such that

$$A = PBP^{-1} \text{ or } P^{-1}BP,$$

then we say A is <u>similar</u> to B (denoted $A \sim B$).

Remark: Can also write $AP = PB$ or $PA = BP$.

"commutative diagram"

$$\begin{tikzcd} V \arrow[r, "A"] \arrow[d, "P"'] & V \arrow[d, "P"] \\ V \arrow[r, "B"'] & V \end{tikzcd}$$

$$A = P^{-1}BP$$

Theorem: For $A \sim B$, $\lambda(A) = \lambda(B)$ and $p_A(\lambda) = p_B(\lambda)$.

$$\det(A - \lambda I) = \det(B - \lambda I).$$

Pf: Let $B = P^{-1}AP$ for some inv. P. Then

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

So $p_B(\lambda) = |B - \lambda I| = |P^{-1}(A - \lambda I)P|$

$$= |P^{-1}|\,|A - \lambda I|\,|P|$$

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

$$= p_A(\lambda).$$

$\Rightarrow$ evals are same. ∎

$$1 = |I| = |P^{-1}P|$$
$$= |P^{-1}|\,|P|$$
$$|P^{-1}| = 1/|P|$$

Diagonalization

Lay - 5.3 Strang - 6.2

Observation: Quick! What is $\begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$?

$$= \begin{pmatrix} 5 \\ 3 \end{pmatrix}$$

How about $\begin{pmatrix} 6 & 0 \\ 0 & -4 \end{pmatrix}\begin{pmatrix} 3 \\ -2 \end{pmatrix}$?

$$= \begin{pmatrix} 18 \\ 8 \end{pmatrix}.$$

Harder: $\begin{pmatrix} 1 & & & \\ & 3 & & \\ & & -4 & \\ & & & 2 \end{pmatrix}\begin{pmatrix} 2 \\ 0 \\ 5 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ -20 \\ 2 \end{pmatrix}.$

Moral: Diagonal matrices are nice for mult.

~~Quest~~ Second obs.: ~~What is~~ Given $\begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix} = D$, $D^2 = ?$

$$D^2 = \begin{pmatrix} 5^2 & 0 \\ 0 & 3^2 \end{pmatrix}$$

$$D^3 = \begin{pmatrix} 5^3 & 0 \\ 0 & 3^3 \end{pmatrix}$$

$$\vdots$$

$$D^k = \begin{pmatrix} 5^k & 0 \\ 0 & 3^k \end{pmatrix}.$$

Question: Can we find a way to compute A^k quickly as in the diagonal case?

A: It depends on A, whether or not A is similar to a diagonal D.

Ex: Here's a case where this is possible.

"Let $A = \begin{pmatrix} 7 & 2 \\ -4 & 1 \end{pmatrix}$. What is A^k?"

Let $P = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix}$, $D = \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix}$

$\left[P^{-1} = \frac{1}{-1} \begin{pmatrix} -2 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix} \right]$.

Check: $PDP^{-1} = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix}$

$= \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} 10 & 5 \\ -3 & -3 \end{pmatrix}$

$= \begin{pmatrix} 7 & 2 \\ -4 & 1 \end{pmatrix} = A$. So $A \sim D$.

Observation: $A^2 = (PDP^{-1})^2 = (PD\cancel{P^{-1})(P}DP^{-1})$

$= PD^2P^{-1}$

$A^3 = (PDP^{-1})(PDP^{-1})(PDP^{-1}) = PD^3P^{-1}$

$\vdots$

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

So $A^k = \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix}^k \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix}$

$$= \begin{pmatrix} 1 & 1 \\ -1 & -2 \end{pmatrix} \begin{pmatrix} 5^k & 0 \\ 0 & 3^k \end{pmatrix} \begin{pmatrix} 2 & 1 \\ -1 & -1 \end{pmatrix}$$

$$\boxed{A^k = \begin{pmatrix} 2\cdot 5^k - 3^k & 5^k - 3^k \\ 2\cdot 3^k - 2\cdot 5^k & 2\cdot 3^k - 5^k \end{pmatrix}}.$$

<u>Def</u>: Let $A \in M_{n\times n}(\mathbb{R})$. If there exists a diagonal matrix D such that A is similar to D, then we call A <u>diagonalizable</u>.

<u>Remark</u>: This means we need P and D where

$$A = P D P^{-1}.$$

<u>Theorem</u>: (Diagonalization) $A \in M_{n\times n}(\mathbb{R})$ is diagonalizable iff A has n linearly independent eigenvectors.

In fact, $A = PDP^{-1}$, w/ diagonal D, iff the columns of P are n lin. ind. evec's of A. In this case, the entries of D are eval's of A that correspond, respectively, to the evec's in P.

E.g. $A = \begin{pmatrix} 7 & 2 \\ -4 & 1 \end{pmatrix}$.

$$|A - \lambda I| = (7-\lambda)(1-\lambda) - (2)(-4)$$
$$= \lambda^2 - 8\lambda + 15 = (\lambda - 3)(\lambda - 5) = 0.$$
$$\lambda = 3, 5$$

are evals of A.

$\underline{\lambda = 3:}$ $(A - 3I)\vec{x} = \vec{0}$ solutions? <u>Find the Null(A-3I).</u>

So $[A-3I \mid \vec{0}] = \left[\begin{array}{cc|c} 4 & 2 & 0 \\ -4 & -2 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 2 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$

$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} x_2 \\ x_2 \end{pmatrix} \Leftarrow \begin{cases} x_1 = -\frac{1}{2} x_2 \\ x_2 = x_2 \end{cases} \Leftarrow \begin{cases} 2x_1 + x_2 = 0 \\ x_2 \text{ free} \end{cases}$

$= \frac{1}{2} x_2 \begin{pmatrix} -1 \\ 2 \end{pmatrix}$

evec. (actually, based on ex. before, take $\vec{v}_3 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$).

$\underline{\lambda = 5:}$ <u>Solve for Null(A-5I).</u>

$[A-5I \mid \vec{0}] = \left(\begin{array}{cc|c} 2 & 2 & 0 \\ -4 & -4 & 0 \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right]$

$\vec{v}_5 = \begin{pmatrix} +1 \\ -1 \end{pmatrix} \leftarrow \begin{cases} x_1 + x_2 = 0 \\ x_2 \text{ free} \end{cases}$

So $D = \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix} \Rightarrow P = [\vec{v}_5 \ \vec{v}_3] \Leftarrow$ ex.

or $D = \begin{pmatrix} 3 & 0 \\ 0 & 5 \end{pmatrix} \Rightarrow P = [\vec{v}_3 \ \vec{v}_5]$.

Then $A = PDP^{-1}$ ... Checked from before!

Proof of thm: observe $AP = A[\vec{v}_1 \cdots \vec{v}_n] = [A\vec{v}_1 \cdots A\vec{v}_n]$.

Also, $PD = [\vec{v}_1 \cdots \vec{v}_n] \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.

If P is comprised of evec's of A w/ $D = \operatorname{diag}(\text{eval's})$,

then $AP = [A\vec{v}_1 \cdots A\vec{v}_n] = [\lambda_1\vec{v}_1 \cdots \lambda_n\vec{v}_n]$

$= PD$.

P^{-1} exists iff $\{\vec{v}_1, \dots, \vec{v}_n\}$ is lin. ind.

So $AP = PD \Rightarrow A = PDP^{-1}$. ∎

Ex: "Let $A = \begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix}$. Is A diagonalizable? If so, decompose A into $A = PDP^{-1}$ for some invertible P and diagonal D."

① Find eval's: $|A - \lambda I| = 0 \xrightarrow{\text{some work}} (1-\lambda)(-2-\lambda)^2 = 0$

$\Rightarrow \lambda(A) = \{1, -2\}$.

$$|A - \lambda I| = \begin{vmatrix} 1-\lambda & 3 & 3 \\ -3 & -5-\lambda & -3 \\ 3 & 3 & 1-\lambda \end{vmatrix}$$

$$= (1-\lambda)\begin{vmatrix} -5-\lambda & -3 \\ 3 & 1-\lambda \end{vmatrix} - (3)\begin{vmatrix} -3 & -3 \\ 3 & 1-\lambda \end{vmatrix} + (3)\begin{vmatrix} -3 & -5-\lambda \\ 3 & 3 \end{vmatrix}$$

$\vdots$

② Get evec* basis for each $Null(A-\lambda I)$: $\lambda = 1$: $[A - 1I \mid \vec{0}]$

$$= \left(\begin{array}{ccc|c} 0 & 3 & 3 & 0 \\ -3 & -6 & -3 & 0 \\ 3 & 3 & 0 & 0 \end{array}\right) \to \vec{v}_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}.$$

$\lambda = -2$: $[A - (-2)I \mid \vec{0}] = \left(\begin{array}{ccc|c} 3 & 3 & 3 & 0 \\ -3 & -3 & -3 & 0 \\ 3 & 3 & 3 & 0 \end{array}\right) \to \vec{v}_{-2,1} = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$

$\vec{v}_{-2,2} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$

③ Make P matrix: $P = [\vec{v}_{-2,1} \ \ \vec{v}_1 \ \ \vec{v}_{-2,2}]$

④ ~~Create~~ Create D from P: $D = \text{diag}(-2, 1, -2)$

$$D = \begin{pmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

Then we claim $A = PDP^{-1}$ ...

