Math 107

Class website — www.math.duke.edu/~cbray/ (click ~~on~~ "Math 107")

Read: — all of main page
— 1st three links under "Course Information"
— familiarize yourself with links

If you think you might want to do a 2nd major in math, talk to me ASAP! Math 107 might be wrong course for you!

Resources

1. Yourself!
2. Book, notes — read before + after class. Not all done in class.
3. Instructor, TA — come to class, section!
4. Classmates — HW in groups; study groups (email list)

Expectations

1. Execution/memorization; 2. Understanding; 3. Apply to new situations
   - Not enough to memorize algorithms!
   - Process/explanations more important than "right answer".
   - Exams not just like HW!
   - "Raise your bar" on comprehension
     - comprehension not binary!
     - higher bar at Duke than h.s.

1.1 – Systems of Linear Equations

How do we solve a system such as:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\
&\vdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m
\end{aligned}$$

Do we know a solution exists? Might there be more than one solution? How many?

<u>Bad Method</u>: Start deriving more and more new equations...

Ex i) Solve

$$\begin{aligned}
x - y &= 1 \\
y - z &= 1 \\
x - z &= 1
\end{aligned}$$

Attempt:

$$\begin{aligned}
2x - y - z &= 2 \\
x + y - 2z &= 2 \\
3x - 3z &= 4 \\
4x + y - 5z &= 6 \\
&\vdots
\end{aligned}$$

What if this never ends? If it does, is that the only solution? Is it even a solution itself??

(N.B.: $x = x+1 \implies x^2 = x^2 + 2x + 1 \implies x = -\frac{1}{2}$

But this is <u>not</u> a solution!)

We want a strategy that is always completely conclusive.

Idea: Think of the system as being a single thing, instead of being made up of individual equations.

Def: Two systems are equivalent if they have identical solution sets.

Def: The following are called elementary equation operations:

1. interchanging two equations
2. adding a multiple of one equation to another
3. multiplying an equation by a nonzero number

These are not just algebraically legal — they are also reversible.

So,

Thm: If S_1 is acted on by elementary operations to yield a new system S_2, then S_1 and S_2 are equivalent.

Ex:

$$S_1 \begin{cases} x_1 + 3x_2 = 5 \\ 2x_1 - x_2 = 3 \end{cases}$$

$$\begin{cases} x_1 + 3x_2 = 5 & (1) \\ -7x_2 = -7 & (2) - 2(1) \end{cases}$$

$$\begin{cases} x_1 + 3x_2 = 5 & (1) \\ x_2 = 1 & -\frac{1}{7}(2) \end{cases}$$

$$S_2 \begin{cases} x_1 = 2 & (1) - 3(2) \\ x_2 = 1 & (2) \end{cases}$$

By our theorem, sol. sets for S_1, S_2 are the same;
but S_2's sol. set is obvious. ☺

Notation:
- "(n)" is a reference to an equation in the previous system, not the original
- Put formula for an equation next to that equation itself (book does differently...)

Matrix Notation

Note, if you fix the order of the variables in the equations, only the coefficients and constants change.

This allows for a convenient short hand.

$$\left\{\begin{array}{l} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{array}\right\} \text{ becomes } \left(\begin{array}{ccc|c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right)$$

This is called an augmented matrix.

The vertical bar is just a reference point, here indicating the "=".

The left side is the coefficient matrix.

The right side is the vector of constants.

Using this notation, "equation operations" are "row operations".

Elementary Row Operations

1. interchanging two rows
2. adding a multiple of one row to another
3. multiplying a row by a nonzero number

Reduced Row Echelon Form

The process of solving a system comes down to applying row operations to make the matrix as convenient as possible... (?)

Def:) A matrix is in <u>row echelon form</u> if each row has more leading zeroes than the preceding row (or is all zeroes).

Def:) The first nonzero entry in a row is called a <u>pivot</u>.

Def:) A matrix (usually a coefficient matrix) is in <u>reduced row echelon form</u> if

1. it is in row echelon form
2. all pivots are 1
3. all other entries in a pivot column are zero.

(A system is in rref if its coefficient matrix is in rref, as above.)

Thm:) For any system of equations (and assuming an ordering of the variables):

1. the matrix <u>can be</u> reduced to rref
2. there is a <u>unique</u> rref
3. there is a "convenient" means of deducing the solutions from the rref.

Ex: $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $\begin{pmatrix} 1 & 3 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$, $\begin{pmatrix} 1 & 2 & 4 & 0 & 5 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}$

Recall that, having fixed an order for the variables, columns correspond to variables.

Def: A variable whose column (in rref) contains a pivot is a <u>pivot variable</u>. The other variables are called <u>free variables</u>.

Why is rref "convenient"? Because, when a matrix is in rref,

$$\boxed{\text{You can always solve for the pivot variables in terms of } \underline{\text{only}} \text{ the free variables}}$$

Ex: $\left(\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 \end{array}\right)$ Solving for the pivot variables (x_1, x_3) in terms of the free variable (x_2), we get:

$$\begin{aligned} x_1 &= 2 - 3x_2 \\ x_3 &= 5 \end{aligned} \implies \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 2-3x_2 \\ x_2 \\ x_3 \end{pmatrix}$$

This always gives <u>exactly</u> the solution set.

Why does this give us the exact solution set?

1. Everything you get is a solution, for every value of the free variables, because all of the "pivot equations" are satisfied — because we solved for the pivot variables in those equations in terms of the free variables.

2. Every solution will be found in this way, because for any solution, the values of the free variables can be used to determine what the pivot variables in that solution would have to be — which we already found when we solved for the pivot variables in terms of the free variables

But watch out for contradictions!

Ex:
$$\left(\begin{array}{ccc|c} 1 & 3 & 0 & 2 \\ 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 7 \end{array}\right)$$
This last row represents the equation " $0 = 7$ "...

Of course there are no values of x_1, x_2, x_3 that make this true.

So the original system has <u>no</u> solutions

Gauss–Jordan Elimination

Usually the easiest way to get to rref is to "fix" the columns one at a time, from left to right.

Ex:
$$\left(\begin{array}{ccc|c} 1 & -2 & -1 & 1 \\ -3 & 8 & 7 & -5 \\ -4 & 11 & 11 & -8 \end{array}\right)$$

The problem with the first column is that the pivot (the 1 in the first row) is not the only nonzero entry.

$$\left(\begin{array}{ccc|c} 1 & -2 & -1 & 1 \\ 0 & 2 & 4 & -2 \\ 0 & 3 & 7 & -4 \end{array}\right) \begin{array}{l} 1 \\ 2 + 3\,1 \\ 3 + 4\,1 \end{array}$$

The next pivot will be in the second column, second row; we have to make it a "1" and eliminate the other entries.

$$\left(\begin{array}{ccc|c} 1 & -2 & -1 & 1 \\ 0 & 1 & 2 & -1 \\ 0 & 3 & 7 & -4 \end{array}\right) \begin{array}{l} 1 \\ 2/2 \\ 3 \end{array}$$

$$\left(\begin{array}{ccc|c} 1 & 0 & 3 & -1 \\ 0 & 1 & 2 & -1 \\ 0 & 0 & 1 & -1 \end{array}\right) \begin{array}{l} 1 + 2\,2 \\ 2 \\ 3 - 3\,2 \end{array}$$

Finally, we fix the last column.

$$\left(\begin{array}{ccc|c} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & -1 \end{array}\right) \begin{array}{l} 1 - 3\,3 \\ 2 - 2\,3 \\ 3 \end{array}$$

Notes

① Sometimes there is no pivot that can be arranged in a given column... No problem!

Ex: $\begin{pmatrix} 1 & 0 & 2 & 7 \\ 0 & 1 & 4 & 8 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 0 & 2 \end{pmatrix}$ can't get a pivot into this column.

② Sometimes you have to, or simply might want to switch rows to get a pivot where it needs to be

Ex: $\begin{pmatrix} 0 & 1 & 2 \\ 5 & 2 & 4 \\ 3 & 0 & 6 \end{pmatrix}$ Could switch ①,②... but better to switch ①,③ since it will avoid, or at least delay, the appearance of fractions.

Ex: $\begin{pmatrix} 5 & 1 & 8 \\ 2 & 3 & 1 \\ 1 & 2 & 3 \end{pmatrix}$ Switching ①,③ not necessary, but gets a 1 to the top left position without fractions.

③ Can do other things to avoid fractions...

$$\begin{pmatrix} 4 & 2 & 7 \\ 8 & 5 & 1 \\ 3 & 0 & 4 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & 3 \\ 8 & 5 & 1 \\ 3 & 0 & 4 \end{pmatrix} \begin{matrix} ① - ③ \\ ② \\ ③ \end{matrix}$$

Combining Row Operations

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

$$\begin{array}{c} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \end{array} \begin{array}{l} 1 \\ 2 - 1 \end{array}$$

$$\begin{array}{c} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{array} \begin{array}{l} 1 - 2 \\ 2 \end{array}$$

It is tempting to combine these two steps into a single step:

$$\left.\right) \begin{array}{l} 1 - 2 \\ 2 - 1 \end{array}$$

But that is not allowed in this case!

In fact in this case doing so would lead to a non-equivalent system!

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$$

$$\begin{array}{c} \begin{pmatrix} 0 & -1 \\ 0 & 1 \end{pmatrix} \end{array} \begin{array}{l} 1 - 2 \\ 2 - 1 \end{array}$$

this further reduces to the non-equivalent: $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$

What went wrong? The point here is that in the original reduction, the "1 − 2" refers to a row 2 that is <u>not the same</u> as the original row 2.

Combining these row operations into a single step blurs this distinction; so, it is <u>not</u> allowed.

As stated, row operations should be done one-at-a-time.

But — there are cases where combining row operations does not lead to any problems. Most importantly:

> If you leave one row *fixed*, and use only that row to adjust other rows, no problems will arise.

Ex:

$$\begin{pmatrix} 1 & 2 & 0 \\ 2 & 3 & 1 \\ 4 & 9 & 6 \end{pmatrix}$$

$$\begin{array}{c} \begin{pmatrix} 1 & 2 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & 6 \end{pmatrix} \end{array} \begin{array}{l} (1) \\ (2) - 2(1) \\ (3) - 4(1) \end{array}$$

this is left fixed

only the fixed (1) is used to adjust the other rows.

Variation:

$$\begin{pmatrix} 2 & 1 & 4 \\ 3 & 2 & 16 \\ 5 & 8 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & \frac{1}{2} & 2 \\ 0 & 1 & 20 \\ 0 & 11 & -18 \end{pmatrix} \begin{array}{l} (1)/2 \\ 2(2) - 3(1) \\ 2(3) - 5(1) \end{array}$$

Here, the first row is only scaled. Had we done that scaling in an initial step, that would only change the multiples needed to use it to adjust the other rows.

Existence, Uniqueness, and Rank

Def:) The <u>rank</u> of a matrix is the number of pivots in the rref.

Rank relates to two natural questions about the solutions to a system of equations.

Existence

We have already seen that a system has solutions iff there are no contradictions.

$$\left(\begin{array}{ccccc|c} 1 & 2 & 0 & 5 & 7 & ? \\ 0 & 0 & 1 & 2 & 8 & ? \\ 0 & 0 & 0 & 0 & 0 & \boxed{?} \\ 0 & 0 & 0 & 0 & 0 & \boxed{?} \end{array}\right)$$

If these are both zero, solutions exist.
If either is nonzero, no solutions.

rows of zeroes in the coefficient matrix

What can we deduce about this if we know <u>only</u> the <u>coefficient matrix</u>?

Def:) A coefficient matrix has the <u>existence property</u> (general existence, universal existence, ...) if solutions must exist no matter what the constants are on the right side.

How can you tell if a matrix has existence?

Thm: A coefficient matrix has the existence property if

(1) there are <u>no</u> rows of zeroes in rref

(equiv:) (2) there is a pivot in every row of rref

(3) rank = # rows

(If there were a row of zeroes, there might be a contradiction!)

Ex: Say $\text{rref}(A) = \begin{pmatrix} 1 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, $\text{rref}(B) = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$

A <u>does</u> have universal existence; B <u>does not</u>.

Related point: A system is called <u>homogeneous</u> if the right side is all zeroes:

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= 0 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= 0 \end{aligned}$$

Note, a homogeneous system always has solutions, even though the coefficient matrix might not have universal existence.

The solution $x_1 = 0, \ldots, x_n = 0$ is called the <u>trivial solution</u>.

Uniqueness

(NB — the uniqueness question should be thought of as independent of the existence question. Ignore potential existence problems; that is, act as if a solution is known to exist, even if that is not known.)

What can we deduce about uniqueness of (possible) solutions if we know only the coefficient matrix?

Def: A coefficient matrix has the uniqueness property if (possible) solutions must be unique.

How can you tell if a matrix has uniqueness?

Thm: A coefficient matrix has the uniqueness property if

(equiv:)

1. there are no free variables in rref
2. there is a pivot in every column of rref
3. rank A = # cols

(If there were free variables, we could not have unique solutions...)

Ex: $\text{rref}(A) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \end{pmatrix}$, $\text{rref}(B) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}$

A does not have uniqueness; B does

Related Facts

Obs:) If A is $m \times n$, then $\text{rank}(A) \leq m, n$.
(rows) (cols)

Thm:) If $m > n$ (more eqs than vars; a "tall & thin" matrix), then A cannot have universal existence.

(There might be RHS constants where solutions exist... but there will be RHS constants where solutions DNE.)

Thm:) If $m < n$ (more vars than eqs; a "short + wide" matrix), then A cannot have uniqueness.

Thm:) A linear system must have either 0, 1, or ∞'ly many solutions.

1.2 – Matrices and Matrix Operations

Previously we have seen matrices used as a shorthand for systems of equations. However, they can be used for other things — and can be thought of as independent things.

Matrix Addition

Say $A = (a_{ij})$, $B = (b_{ij})$ have same dimensions.