Check: $P^{-1} = \ldots$ screw that! Check $PD = AP$ instead!

$$AP = \begin{pmatrix} 1 & 3 & 3 \\ -3 & -5 & -3 \\ 3 & 3 & 1 \end{pmatrix} \begin{pmatrix} -1 & 1 & -1 \\ 1 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 2 \\ -2 & -1 & 0 \\ 0 & 1 & -2 \end{pmatrix}. \checkmark$$

$$PD = \begin{pmatrix} -1 & 1 & -1 \\ 1 & -1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} -2 & & \\ & 1 & \\ & & -2 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 2 \\ -2 & -1 & 0 \\ 0 & 1 & -2 \end{pmatrix}. \checkmark$$

Yes, A is diagonalizable.

When doesn't this work?

Ex: $A = \begin{pmatrix} 2 & 4 & 3 \\ -4 & -6 & -3 \\ 3 & 3 & 1 \end{pmatrix}$.

① $|A - \lambda I| = 0 \iff (1-\lambda)(-2-\lambda)^2 = 0$

$\Downarrow$

$\lambda = 1, -2$

(alg) mult 1   mult 2.

② $Null(A - I) = span\left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\}$.

$Null(A + 2I) = span\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \right\}$... Can't get enough evec's in basis for E_{-2}.

Stop... this means A is not diagonalizable since we can't create an invertible P that contains evec's of A.

Def: The <u>geometric multiplicity</u> of an eval λ of A is the dimension of $Null(A - \lambda I)$.

Theorem: A matrix A is diagonalizable iff ~~the alg. mult~~ for each eval λ of A, alg. mult (λ) = geo. mult (λ),

Remark: $a(\lambda) \geq g(\lambda) \geq 1$ always!

A diagonalizable iff $a(\lambda_1) = g(\lambda_1), \dots, a(\lambda_k) = g(\lambda_k)$.

Theorem: $A \in M_{n\times n}(\mathbb{R})$ is diagonalizable if it has n distinct eval's.

Pf: Each eval has <u>at least</u> one evec.

So $\displaystyle\sum_{i=1}^{k} g(\lambda_i) \geq n$

But $\displaystyle\sum_{i=1}^{k} a(\lambda_i) = \underbrace{n \geq \sum_{i=1}^{k} g(\lambda_i) \geq n}_{\sum_{i=1}^{k} g(\lambda_i) = n}.$

$\displaystyle\sum_{i=1}^{k} g(\lambda_i) = n. \Rightarrow g(\lambda_i) = 1$ for all i.

λ_i distinct $\Rightarrow a(\lambda_i) = 1$ for all i.

So $a(\lambda_i) = 1 = g(\lambda_i)$ for all i.

Wake-up Call! "Diagonalize (if possible) $A = \begin{bmatrix} 1 & 3 \\ 1 & -1 \end{bmatrix}$."

① $\lambda = ?$: $|A - \lambda I| = p(\lambda) = (1-\lambda)(-1-\lambda) - (1)(3) = 0$

$$(\lambda^2 - 1) - 3 = 0$$
$$\lambda^2 - 4 = 0$$
$$\lambda^2 = 4$$
$$\lambda = \pm 2.$$

② Find E_2, E_{-2} : $E_2 = Null(A - 2I)$

$$\left[\begin{array}{cc|c} 1-2 & 3 & 0 \\ 1 & -1-2 & 0 \end{array}\right] = \left[\begin{array}{cc|c} -1 & 3 & 0 \\ 1 & -3 & 0 \end{array}\right]$$

$$\begin{cases} x_1 = 3x_2 \\ x_2 \text{ free} \end{cases} \to \vec{v}_2 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}.$$

$E_{-2} = Null(A + 2I) \quad \ldots \quad \vec{v}_{-2} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$

③ Create P : P

$$P_1 = [\vec{v}_2 \ \vec{v}_{-2}] = \begin{pmatrix} 3 & -1 \\ 1 & 1 \end{pmatrix} \qquad P_2 = [\vec{v}_{-2} \ \vec{v}_2] = \begin{pmatrix} -1 & 3 \\ 1 & 1 \end{pmatrix}.$$

④ Create D : $D_1 = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$ $\quad D_2 = \begin{pmatrix} -2 & 0 \\ 0 & 2 \end{pmatrix}.$

Dot Products, Orthogonality, Projections, & Gram-Schmidt

§ Inner product, Length, & Orthogonality

Lay - 6.1 Strang - 1.2, 4.1

Question: How "far" is 1 from t in $\mathbb{P}_2^f$?

Answer: Can't tell you until I create a notion of distance for that vector space. This is the idea of an inner product ... Our emphasis will be on the standard inner product on $\mathbb{R}^n$ - namely, dot product.

Def: Given $\vec{u} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$, $\vec{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$, define

$$\vec{u} \cdot \vec{v} := \vec{u}^T \vec{v} = \sum_{i=1}^{n} u_i \cdot v_i = u_1 v_1 + \cdots + u_n v_n .$$

This is the dot product on $\mathbb{R}^n$.

Ex: $\vec{u} = \begin{pmatrix} 2 \\ -5 \\ -1 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} 3 \\ 2 \\ -3 \end{pmatrix}$

$$\vec{u} \cdot \vec{v} = 2 \cdot 3 + (-5) \cdot 2 + (-1)(-3) = 6 - 10 + 3 = -1$$

$$\vec{v} \cdot \vec{u} = 3 \cdot 2 + 2(-5) + (-3)(-1) = -1 .$$

Theorem: $\vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^n$, $c \in \mathbb{R}$.

(a) $\vec{u} \cdot \vec{v} = \vec{v} \cdot \vec{u}$ (symmetric)

(b) $(\vec{u} + \vec{v}) \cdot \vec{w} = \vec{u} \cdot \vec{w} + \vec{v} \cdot \vec{w}$

(c) $(c\vec{u}) \cdot \vec{v} = c(\vec{u} \cdot \vec{v}) = \vec{u} \cdot (c\vec{v})$

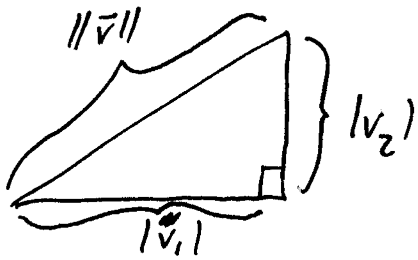
(d) $\vec{u} \cdot \vec{u} \geq 0$, $\vec{u} \cdot \vec{u} = 0$ iff $\vec{u} = \vec{0}$.

Def: The <u>length</u> of a vector $\vec{v} \in \mathbb{R}^n$ is given as

$$\|\vec{v}\| := \sqrt{\vec{v}\cdot\vec{v}} = \sqrt{v_1^2 + \cdots + v_n^2}$$

It is also called the <u>norm</u> of $\vec{v}$.

Picture: (in $\mathbb{R}^2$)



$$\|\vec{v}\|^2 = |v_1|^2 + |v_2|^2$$

(Pythagorean thm).

Notice: $\|c\vec{v}\| = |c|\,\|\vec{v}\| \qquad (\sqrt{c^2} = |c|)$.

Def: A vector w/ length 1 is called <u>unit</u>. We denote for any nonzero $\vec{v} \in \mathbb{R}^n$ its direction $\hat{v}$ as

$$\hat{v} = \frac{1}{\|\vec{v}\|}\vec{v} \qquad (\hat{v} \text{ is called the } \underline{\text{normalized}}\ \vec{v}).$$

$$\left(= \frac{\vec{v}}{\|\vec{v}\|} \right).$$

Ex: $\vec{u} = \begin{pmatrix} 2 \\ -5 \\ -1 \end{pmatrix}, \quad \vec{v} = \begin{pmatrix} 3 \\ 2 \\ -3 \end{pmatrix}, \quad \vec{w} = \begin{pmatrix} 1 \\ -2 \\ 2 \\ 0 \end{pmatrix}$

$\|\vec{u}\| = \sqrt{4+25+1} = \sqrt{30} \qquad \|\vec{v}\| = \sqrt{9+4+9} = \sqrt{22}$

$\|\vec{w}\| = \sqrt{1+4+4} = 3.$

$\Downarrow$

$$\hat{w} = \frac{1}{3}\vec{w} = \begin{pmatrix} 1/3 \\ -2/3 \\ 2/3 \\ 0 \end{pmatrix}.$$

Def: For $\vec{u}, \vec{v} \in \mathbb{R}^n$, the <u>distance between $\vec{u}$ and $\vec{v}$</u> ($dist(\vec{u},\vec{v})$) is the length of $\vec{u}-\vec{v}$. That is

$$dist(\vec{u},\vec{v}) = \|\vec{u}-\vec{v}\|.$$

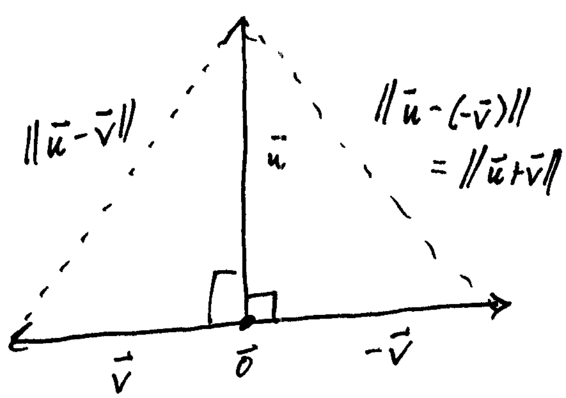
Ex: $\vec{u} = \begin{pmatrix} 2 \\ -5 \\ -1 \end{pmatrix}$, $\vec{v} = \begin{pmatrix} 3 \\ 2 \\ -3 \end{pmatrix}$.

$$dist(\vec{u},\vec{v}) = \left\| \begin{pmatrix} 2 \\ -5 \\ -1 \end{pmatrix} - \begin{pmatrix} 3 \\ 2 \\ -3 \end{pmatrix} \right\|$$

$$= \left\| \begin{pmatrix} -1 \\ -7 \\ 2 \end{pmatrix} \right\|$$

$$= \sqrt{1+49+4} = \boxed{3\sqrt{6}}.$$

Picture:



For "perpendicularity", $\|\vec{u}-\vec{v}\|^2 = \|\vec{u}+\vec{v}\|^2$.

Consider $\|\vec{u}+\vec{v}\|^2 = (\vec{u}+\vec{v})\cdot(\vec{u}+\vec{v}) = \vec{u}\cdot\vec{u} + 2\vec{u}\cdot\vec{v} + \vec{v}\cdot\vec{v}$.

$$= \|\vec{u}\|^2 + 2\vec{u}\cdot\vec{v} + \|\vec{v}\|^2$$

Similarly, $\|\vec{u}-\vec{v}\|^2 = (\vec{u}-\vec{v})\cdot(\vec{u}-\vec{v}) = \vec{u}\cdot\vec{u} - 2\vec{u}\cdot\vec{v} + \vec{v}\cdot\vec{v}$.

$$= \|\vec{u}\|^2 - 2\vec{u}\cdot\vec{v} + \|\vec{v}\|^2.$$

So $\|\vec{u}\|^2 + 2\vec{u}\cdot\vec{v} + \|\vec{v}\|^2 = \|\vec{u}\|^2 - 2\vec{u}\cdot\vec{v} + \|\vec{v}\|^2$

$$4\vec{u}\cdot\vec{v} = 0 \implies \vec{u}\cdot\vec{v} = 0$$

Def: Two vectors $\vec{u}, \vec{v} \in \mathbb{R}^n$ are <u>orthogonal</u> (perpendicular) if $\vec{u} \cdot \vec{v} = 0$.

Theorem: (Pythagorean) $\vec{u}, \vec{v} \in \mathbb{R}^n$ are orthogonal iff

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

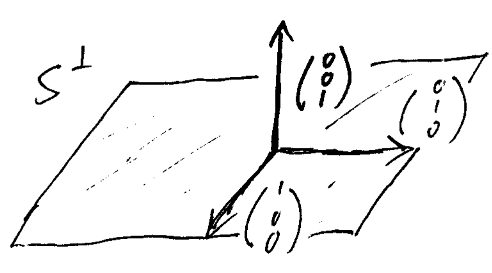
(law of cosines : add $-2\|\vec{u}\|\|\vec{v}\| \cos\theta$).

Def: Let $S \subseteq \mathbb{R}^n$ (just a nonempty subset) of vectors. We say that $\vec{v}$ is <u>orthogonal to S</u> if $\vec{v}$ is orthogonal to each element of S.

The set of all vectors orthogonal to S is called the <u>orthogonal complement of S</u> (denoted $S^\perp$) in $\mathbb{R}^n$.

Picture: Let $S = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \subseteq \mathbb{R}^3$.

Then $S^\perp$ is the xy-plane : $S^\perp = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$.



How do we know?

$\vec{x} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0$ for $\vec{x} \in S^\perp$.

So $\begin{pmatrix} x \\ y \\ z \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0$

$\{ 0x + 0y + z = 0$

$\begin{cases} z = 0 \\ x, y \text{ free} \end{cases}$

$$\vec{x} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = x \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

spanning set (basis) for $S^\perp$.

Let's see that again ... in slow motion:

Ex: Let $S = \left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -4 \\ 2 \\ 0 \end{pmatrix} \right\}$. ~~The~~ To find $S^\perp$, we need

$$\vec{x} \cdot \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} = 0 \quad \text{and} \quad \vec{x} \cdot \begin{pmatrix} -4 \\ 2 \\ 0 \end{pmatrix} = 0 .$$

Hence, $\begin{cases} x + y + 2z = 0 \\ -4x + 2y = 0 \end{cases} \rightarrow \left[\begin{array}{ccc|c} 1 & 1 & 2 & 0 \\ -4 & 2 & 0 & 0 \end{array}\right]$

$$\swarrow$$

$$\vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = z \begin{pmatrix} -4/3 \\ -2/3 \\ 1 \end{pmatrix} \leftarrow \begin{cases} x = -4/3 \, z \\ y = -2/3 \, z \\ z \text{ free} \end{cases}$$

So $S^\perp = \text{span}\left\{ \begin{pmatrix} -4/3 \\ -2/3 \\ 1 \end{pmatrix} \right\} = \text{span}\left\{ \begin{pmatrix} -4 \\ -2 \\ 3 \end{pmatrix} \right\}$.

Facts: Let $W \subseteq \mathbb{R}^n$ be a subspace of $\mathbb{R}^n$.

① $\vec{x} \in W^\perp$ iff $\vec{x} \perp \vec{v}_i$ for each $\vec{v}_i$ with $W = \text{span}\{\vec{v}_1, \dots, \vec{v}_p\}$

② $W^\perp \subseteq \mathbb{R}^n$ is a subspace.

③ $(W^\perp)^\perp = W$ ... note: $(S^\perp)^\perp \neq S$ for subsets of $\mathbb{R}^n$.

Ex: $S = \left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \subseteq \mathbb{R}^3$.

$S^\perp = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$.

$(S^\perp)^\perp = \text{span}\left\{ \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\} \neq S$.

Theorem: Let $A \in M_{m\times n}(\mathbb{R})$. Then

(**) $(\text{Row } A)^{\perp} = \text{Null}(A)$ and $(\text{Col } A)^{\perp} = \text{Null}(A^T)$.

$\subseteq \mathbb{R}^n$ $\quad$ $\text{Col}(A) \subseteq \mathbb{R}^m$ $\quad$ $\subseteq \mathbb{R}^m$

$\text{Row}(A) \subseteq \mathbb{R}^n$

Pf: $A\vec{x} = [\vec{a}_1 \cdots \vec{a}_n]\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{bmatrix} a_1^T \cdot \vec{x} \\ \vdots \\ a_n^T \cdot \vec{x} \end{bmatrix}$ ~~$= \begin{bmatrix} \vec{a}_1 \cdot \vec{x} \\ \vdots \\ \vec{a}_n \cdot \vec{x} \end{bmatrix}$~~.

$(\text{Col}(A))^{\perp}$:

So $A\vec{x} = \vec{0} \implies \vec{a}_i^T \cdot \vec{x} = 0$ for all $1 \le i \le n$.

Also, $\text{Row}(A) = \text{span}\{\vec{a}_1^T, \ldots, \vec{a}_n^T\}$.

$\Downarrow$

$\vec{x} \in$ ~~$(\text{Col } A)^T$~~.

$\implies$ ~~$\vec{x} \in \text{Null} = (\text{Row}(A))^{\perp}$~~.

$\vec{x} \in \text{Null}(A^T) = (R$

Pf:

$$A\vec{x} = \begin{bmatrix} \vec{a}_1^{\,T} \\ \vdots \\ \vec{a}_m^{\,T} \end{bmatrix} \vec{x} = \begin{bmatrix} \vec{a}_1 \vec{x} \\ \vdots \\ \vec{a}_m \vec{x} \end{bmatrix} = \begin{bmatrix} \vec{a}_1^{\,T} \cdot \vec{x} \\ \vdots \\ \vec{a}_m^{\,T} \cdot \vec{x} \end{bmatrix} \in \mathbb{R}^m.$$

$\vec{a}_i \in Row(A)$

$$A^T \vec{y} = \begin{bmatrix} \vec{b}_1^{\,T} \\ \vdots \\ \vec{b}_m^{\,T} \end{bmatrix} \vec{x} = \begin{bmatrix} \vec{b}_1 \vec{x} \\ \vdots \\ \vec{b}_n \vec{x} \end{bmatrix} = \begin{bmatrix} \vec{b}_1^{\,T} \cdot \vec{x} \\ \vdots \\ \vec{b}_n^{\,T} \cdot \vec{x} \end{bmatrix} \in \mathbb{R}^n.$$

$b_i^T \in Col(A)$.

So $A\vec{x} = \vec{0} \Rightarrow \vec{a}_i \cdot \vec{x} = 0$ for all $1 \le i \le m$.

So All rows of A are $\perp$ to $\vec{x}$. ✓

Then $(Row(A))^\perp = Null(A)$.

Similarly, $A^T \vec{y} = \vec{0} \Rightarrow \vec{b}_i \cdot \vec{y} = 0$ for all $1 \le i \le n$.

That is, all columns are $\perp$ to $\vec{x}$.

So $(Col(A))^\perp = Null(A^T)$. ∎

Remark: These 4 spaces ($Row\ A$, $Col\ A$, $Null\ A$, $Null\ A^T$) are the 4 subspaces of a matrix transformation as alluded to in Strang.

Orthogonal sets

Lay - 6.2 Strang - 4.2

Motivation:
1. Dot products between $\perp$ elements are zero.
2. Makes "projections" nicer.
3. Mimic rectangular $\mathbb{R}^n$ bases.
4. Decompose vectors easily.

Def: A set $\{\bar{u}_1, \dots, \bar{u}_p\} \subseteq \mathbb{R}^n$ is <u>orthogonal</u> if $\bar{u}_i \perp \bar{u}_j$ ~~...~~ for all $i \neq j$.