Then $C = A + B$ where $C = (c_{ij})$

is defined by $c_{ij} = a_{ij} + b_{ij}$

Ex: $\begin{pmatrix} 1 & 3 \\ 7 & 2 \end{pmatrix} + \begin{pmatrix} 2 & -1 \\ -3 & 5 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ 4 & 7 \end{pmatrix}$

Scalar Multiplication

Say $A = (a_{ij})$. Then

$C = kA$ where $C = (c_{ij})$

is defined by $c_{ij} = k\, a_{ij}$

Ex: $5 \begin{pmatrix} 3 & 2 \\ 1 & -4 \end{pmatrix} = \begin{pmatrix} 15 & 10 \\ 5 & -20 \end{pmatrix}$

Matrix Multiplication

Need ① # cols of A = # rows of B

(equiv)

② # elts of row of A = # elts of col of B

$$A = \begin{pmatrix} \text{—}\ A_1\ \text{—} \\ \vdots \\ \text{—}\ A_m\ \text{—} \end{pmatrix} = \begin{pmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{pmatrix} \quad \text{is } m \times n$$

$$B = \begin{pmatrix} \text{—}\ B_1\ \text{—} \\ \vdots \\ \text{—}\ B_n\ \text{—} \end{pmatrix} = \begin{pmatrix} | & & | \\ b_1 & \cdots & b_p \\ | & & | \end{pmatrix} \quad \text{is } n \times p$$

Then $C = AB$, $C = (c_{ij})$ is defined by a dot product:

$$c_{ij} = A_i \cdot b_j$$

Note the dimensions:

$$\underset{(m \times p)}{C} = \underset{(m \times n)}{A} \; \underset{(n \times p)}{B}$$

Ex:
$$\begin{pmatrix} 1 & 3 & 1 \\ 2 & 0 & -1 \end{pmatrix} \begin{pmatrix} 2 & 3 \\ 5 & 1 \\ 4 & -2 \end{pmatrix} = \begin{pmatrix} 21 & 4 \\ 0 & 8 \end{pmatrix}$$

Alternatives:

① Cols of AB are linear combinations of cols of A, using cols of B as coeffs.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$$

$$= 5 \begin{pmatrix} 1 \\ 3 \end{pmatrix} + 7 \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 19 \\ 43 \end{pmatrix}$$

② Rows of AB are linear combinations of rows of B, using rows of A as coeffs.

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 5 & 6 \\ 7 & 8 \end{pmatrix} = \begin{pmatrix} 19 & 22 \\ 43 & 50 \end{pmatrix}$$

$$3 \begin{pmatrix} 5 & 6 \end{pmatrix} + 4 \begin{pmatrix} 7 & 8 \end{pmatrix} = \begin{pmatrix} 43 & 50 \end{pmatrix}$$

③ Sigma notation:

$A = (a_{ij})$, $B = (b_{ij})$, $C = (c_{ij}) = AB$, then

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

Properties

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

(this is inconvenient to prove here; but **easy** to prove with "linear trans", which we will study later.)

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

$$k(AB) = (kA)B = A(kB)$$

Defs) $O_{m\times n} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \cdots & 0 \end{pmatrix}$ $(m\times n)$

$I_n = \begin{pmatrix} 1 & & & O \\ & 1 & & \\ & & \ddots & \\ O & & & 1 \end{pmatrix}$ $(n\times n)$

Obs.) ① $A + O = A = O + A$

② If A is $m\times n$, then

$$I_m A = A = A I_n$$

(think about matrix prod as l.c.'s of rows, cols)

③ Can view a system of equations

$$\begin{aligned} a_{11}x_1 + \cdots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

as $AX = B$

where $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$ $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ $B = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$

Thm:) Suppose that X_p ("particular solution") is any solution to $AX=B$.

Then X is a solution iff $X = X_p + X_H$, where

X_H is a solution to $AX=0$ ("homogeneous solution").

Pf:) First observe that if $X = X_p + X_H$, then

$$AX = AX_p + AX_H = B + 0 = B$$

So X is a solution.

On the other hand, if X is a solution, then

$$A(X - X_p) = AX - AX_p = B - B = 0$$

which means $X_H = X - X_p$ is a homogeneous solution

and thus $X = X_p + X_H$.

Interpretation: If you know any particular solution, and all homogeneous solutions, then the sum is the complete set of sols.

Ex:) Consider $\left(\begin{array}{ccc|c} 1 & 2 & -2 & -5 \\ -2 & -4 & 5 & 13 \\ -5 & -10 & 7 & 16 \end{array}\right)$

Can check $\begin{pmatrix} -2y \\ y \\ 0 \end{pmatrix}$ is all homogeneos sols, and $\begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}$ is a particular sol.

So the complete set of sols is

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

Inverses of Matrices

For an $n \times n$ matrix A, we say B is an <u>inverse</u> if

$$AB = BA = I$$

<u>Ex:</u> $A = \begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$ $\quad B = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}$

Note $\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}\begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -3 & 2 \end{pmatrix}\begin{pmatrix} 2 & 1 \\ 3 & 2 \end{pmatrix}$

If A has an inverse, say A is <u>invertible</u>.

<u>Thm:</u> If A is invertible, then the inverse is <u>unique</u>.

<u>Pf:</u> Say B, C are both inverses for A. Then

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

So $B = C$. (Given this uniqueness, we write the inverse as A^{-1}.)

<u>Thm:</u> If A is $n \times n$ and $AB = I$, then $BA = I$ (so A is invertible). (Note, A must be square!)

<u>Pf:</u> Say $AB = I$. First we will show A has the uniqueness property.

Note that $I\vec{x} = \vec{c}$ has existence for all $\vec{c}$.

Then $AB\vec{x} = \vec{c}$ has existence ....

and so $A\vec{y} = \vec{c}$ has existence .... (choose $\vec{y} = B\vec{x}$).

So A has a pivot in every row (of its rref).

Since A is <u>square</u>, A must have a pivot in every column, and thus A has uniqueness.

Having established uniqueness for A, we argue as follows:

$$AB = I \Rightarrow ABA = A \Rightarrow ABA\vec{x} = A\vec{x}$$

Since A has uniqueness, this means that $BA\vec{x} = \vec{x}$

As this is true for any $\vec{x}$, we must have $BA = I$.

<u>Elementary Matrices</u>

An elementary matrix is the result of applying a row operation to the identity matrix.

Momentarily let us define a "row operation matrix" to be a matrix F such that FA is the result of a row operation on A. (Note, all row operations are l.c.'s of rows of A, and so can be realized by some F.)

Ex:

$$\begin{pmatrix} 1 & 0 & 0 \\ 5 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 2 & 1 \\ 3 & 4 & 3 & 7 \\ 2 & -3 & -2 & -1 \end{pmatrix} = \begin{array}{cc} \begin{pmatrix} 1 & 0 & 2 & 1 \\ 8 & 4 & 13 & 12 \\ 2 & -3 & -2 & -1 \end{pmatrix} & \begin{array}{l} 1 \\ 2+5\cdot 1 \\ 3 \end{array} \end{array}$$

Ex:

$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{pmatrix} = \begin{array}{cc} \begin{pmatrix} 5 & 6 \\ 3 & 4 \\ 1 & 2 \end{pmatrix} & \begin{array}{l} 3 \\ 2 \\ 1 \end{array} \end{array}$$

Note:

1. For every row operation, each new row is a linear combination of the previous rows (of A).
2. When left multiplying a matrix A by a matrix F (as in FA), the rows of the product are linear combinations of the rows of A.
3. This similarity shows these row operation matrices always exist.

So:

$$A \xrightarrow{\text{row op.}} B$$

$$FA = B$$

same!

Thm: If E is obtained by a row operation on I, then the result of applying the same row operation on a matrix A is EA.

Pf: Let F be the row operation matrix for this operation; then

$$E = FI$$

so

$$E = F$$

Applying this operation to A gives

$$FA = EA.$$

Note, a different way to view this is simply that elementary matrices and "row operation matrices" are the same thing. We will refer to these henceforth as elementary matrices.

Thm: A is invertible $\iff$ $\text{rref}(A) = I$ (A is "nonsingular")

Pf: ($\Rightarrow$) Say A is invertible. We need to show there is a pivot in every row (since A is square, it will follow that there is a pivot in every column.)

That is, we need to show existence for $A\vec{x} = \vec{b}$.

This is easy, since $\vec{x} = A^{-1}\vec{b}$ is clearly a solution.

($\Leftarrow$) Say rref(A) is the identity. We can realize the row reduction as

$$E_k \cdots E_1 A = I$$

Writing $E = E_k \cdots E_1$, this becomes

$$EA = I$$

So $E = A^{-1}$ and A is invertible.

This also leads us to a method for finding A^{-1}...

$$E = A^{-1} \implies EI = A^{-1}$$

So if we apply E to the augmented matrix $(A \mid I)$, we get

$$E(A \mid I) = (EA \mid EI)$$
$$= (I \mid A^{-1}).$$

That is — apply row ops to I <u>in parallel</u> as you do so for A; and if A reduces to I, then I reduces to A^{-1}!

Ex: Is $\begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}$ invertible, and if so, what is the inverse?

$$\left(\begin{array}{cc|cc} 3 & 1 & 1 & 0 \\ 5 & 2 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{cc|cc} 3 & 1 & 1 & 0 \\ -1 & 0 & -2 & 1 \end{array}\right) \begin{array}{l} ① \\ ② - 2① \end{array}$$

$$\left(\begin{array}{cc|cc} 1 & 0 & 2 & -1 \\ 3 & 1 & 1 & 0 \end{array}\right) \begin{array}{l} -② \\ ① \end{array}$$

$$\left(\begin{array}{cc|cc} 1 & 0 & 2 & -1 \\ 0 & 1 & -5 & 3 \end{array}\right) \begin{array}{l} ① \\ ② - 3① \end{array}$$

This is the inverse (must have I on left!)

This tells us $\begin{pmatrix} 3 & 1 \\ 5 & 2 \end{pmatrix}$ <u>is</u> invertible.

Other facts: ① $(AB)^{-1} = B^{-1}A^{-1}$

② If A not square, $AB = I \not\Rightarrow BA = I$

Ex: $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$

$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \neq I_3$

Another application of elementary matrices involves solving the system $A\vec{x} = \vec{b}$ with several different $\vec{b}$ vectors.

Note, $A\vec{x} = \vec{b}_1$ and $A\vec{x} = \vec{b}_2$ require doing the <u>same</u> row operations to solve (or to determine if there is no solution). Why do all this work multiple times?

Suppose the row reduction of A is represented by

$$EA = R$$

Then solving $\quad A\vec{x} = \vec{b} \quad$ involves row reducing, with E...

$$EA\vec{x} = E\vec{b}$$

$$R\vec{x} = E\vec{b}$$

This is equivalent, and already row reduced.

Ex i) Solve $A\vec{x} = \vec{b}$ with $\vec{b} = \vec{b}_1, \vec{b}_2,$ and $\vec{b}_3$.

1) Reduce $(A \mid I)$ to $(R \mid E)$

2) Solve $A\vec{x} = \vec{b}_1$ with $R\vec{x} = E\vec{b}_1$

Solve $A\vec{x} = \vec{b}_2$ with $R\vec{x} = E\vec{b}_2$

Solve $A\vec{x} = \vec{b}_3$ with $R\vec{x} = E\vec{b}_3$

⋮

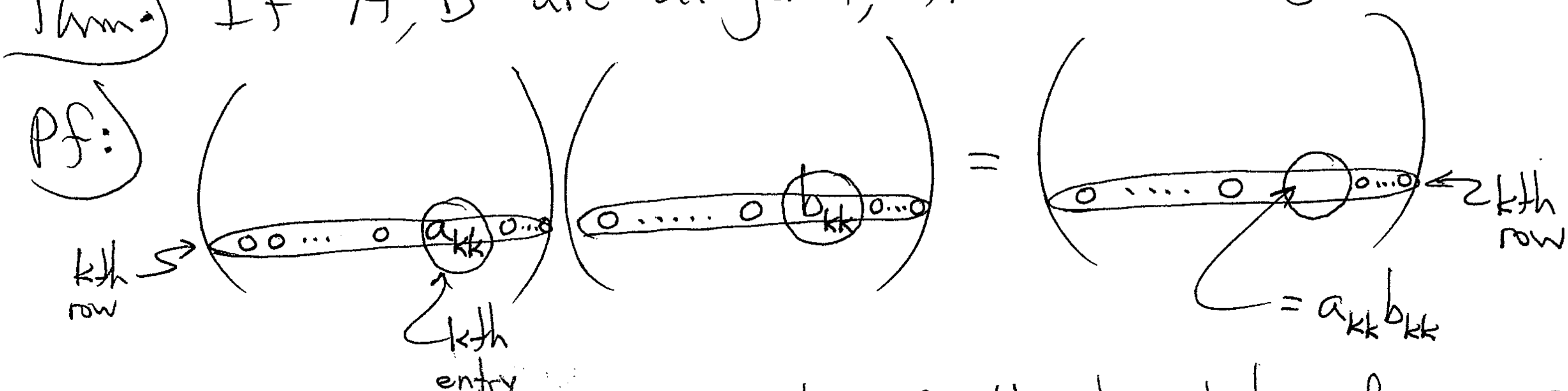
Special Matrices

Def: A is a <u>diagonal matrix</u> if all non-diagonal entries are zero, (and A is square).

Ex: $A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$ is a diagonal matrix

Thm: If A, B are diagonal, then AB is diagonal.

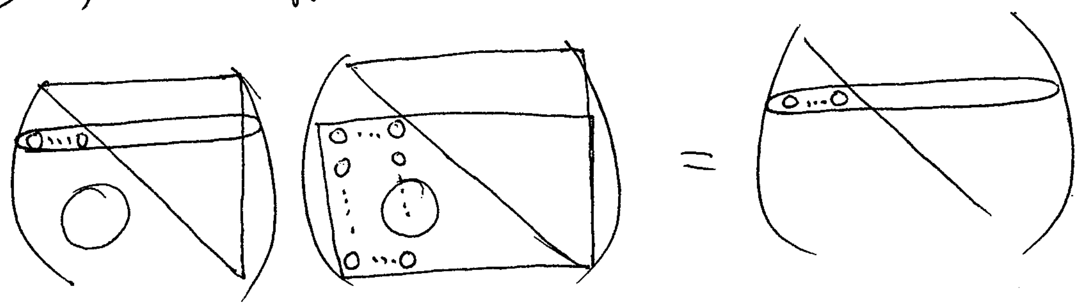
Pf:



Def: A (square) is <u>upper triangular</u> if all entries below diag are zero.
- - - - - - <u>lower</u> - - - - - - - - - - - - above - - - - -

Thm: A, B both upper triangular $\Rightarrow$ AB is upper triangular.