E.g. Sets like

$$S = \left\{ \underset{\bar{u}_1}{\underset{\shortparallel}{\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}}}, \underset{\bar{u}_2}{\underset{\shortparallel}{\begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}}} \right\} \subseteq \mathbb{R}^3$$

Check: $\bar{u}_1 \cdot \bar{u}_2$
$= 1 \cdot 0 + 2 \cdot 1 + (-1) \cdot 2$
$= 2 - 2 = 0.$ ✓

and

$$S_0 = \left\{ \underset{\bar{u}_1}{\underset{\shortparallel}{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}}, \underset{\bar{u}_2}{\underset{\shortparallel}{\begin{pmatrix} -1 \\ 1 \end{pmatrix}}}, \underset{\bar{u}_3}{\underset{\shortparallel}{\begin{pmatrix} 0 \\ 0 \end{pmatrix}}} \right\} \subseteq \mathbb{R}^2.$$

Check:
$\bar{u}_1 \cdot \bar{u}_2 = 1 \cdot (-1) + 1 \cdot 1 = 0$ ✓
$\bar{u}_1 \cdot \bar{u}_3 = \bar{u}_1 \cdot \vec{0} = 0$ ✓
$\bar{u}_2 \cdot \bar{u}_3 = \bar{u}_2 \cdot \vec{0} = 0$ ✓.

Theorem: If $S = \{\bar{u}_1, \dots, \bar{u}_p\} \subseteq \mathbb{R}^n$ is orthogonal and no $\bar{u}_i \neq \vec{0}$ for any i, then S is lin. ind. <u>and</u> is a basis for span(S).

Pf: Suppose $c_1 \bar{u}_1 + \cdots + c_p \bar{u}_p = \vec{0}$ for some scalars $c_1, \dots, c_p$.
Then consider $\bar{u}_1 \cdot \left(\sum_{i=1}^{p} c_i \bar{u}_i \right) = 0$ (since is $\vec{0}$).

$$\sum_{i=1}^{p} c_i (\bar{u}_1 \cdot \bar{u}_i) = 0 = c_1 \bar{u}_1 \cdot \bar{u}_1 + c_2 \overset{0}{\cancel{\bar{u}_1 \cdot \bar{u}_2}} + \cdots$$

$$\Rightarrow \quad c_1 \, (\vec{u}_1 \cdot \vec{u}_1) = 0$$

$$c_1 = 0 \qquad \cancel{\vec{u}_1 \cdot \vec{u}_1 = 0}.$$

Similarly,

$$\vec{u}_i \cdot \vec{u}_j = \begin{cases} 0, & i \neq j \\ \|u_i\|^2, & i = j. \end{cases}$$

$$\Rightarrow \quad c_i = 0 \text{ for all } 2 \leq i \leq p. \quad (1 \leq i \leq p)$$

<u>Def</u>: An <u>orthogonal basis</u> for a subspace $W \subseteq \mathbb{R}^n$ is a basis for W that is also an orthogonal set.

<u>Theorem</u>: Let $\{\vec{u}_1, \ldots, \vec{u}_p\} \subseteq W \subseteq \mathbb{R}^n$ be an orthogonal basis for W. For each $\vec{y} \in W$, there is a lin. comb. of u_i's that gives $\vec{y}$ as

$$\vec{y} = \sum_{i=1}^{p} \left(\frac{1}{\|\vec{u}_i\|^2} \vec{y} \cdot \vec{u}_i \right) \vec{u}_i$$

<u>Pf</u>: (This also gives uniqueness of the lin. comb. weights as $\{\vec{u}_1, \ldots, \vec{u}_p\}$ is a basis.)

Suppose $\vec{y} = \sum c_i \vec{u}_i$. Then $\vec{y} \cdot \vec{u}_1 = c_1 (\vec{u}_1 \cdot \vec{u}_1)$

$$c_1 = \frac{\vec{y} \cdot \vec{u}_1}{\vec{u}_1 \cdot \vec{u}_1} = \frac{\vec{y} \cdot \vec{u}_1}{\|\vec{u}_1\|^2}$$

Similar for all other c_i.

<u>Def</u>: $\dfrac{\vec{y} \cdot \vec{u}_i}{\|\vec{u}_i\|^2} \vec{u}_i$ is the <u>orthogonal projection</u> of $\vec{y}$ onto span $\{\vec{u}_i\}$.)

<u>Ex</u>: " $S = \left\{ \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} -1/2 \\ -2 \\ 7/2 \end{pmatrix} \right\}$ is orthogonal basis for $\mathbb{R}^3$.

Express $\bar{y} = \begin{pmatrix} 6 \\ 1 \\ -8 \end{pmatrix}$ in this basis."

$\underline{\vec{u}_1}$: $\bar{y} \cdot \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} = 6 \cdot 3 + 1 \cdot 1 + (-8) \cdot 1 = 19 - 8 = 11$. $\|\vec{u}_1\|^2 = 11$.

$\underline{\bar{u}_2}$: $\bar{y} \cdot \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} = -12$. $\|\vec{u}_2\|^2 = 6$.

$\underline{\vec{u}_3}$: $\bar{y} \cdot \begin{pmatrix} -1/2 \\ -2 \\ 7/2 \end{pmatrix} = -33$. $\|\vec{u}_3\|^2 = 33/2$

By formula/Theorem above,

$$c_1 = \frac{\bar{y} \cdot \vec{u}_1}{\|\vec{u}_1\|^2} = 11/11 = 1, \quad c_2 = -12/6 = -2, \quad c_3 = \frac{-33}{(33/2)} = -2.$$

Hence, $\bar{y} = 1\vec{u}_1 - 2\vec{u}_2 - 2\vec{u}_3$.

$$= \begin{pmatrix} 3 \\ 1 \\ 1 \end{pmatrix} - 2\begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} - 2\begin{pmatrix} -1/2 \\ -2 \\ 7/2 \end{pmatrix} = \begin{pmatrix} 3+2+1 \\ 1-4+4 \\ 1-2-7 \end{pmatrix} = \begin{pmatrix} 6 \\ 1 \\ -8 \end{pmatrix} \checkmark$$

<u>Remark</u>: No row reduction necessary!

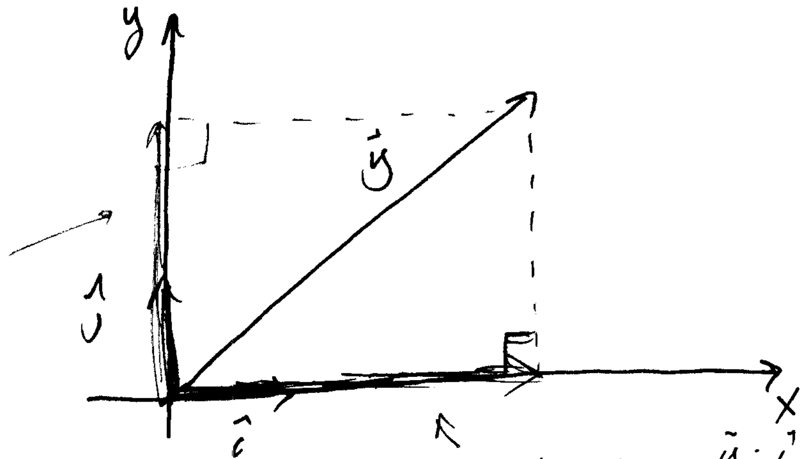
In fact, <u>orthogonal expansion</u> is $O(n^2)$!

Beats RREF, which is $O(n^3)$!

Question: So, what did we do exactly:

Picture:

$$\text{proj}_{\hat{j}}\, \vec{y} = \frac{\vec{y}\cdot\hat{j}}{\|\hat{j}\|^2}\,\hat{j}$$



$\{\hat{i}, \hat{j}\}$ is an orth. basis for $\mathbb{R}^2$

$$\text{proj}_{\hat{i}}\, \vec{y} = \frac{\vec{y}\cdot\hat{i}}{\|\hat{i}\|^2}\,\hat{i}$$

(☆☆) Remark: Dot products measure "how much of one vector is in the subspace spanned by the other."

Warm-up: ① Given $\{\vec{u}_1, \dots, \vec{u}_p\} \subseteq W \subseteq \mathbb{R}^n$, $W = \operatorname{span}\{\vec{u}_1, \dots, \vec{u}_p\}$ is an orthogonal basis for W, ~~what~~ for any $\vec{y} \in \mathbb{R}^n$ define

$$\operatorname{proj}_W \vec{y} = \dots$$

② What can you say about what set $\vec{y} - \operatorname{proj}_W \vec{y}$ lies in?

---

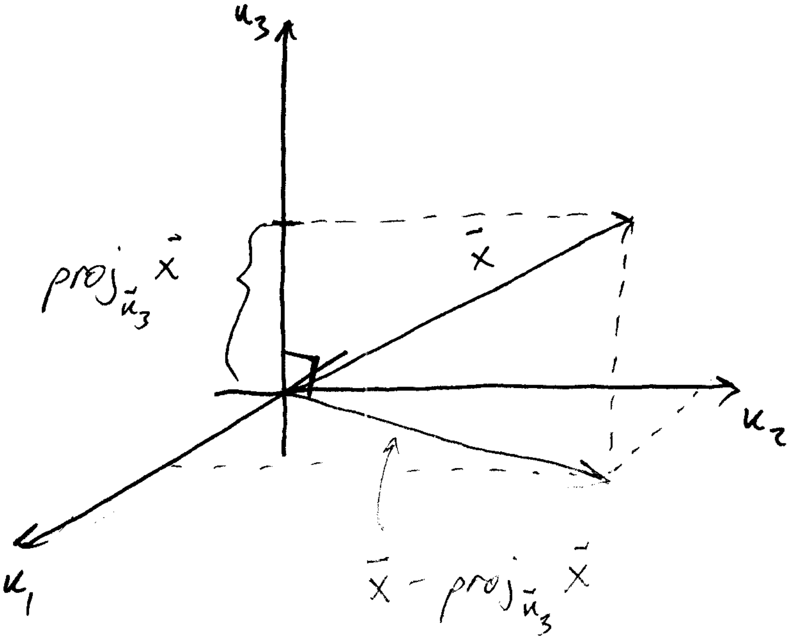
① $$\operatorname{proj}_W \vec{y} = \sum_{i=1}^{p} \underbrace{\frac{\vec{y} \cdot \vec{u}_i}{\|\vec{u}_i\|^2}}_{\substack{\text{how much of} \\ \vec{y} \text{ is in } \operatorname{span}\{\vec{u}_i\}.}} \vec{u}_i$$

② Tell me about $\vec{y} - \operatorname{proj}_W \vec{y}$. $\longrightarrow$ $(\vec{y} - \operatorname{proj}_W \vec{y}) \in W^\perp$.

$\vec{y} \in W \Rightarrow \vec{y} - \operatorname{proj}_W \vec{y} = \vec{0}$.

Check:
$$\vec{u}_i \cdot (\vec{y} - \operatorname{proj}_W \vec{y}) = \vec{u}_i \cdot \left(\vec{y} - \sum_{k=1}^{p} \frac{\vec{y} \cdot \vec{u}_k}{\|\vec{u}_k\|^2} \vec{u}_k\right)$$
$$= \vec{u}_i \cdot \vec{y} - \frac{\vec{y} \cdot \vec{u}_i}{\|\vec{u}_i\|^2} \vec{u}_i \cdot \vec{u}_i$$
$$= 0. \checkmark$$

Moral: Decompose $\vec{y}$ into "rectangular" components in W. Take away those W pieces, left w/ element of $W^\perp$.



<u>Remark</u>: Taking away a piece of $\bar{x}$ in the $\{\bar{u}_1, \bar{u}_2, \bar{u}_3\}$ orth. basis leaves a vector that is orthogonal to all the others.

<u>Ex</u>: "Let $\bar{y} = \begin{pmatrix} 7 \\ 6 \end{pmatrix}$, $\bar{u} = \begin{pmatrix} 4 \\ 2 \end{pmatrix}$. Write $\bar{y}$ as the sum of two vectors: one in span $\{\bar{u}\}$, the other in $(\text{span}\{\bar{u}\})^{\perp}$."

Consider $\bar{y} - proj_{\bar{u}}\bar{y}$ as orthogonal complement candidate.

So $\bar{y} - proj_{\bar{u}}\bar{y} = \begin{pmatrix} 7 \\ 6 \end{pmatrix} - \frac{40}{20}\begin{pmatrix} 4 \\ 2 \end{pmatrix} = \begin{pmatrix} 7 \\ 6 \end{pmatrix} - \begin{pmatrix} 8 \\ 4 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}$.

Notice we also get $proj_{\bar{u}}\bar{y}$

So $\begin{pmatrix} 7 \\ 6 \end{pmatrix} = \underbrace{\begin{pmatrix} 8 \\ 4 \end{pmatrix}}_{\text{in span}\{\bar{u}\}} + \underbrace{\begin{pmatrix} -1 \\ 2 \end{pmatrix}}_{\text{in the } \perp}$

IDEA: Let's make the computation easier.

How? Fix the length of all $\bar{u}_i$ in the orth basis.

Take $\|\bar{u}_i\|$ to 1 ... namely, replace

$\{\bar{u}_1, \dots, \bar{u}_p\}$ by $\{\hat{u}_1, \dots, \hat{u}_p\}$.

$\dfrac{\bar{u}_i}{\|\bar{u}_i\|} = \hat{u}_i$ has unit length.

Def: An orthogonal set $S = \{\bar{u}_1, \dots, \bar{u}_p\} \subseteq \mathbb{R}^n$ is <u>orthonormal</u> if in addition to mutual orthogonality between $\bar{u}_i$'s, each $\bar{u}_i$ is unit. If S is a basis for some $W \subseteq \mathbb{R}^n$, S is called an <u>orthonormal basis</u> for W.

E.g. " $\left\{\bar{u}_1 = \begin{pmatrix}1\\1\end{pmatrix}, \bar{u}_2 = \begin{pmatrix}-1\\+1\end{pmatrix}\right\}$. This a basis for $\mathbb{R}^n$.

Let $\bar{y} = \begin{pmatrix}5\\-4\end{pmatrix}$. what is $\text{proj}_{\hat{u}_1} \bar{y}$ and $\text{proj}_{\hat{u}_2} \bar{y}$? "

Create $\left\{\hat{u}_1 = \begin{pmatrix}1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix}, \hat{u}_2 = \begin{pmatrix}-1/\sqrt{2}\\1/\sqrt{2}\end{pmatrix}\right\}$.

Then $\text{proj}_{\hat{u}_1} \bar{y} = (\bar{y} \cdot \hat{u}_1)\hat{u}_1 = \frac{1}{\sqrt{2}}\hat{u}_1$.

$\text{proj}_{\hat{u}_2} \bar{y} = (\bar{y} \cdot \hat{u}_2)\hat{u}_2 = -\frac{9}{\sqrt{2}}\hat{u}_2$.

Thought: If $A = \begin{pmatrix} \vec{a}_1^{\,T} \\ \vec{a}_2^{\,T} \\ \vdots \\ \vec{a}_m^{\,T} \end{pmatrix}$, $B = [\vec{b}_1 \cdots \vec{b}_m]$,

then $$\underline{\underline{AB = (\vec{a}_i \cdot \vec{b}_j)}}$$

Theorem: An $m \times n$ matrix U has orthonormal columns iff
$$\underset{(n\times m)}{U^T}\,\underset{(m\times n)}{U} = \underset{(n\times n)}{I_n}.$$

"Pf": $\hat{u}_i \cdot \hat{u}_j = \delta_{ij} = \begin{cases} 0, & i \neq j \\ 1, & i = j \end{cases}$. ("Kronecker delta")

$$I_n = (\delta_{ij})_{n\times n}.$$

$\blacksquare$

Remark: Switches p.o.v. from algebra (lin. comb.'s of col's) to ~~alt~~ geometry (dot products of rows w/ col's).

Theorem: Let $U \in M_{m\times n}(\mathbb{R})$ w/ o.n. columns, and let $\vec{x}, \vec{y} \in \mathbb{R}^n$.

(a) $\|U\vec{x}\|^2 = (U\vec{x})^T(U\vec{x}) = x^T U^T U \vec{x} = \vec{x}^T\vec{x} = \|\vec{x}\|^2$.

Shortly, $\|U\vec{x}\| = \|\vec{x}\|$ (isometry).

(b) $(U\vec{x})\cdot(U\vec{y}) = \vec{x}\cdot\vec{y}$ (conformal - preserves angles).

(c) $(U\vec{x})\cdot(U\vec{y}) = 0$ iff $\vec{x}\cdot\vec{y} = 0$ (definite/nondegenerate).

Ex: "Verify that $U = \begin{bmatrix} 1/\sqrt{2} & 2/3 \\ 1/\sqrt{2} & -2/3 \\ 0 & 1/3 \end{bmatrix}$ has o.n. col's and that $\|U\vec{x}\| = \|\vec{x}\|$ for $\vec{x} = \begin{pmatrix} \sqrt{2} \\ 3 \end{pmatrix}$."

$\qquad\qquad\qquad\qquad\qquad\quad$ " $\hat{u}_1$ $\quad$ " $\hat{u}_2$

Check - $U^T U = \begin{pmatrix} \hat{u}_1^T \\ \hat{u}_2^T \end{pmatrix} \begin{pmatrix} \hat{u}_1 & \hat{u}_2 \end{pmatrix} = \begin{pmatrix} \hat{u}_1 \cdot \hat{u}_1 & \hat{u}_1 \cdot \hat{u}_2 \\ \hat{u}_2 \cdot \hat{u}_1 & \hat{u}_2 \cdot \hat{u}_2 \end{pmatrix}$.

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2 \checkmark.$$

$$\|\vec{x}\| = \sqrt{(\sqrt{2})^2 + 3^2} = \sqrt{11}$$

$$U\vec{x} = \sqrt{2}\,\hat{u}_1 + 3\hat{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \to \|U\vec{x}\|$$

$$= \sqrt{3^2 + 1 + 1}$$

$$= \sqrt{11}. \checkmark.$$

Orthogonal Projections

Lay - 6.3 $\quad$ Strang - 4.3

IDEA: Regrouping - $\vec{x} = \sum_{i=1}^{p} c_i \cdot \hat{u}_i$ for some onb $\{\hat{u}_1, \dots, \hat{u}_p\}$.

$$= \left(\sum_{i=1}^{k} c_i \cdot \hat{u}_i\right) + \left(\sum_{i=k+1}^{p} c_i \cdot \hat{u}_i\right).$$

$$= \vec{x}_1 + \vec{x}_2.$$

Recall: $\vec{x}_1 \in \text{span}\{\hat{u}_1, \dots, \hat{u}_k\}$, $\vec{x}_2 \in \text{span}\{\vec{u}_{k+1}, \dots, \vec{u}_p\}$.

So $\underline{\vec{x}_1 \perp \vec{x}_2}$. !!! ** :-)

Theorem (Orth. Decomp.): Let $W \subseteq \mathbb{R}^n$ be a subspace. Then each $\vec{y} \in \mathbb{R}^n$ can be written uniquely as $\vec{y} = \hat{y} + \vec{z}$ where $\hat{y} \in W$ and $z \in W^\perp$. In fact, if $\{\vec{u}_1, \dots, \vec{u}_p\}$ is an ~~ons~~ orth. basis for W, then

$$\hat{y} = \sum_{i=1}^{p} \frac{\vec{y} \cdot \vec{u}_i}{\|\vec{u}_i\|^2} \vec{u}_i$$

and $\vec{z} = \vec{y} - \hat{y}$.

(**Def:** $\hat{y}$ is the orthogonal projection of $\vec{y}$ onto W.)