Pf:



Def: The <u>transpose</u>, A^T, of a matrix A is defined by

$$\left(A^T\right)_{ij} = A_{ji}$$

Ex: $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$ $\qquad A^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$

Thm: $(AB)^T = B^T A^T$

Pf: $\left((AB)^T\right)_{ij} = (AB)_{ji} = A_j \cdot b_i$

$$\left(B^T A^T\right)_{ij} = \left(i\text{th row of } B^T\right) \cdot \left(j\text{th col. of } A^T\right)$$

$$= \left(i\text{th col. of } B\right) \cdot \left(j\text{th row of } A\right)$$

$$= b_i \cdot A_j$$

Thm: $\left(A^T\right)^{-1} = \left(A^{-1}\right)^T$

Pf: $\left(A^{-1}\right)^T A^T = \left(A A^{-1}\right)^T = I^T = I$

Def: A (square) is <u>symmetric</u> if

$$A^T = A$$

Ex: $\begin{pmatrix} 1 & 3 & 7 \\ 3 & 10 & 2 \\ 7 & 2 & 4 \end{pmatrix}$ is symmetric $\left(a_{ij} = a_{ji} \text{ for all } i,j\right)$

Thm: A^TA, AA^T are symmetric

Pf: $(A^TA)^T = (A)^T(A^T)^T = A^TA$

$(AA^T)^T = (A^T)^T(A)^T = AA^T$

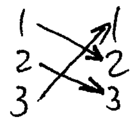
Thm: If A is invertible and symmetric, then A^{-1} is symm.

Pf: $(A^{-1})^T = (A^T)^{-1} = A^{-1}$

Determinants

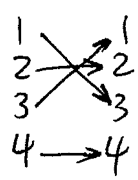
First we must define some tools.

Def: A <u>permutation</u> is a function $\alpha : \{1, \dots, n\} \to \{1, \dots, n\}$ that is 1-1 and onto.

Ex:  is a permutation of $\{1,2,3\}$

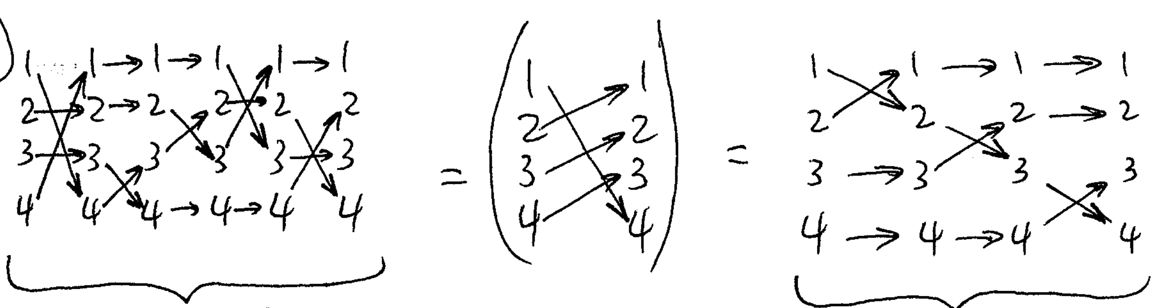
Def: The <u>symmetric group on n elements</u>, written S_n, is the set of all permutations on $\{1, \dots, n\}$.

Def: A <u>transposition</u> is a permutation that fixes all but two elements of $\{1, \dots, n\}$, which it switches.

Ex:  is a transposition in S_4

Thm: If a permutation α is written as a composition of transpositions in two different ways, the numbers of transpositions will be either both odd or both even.

Ex:



5 transpositions $\qquad\qquad$ 3 transpositions

Def:) An odd permutation is one requiring an odd number of transpositions. An even permutation .... even .... .

Def:) $\operatorname{sgn}(\alpha) = (-1)^k$, where k is the number of transpositions. (well defined!)

We are now ready to define determinant.

Def:) The determinant of an $n \times n$ matrix A (with elements (a_{ij})) is

$$\det A = \sum_{\alpha \in S_n} \operatorname{sgn}(\alpha)\, a_{1\alpha(1)}\, a_{2\alpha(2)} \cdots a_{n\alpha(n)}$$

Notice that each term in this summation is a product of elements from A taken with exactly one in each row and one in each column.

Ex! Compute the determinant of $A = \begin{pmatrix} -1 & 3 & 5 \\ 2 & 1 & -2 \\ 1 & 0 & 4 \end{pmatrix}$

There are six permutations in S_3

$\begin{matrix} 1 \to 1 \\ 2 \to 2 \\ 3 \to 3 \end{matrix} \rightsquigarrow \begin{pmatrix} \cdot & & \\ & \cdot & \\ & & \cdot \end{pmatrix}$ even $\rightsquigarrow (+)(-1)(1)(4)$

$\begin{matrix} 1 \to 2 \\ 2 \to 1 \\ 3 \to 3 \end{matrix} \rightsquigarrow \begin{pmatrix} & \cdot & \\ \cdot & & \\ & & \cdot \end{pmatrix}$ odd $\rightsquigarrow (-)(3)(2)(4)$

$\begin{matrix} 1 \to 3 \\ 2 \to 2 \\ 3 \to 1 \end{matrix} \rightsquigarrow \begin{pmatrix} & & \cdot \\ & \cdot & \\ \cdot & & \end{pmatrix}$ odd $\rightsquigarrow (-)(5)(1)(1)$

$\begin{matrix} 1 \to 1 \\ 2 \to 3 \\ 3 \to 2 \end{matrix} \rightsquigarrow \begin{pmatrix} \cdot & & \\ & & \cdot \\ & \cdot & \end{pmatrix}$ odd $\rightsquigarrow (-)(-1)(-2)(0)$

$\begin{matrix} 1 \to 2 \\ 2 \to 3 \\ 3 \to 1 \end{matrix} \rightsquigarrow \begin{pmatrix} & \cdot & \\ & & \cdot \\ \cdot & & \end{pmatrix}$ even $\rightsquigarrow (+)(3)(-2)(1)$

$\begin{matrix} 1 \to 3 \\ 2 \to 1 \\ 3 \to 2 \end{matrix} \rightsquigarrow \begin{pmatrix} & & \cdot \\ \cdot & & \\ & \cdot & \end{pmatrix}$ even $\rightsquigarrow (+)(5)(2)(0)$

sum $= -39$

Note, a square matrix can be viewed as a function, by

$$T(\vec{x}) = A\vec{x}, \quad T: \mathbb{R}^n \to \mathbb{R}^n$$

An important observation is that the <u>columns</u> of A are the <u>images</u> of the standard basis vectors:

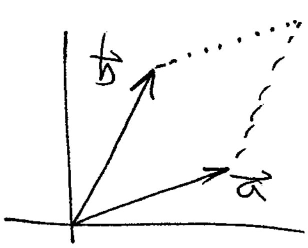
$$T(\vec{e}_i) = A\vec{e}_i = A\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \leftarrow i\text{th coord.} = i\text{th col. of } A = \vec{a}_i$$

So

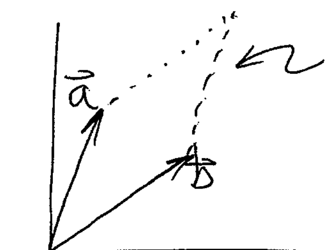
$$A = \begin{pmatrix} | & & | \\ \vec{a}_1 & \cdots & \vec{a}_n \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ T(\vec{e}_1) & \cdots & T(\vec{e}_n) \\ | & & | \end{pmatrix}$$

Students will recall from Math 103 that determinants relate to areas and volumes:

Thm i) If $A = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$, then the determinant $\det A$ is the "signed area" of the llgram defined by $\vec{a}, \vec{b}$ (area if $\vec{a}, \vec{b}$ are in counterclockwise order, $-$area if $\vec{a}, \vec{b}$ are in clockwise order.)



counterclockwise, "signed area" positive

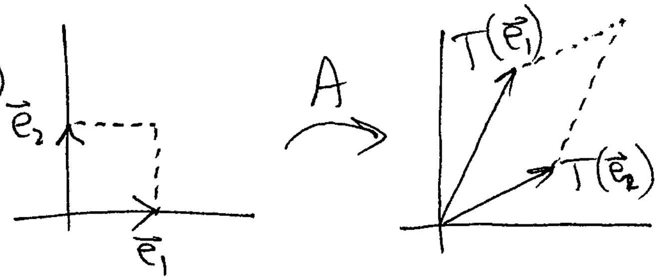


clockwise, "signed area" is negative

Thm: If $A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$, then $\det A$ is the "signed volume" of the $\parallel$piped defined by $\vec{a}, \vec{b}, \vec{c}$ (volume if $\vec{a}, \vec{b}, \vec{c}$ in RH order, $-$volume if LH order)

Given our point of view of matrices as functions, we can also note that these $\parallel$gram, $\parallel$piped are the images of the unit square and unit cube, which have area $= 1$, volume $= 1$. We can then interpret the determinant as a "stretching factor", whose sign indicates a "flip" or "inversion".

Ex:

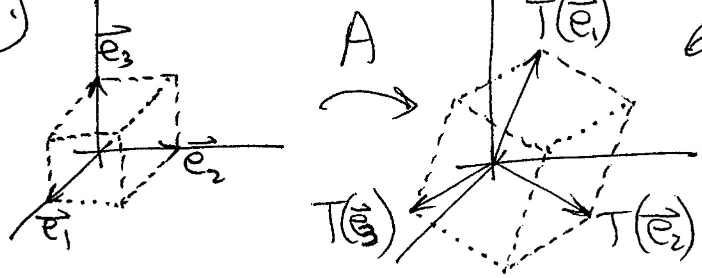


$\det A = -2$

"2" $\Rightarrow$ areas stretched by 2

"$-$" $\Rightarrow$ areas "flipped"

Ex:



$\det A = -3$

"3" $\Rightarrow$ vols stretched by 3

"$-$" $\Rightarrow$ vol. "inverted"

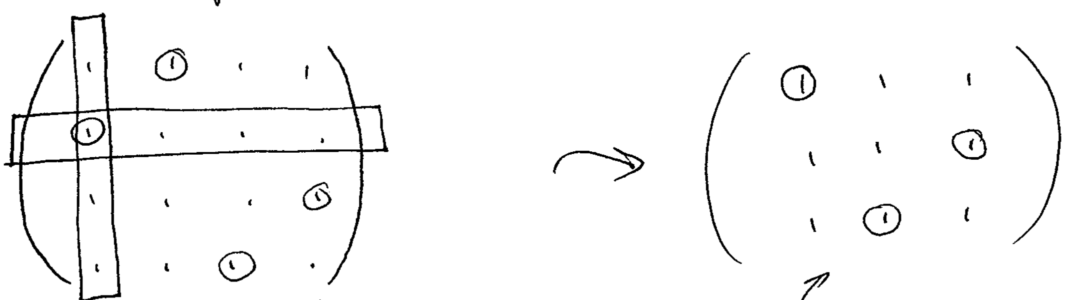
We can interpret many facts about determinants from this geometric point of view.

Cofactor Expansions

Let M_{ij} denote the <u>ij-minor</u> of A, which is the matrix obtained from A by removing the ith row and jth column.

Note, any permutation of S_n with $\alpha(i)=j$ can be thought of as yielding a permutation of S_{n-1} if we remove i from the inputs and j from the outputs.

Ex:



$\alpha \in S_n$ $\qquad$ $\sigma \in S_{n-1}$

$i=2,\ j=1$

So there is a correspondence $\{S_n \mid i \to j\} \leftrightarrow \{S_{n-1}\}$

It can be shown that the signs of these permutations are related by the values of i and j, by the factor $(-1)^{i+j}$. (We will not show this here.)

<u>Notation</u>: We will use a'_{kl} to refer to the kl entry in the matrix M_{ij} (where the values of i and j are clear).

If we group the terms in $\det A = \sum_{\alpha \in S_n}$ by the value of $\alpha(i)$ then, we can rewrite the summation as

$$\det A = \sum_{\alpha \in S_n} \operatorname{sgn}(\alpha)\, a_{1\alpha(1)} \cdots a_{n\alpha(n)}$$

$$= \sum_{j=1}^{n} \sum_{\substack{\alpha \in S_n \\ \text{with } i \to j}} \operatorname{sgn}(\alpha)\, a_{1\alpha(1)} \cdots a_{n\alpha(n)}$$

$$= \sum_{j=1}^{n} \sum_{\sigma \in S_{n-1}} (-1)^{i+j} \operatorname{sgn}(\sigma)\, a_{ij}\, a'_{1\sigma(1)} \cdots a'_{(n-1)\sigma(n-1)}$$

$$= \sum_{j=1}^{n} a_{ij} (-1)^{i+j} \sum_{\sigma \in S_{n-1}} \operatorname{sgn}(\sigma)\, a'_{1\sigma(1)} \cdots a'_{(n-1)\sigma(n-1)}$$

$$= \sum_{j=1}^{n} a_{ij} (-1)^{i+j} \det(M_{ij})$$

This is true for any value of i — so, you can use this formula to compute $\det(A)$ "along" any row.

Ex: $A = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 0 & 2 \\ 5 & 7 & 1 \end{pmatrix}$ We compute $\det A$ along 2nd row:

$$\det A = (1)(-1)^{2+1} \det\begin{pmatrix} 3 & 1 \\ 7 & 1 \end{pmatrix} + (0)(-1)^{2+2} \det\begin{pmatrix} 2 & 1 \\ 5 & 1 \end{pmatrix} + (2)(-1)^{2+3} \det\begin{pmatrix} 2 & 3 \\ 5 & 7 \end{pmatrix}$$

$$= 6$$

Writing $C_{ij} = (-1)^{i+j} \det M_{ij}$, this becomes $\det A = \sum_{j=1}^{n} a_{ij} C_{ij}$

this is called the "ij-cofactor".

More about determinants:

Thm: $\det(A^T) = \det(A)$

Pf: $\det(A^T) = \sum_{\alpha} \text{sgn}(\alpha)\, a^T_{1\alpha(1)} \cdots a^T_{n\alpha(n)}$

$= \sum_{\alpha} \text{sgn}(\alpha)\, a_{\alpha(1)1} \cdots a_{\alpha(n)n}$

$= \sum_{\alpha} \text{sgn}(\alpha)\, a_{1\alpha^{-1}(1)} \cdots a_{n\alpha^{-1}(n)}$

$= \sum_{\sigma} \text{sgn}(\sigma^{-1})\, a_{1\sigma(1)} \cdots a_{n\sigma(n)}$ (where $\sigma = \alpha^{-1}$)

$= \sum_{\sigma} \text{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{n\sigma(n)}$ (since $\text{sgn}(\sigma^{-1}) = \text{sgn}(\sigma)$)

$= \det(A)$

This means that in addition to computing by cofactors along rows, we can also use any column.