Pf: Check $\vec{z} \in W^\perp$:

$$\vec{z} \cdot \vec{u}_k = (\vec{y} - \hat{y}) \cdot \vec{u}_k = \vec{y} \cdot \vec{u}_k - \vec{u}_k \cdot \sum_{i=1}^{p} \frac{\vec{y} \cdot \vec{u}_i}{\|\vec{u}_i\|^2} \vec{u}_i$$

$$= \vec{y} \cdot \vec{u}_k - \sum_{i=1}^{p} \frac{\vec{y} \cdot \vec{u}_i}{\|\vec{u}_i\|^2} \vec{u}_k \cdot \vec{u}_i$$

$$= \vec{y} \cdot \vec{u}_k - \frac{\vec{y} \cdot \vec{u}_k}{\|\vec{u}_k\|^2} \vec{u}_k \cdot \vec{u}_k = 0. \checkmark$$

Uniqueness of $\hat{y}$ and z:

Let $\vec{y} = \hat{y} + \vec{z} = \hat{y}_1 + \vec{z}_1$ (WTS $\hat{y} = \hat{y}_1$).

Then $\vec{y} - \vec{y} = (\hat{y} + \vec{z}) - (\hat{y}_1 + \vec{z}_1) = \vec{0}$.

$$\underbrace{\hat{y} - \hat{y}_1}_{\text{in } W} = \underbrace{\vec{z}_1 - \vec{z}}_{\text{in } W^\perp}.$$

$\Rightarrow$ $\hat{y} - \hat{y}_1 = \vec{0}$

and $\vec{z}_1 - \vec{z} = \vec{0}$. ∎

<u>Ex</u>: "Let $W = \operatorname{span}\left\{ \vec{u}_1 = \begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix}, \vec{u}_2 = \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} \right\}$, $\vec{y} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$.

Write $\vec{y} = \hat{y} + \vec{z}$ w/ $\hat{y} \in W$, $\vec{z} \in W^\perp$."

$$\hat{y} = \frac{\vec{y} \cdot \vec{u}_1}{\|\vec{u}_1\|^2} \vec{u}_1 + \frac{\vec{y} \cdot \vec{u}_2}{\|\vec{u}_2\|^2} \vec{u}_2$$

$$= \frac{9}{30} \vec{u}_1 + \frac{3}{6} \vec{u}_2 = \frac{3}{10} \begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} -2 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2/5 \\ 2 \\ 1/5 \end{pmatrix} \quad (*).$$

So $\vec{z} = \vec{y} - \hat{y} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} - \begin{pmatrix} -2/5 \\ 2 \\ 1/5 \end{pmatrix} = \begin{pmatrix} 7/5 \\ 0 \\ 14/5 \end{pmatrix} \quad (\#)$.

Then
$$\boxed{\vec{y} = \underbrace{\begin{pmatrix} -2/5 \\ 2 \\ 1/5 \end{pmatrix}}_{(*)} + \underbrace{\begin{pmatrix} 7/5 \\ 0 \\ 14/5 \end{pmatrix}}_{(\#)}.}$$

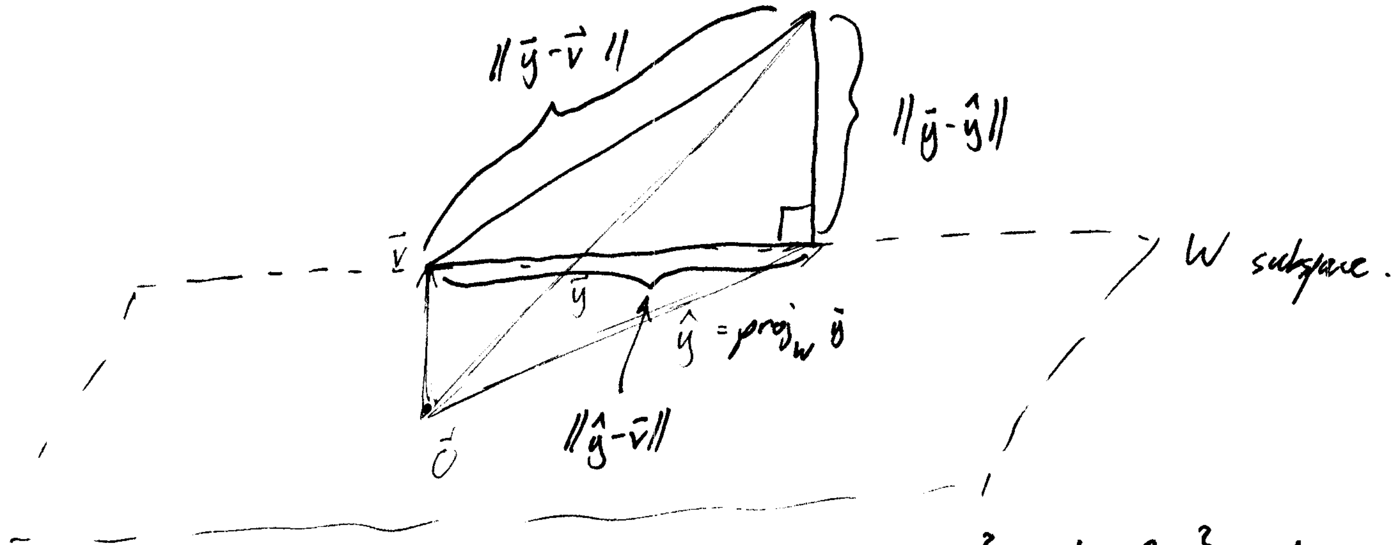
<u>Fact</u>: If $\vec{y} \in \operatorname{span}\{\vec{u}_1, \dots, \vec{u}_p\} = W$, then $\operatorname{proj}_W \vec{y} = \vec{y}$.

<u>Theorem</u>: Let $W \subseteq \mathbb{R}^n$, $\vec{y} \in \mathbb{R}^n$, $\hat{y} = \operatorname{proj}_W \vec{y}$. Then $\hat{y}$ is the closest point in W to $\vec{y}$. That is,

$$\|\vec{y} - \hat{y}\| < \|\vec{y} - \vec{v}\|$$

for all $\vec{v} \in W - \{\hat{y}\}$.

<u>Pf</u>: $\vec{y} - \hat{y}$ makes one leg of any other right triangle w/ leg $\vec{y} - \vec{v}$ ($\vec{v} \in W$, not $\hat{y}$).



So $\|\hat{y}-\vec{v}\|^2 + \|\vec{y}-\hat{y}\|^2 = \|\vec{y}-\vec{v}\|^2$.

~~Then~~ $\cancel{\|\vec{y}-\vec{v}\| = \sqrt{\|}}$

$$\|\vec{y}-\hat{y}\|^2 < \|\hat{y}-\vec{v}\|^2 + \|\vec{y}-\hat{y}\|^2 = \|\vec{y}-\vec{v}\|^2.$$

$$\Rightarrow \|\vec{y}-\hat{y}\| < \|\vec{y}-\vec{v}\|. \quad \blacksquare$$

<u>Def</u>: $\hat{y}$ is called a <u>"solution of least-squares."</u>

<u>Def</u>: In light of the previous thm, $\|\vec{y}-\hat{y}\|$ is called the <u>distance between $\hat{y}$ and W</u>, denoted $\text{dist}(\vec{y}, W)$.

<u>Ex</u>: From before, $\vec{y} = \begin{pmatrix} 7 \\ 6 \end{pmatrix}$, $W = \text{span}\left\{ \begin{pmatrix} 4 \\ 2 \end{pmatrix} \right\}$.

We showed $\hat{y} = \text{proj}_W \vec{y} = \begin{pmatrix} 8 \\ 4 \end{pmatrix}$. So

$$\text{dist}(\vec{y}, W) = \|\vec{y}-\hat{y}\| = \left\| \begin{pmatrix} 7 \\ 6 \end{pmatrix} - \begin{pmatrix} 8 \\ 4 \end{pmatrix} \right\| = \|\vec{z}\|$$

$$= \left\| \begin{pmatrix} -1 \\ 2 \end{pmatrix} \right\|$$

$$= \boxed{\sqrt{5}}$$

<u>Theorem</u>: If $\{\hat{u}_1, \dots, \hat{u}_p\} \subseteq W \subseteq \mathbb{R}^n$ is an onb for W, then

onb = orthonormal basis.

$$\text{proj}_W \vec{y} = \sum_{i=1}^{p} (\vec{y} \cdot \hat{u}_i)\, \hat{u}_i .$$

$$= U\left(U^T \vec{y}\right) \quad \text{if } U = [\hat{u}_1 \cdots \hat{u}_p].$$

lin. comb.'s

~~dot~~ dot prod's

<u>Def</u>: The matrix UU^T is called the **projection matrix** of $\mathbb{R}^n$ onto W.

The Gram-Schmidt Process

Lay - 6.4 Strang - 4.4

IDEA: Let $\beta = \{\vec{x}_1, \ldots, \vec{x}_p\} \subseteq W \subseteq \mathbb{R}^n$ be a basis for W.

Maybe we don't like β, or maybe we're more interested in the geometry of W than its algebra.

Perhaps an orth. basis would be better suited to our interests. Don't have to start from scratch to create this other basis.

$\star$ Use β to construct an orthogonal basis for W.

~~Def~~ This is Gram-Schmidt orthogonalization.

Theorem (Gram-Schmidt): Given a basis $\beta = \{\vec{x}_1, \ldots, \vec{x}_p\} \subseteq W \subseteq \mathbb{R}^n$ for W, define

$$\vec{v}_1 = \vec{x}_1, \quad \text{set } S_1 = \text{span}\{\vec{v}_1\}.$$

$$\vec{v}_2 = \vec{x}_2 - \text{proj}_{S_1} \vec{x}_2, \quad \text{set } S_2 = \text{span}\{\vec{v}_1, \vec{v}_2\}.$$

$$\vec{v}_3 = \vec{x}_3 - \text{proj}_{S_2} \vec{x}_3, \quad \text{set } S_3 = \text{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}.$$

$$\vdots$$

$$\vec{v}_p = \vec{x}_p - \text{proj}_{S_{p-1}} \vec{x}_p, \quad \text{set } S_p = \text{span}\{\vec{v}_1, \ldots, \vec{v}_p\} = W.$$

Then $\{\vec{v}_1, \ldots, \vec{v}_p\} = \gamma$ is an orthogonal basis for W.

Moreover, $\text{span}\{S_k) = \text{span}\{\vec{x}_1, \ldots, \vec{x}_k\}$.

Ex: "Let $\beta = \left\{ \vec{x}_1 = \begin{pmatrix} 1\\1\\1\\1 \end{pmatrix}, \vec{x}_2 = \begin{pmatrix} 0\\1\\1\\1 \end{pmatrix}, \vec{x}_3 = \begin{pmatrix} 0\\0\\1\\1 \end{pmatrix} \right\}$.

This is a basis for $W = \text{span}(\beta)$. Create an orth. basis for W using β."

① Let $\vec{v}_1 = \vec{x}_1$. Set $S_1 = \text{span}\{\vec{v}_1\}$.

② Get $\vec{v}_2$: $\vec{v}_2 = \vec{x}_2 - \text{proj}_{S_1} \vec{x}_2 = \vec{x}_2 - \dfrac{\vec{x}_2 \cdot \vec{v}_1}{\|\vec{v}_1\|^2} \vec{v}_1$

$$= \begin{pmatrix} 0\\1\\1\\1 \end{pmatrix} - \frac{3}{4}\begin{pmatrix} 1\\1\\1\\1 \end{pmatrix} = \begin{pmatrix} -3/4\\1/4\\1/4\\1/4 \end{pmatrix}.$$

Set $S_2 = \text{span}\{\vec{v}_1, \vec{v}_2\}$.

③ Get $\vec{v}_3$: $\vec{v}_3 = \vec{x}_3 - \text{proj}_{S_2} \vec{x}_3 = \vec{x}_3 - \left(\dfrac{\vec{x}_3 \cdot \vec{v}_1}{\|\vec{v}_1\|^2} \vec{v}_1 + \dfrac{\vec{x}_3 \cdot \vec{v}_2}{\|\vec{v}_2\|^2} \vec{v}_2 \right)$

$$= \begin{pmatrix} 0\\0\\1\\1 \end{pmatrix} - \left(\frac{1}{2}\begin{pmatrix} 1\\1\\1\\1 \end{pmatrix} + \frac{2}{3}\begin{pmatrix} -3/4\\1/4\\1/4\\1/4 \end{pmatrix} \right)$$

$$= \begin{pmatrix} 0\\-2/3\\1/3\\1/3 \end{pmatrix}. \quad S_3 = \text{span}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = W.$$

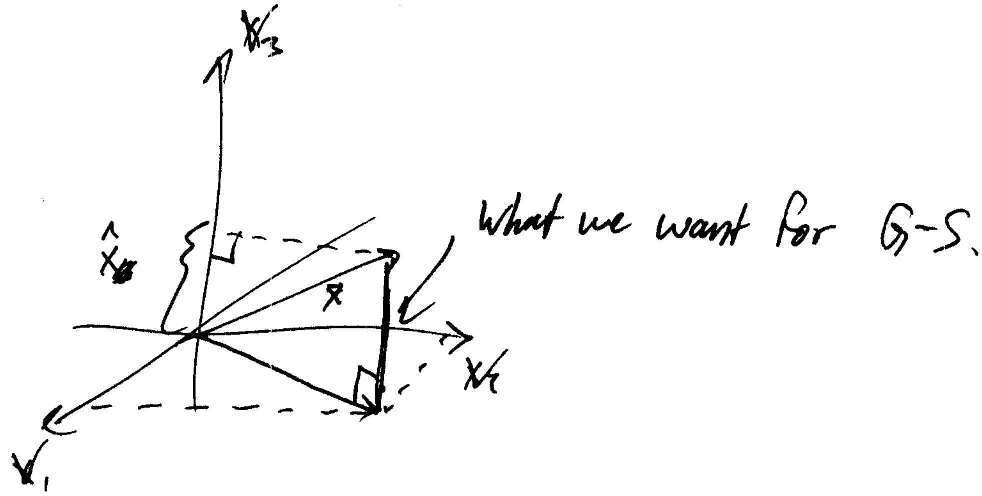
Check: Is $\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \gamma$ orthogonal?

$$\vec{v}_1 \cdot \vec{v}_2 = \begin{pmatrix} 1\\1\\1\\1 \end{pmatrix} \cdot \begin{pmatrix} -3/4\\1/4\\1/4\\1/4 \end{pmatrix} = -3/4 + 1/4 + 1/4 + 1/4 = 0$$

$$\vec{v}_1 \cdot \vec{v}_3 = \begin{pmatrix} 1\\1\\1\\1 \end{pmatrix} \cdot \begin{pmatrix} 0\\-2/3\\1/3\\1/3 \end{pmatrix} = 0, \quad \vec{v}_2 \cdot \vec{v}_3 = \begin{pmatrix} -3/4\\1/4\\1/4\\1/4 \end{pmatrix} \cdot \begin{pmatrix} 0\\-2/3\\1/3\\1/3 \end{pmatrix} = 0 \checkmark$$

Why did this work?

Recall:



Remarks: ① ~~We~~ This algorithm only gives us an orthogonal basis. But we can o.n.b. by normalizing each $\bar{v}_i$ to $\hat{v}_i$.

② We can simplify the computations by scaling the $\bar{v}_i$'s along the way.

E.g. $\bar{v}_2 = \bar{x}_2 - \text{proj}_{S_1} \bar{x}_2 = \begin{pmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix}$.

Use ~~insted~~ instead $\bar{v}_2 = \begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix}$

Announcements: ① Today is our last lecture day!

- Monday = Holiday
- Wednesday = Review session
- Friday = FINAL EXAM (full 2 hrs) plan for it!

② Class Evaluation

- Email from school already set : follow link
- Due (I think) next Wednesday (9/9).

③ FINAL : Covers everything up through 6.5

- 8 questions : 1 is definitions (from lecture)
  1 is proof
  1 is T-F. (from ch. 4,5, or 6)

Gram-Schmidt

Comes as a recursive algorithm:

Given: A basis β for $W = \operatorname{span}\beta \underset{\text{proper}}{\subseteq} \mathbb{R}^n$.

Output: An orthogonal basis γ for W such that if $\beta = \{\bar{x}_1, \dots, \bar{x}_p\}$, $\gamma = \{\bar{v}_1, \dots, \bar{v}_p\}$,

then $S_k = \operatorname{span}\{\bar{v}_1, \dots, \bar{v}_k\} = \operatorname{span}\{\bar{x}_1, \dots, \bar{x}_k\}$ $(1 \le k \le p)$.

(in particular, $S_p = W$).

Steps:

1. Set $\bar{v}_1 = \bar{x}_1$. Set $S_1 = \operatorname{span}\{\bar{v}_1\}$.
2. $(k{+}{+})$ Set $\bar{v}_k = \bar{x}_k - \operatorname{proj}_{S_{k-1}} \bar{x}_k$. Set $S_k = \operatorname{span}\{\bar{v}_1, \dots, \bar{v}_k\}$

   Repeat 2 until done. (until $k = p$)
3. Output $\gamma = \{\bar{v}_1, \dots, \bar{v}_p\}$.

   satisfy the above.

Ex: $\bar{v}_k = \bar{x}_k - \left(\frac{\bar{x}_k \cdot \bar{v}_1}{\|\bar{v}_1\|^2} \bar{v}_1 + \cdots + \frac{\bar{x}_k \cdot \bar{v}_{k-1}}{\|\bar{v}_{k-1}\|^2} \bar{v}_k \right).$

Ex: Let $\beta = \left\{ \bar{x}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \bar{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \bar{x}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right\}.$

① $\bar{v}_1 = \bar{x}_1$. $S_1 = \text{span}\{\bar{v}_1\}$.

② (a) Get $\bar{v}_2$: $\bar{v}_2 = \bar{x}_2 - \frac{\bar{x}_2 \cdot \bar{v}_1}{\|\bar{v}_1\|^2} \bar{v}_1$.

$$= \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} - \frac{3}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -3/4 \\ 1/4 \\ 1/4 \\ 1/4 \end{pmatrix}.$$

Use instead $\bar{v}_2 = \begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix}$.

(b) Get $\bar{v}_3$: $\bar{v}_3 = \bar{x}_3 - \left(\frac{\bar{x}_3 \cdot \bar{v}_1}{\|\bar{v}_1\|^2} \bar{v}_1 + \frac{\bar{x}_3 \cdot \bar{v}_2}{\|\bar{v}_2\|^2} \bar{v}_2 \right)$

$$= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \left(\frac{2}{4} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + \frac{2}{12} \begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix} \right)$$

$$= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \left(\begin{pmatrix} 1/2 \\ 1/2 \\ 1/2 \\ 1/2 \end{pmatrix} + \begin{pmatrix} -1/2 \\ 1/6 \\ 1/6 \\ 1/6 \end{pmatrix} \right)$$

$$= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 2/3 \\ 2/3 \\ 2/3 \end{pmatrix} = \begin{pmatrix} 0 \\ -2/3 \\ 1/3 \\ 1/3 \end{pmatrix}.$$

Use instead $\bar{v}_3 = \begin{pmatrix} 0 \\ -2 \\ 1 \\ 1 \end{pmatrix}$.