Thm: The determinant of a triangular (upper or lower) matrix is the product of the diagonal entries.

Pf: Expand by cofactors along 1st row (for lower tri.) or 1st col. (for upper tri.) successively.

Thm: The determinant of a diagonal matrix is the product of the diagonal entries.

This is a corollary of the previous thm. Also — corresponds to the volume of a box in 3x3 case...

Thm:) If there is a row or column of zeroes, then the determinant is zero.

Pf:) Expand by cofactors along that row or column.

Also, note that a column of zeroes in a 3×3 matrix would mean that the corresponding $\parallel$piped is "flat", so the volume would be zero.

Multilinearity

A function $T: \mathbb{R}^n \to \mathbb{R}^n$ is <u>linear</u> if

$$T(a\vec{x}+b\vec{y}) = aT(\vec{x}) + bT(\vec{y}) \quad \text{(for all } a,b,\vec{x},\vec{y}\text{)}$$

For example, a matrix function $T(\vec{x}) = A\vec{x}$ is linear.

A function with several vector inputs is <u>multilinear</u> if it is linear in each input; that is,

$$S(\vec{v}_1, \ldots, a\vec{v}_k + b\vec{w}_k, \ldots, \vec{v}_n) = a\,S(v_1, \ldots, \vec{v}_k, \ldots, \vec{v}_n) + b\,S(v_1, \ldots, \vec{w}_k, \ldots, \vec{v}_n)$$

(for all $a, b, \vec{v}_k, \vec{w}_k$, and k)

Remember that a matrix can be thought of as a list of vectors (row or column)... So, det can be thought of as applying to a list of vectors.

Thm: Determinant is multilinear

Pf: We need to show

$$\det\begin{pmatrix} \vec{v}_1 \\ \vdots \\ a\vec{v}_k + b\vec{w}_k \\ \vdots \\ \vec{v}_n \end{pmatrix} = a \det\begin{pmatrix} \vec{v}_1 \\ \vdots \\ \vec{v}_k \\ \vdots \\ \vec{v}_n \end{pmatrix} + b \det\begin{pmatrix} \vec{v}_1 \\ \vdots \\ \vec{w}_k \\ \vdots \\ \vec{v}_n \end{pmatrix}$$

This is easily shown by cofactor expansion along kth row.

Ex:

$$\det\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{pmatrix} = 7 \det\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 1 & 0 & 0 \end{pmatrix} + 8 \det\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 0 & 1 & 0 \end{pmatrix}$$

$$= 7 \det\begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix}$$

$$= 8(-1) \det\begin{pmatrix} 1 & 3 \\ 4 & 6 \end{pmatrix}$$

$$= (7)\det\begin{pmatrix} 2 & 3 \\ 5 & 6 \end{pmatrix} + 8(-1)\det\begin{pmatrix} 1 & 3 \\ 4 & 6 \end{pmatrix}$$

Note, multilinearity also on column vectors.

Antisymmetry

Thm: If you switch two columns (or two rows) of a matrix, the determinant changes by a factor of (-1).

Pf: Let A' be the result of switching columns i and j of the matrix A. Then

$$\det A' = \sum_{\alpha \in S_n} \operatorname{sgn}(\alpha)\, a'_{1\alpha(1)} \cdots a'_{n\alpha(n)}$$

$$= \sum_{\alpha \in S_n} \operatorname{sgn}(\alpha)\, a_{1\sigma(1)} \cdots a_{n\sigma(n)}$$

(where $\sigma = \tau \circ \alpha$, where τ is the transposition that switches i and j.)

$$= \sum_{\sigma \in S_n} \operatorname{sgn}(\tau \circ \sigma)\, a_{1\sigma(1)} \cdots a_{n\sigma(n)} \qquad (\text{Note } \tau \circ \sigma = \alpha)$$

Because τ is a single transposition, $\operatorname{sgn}(\tau \circ \sigma) = -\operatorname{sgn}(\sigma)$.

So

$$\det A' = -\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, a_{1\sigma(1)} \cdots a_{n\sigma(n)}$$

$$= -\det A$$

The corresponding fact about switching rows can be seen by a similar argument (with $\sigma = \alpha \circ \tau$) — or by simply using the established fact that $\det(A^T) = \det(A)$.

This antisymmetry can be interpreted geometrically for 3×3 matrices. If you switch two columns, they define a parallelepiped of the same volume — but the handedness switches, thus giving the extra (-1).

We also get a nice corollary.

Cor:) If two rows (or two columns) of A are the same, then the determinant is zero.

Pf:) If A' is the result of switching those two rows (or columns), then we have

$$\det A' = -\det A$$

But $A' = A$ also, giving us

$$\det A' = \det A$$

So we must have $\det A = 0$.

This corollary can also be viewed geometrically in $\mathbb{R}^3$ — if two of the vectors defining a llpiped are the same, then that llpiped is flat, and thus has no volume.

Row Operations and Elementary Matrices

What happens to a determinant if we add a scalar multiple of one row to another row? As it turns out, nothing!

$$\det\begin{pmatrix} \vec{r}_1 \\ \vdots \\ \vec{r}_j + c\vec{r}_i \\ \vdots \\ \vec{r}_n \end{pmatrix} = \det\begin{pmatrix} \vec{r}_1 \\ \vdots \\ \vec{r}_j \\ \vdots \\ \vec{r}_n \end{pmatrix} + c\det\begin{pmatrix} \vec{r}_1 \\ \vdots \\ \vec{r}_i \\ \vdots \\ \vec{r}_n \end{pmatrix}$$

but in the last matrix the ith and jth rows are the same, so that determinant is zero.

We can now summarize row operations and determinants:

Row Op.	Effect on det	Corr. El. Mat.	det of El. Mat.
switch 2 rows	factor of (-1) (by antisymmetry)	$\begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}$	(-1)
mult. row by $c \neq 0$	factor of (c) (by multilinearity)	$\begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & c & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix}$	(c)
add multiple of 1 row to another	nothing	$\begin{pmatrix} 1 & & c & \\ & \ddots & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}$	(1)

We can then use row reductions to compute determinants.

Ex: $\begin{pmatrix} 0 & 1 & 3 \\ 1 & 0 & 2 \\ 0 & 2 & 10 \end{pmatrix} = A$

$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 2 & 10 \end{pmatrix} \begin{matrix} 2 \\ 1 \\ 3 \end{matrix}$ $\quad \det = -\det(A)$

$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 4 \end{pmatrix} \begin{matrix} 1 \\ 2 \\ 3 - 2\,2 \end{matrix}$ $\quad \det = -\det(A)$

$\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} 1 \\ 2 \\ 3/4 \end{matrix}$ $\quad \det = -\frac{1}{4}\det(A)$

$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{matrix} 1 - 2\,3 \\ 2 - 3\,3 \\ 3 \end{matrix}$ $\quad \det = -\frac{1}{4}\det(A)$

$\det = 1$

Since $1 = -\frac{1}{4}\det(A)$, we have $\det(A) = -4$

Note also — for each of the row operations, the effect on the determinant is the same as the determinant of the corresponding elementary matrix. So,

Thm: For any elementary matrix E and any matrix A,

$$\det(EA) = \det(E)\det(A)$$

By successive applications, we see that this also applies to products of elementary matrices.

This allows us to connect determinants to non-singularity and invertibility (we already know these last two are the same.)

Thm: A is nonsingular (invertible) $\iff$ $\det A \neq 0$

Pf: The row reduction of A can be represented by

$$EA = R$$

and thus

$$\det(E)\det(A) = \det(R)$$

($\Rightarrow$) If A is nonsingular, $R = I$ and so $\det(R) = 1 \neq 0$.
So $\det(A) \neq 0$.

($\Leftarrow$) If $\det(A) \neq 0$, of course so is $\det(E)$, so $\det(R) \neq 0$.
But this means R cannot have a row of zeroes; so, A is nonsingular.

This allows us to prove a very surprising result about determinants of products.

Thm:) $\det(AB) = \det(A)\det(B)$

Pf:) If A is invertible then $EA = I$, so $A = E^{-1}$ is a product of elementary matrices — for which we already know the result.

If A is not invertible, we know $\det(A) = 0$, so we need only show $\det(AB) = 0$...

In this case we have $EA = R$ where R has a row of zeroes. So $RB = EAB$ has a row of zeroes and thus determinant zero. But

$$0 = \det(EAB) = \det(E)\det(AB)$$

Since $\det(E)$ is not zero, we must have $\det(AB) = 0$.

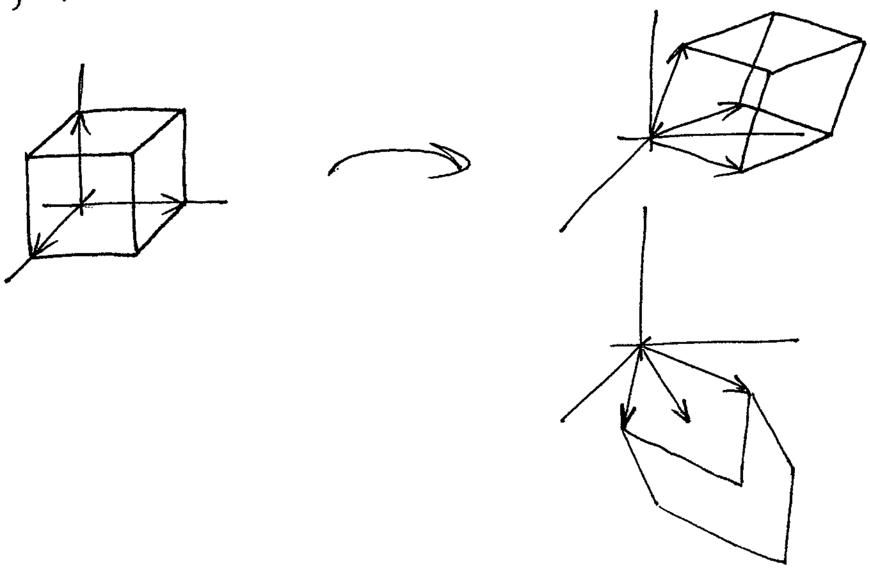
We then have the following result about inverse matrices.

Thm:) $\det(A^{-1}) = 1/\det(A)$

Pf:) $\det(A^{-1})\det(A) = \det(A^{-1}A) = \det(I) = 1$

Geometrically, we can see these results by interpreting matrices as functions $T: \mathbb{R}^n \to \mathbb{R}^n$ as previously discussed.

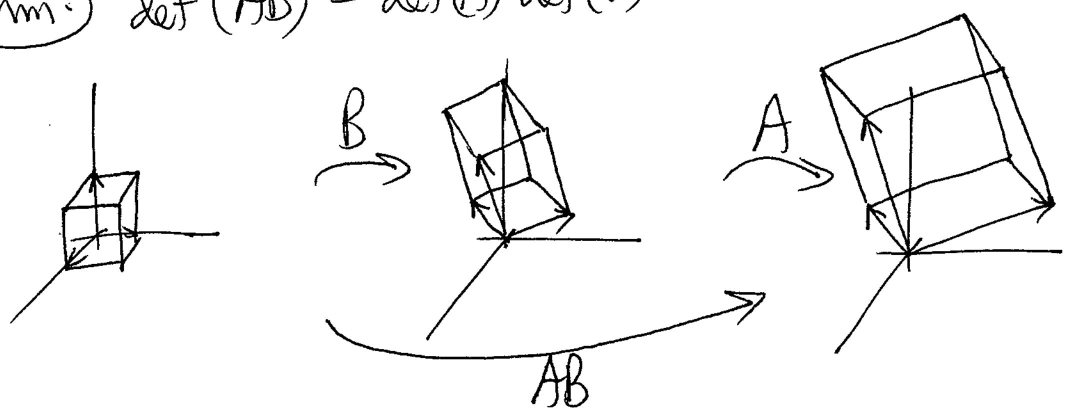
Thm: A invertible $\iff$ $\det A \neq 0$



det $\neq 0$
$\iff$ ||piped "full"
$\iff$ T 1-1/onto

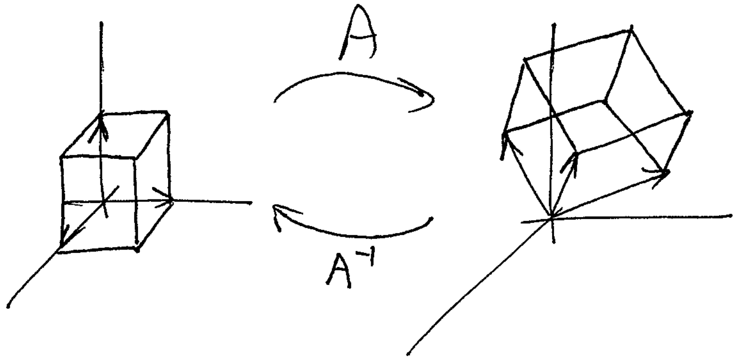
det $= 0$
$\iff$ ||piped "flat"
$\iff$ T not 1-1/onto

Thm: $\det(AB) = \det(A)\det(B)$



stretching factors multiply

Thm: $\det(A^{-1}) = 1/\det(A)$



The inverse has to "unstretch" by however much A stretches.

More Applications of Determinants

Thm: Consider the system $A\vec{x} = \vec{b}$, with A square + invtbl,

$$A = \begin{pmatrix} | & & | \\ \vec{a}_1 & \cdots & \vec{a}_n \\ | & & | \end{pmatrix} \quad \text{and} \quad \vec{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

Define A_i to be the matrix $A_i = \begin{pmatrix} | & & | & & | \\ \vec{a}_1 & \cdots & \vec{b} & \cdots & \vec{a}_n \\ | & & | & & | \end{pmatrix}$

$\vec{b}$ replaces ith col. of A

Then $$x_i = \frac{\det A_i}{\det A}$$

Pf: Note that $\vec{b} = x_1\vec{a}_1 + \dots + x_n\vec{a}_n$. By multilinearity of ith column then, we have

$$\det(A_i) = x_1 \det\begin{pmatrix} | & & | & & | \\ \vec{a}_1 & \cdots & \vec{a}_1 & \cdots & \vec{a}_n \\ | & & | & & | \end{pmatrix} + \dots + x_i \det\begin{pmatrix} | & & | & & | \\ \vec{a}_1 & \cdots & \vec{a}_i & \cdots & \vec{a}_n \\ | & & | & & | \end{pmatrix}$$
$$+ \dots + x_n \det\begin{pmatrix} | & & | & & | \\ \vec{a}_1 & \cdots & \vec{a}_n & \cdots & \vec{a}_n \\ | & & | & & | \end{pmatrix}$$

Most of the terms on the right are zero since there are repeated columns; only the ith term is nonzero, giving us

$$\det(A_i) = x_i \det(A)$$

and thus $$x_i = \frac{\det(A_i)}{\det(A)}$$

This result is called "Cramer's rule".

Cramer's rule is usually most useful when needing only one of the variables.

Def:) The cofactor matrix of A is the matrix C whose entries are the corresponding cofactors of A.

$$C = (c_{ij}) \quad , \quad c_{ij} = (-1)^{i+j} \det M_{ij}$$

Recall that the cofactor expansion of determinant is

$$\det A = \sum_{j=1}^{n} a_{ij} C_{ij}$$

Def:) The adjoint of A is the transpose of the cofactor matrix.

$$\operatorname{adj}(A) = C^T$$

Thm:) $A^{-1} = \dfrac{\operatorname{adj}(A)}{\det(A)}$

Pf:) We compute $A \operatorname{adj}(A)$; the ij element is

$$\sum_{k=1}^{n} a_{ik} \left(\operatorname{adj}(A)\right)_{kj}$$

$$= \sum_{k} a_{ik} C_{jk}$$

We consider separately the cases $i=j$ and $i \neq j$.

$(i=j)$ The ij element of the product becomes

$$\sum_k a_{ik} C_{ik}$$

This is exactly the cofactor expansion for $\det(A)$ along the ith row.

$(i \neq j)$ The ij element of the product becomes exactly the cofactor expansion along the jth row for the matrix A' obtained by replacing the jth row by the ith row.

$$A' = \begin{pmatrix} \vdots \\ \text{———} A_i \text{———} \\ \vdots \\ \text{———} A_i \text{———} \\ \vdots \end{pmatrix} \begin{matrix} \\ \leftarrow i\text{th row} \\ \\ \leftarrow j\text{th row} \\ \\ \end{matrix}$$

(because in the jth row, the cofactors of A' are the same as the cofactors of A.)

Given these above two results, we have

$$A \operatorname{adj}(A) = \det(A) I$$

so we conclude

$$A^{-1} = \frac{\operatorname{adj}(A)}{\det(A)}$$

2.3 (i) – Linear Independence & Span in $\mathbb{R}^n$

Def:) The <u>span</u> of a collection of vectors $\{\vec{v}_1, \dots, \vec{v}_n\}$ is the set of all linear combinations of those vectors.

Ex:) The span of $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$ is the xy-plane in $\mathbb{R}^3$.

Ex:) The span of $\left\{ \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \end{pmatrix} \right\}$ is all of $\mathbb{R}^2$.

Why? Show $c_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} + c_2 \begin{pmatrix} 3 \\ 2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ can be solved for c_1, c_2 for any $b_1, b_2 \dots$

This is a 2×2 system with $A = \begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix}$ (nonsing!).

<u>Terminology</u>: We say that $\{\vec{v}_1, \dots, \vec{v}_n\}$ "span" a set V in $\mathbb{R}^n$ if $V \subset \text{span}\{\vec{v}_1, \dots, \vec{v}_n\}$.

Def:) We say that the set of vectors $\{\vec{v}_1, \dots, \vec{v}_n\}$ ~~are~~ is <u>linearly dependent</u> if there is a nontrivial (coeffs not all zero) linear combination that equals zero.

That is, $\exists\ c_1, \dots, c_n$, not all zero, with

$$c_1\vec{v}_1 + \dots + c_n\vec{v}_n = \vec{0}$$

Def:) We say $\{\vec{v}_1, \dots, \vec{v}_n\}$ is <u>linearly independent</u> if not l.d.

Equiv.) $\{\vec{v}_1, \dots, \vec{v}_n\}$ are l.i. iff

$$\left(c_1\vec{v}_1 + \dots + c_n\vec{v}_n = \vec{0}\right) \Longrightarrow \left(c_1, \dots, c_n = 0\right)$$

Ex:) ~~Are~~ Is $\left\{\begin{pmatrix}2\\5\end{pmatrix}, \begin{pmatrix}3\\1\end{pmatrix}\right\}$ l.d. or l.i.?

We consider $c_1\begin{pmatrix}2\\5\end{pmatrix} + c_2\begin{pmatrix}3\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$

$$2c_1 + 3c_2 = 0$$
$$5c_1 + 1c_2 = 0$$

$$\begin{pmatrix}2 & 3\\5 & 1\end{pmatrix}\begin{pmatrix}c_1\\c_2\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

The matrix $\begin{pmatrix}2 & 3\\5 & 1\end{pmatrix}$ is nonsingular $(\det = -13 \neq 0)$, so we have uniqueness, so $c_1 = 0$ $c_2 = 0$ is the only sol. So, the set is <u>l.i.</u>.

Thm:) $\left(\{\vec{v}_1, \dots, \vec{v}_n\} \text{ is l.d.}\right) \Longleftrightarrow$ (one of the vectors is a l.c. of the others)

Pf:) ($\Longleftarrow$) If $\vec{v}_i = c_1\vec{v}_1 + \dots \quad \dots + c_n\vec{v}_n$ (no ith term!), then

$$c_1\vec{v}_1 + \dots + (-1)\vec{v}_i + \dots + c_n\vec{v}_n = \vec{0}$$

This is a nontriv. l.c. $(c_i = -1 \neq 0)$, so vectors are l.d.

($\Rightarrow$) Say $c_1\vec{v}_1 + \dots + c_i\vec{v}_i + \dots + c_n\vec{v}_n = \vec{0}$, with $c_i \neq 0$. Then we can solve for $\vec{v}_i$:

$$\vec{v}_i = \left(-\frac{c_1}{c_i}\right)\vec{v}_1 + \dots \qquad \dots + \left(-\frac{c_n}{c_i}\right)\vec{v}_n$$

(no ith term)

So $\vec{v}_i$ is a l.c. of the others.

Note, in the above theorem, the l.c. does <u>not</u> have to be nontrivial!

Ex: $\left\{\begin{pmatrix}0\\0\end{pmatrix}, \begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}\right\}$ is l.d. b/c

$$1\begin{pmatrix}0\\0\end{pmatrix} + 0\begin{pmatrix}1\\0\end{pmatrix} + 0\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$$

nontrivial!

Alt: $\begin{pmatrix}0\\0\end{pmatrix} = 0\begin{pmatrix}1\\0\end{pmatrix} + 0\begin{pmatrix}0\\1\end{pmatrix}$

trivial!! Okay!

Because there is only one condition to check, the definition is usually easier to work with than this result. But, this result is great for intuition, and sometimes useful in proofs.

Thm! $\left(\{\vec{v}_1, \dots, \vec{v}_n\} \text{ is l.d.}\right)$

iff

$\left(\begin{array}{l} \text{there is a vector } \vec{v}_i \text{ such that} \\ \operatorname{span}\{\vec{v}_1, \dots, \vec{v}_n\} = \operatorname{span}\{\vec{v}_1, \dots \underset{\text{(no ith vector)}}{\,} \dots, \vec{v}_n\} \end{array}\right)$

Pf. ($\Rightarrow$) If $\{\vec{v}_1, \dots, \vec{v}_n\}$ is l.d., then for some $\vec{v}_i$,

$$\vec{v}_i = d_1\vec{v}_1 + \dots \underset{\text{(no ith term)}}{\,} \dots + d_n\vec{v}_n$$

Then
$$\begin{aligned} c_1\vec{v}_1 + \dots + c_n\vec{v}_n &= c_1\vec{v}_1 + \dots + c_i\left(d_1\vec{v}_1 + \dots \underset{\text{(no ith term)}}{\,} \dots + d_n\vec{v}_n\right) \\ &\quad + \dots + c_n\vec{v}_n \\ &= (c_1 + c_i d_1)\vec{v}_1 + \dots \underset{\text{(no ith term)}}{\,} \dots + (c_n + c_i d_n)\vec{v}_n \end{aligned}$$

So every l.c. of $\{\vec{v}_1, \dots, \vec{v}_n\}$ can be rewritten without $\vec{v}_i$. So, $\operatorname{span}\{\vec{v}_1, \dots, \vec{v}_n\} = \operatorname{span}\{\vec{v}_1, \dots \underset{\text{(no ith vector)}}{\,} \dots, \vec{v}_n\}$

($\Leftarrow$) If the spans are equal, then

$$\vec{v}_i = 0\vec{v}_1 + \dots + 1\vec{v}_i + \dots + 0\vec{v}_n$$

is in the span with $\vec{v}_i$ removed. So $\vec{v}_i$ is a l.c. of the other vectors, and so $\{\vec{v}_1, \dots, \vec{v}_n\}$ is l.d.

2.1 — Vector Spaces

Students are already familiar with Euclidean vector spaces like $\mathbb{R}^2$ and $\mathbb{R}^3$.

Here we will generalize this idea to some extent.

Def:) A vector **space** (over the reals) is a set V, with an operation of addition and an operation of (real) scalar multiplication, such that:

1. $u + v = v + u$ for all $u, v \in V$.
2. $u + (v+w) = (u+v) + w$ . . . .
3. There is a "0", with $0 + v = v$ for all $v \in V$.
4. For each $v \in V$ there is a "$-v$", with $v + (-v) = 0$.
5. $c(u+v) = cu + cv$ for all $u, v \in V$, $c \in \mathbb{R}$
6. $c(dv) = (cd)v$ for all $v \in V$, $c, d \in \mathbb{R}$
7. $(c+d)v = cv + dv$ . . . .
8. $1 \cdot v = v$

The reals $\mathbb{R}$ are something called a "field". There is a more general thing called a "vector space over a field", defined for a general field. We will see some of these later.

We refer to the elements in a vector space as "vectors", even if they are not traditional vectors.

Ex:) $\mathbb{R}^n$ is a vector space, for any n.

Ex:) $M_{m\times n}(\mathbb{R})$, the set of $m \times n$ matrices with real entries, is a vector space. (Check!)

Ex:) $F(a,b)$, the set of real valued fns on (a,b), is a vector space, using

$f+g$ is defined by $$(f \overset{\text{vector addition}}{+} g)(x) = f(x) \overset{\text{real addition}}{+} g(x)$$

cf . . . . . $$(\underset{\text{scalar vector prod.}}{c} f)(x) = \underset{\text{real multiplication.}}{c}\, f(x)$$

Ex:) $W = \{ \vec{x} \in \mathbb{R}^3 \mid x+y+z = 1 \}$ is <u>not</u> a vector space with the usual addition.

In fact, the usual vector addition on $\mathbb{R}^3$ is not even an addition operation on this W:

these are in W $\longrightarrow$ $\begin{pmatrix}1\\0\\0\end{pmatrix} + \begin{pmatrix}0\\1\\0\end{pmatrix} = \begin{pmatrix}1\\1\\0\end{pmatrix}$ $\longleftarrow$ this is <u>not</u> in W.

Ex: $V = \left\{ \vec{x} \in \mathbb{R}^3 \;\middle|\; x+y+z=1 \right\}$, with operations defined by:

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} x_1+x_2-1 \\ y_1+y_2 \\ z_1+z_2 \end{pmatrix}$$

$$c \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} cx - c + 1 \\ cy \\ cz \end{pmatrix}$$

is a vector space.

$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ is the "0" vector

$\begin{pmatrix} 2-x \\ -y \\ -z \end{pmatrix}$ is the "$-\vec{x}$" vector

(Check other conditions!)

Facts:
1. The zero vector of V is unique.
2. The negative of a vector $v \in V$ is unique.
3. $0 \cdot \vec{v} = \vec{0}$
4. $c \cdot \vec{0} = \vec{0}$
5. $(-1)\vec{v} = -\vec{v}$

2.2 - Subspaces

Def: $W \subset V$ is a <u>subspace</u> of V if W, using the same operations as V, is a vector space.

Ex: $\left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \middle| \; x, y \in \mathbb{R} \right\}$ is a subspace of $\mathbb{R}^3$. (Check)

Note that in order for W to be a subspace of V, we need for the operations on V to be operations on W.

That is, we need

$$\vec{w}_1, \vec{w}_2 \in W \implies \vec{w}_1 + \vec{w}_2 \in W$$
$$\vec{w} \in W \implies c\vec{w} \in W$$

<u>Terminology</u>: W must be "closed under addition" and "closed under scalar multiplication".

It turns out that this is enough!

Thm: If W is a nonempty subset of V that is closed under addition and scalar multiplication, then W is a subspace of V.

Pf:) Most of the conditions are on the operations — and thus they are already satisfied because V is a vector space.

In addition, we can confirm that $\vec{0} \in W$ because W is closed under scalar multiplication, and

$$\vec{0} = 0\vec{w}$$

where $\vec{w}$ is any element of the (nonempty) set W.

We know $\vec{w}$ has an additive inverse, namely

$$-\vec{w} = (-1)\vec{w}$$

Ex:) $\left\{ f \in F(a,b) \;\middle|\; f(m) = 0 \quad \left(\text{where } m = \frac{a+b}{2}\right) \right\}$

This is a subspace of $F(a,b)$ because

$$f(m) = 0, \; g(m) = 0 \quad \Rightarrow \quad (f+g)(m) = f(m) + g(m) = 0 \checkmark$$

$$f(m) = 0 \quad \Rightarrow \quad (cf)(m) = c \cdot f(m) = 0 \checkmark$$

So this set is closed under both operations, and so it is a subspace.

Ex: Let $C(a,b) = \{ \text{continuous fns on } (a,b) \}$

$D(a,b) = \{ \text{diff'bl fns on } (a,b) \}$

We know every diff'bl fn is continuous; and $D(a,b)$ is closed under addition and scalar mult.

So $D(a,b)$ is a subspace of $C(a,b)$

Ex: $D^n(a,b) = \{ f \mid f \text{ has an } n\text{th deriv.} \}$

$C^n(a,b) = \{ f \mid f^{[n]} \text{ is continuous} \}$

These are all subspaces of $F(a,b)$, and

$$C = C^0 \supset D^1 \supset C^1 \supset D^2 \supset C^2 \supset \cdots$$

Ex: $C^\infty = \{ f \mid f \text{ has } \infty\text{'ly many derivatives} \}$ is a subspace of <u>all</u> of the above.