$\gamma = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} -3 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -2 \\ 1 \\ 1 \end{pmatrix} \right\}$ $\leftarrow$ orthogonal set spanning span β.

ONB for span $\beta = \left\{ \frac{1}{2}\begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \frac{1}{\sqrt{12}}\begin{pmatrix}-3\\1\\1\\1\end{pmatrix}, \frac{1}{\sqrt{6}}\begin{pmatrix}0\\-2\\1\\1\end{pmatrix} \right\}$.

<u>Recall</u>: LU factorization

Takes matrix $A \in M_{m\times n}(\mathbb{R})$,

spits out L, U such that

① L is low-tri. (encodes elem. row op's)

② U is up-tri. (encodes $REF(A)$)

③ $A = LU$.

Moral - factorizations encapsulate information (usually from an algorithm).

<u>Theorem</u>: (QR factorization) If $A \in M_{m\times n}(\mathbb{R})$ w/ lin. ind columns, then A can be factored as $A = QR$, where

① $Q \in M_{m\times n}(\mathbb{R})$ whose columns form an ONB for $Col(A)$ (i.e. $Q^TQ = I_n$)

and ② $R \in \cancel{M_{m\times n}(\mathbb{R})}\ M_{n\times n}(\mathbb{R})$ an invertible up-tri. matrix w/ positive entries on its diagonal.

<u>Pf</u>: Gram-Schmidt to get Q. $Q = [\hat{v}_1 \cdots \hat{v}_n]$,

$$A = [Q\bar{r}_1 \ \cdots \ Q\bar{r}_n] \Rightarrow r_{k,i} = \frac{\bar{x}_k \cdot \hat{v}_i}{\|\hat{v}_i\|^2} = \bar{x}_k \cdot \hat{v}_i$$

$\blacksquare$

Ex: "From before, let $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$. Find a QR factorization for A."

$$Q = \begin{bmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{bmatrix}.$$

ONB for $Col(A)$.

To find R, notice that $Q^T Q = I_3$.

Then $A = QR \Rightarrow Q^T A = Q^T Q R$

$$Q^T A = R$$

So
$$R = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ -3/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} & 1/\sqrt{12} \\ 0 & -2/\sqrt{6} & 1/\sqrt{6} & 1/\sqrt{6} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} \end{bmatrix}$$

Then
$$A = \underbrace{\begin{bmatrix} 1/2 & -3/\sqrt{12} & 0 \\ 1/2 & 1/\sqrt{12} & -2/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \\ 1/2 & 1/\sqrt{12} & 1/\sqrt{6} \end{bmatrix}}_{Q} \underbrace{\begin{bmatrix} 2 & 3/2 & 1 \\ 0 & 3/\sqrt{12} & 2/\sqrt{12} \\ 0 & 0 & 2/\sqrt{6} \end{bmatrix}}_{R} = \underbrace{\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}}_{A}.$$

Least-Squares Problems

Lay - 6.5 Strang - 4.3

Motivating Problem: When $A\bar{x}=\bar{b}$ has no solution, how "close" can we get to a solution ~~to~~ while staying in $Col(A)$?

What does "close" mean? $\|\cdot\| = \sqrt{\text{dot prod.}}$

Def: If $A \in M_{m\times n}(\mathbb{R})$, $\bar{b} \in \mathbb{R}^m$, a least-squares solution of $A\bar{x}=\bar{b}$ is an $\hat{x} \in \mathbb{R}^n$ such that

$$\|\bar{b}-A\hat{x}\| \leq \|\bar{b}-A\bar{x}\| \text{ for all } \bar{x}\in\mathbb{R}^n.$$

We're looking for $\hat{x}$ s.t. $\bar{b} = \underbrace{\hat{b}}_{``A\hat{x}"} + \bar{e}$ w/ $A\bar{x}=\bar{b}$.

$\bar{e}$ is typically called the error vector.

($\|\bar{e}\|$ is the error)

Observations: If $A\bar{x}=\bar{b}$ has no solution, $\bar{b} \notin Col(A)$.

Put $\bar{b}$ into $Col(A)$ by "casting a shadow."

$$\hat{b} = \text{proj}_{Col(A)}\bar{b} \in Col(A).$$

Then $A\hat{x}=\hat{b}$ has a solution.

This solution is as close to $\bar{b}$ as we can get in $Col(A)$.

Reframe: $(\bar{b} - \hat{b}) \perp \mathrm{Col}(A)$ (by thm. before).

Hence, ~~$A^T \hat{b} = \bar{0}$~~ $A^T(\bar{b} - \hat{b}) = \bar{0}$.

Notice $A^T(\bar{b} - A\hat{x}) = \bar{0}$

$\Rightarrow A^T \bar{b} = A^T A \hat{x}$. $(*)$

This is the <u>normal equation</u> for $A\bar{x} = \bar{b}$.

<u>Theorem:</u> Hence, a least-squares solution of $A\bar{x} = \bar{b}$ coincides w/ the nonempty set of solutions of the normal equation $A^T \bar{b} = A^T A \hat{x}$.

~~Theorem?~~ <u>Pf:</u> Just showed in $(*)$ that LSS to $A\bar{x} = \bar{b}$ must also solve normal equation. ✓

Suppose $\hat{x}$ solves $A^T \bar{b} = A^T A \hat{x}$. Then

$A^T(\bar{b} - A\hat{x}) = \bar{0}$.

This shows that $(\bar{b} - A\hat{x}) \in (\mathrm{Col}(A))^\perp$. Hence,

$\bar{b} = A\hat{x} + (\bar{b} - A\hat{x})$ decomposes uniquely by using $\mathrm{Col}(A)$ and $(\mathrm{Col}(A))^\perp$.

Since projections minimize distances to subspaces, $A\hat{x} = \hat{b}$ is an LSS.

∎

Ex: "Find a LSS solution for $A\bar{x}=\bar{b}$ when when

$$A = \begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix} \text{ and } \bar{b} = \begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix}."$$

Check... Is $A\bar{x}=\bar{b}$ consistent? $RREF(A|\bar{b}) = I_3$, so no!

Solve normal equation: $A^T A \hat{x} = A^T \bar{b}$.

$$\text{So } A^T A = \begin{pmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix}\begin{pmatrix} 4 & 0 \\ 0 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 17 & 1 \\ 1 & 5 \end{pmatrix}, \quad A^T\bar{b} = \begin{pmatrix} 4 & 0 & 1 \\ 0 & 2 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix} = \begin{pmatrix} 19 \\ 11 \end{pmatrix}.$$

$$\text{Solve } \left(\begin{array}{cc|c} 17 & 1 & 19 \\ 1 & 5 & 11 \end{array}\right) \sim \left(\begin{array}{cc|c} 1 & 0 & 1 \\ 0 & 1 & 2 \end{array}\right) \Rightarrow \hat{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

is <u>the</u> LSS for $A\bar{x}=\bar{b}$.

~~$\bar{x} = \hat{x} + \bar{e}$~~ $\quad \bar{b} = \hat{b} + \bar{e}$

$$\bar{b} - \hat{b} = \bar{e}$$

$$\begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix} - A\hat{x} = \begin{pmatrix} 2 \\ 0 \\ 11 \end{pmatrix} - \begin{pmatrix} 4 \\ 4 \\ 3 \end{pmatrix} = \begin{pmatrix} -2 \\ -4 \\ 8 \end{pmatrix}$$

$$\|\bar{e}\| = \sqrt{4+16+64} = \sqrt{84} = \boxed{2\sqrt{41}}.$$

error.

Moral: ~~If $A\bar{x}=\bar{b}$ w/~~ LSS of $A\bar{x}=\bar{b}$ w/ $\bar{b} \in Col(A)$ is <u>ANY</u> solution to $A\bar{x}=\bar{b}$.

<u>Theorem</u>: The matrix A^TA is invertible iff the columns of A are lin. ind. In this case, $A\bar{x}=\bar{b}$ has a <u>unique</u> LSS $\hat{x}$ given by

$$\hat{x} = (A^TA)^{-1}A^T\bar{b}$$

(<u>Def</u>: The vector $\bar{b} - A\hat{x} = \bar{e}$ is called the <u>error vector</u> and $\|\bar{e}\|$ is the <u>error</u> of $\hat{x}$.)

<u>Alternatively</u>: <u>Theorem</u>: Given an $A \in M_{m\times n}(\mathbb{R})$ w/ lin. ind. columns, let $A = QR$ be a fact. of A. Then for each $\bar{b} \in \mathbb{R}^m$, $A\bar{x}=\bar{b}$ has a unique LSS to $A\bar{x}=\bar{b}$ given by

$$\hat{x} = R^{-1}Q^T\bar{b}.$$

(Usually solved in computer as $R\hat{x} = Q^T\bar{b}$).

<u>Pf</u>: From thm above,

$$\hat{x} = (A^TA)^{-1}A^T\bar{b}$$
$$= ((QR)^T(QR))^{-1}(QR)^T\bar{b}$$
$$= (R^TQ^TQR)^{-1}(R^TQ^T)\bar{b}$$
$$\cancel{= R^TR \cdot R^TQ^T\bar{b}}$$

... this way sucks.

Don't do it this way... involves inverse.

Consider $A\hat{x} = QR\hat{x} = QR(R^{-1}Q^T\bar{b}) = \underbrace{QQ^T}_{\parallel}\bar{b} = \hat{b}.$

projection matrix onto $Col(Q) = Col(A)$.