Ex: $P = \{ \text{all polynomials} \}$ is a subspace of C^∞.

Ex: For any set of vectors $\{\vec{v}_1, \dots, \vec{v}_n\} \in V$, the span,

$$\text{span}\{\vec{v}_1, \dots, \vec{v}_n\}$$

is a subspace of V.

Ex: For any matrix A, the set of homogeneous solutions

$$\{\vec{x} \mid A\vec{x} = \vec{0}\}$$

is a subspace of $\mathbb{R}^n$.

Check: $A(\vec{x}_1 + \vec{x}_2) = A\vec{x}_1 + A\vec{x}_2 = \vec{0} + \vec{0} = \vec{0}$ ✓

$A(c\vec{x}) = cA\vec{x} = c\vec{0} = \vec{0}$ ✓

So the set is closed under addition and scalar multiplication.

2.3 (ii) – Linear Independence + Bases

Our previous definitions and results about linear dependence and span apply also to general vector spaces.

Ex: Are $p_1(x) = x^2 - 1$ and $p_2(x) = x + 5$ l.i. or l.d.?

Suppose $c_1 p_1 + c_2 p_2 = 0$; then

$$c_1(x^2 - 1) + c_2(x+5) = 0$$

$$(c_1)x^2 + (c_2)x + (5c_2 - c_1) = 0$$

$$\Rightarrow \begin{aligned} c_1 &= 0 \\ c_2 &= 0 \\ 5c_2 - c_1 &= 0 \end{aligned} \Rightarrow \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 5 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Rightarrow c_1 = 0 \text{ and } c_2 = 0.$$

So p_1, p_2 are l.i..

Def: The set of vector $\{\vec{v}_1, \dots, \vec{v}_n\}$ in V are a **basis** for V if both:

1. $\{\vec{v}_1, \dots, \vec{v}_n\}$ is l.i.

and

2. $\{\vec{v}_1, \dots, \vec{v}_n\}$ span V

Because they span V, we can say that every vector in V can be represented by the vectors in $\{\vec{v}_1, \dots, \vec{v}_n\}$

Because they are l.i., we cannot remove any of these vectors and still span. So, none of the vectors are "unnecessary".

Ex: Are $\left\{\begin{pmatrix} 3 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 5 \end{pmatrix}\right\}$ a basis for $\mathbb{R}^2$?

To see that they span, we need existence in

$$c_1\begin{pmatrix} 3 \\ 2 \end{pmatrix} + c_2\begin{pmatrix} 1 \\ 5 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \iff \begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$$

To see that they are l.i., we need uniqueness in

$$c_1\begin{pmatrix} 3 \\ 2 \end{pmatrix} + c_2\begin{pmatrix} 1 \\ 5 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \iff \begin{pmatrix} 3 & 1 \\ 2 & 5 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

The matrix is nonsingular, so we have both conditions.

Ex:) The "standard basis vectors" in $\mathbb{R}^n$

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \right\}$$

are a basis for $\mathbb{R}^n$.

Similarly to previous example, this is because

$$\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

is nonsingular.

Notation reminder:

$$\vec{e}_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$ ← 1 is in the ith entry

Ex:) Are $\left\{ \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \right\}$ a basis for $\{\vec{x} \mid x+y+z=0\}$?

For l.i., need uniqueness in $\begin{pmatrix} 1 & 1 \\ 2 & 1 \\ -3 & -2 \end{pmatrix}$. This works. (Check!)

For span, we do not need a pivot in every row though, because not all $\vec{b}$ are required...

Option 1: Check for existence in

$$c_1\begin{pmatrix}1\\2\\-3\end{pmatrix} + c_2\begin{pmatrix}1\\1\\-2\end{pmatrix} = \underbrace{\begin{pmatrix}b_1\\b_2\\-b_1-b_2\end{pmatrix}}_{\substack{\text{These are all of the vectors}\\ \text{in } \{x+y+z=0\}}}$$

$$\left(\begin{array}{cc|c}1 & 1 & b_1\\ 2 & 1 & b_2\\ -3 & -2 & -b_1-b_2\end{array}\right)$$

Option 2: Note solutions to $x+y+z=0$ are

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

So $\begin{pmatrix}-1\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\0\\1\end{pmatrix}$ span the space.

If these are l.c.'s of $\begin{pmatrix}1\\2\\-3\end{pmatrix}, \begin{pmatrix}1\\1\\-2\end{pmatrix}$, then these span the space too.

So, need to check for existence in

$$\left(\begin{array}{cc|c}1 & 1 & -1\\ 2 & 1 & 1\\ -3 & -2 & 0\end{array}\right) \text{ and } \left(\begin{array}{cc|c}1 & 1 & -1\\ 2 & 1 & 0\\ -3 & -2 & 1\end{array}\right)$$

We will soon find an easier way to answer these kinds of questions.

Thm:) $\{\vec{v}_1, \ldots, \vec{v}_n\}$ in V are a basis for V

iff

each vector in $\vec{v}$ is uniquely expressible as a l.c. of $\{\vec{v}_1, \ldots, \vec{v}_n\}$ in V.

Pf:) ($\Rightarrow$) We assume $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a basis for V, so we know we have span and l.i.

Because they span V, we know we can write

$$\vec{v} = c_1\vec{v}_1 + \ldots + c_n\vec{v}_n$$

for any $\vec{v} \in V$.

If we can also write

$$\vec{v} = d_1\vec{v}_1 + \ldots + d_n\vec{v}_n$$

then $\vec{v} - \vec{v} = (c_1 - d_1)\vec{v}_1 + \ldots + (c_n - d_n)\vec{v}_n = \vec{0}$

Because of l.i., this means $c_1 = d_1, \ldots, c_n = d_n$.

So the expression is unique.

($\Leftarrow$) If every vector is uniquely expressible, then certainly every vector is expressible, so $\{\vec{v}_1, \ldots, \vec{v}_n\}$ span V.

By uniqueness, we have

$$c_1\vec{v}_1 + \ldots + c_n\vec{v}_n = \vec{0} \quad \Longrightarrow \quad c_1 = 0, \ldots, c_n = 0.$$

So the vectors are l.i., and thus are a basis.

So a basis is not only an "efficient" way to represent elements of a vector space (can't remove vectors from a basis), it is also a way to do so unambiguously

Def: Say $\alpha = \{\vec{v}_1, \dots, \vec{v}_n\}$ is a basis for V, and $\vec{v} \in V$. Then the coordinates of $\vec{v}$ relative to the basis α are

$$[\vec{v}]_\alpha = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

where $\vec{v} = c_1\vec{v}_1 + \dots + c_n\vec{v}_n$

<u>Note</u>: What we usually refer to as the "coordinates" of a vector $\vec{v}$ are the "coordinates relative to the standard basis, because, for example,

$$\begin{pmatrix} a \\ b \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \end{pmatrix} + b\begin{pmatrix} 0 \\ 1 \end{pmatrix} = a\vec{e}_1 + b\vec{e}_2$$

Ex:) Let β be the basis $\{x^2-x+1,\ x-1,\ 1\}$ for the vector space of polynomials of degree ≤ 2.

Say f is the vector $f(x) = 3x^2+2x+5$.

What is $[f]_\beta$?

We need to write

$$c_1\left(x^2-x+1\right) + c_2\left(x-1\right) + c_3\left(1\right) = 3x^2+2x+5$$

$$\begin{aligned} c_1 - c_2 + c_3 &= 5 &&\text{(const. terms)} \\ -c_1 + c_2 &= 2 &&\text{(x terms)} \\ c_1 &= 3 &&\text{(}x^2\text{ terms)} \end{aligned}$$

The unique solution is $c_1 = 3$, $c_2 = 5$, $c_3 = 7$

2.4 — Dimension & Matrix Subspaces

Dimension

Thm: If $\{\vec{v}_1, \dots, \vec{v}_n\}$ is a basis for a v.s. V, then any set $\{\vec{w}_1, \dots, \vec{w}_m\}$ of $m > n$ vectors is l.d.

Pf: Since $\{\vec{v}_1, \dots, \vec{v}_n\}$ is a basis, we can write each $\vec{w}_i$ as a l.c.:

$$
\begin{array}{cccc}
\vec{w}_1 & \vec{w}_2 & & \vec{w}_m \\
\parallel & \parallel & \cdots & \parallel \\
a_{11}\vec{v}_1 & a_{12}\vec{v}_1 & & a_{1m}\vec{v}_1 \\
+ & + & & + \\
a_{21}\vec{v}_2 & a_{22}\vec{v}_2 & & a_{2m}\vec{v}_2 \\
+ & + & & + \\
\vdots & \vdots & & \vdots \\
+ & + & & + \\
a_{n1}\vec{v}_n & a_{n2}\vec{v}_n & & a_{nm}\vec{v}_n
\end{array}
$$

Let A be the matrix of these coefficients:

$$
A = \begin{pmatrix}
a_{11} & a_{12} & \cdots & a_{1m} \\
a_{21} & a_{22} & \cdots & a_{2m} \\
\vdots & \vdots & & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nm}
\end{pmatrix}
$$

Lemma:
$$c_1\vec{w}_1 + \cdots + c_m\vec{w}_m = d_1\vec{v}_1 + \cdots + d_n\vec{v}_n$$
iff
$$A\vec{c} = \vec{d}$$

Pf of Lemma:
$$c_1\vec{w}_1 + \cdots + c_m\vec{w}_m$$
$$= c_1\begin{pmatrix} a_{11}\vec{v}_1 \\ + \\ a_{21}\vec{v}_2 \\ + \\ \vdots \\ + \\ a_{n1}\vec{v}_n \end{pmatrix} + c_2\begin{pmatrix} a_{12}\vec{v}_1 \\ + \\ a_{22}\vec{v}_2 \\ + \\ \vdots \\ + \\ a_{n2}\vec{v}_n \end{pmatrix} + \cdots + c_m\begin{pmatrix} a_{1m}\vec{v}_1 \\ + \\ a_{2m}\vec{v}_2 \\ + \\ \vdots \\ + \\ a_{nm}\vec{v}_n \end{pmatrix}$$
$$\begin{aligned} = \ & (c_1a_{11} + c_2a_{12} + \cdots + c_ma_{1m})\vec{v}_1 \\ + \ & (c_1a_{21} + c_2a_{22} + \cdots + c_ma_{2m})\vec{v}_2 \\ & \vdots \\ + \ & (c_1a_{n1} + c_2a_{n2} + \cdots + c_ma_{nm})\vec{v}_n \end{aligned} = \begin{matrix} d_1\vec{v}_1 \\ + \\ d_2\vec{v}_2 \\ + \\ \vdots \\ + \\ d_n\vec{v}_n \end{matrix}$$

So
$$\begin{aligned} a_{11}c_1 + a_{12}c_2 + \cdots + a_{1m}c_m &= d_1 \\ a_{21}c_1 + a_{22}c_2 + \cdots + a_{2m}c_m &= d_2 \\ &\vdots \\ a_{n1}c_1 + a_{n2}c_2 + \cdots + a_{nm}c_m &= d_n \end{aligned}$$

and thus $A\vec{c} = \vec{d}$

(back to main proof...)

If $m>n$, then A has more columns than rows. So,

$$A\vec{c} = \vec{0}$$

has a nontrivial solution. By the lemma, this means there are nontrivial $c_1, \dots, c_m$ with

$$c_1\vec{w}_1 + \cdots + c_m\vec{w}_m = 0\vec{v}_1 + \cdots + 0\vec{v}_n = \vec{0}$$

So $\{\vec{w}_1, \dots, \vec{w}_m\}$ is l.d. $\blacksquare$

This theorem allows us to prove another important result.

Thm: If $\{\vec{v}_1, \dots, \vec{v}_n\}$, $\{\vec{u}_1, \dots, \vec{u}_m\}$ are both bases for V, then $n=m$.

Pf: If $m>n$, the previous thm implies $\{\vec{u}_1, \dots, \vec{u}_m\}$ l.d. $\times$

If $n>m$, ... $\{\vec{v}_1, \dots, \vec{v}_n\}$ l.d. $\times$

So we must have $n=m$.

So for any vector space, the number of elements in a basis is the same for every basis.

That is — this number is an <u>invariant</u>. It is intrinsic to the space V, not just an individual basis.

Def: If V has a basis with n elements, then we say the <u>dimension</u> of V is n, and write

$$\dim(V) = n$$

Ex: $\dim(\mathbb{R}^n) = n$ (consider standard basis).

Ex: $\dim(M_{m \times n}) = mn$ (find a basis!)

Nontrivial vector spaces with no finite basis are called "infinite dimensional".

Ex: $C^0(a,b)$ is infinite dimensional.

Facts proved in the book:

① If $\{\vec{v}_1, \dots, \vec{v}_k\}$ is l.i. in V with $\dim(V) = n$ then one can find vectors $\vec{v}_{k+1}, \dots, \vec{v}_n$ with $\{\vec{v}_1, \dots, \vec{v}_n\}$ a basis for V.

② If $\{\vec{v}_1, \dots, \vec{v}_k\}$ span V, then there is a subset that forms a basis for V.

③ If $\dim(V) = n$ and $\{\vec{v}_1, \dots, \vec{v}_n\}$ is l.i., then it is a basis.

④ If $\dim(V) = n$ and $\{\vec{v}_1, \dots, \vec{v}_n\}$ spans V, then it is a basis.

Ex: $\{x^2 - 2x + 7, \ x+1, \ 3\}$ is l.i. in P_2 and $\dim(P_2) = 3$, so this is a basis for P_2.

Nullspace

(sometimes called the "kernel")

Def: The <u>nullspace</u> of a matrix A is the set of homogeneous solutions

$$NS(A) = \{ \vec{x} \mid A\vec{x} = \vec{0} \}$$

We already know how to solve systems; but in fact this method also gives us a <u>basis</u> for NS.

Ex: Suppose A has rref given by

$$\text{rref}(A) = \begin{pmatrix} 1 & 3 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Then the homogeneous solutions are

$$\vec{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -3y \\ y \\ 0 \end{pmatrix} = y \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix}$$

This is a 1-dim. sol. set, with basis $\left\{ \begin{pmatrix} -3 \\ 1 \\ 0 \end{pmatrix} \right\}$

Ex: Suppose A has rref given by

$$\text{rref}(A) = \begin{pmatrix} 1 & 4 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

Then the homogeneous solutions are given by

$$\vec{x} = \begin{pmatrix} x \\ y \\ z \\ w \end{pmatrix} = \begin{pmatrix} -4y-2w \\ y \\ -3w \\ w \end{pmatrix} = y\begin{pmatrix} -4 \\ 1 \\ 0 \\ 0 \end{pmatrix} + w\begin{pmatrix} -2 \\ 0 \\ -3 \\ 1 \end{pmatrix}$$

This is a 2-dim. sol. set, with basis $\left\{ \begin{pmatrix} -4 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ -3 \\ 1 \end{pmatrix} \right\}$

Note, these basis vectors are independent because of the 1's in the positions corresponding to the corresponding free variables, with 0's in the other free variable positions.

Row Space

Def: The <u>row space</u> of A is the span of the row vectors of A. It is written $RS(A)$.

Thm: If E is an elementary matrix, then

$$RS(A) = RS(EA)$$

Pf: Every row of each matrix is a l.c. of rows of the other matrix, so their spans are equal.

Given the above thm, we can also conclude

$$RS(A) = RS(\text{rref}(A))$$

Conveniently, the non zero rows of $\text{rref}(A)$ are a basis for $RS(\text{rref}(A))$. So, they are also a basis for $RS(A)$.

Ex: $A = \begin{pmatrix} 1 & 2 & -1 & 3 & 0 \\ 1 & 1 & 0 & 4 & 1 \\ 1 & 4 & -3 & 1 & -2 \end{pmatrix} \Rightarrow \text{rref}(A) = \begin{pmatrix} 1 & 0 & +1 & 5 & 2 \\ 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

Clearly $\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 5 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \\ -1 \\ -1 \end{pmatrix} \right\}$ is a basis for $RS(\text{rref}(A))$, so they are also for $RS(A)$.

Note, we have a basis vector for $RS(A)$ for every pivot...

And we had a basis vector for $NS(A)$ for every free variable...

So, $\dim(NS(A)) + \dim(RS(A)) = \#$ cols of A

We also have the following:

Thm: $RS(A) \perp NS(A)$

Pf: If $\vec{r} \in RS(A)$, then $\vec{r} = c_1\vec{r}_1 + \cdots + c_k\vec{r}_k$

If $\vec{n} \in NS(A)$, then $A\vec{n} = 0$, so $\vec{r}_i \cdot \vec{n} = 0$ for all i.

Then we can compute $\vec{r} \cdot \vec{n}$ by

$$
\begin{aligned}
\vec{r} \cdot \vec{n} &= (c_1\vec{r}_1 + \cdots + c_k\vec{r}_k) \cdot \vec{n} \\
&= c_1(\vec{r}_1 \cdot \vec{n}) + \cdots + c_k(\vec{r}_k \cdot \vec{n}) \\
&= 0
\end{aligned}
$$

So every vector in $RS(A)$ is $\perp$ to every vector in $NS(A)$.

It can also be shown that

① $NS(A) = RS(A)^{\perp}$ (the set of all vectors $\perp$ to the row space.)

② $RS(A) = NS(A)^{\perp}$ (the set of all vectors $\perp$ to the null space.)

③ If $\{\vec{v}_1, \dots, \vec{v}_k\}$ is a basis for $RS(A)$ and $\{\vec{w}_1, \dots, \vec{w}_{n-k}\}$ is a basis for $NS(A)$, then $\{\vec{v}_1, \dots, \vec{v}_k, \vec{w}_1, \dots, \vec{w}_{n-k}\}$ is a basis for $\mathbb{R}^n$.

Column Space

Def: The <u>column space</u> of A is the span of the column vectors of A. It is written $CS(A)$.

Thm: If E is an elementary matrix, then any relation between the column vectors of A is also a relation between the column vectors of EA.

Pf: A relation is a nontrivial l.c. that equals zero; interpreting matrix-vector products as a l.c. of cols, we can write the relations in question as

$$A\vec{c} = \vec{0} \quad \text{and} \quad (EA)\vec{c} = \vec{0}$$

Since every elementary matrix has uniqueness,

$$(EA)\vec{c} = \vec{0} \iff E(A\vec{c}) = \vec{0} \iff A\vec{c} = \vec{0}$$

Said differently, this theorem says that row operations "preserve relations between column vectors".

Note, the pivot columns of rref(A) are a basis for rref(A).

<u>Ex:</u>
$$\begin{pmatrix} 1 & 0 & 1 & 5 & 2 \\ 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \text{rref}(A)$$

$\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \right\}$ are a basis for $CS(\text{rref}(A))$

This is because they are <u>independent</u>
(this is a statement about rels. btw col. vectors)
and they <u>span</u> the entire column space,
(this is also a statement about rels btw col. vects)

Combining the theorem with this observation about bases for $CS(\text{rref}(A))$, we get a result about bases for $CS(A)$.

<u>Thm:</u> The columns of A corresponding to pivot columns in rref(A) form a basis for $CS(A)$.

Ex: $A = \begin{pmatrix} 1 & 2 & -1 & 3 & 0 \\ 1 & 1 & 0 & 4 & 1 \\ 1 & 4 & -3 & 1 & -2 \end{pmatrix}$

$\text{rref}(A) = \begin{pmatrix} 1 & 0 & 1 & 5 & 2 \\ 0 & 1 & -1 & -1 & -1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

There are pivots in the first two columns of rref(A),
so $\left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix} \right\}$ is a basis for $CS(A)$.

Alt: Note $CS(A) = RS(A^T)$. So you can also get a basis for $CS(A)$ by row reducing A^T... but, this involves doing a different row reduction.

Thm: $\dim(RS(A)) = \text{rank}(A) = \dim(CS(A))$

Wronskians

How can you tell if functions (vectors in $F(a,b)$) are l.i. or l.d.?

With polynomials this is easy because we have a finite basis, and can deal with coefficients.

If the functions are not polynomials, this won't work.

Consider $f_1, \dots, f_n \in D^{n-1}(a,b)$. We need to consider possible solutions to

$$c_1 f_1(x) + \dots + c_n f_n(x) = 0$$

(Note, on the right side is the zero function — which has value 0 <u>for all</u> $x \in (a,b)$.)

We consider the following argument. Suppose $f_1, \dots, f_n$ are <u>dependent</u>:

$$\{f_1, \dots, f_n\} \text{ is l.d.} \implies c_1 f_1 + \dots + c_n f_n = 0 \quad \left(\begin{matrix}\text{nontriv.}\\ c_i\end{matrix}\right)$$

$$\implies c_1 f_1' + \dots + c_n f_n' = 0$$

$$\vdots$$

$$\implies c_1 f_1^{[n-1]} + \dots + c_n f_n^{[n-1]} = 0$$

$$\Rightarrow \begin{pmatrix} f_1(x) & \cdots & f_n(x) \\ f_1'(x) & & f_n'(x) \\ \vdots & & \vdots \\ f_1^{[n-1]}(x) & \cdots & f_n^{[n-1]}(x) \end{pmatrix} \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

has nontrivial solutions.

This is a square matrix, so nontrivial solutions means that its determinant must be zero.

$$\Rightarrow \det \begin{pmatrix} f_1(x) & \cdots & f_n(x) \\ \vdots & & \vdots \\ f_1^{[n-1]}(x) & \cdots & f_n^{[n-1]}(x) \end{pmatrix} = 0$$

We call this determinant the <u>Wronskian</u> of $\{f_1, \dots, f_n\}$.

Note that the Wronskian is a <u>function of x</u>.

Thm: $\{f_1, \dots, f_n\}$ is l.d. $\Rightarrow$ $W(x) = 0$ <u>for all x</u>

Thm: If $W(x) \neq 0$ for <u>any</u> x, then $\{f_1, \dots, f_n\}$ is l.i.

Ex: Is $\{x, x^2, x^3\}$ l.i.?

$$W(x) = \det\begin{pmatrix} x & x^2 & x^3 \\ 1 & 2x & 3x^2 \\ 0 & 2 & 6x \end{pmatrix} = x(6x^2) - 1(4x^3)$$

$$= 2x^3$$

This happens to be zero for some x, but the point is that this is not the zero function.

So, these functions are l.i.

WARNING: $W(x) = 0 \not\Rightarrow \{f_1, \dots, f_n\}$ l.d.

The argument $(\text{l.d.} \Rightarrow W = 0)$ appears on first glance to be reversible, but it is not. Specifically:

$$\left(\begin{array}{c} A\vec{c} = \vec{0} \quad (\text{for all } x) \\ \text{has a nontriv. sol'n } \vec{c} \end{array}\right) \Longrightarrow \left(\det(A) = 0 \quad (\text{for all } x)\right)$$

This step is not reversible, because $\det(A) = 0$ implies only that for every $\vec{x}$ there is a nontrivial $\vec{c}$ — not that the same $\vec{c}$ will work for each $\vec{x}$.

The book shows a nice counterexample.

In a restricted category though, the result does reverse.

Def:) A function $f: \mathbb{R}^1 \to \mathbb{R}^1$ is analytic (or, "real analytic") if its Taylor series converges to itself.

Ex:) In single variable calculus courses it is shown that e^x, $\sin x$, $\cos x$ all have Taylor series that converge to themselves. So, these functions are analytic.

Fact:) Sums, products, compositions, fractions (with nonzero denoms!) of analytic functions are analytic.

Thm:) If $f_1, \dots, f_n$ are analytic and the Wronskian is identically zero, then $\{f_1, \dots, f_n\}$ is l.d..

3.1–3.4 Differential Equations, Intro/review

Students should already have seen some differential equations in Math 32 (or equiv.).

Ex: $\frac{dy}{dx} = f(x)$; $\frac{dy}{dx} = ky$; $\frac{dy}{dx} = \frac{f(x)}{g(y)}$

Note, DE's do not always have solutions. And they do not always have unique solutions even with an initial condition...

Ex: Consider the initial value problem

$$\frac{dy}{dx} = 3y^{2/3}, \quad y(0) = 0$$

Note that $y(x) = 0$ is a sol'n. But, so is $y(x) = x^3$!

Under some conditions, we can ensure this does not happen.

Thm: Consider the IVP: $\frac{dy}{dx} = f(x,y) \quad y(x_0) = y_0$

If f <u>and</u> $\frac{\partial f}{\partial y}$ are continuous near (x_0, y_0), then, sufficiently near x_0, the IVP has a unique solution.

Some techniques:

Separable: Some DE's can be put into the form

$$g(y)\,dy = f(x)\,dx$$

Integrate both sides and solve.

Exact: Some DE's can be put into the form

$$\frac{\partial F(x,y)}{\partial x} + \frac{\partial F(x,y)}{\partial y}\frac{dy}{dx} = 0$$

Rewrite as $\quad \frac{d}{dx} F(x,y) = 0$

and integrate to get $\quad F(x,y) = C$

Ex: $\quad (2x\cos y) - (x^2\sin y + 2y)\frac{dy}{dx} = 0$

This is exact, with $F(x,y) = x^2\cos y - y^2$

Integrating Factor: If we have $y' + p(x)y = q(x)$,

multiply both sides by $\mu(x) = e^{\int p(x)\,dx}$

(note that this gives us $\mu'(x) = \mu(x)\,p(x)$

We get

$$\mu y' + \mu p y = \mu q$$

$$\mu y' + \mu' y = \mu q$$

$$(\mu y)' = \mu q$$

$$\vdots$$

3.6 – Modeling with DE's

Many real situations can be modeled with DE's. These are some of the simplest.

Ex:) An "interest rate" is really a factor multiplied by principal to give the rate that interest is paid.

For example, if you have \$1,000 in an account, with an interest rate of 6%/yr, then at <u>that instant</u>, the rate of change of your balance is $(6\%/\text{yr})(\$1{,}000) = \$60/\text{yr}$. This is an <u>instantaneous rate</u>.

This generalizes to the DE:

$$\frac{dB}{dt} = rB \quad \leftarrow \left(\begin{matrix}\text{Natural Growth}\\ \text{Equation}\end{matrix}\right)$$

With an initial balance $B(0) = B_0$, the solution is

$$B(t) = B_0 e^{rt}$$

(sometimes this is referred to as "continuous compounding", but it is better to think of this as simply a proper interpretation of rate.)

Ex: Suppose you have an interest rate r and an additional saving rate of $\$S/\text{yr}$. Then we have

$$\frac{dB}{dt} = rB + S$$

Trick: let $z = rB + S$, so $z' = rB'$.

Then $z' = rz$, which is N.G.E.

Ex: Say we have a (full) 1000 g tank. 5 g/min flows in, mixes perfectly, and flows out at the same rate. The input has 3 lbs/gal concentration of salt. How does the concentration in the tank change over time?

Let $C(t)$ = concentration, and $Q(t)$ = quantity of salt.

Then
$$\frac{dQ}{dt} = \left(\frac{3\text{ lbs}}{\text{gal}}\right)\left(\frac{5\text{ gal}}{\text{min}}\right) - \left(C(t)\right)\left(\frac{5\text{ gal}}{\text{min}}\right)$$

$$= 15\,\frac{\text{lbs}}{\text{min}} - \frac{5\,Q}{1000\text{ gal}}\,\frac{\text{gal}}{\text{min}}$$

Solve with same trick above, for $Q(t)$. Then $C = \frac{Q}{1000}$.

See also Beale's notes on first order DE's.

4.1 – Higher Order Linear DE's

Def: $$g_n(x)y^{[n]} + g_{n-1}(x)y^{[n-1]} + \dots + g_1(x)y' + g_0(x)y = g(x)$$

is an nth order linear DE.

– "<u>nth order</u>" because highest deriv. is nth

– "<u>linear</u>" because LHS is a linear operator on y. If we write the LHS as $L(y)$, note that

$$L(ay_1 + by_2) = aL(y_1) + bL(y_2)$$

– If $g(x)$ is the zero fn, then the DE is "<u>homogeneous</u>".

– If we prescribe values for $y, y', \dots, y^{[n-1]}$ at some value x_0, this becomes an "<u>initial value problem</u>".

Thm: If $g_0, g_1, \dots, g_{n-1}, g_n$ are continuous, g is continuous, and $g_n \neq 0$ for all x, then the IVP

$$g_n y^{[n]} + \dots + g_0 y = g(x)$$
$$y(x_0) = k_0, \quad \dots, \quad y^{[n-1]}(x_0) = k_{n-1}$$

has a unique solution.

(this thm is not proved in Math 107)

Thm: Consider the homogeneous linear DE

$$q_n y^{[n]} + \cdots + q_0 y = 0$$

with q_i all continuous, and $q_n \neq 0$ for all x.
The set of solutions is an n-dim. subspace of C^n.

Pf: If $y(x)$ is a solution it is clearly n-diff. And, solving for $y^{[n]}$ in the equation shows $y^{[n]}$ must be cont., so $y \in C^n$.

Let's rewrite the DE with a lin. diff. operator, as

$$L(y) = 0$$

If y_1 and y_2 are solutions, then

$$L(y_1 + y_2) = L(y_1) + L(y_2) = 0$$

so $y_1 + y_2$ is a solution. And

$$L(cy_1) = cL(y_1) = 0$$

so cy_1 is a solution.

So the set of solutions is a subspace of C^n.

To show the dimension is n, we will find a basis.

Consider the related IVP:

$$L(y) = 0$$

$$y(x_0) = k_0, \; y'(x_0) = k_1, \ldots, y^{[n-1]}(x_0) = k_{n-1}$$

(we will write $\vec{k} = (k_0, \ldots, k_{n-1})$ for convenience.)

Let y_1 be the unique solution with $\vec{k} = \vec{e}_1$

... y_2 - - - - - - - - - - - - $\vec{k} = \vec{e}_2$

$\vdots$

... y_n - - - - - - - - - - - - $\vec{k} = \vec{e}_n$

We will consider the collection of sols $\{y_1, \ldots, y_n\}$.

<u>Independence</u>: At x_0, the Wronskian is

$$W(x_0) = \det\begin{pmatrix} | & & | \\ \vec{e}_1 & \cdots & \vec{e}_n \\ | & & | \end{pmatrix} = 1 \neq 0$$

So these fns are independent.

<u>Span</u>: If f is a solution with $f(x_0) = c_1, \ldots, f^{[n-1]}(x_0) = c_n$, then it is the sol. to the IVP with $\vec{k} = \vec{c}$.

But $c_1 y_1 + \cdots + c_n y_n$ is also...

So every sol. to $L(y) = 0$ is a l.c. of $\{y_1, \ldots, y_n\}$.

■

A basis for the set of solutions to a hom. lin. DE is called a "<u>fundamental set of solutions</u>".

Ex:) Consider the DE.

$$y'' + y = 0$$

This is hom. and linear, and 2nd order.

So we know that the set of solutions is 2-dim.

Observe that $\sin x$ and $\cos x$ are solutions...

So, <u>every</u> solution is of the form

$$y = c_1 \sin x + c_2 \cos x$$

and $\{\sin x, \cos x\}$ is a fundamental set of sols.

If we have a nonhomogeneous linear DE, we only need to find <u>one</u> particular solution and the homogeneous solutions in order to have all of the solutions.

Thm:) Suppose y_p is a solution to the linear DE

$$L(y) = g(x)$$

Suppose that $\{y_1, \ldots, y_n\}$ are a fund. set of sols to the associated hom. linear DE

$$L(y) = 0$$

Then the general solution to the non hom. lin. DE is

$$y = y_p + (c_1 y_1 + \cdots + c_n y_n)$$

Pf:)
$$\begin{aligned} L(y) &= L(y_p) + c_1 L(y_1) + \ldots + c_n L(y_n) \\ &= g(x) + 0 + \ldots + 0 \\ &= g(x). \end{aligned}$$

So every function of this form is a solution.

On the other hand, if y is any solution, we have
$$L(y - y_p) = L(y) - L(y_p) = g(x) - g(x) = 0$$

So $y - y_p$ is a solution to the associated hom. lin. DE, and thus
$$y - y_p = c_1 y_1 + \cdots c_n y_n$$

So
$$y = y_p + (c_1 y_1 + \ldots + c_n y_n)$$

Ex: What is the complete set of solutions to the DE

$$y'' + y = 2e^x$$

First, note that $y_p = e^x$ is a solution.

Second, we already know from prev. example that $\{\sin x, \cos x\}$ are a fund. set of sols to the associated hom. DE

$$y'' + y = 0$$

So, every solution to the given equation is

$$y = e^x + C_1 \sin x + C_2 \cos x$$

Note that there is a very strong similarity between this result about solutions to nonhom. DE's and an old result about solutions to nonhom. systems of equations!

Recall that

$$\Big(w(x) \neq 0 \text{ for } \underline{\text{any}}\ x \Big) \Rightarrow \Big(\{f_1, \dots, f_n\} \text{ is l.i.} \Big)$$

But, unfortunately

$$\Big(w(x) = 0 \text{ for } \underline{\text{all}}\ x \Big) \not\Rightarrow \Big(\{f_1, \dots, f_n\} \text{ is l.d.} \Big)$$

It turns out that solutions to hom. lin. DE's are sufficiently nice that this actually works out for these fns.

In fact, we get an even stronger fact.

Thm:) Let $Y_1, \dots, Y_n$ be solutions to the nth order lin. hom. DE (satisfying the conditions of the uniqueness/exist. thm.)

$$L(Y) = 0$$

Then

$$\Big(w(x_0) = 0 \text{ for } \underline{\text{any}}\ x_0 \Big) \Rightarrow \Big(\{Y_1, \dots, Y_n\} \text{ is l.d.} \Big)$$

Pf:) If $w(x_0) = 0$, then there is a <u>nontrivial</u> sol'n $C_1, \dots, C_n$ to

$$C_1 Y_1(x_0) + \cdots + C_n Y_n(x_0) = 0$$

$$\vdots$$

$$C_1 Y_1^{[n-1]}(x_0) + \cdots + C_n Y_n^{[n-1]}(x_0) = 0$$

Consider the function

$$u(x) = C_1 Y_1(x) + \cdots + C_n Y_n(x)$$

The above equations then tell us that

$$u(x_0) = 0, \quad u'(x_0) = 0, \quad \ldots, \quad u^{[n-1]}(x_0) = 0$$

So u is ~~a~~ the solution to the IVP

$$L(y) = 0$$

$$y(x_0) = 0, \quad y'(x_0) = 0, \quad \ldots, \quad y^{[n-1]}(x_0) = 0$$

But of course the zero function is the sol. to that IVP.

So we have $\quad u(x) = 0$

$$\Rightarrow c_1 y_1(x) + \cdots + c_n y_n(x) = 0$$

So $\{y_1, \ldots, y_n\}$ is l.d.

Recall that for arbitrary functions, the problem was that the c's that work at one value of x are not necessarily the same as the c's that work at another x...

In this proof, it is the uniqueness of sols to IVP's that allows us to get around that problem.

Note the following easy corollary:

Cor:) If $\{y_1, \dots, y_n\}$ are solutions to $L(y) = 0$, (nth order, satisfying the conditions of the uniqueness/exist. thm.)

then either

① W is always 0 and the fns are l.d.

or

② W is never 0 and the fns are l.i.

Pf:) ② follows from old result about Wronskian.

If ② does not happen then W is 0 for some x_0. Our previous thm tells us then that the fns are l.d., and again our old result about the Wronskian tells us that W is always 0.

4.2 - Homogeneous Const. Coeff. Linear D.E.'s

If the coefficient functions $g_i(x)$ are <u>constants</u>, we call the DE. a "constant coefficient" equation.

This is very common.

Ex:) <u>Forced harmonic motion</u> What is the position function for a mass on a spring moving in a resistive medium?

Model:
$$\underbrace{ay''}_{F=ma} + \underbrace{by'}_{\text{damping, friction}} + \underbrace{cy}_{\text{Hooke's Law}} = 0$$

Ex:) <u>LRC circuits</u>: Suppose you have a circuit with an inductor, resistor, and capacitor, and an applied voltage. What is current as a fn of time?

Model:
$$LI'' + RI' + \frac{1}{C}I = E'$$

inductance resistance capacitance voltage

Our equation has the form

$$a_n y^{[n]} + \cdots + a_0 y = 0$$

Consider solutions of the form

$$y(x) = e^{\lambda x}$$

The equation becomes

$$\left(a_n \lambda^n + \cdots + a_0\right) e^{\lambda x} = 0$$

This gives us solutions <u>iff</u>

$$a_n \lambda^n + \cdots + a_0 = 0$$

This is called the <u>characteristic polynomial</u>. We consider three cases of types of roots – distinct real, repeated real, complex.

But we must know how to find roots of polynomials!

Finding Roots

(1) Quadratic equation

(2) Factor Theorem: a is a root of polynomial $p(x)$ iff $(x-a)$ is a factor of polynomial $p(x)$.

(see proof on class webpage)

We often find roots by factoring!

(3) Rational root theorem

If $f(x) = a_n x^n + \dots + a_1 x + a_0$ has integer coeffs and if $r = p/q$ is a rational root, then

$p \mid a_0$ and $q \mid a_n$

(4) Polynomial division — If $p(x) = (x-a)\,q(x)$, then you can look for factors/roots of q instead of p...

(5) Intermediate value theorem — If $p(a)$ and $p(b)$ have opposite signs, there must be a root between a and b.

Ex:) Solve the D.E.

$$y'' - 3y' + 2y = 0$$

The char. poly. is

$$\lambda^2 - 3\lambda + 2 = 0$$

$$\Rightarrow (\lambda - 2)(\lambda - 1) = 0$$

$$\Rightarrow \lambda = 1, 2$$

So e^x, e^{2x} are solutions.

Also — as these are indep. and we know there is a 2-dim space of solutions, we have that

$\{e^x, e^{2x}\}$ is a fund. set of sols.

Ex:) Solve

$$y'''' + 8y''' + 3y'' - 32y' - 28y = 0$$

Char. Poly. is

$$p(\lambda) = \lambda^4 + 8\lambda^3 + 3\lambda^2 - 32\lambda - 28 = 0$$

-1 is a root! And

$$p(\lambda) = (\lambda + 1)(\lambda^3 + 7\lambda^2 - 4\lambda - 28)$$

So now we can look for roots of

$$g_1(\lambda) = \lambda^3 + 7\lambda^2 - 4\lambda - 28$$

Note that 2 is a root, and

$$q_1(\lambda) = (\lambda-2)(\lambda^2+9\lambda+14)$$

Then $q_2(\lambda) = \lambda^2+9\lambda+14$ is easy to factor,

$$q_2(\lambda) = (\lambda+7)(\lambda+2)$$

So we have

$$p(\lambda) = (\lambda+1)(\lambda-2)(\lambda+7)(\lambda+2)$$

and then solutions $e^{-x}, e^{2x}, e^{-7x}, e^{-2x}$

And $\{e^{-x}, e^{2x}, e^{-7x}, e^{-2x}\}$ is a fund. set of sols because they are 4 fns that are l.i., in a 4-dim. set of sols.

Distinct Real Roots

If roots are distinct, then:

$$
\begin{aligned}
\#\text{ of solutions } e^{\lambda x} &= \#\text{ of roots} \\
&= \text{degree of char. poly} \\
&= \text{order of CCLDE} \\
&= \text{dim of space of sols} \\
&= \#\text{ of fns in a fund sol. set}
\end{aligned}
$$

These functions are l.i., so they are a fund. sol. set.

Real Repeated Roots

If a root is repeated, then we don't get enough solutions...

Ex:) Solve the CCLDE:

$$y'' - 2y' + y = 0$$

The char. poly. is $\lambda^2 - 2\lambda + 1 = 0$

$$(\lambda - 1)^2 = 0 \Rightarrow \lambda = 1$$

So we know e^{1x} is a solution; but we also know that we still need to find one more solution...

One might suspect that since 1 is a root "twice",
the other solution might be related to e^{1x}

We try a function of the form

$$y = u(x)\, e^{1x}$$

Plugging this in gives us

$$(ue^x)'' - 2(ue^x)' + (ue^x) = 0$$

$$\vdots$$

$$u'' e^x = 0$$

$$\Rightarrow u'' = 0$$

Conveniently $u(x) = x$ satisfies this. So

$$y = xe^x$$

is another solution.

Observe, 1 was repeated once (multiplicity = 2),
and we needed (and found) one new solution.

This good news generalizes.

Def: If the complete factorization of $p(x)$ is

$$p(x) = (x-r_1)^{m_1} \cdots (x-r_k)^{m_k}$$

then we say that each root r_i has <u>multiplicity</u> m_i.

Thm: Let $p(\lambda)$ be the char. poly. of a CCLDE, and r a root of multiplicity m. Then

$$e^{rx}, xe^{rx}, \ldots, x^{m-1}e^{rx}$$

are solutions to the homogeneous CCLDE.

We then get the m solutions that we "should" have for this root.

Will these solutions be independent?

Ex:) Consider $\{e^{-x}, e^{2x}, xe^{2x}, e^{3x}, xe^{3x}, x^2e^{3x}\}$.

Suppose $c_1e^{-x} + c_2e^{2x} + c_3xe^{2x} + c_4e^{3x} + c_5xe^{3x} + c_6x^2e^{3x} = 0$

Multiplying through by e^{-3x} and taking $\lim\limits_{x\to\infty}$, we get

$$\lim_{x\to\infty}\left(\underbrace{c_1e^{-4x} + c_2e^{-2x} + c_3xe^{-2x}}_{\text{these limits are zero}} + c_4 + c_5x + c_6x^2\right) = 0$$

$$\lim_{x\to\infty}\left(c_4 + c_5x + c_6x^2\right) = 0$$

$$\Rightarrow c_4, c_5, c_6 = 0$$

So our original linear combination becomes

$$c_1e^{-x} + c_2e^{2x} + c_3xe^{2x} = 0$$

Now multiplying through by e^{-2x} and taking $\lim\limits_{x\to\infty}$, we get

$$\lim_{x\to\infty}\left(c_1e^{-3x} + c_2 + c_3x\right) = 0$$

this limit is zero

$$\lim_{x\to\infty}\left(c_2 + c_3x\right) = 0$$

$$\Rightarrow c_2, c_3 = 0$$

This leaves us with just

$$c_1e^{-x} = 0$$

$$\Rightarrow c_1 = 0$$

So $c_1, \ldots, c_6$ are all zero, and thus the functions are l.i..

This method always works for these sorts of functions!

Thm:) Consider the CCLDE $L(y)=0$, with char. poly. $p(\lambda)$.

If $p(\lambda)$ has all real roots, giving solutions of the forms

$$e^{r_i x} \quad \text{and} \quad x^k e^{r_j x}$$

then these solutions are linearly independent.

Pf:) Suppose $c_1 y_1 + \cdots + c_m y_m = 0$. Let r_1 be the greatest root of $p(\lambda)$, and consider the fact that

$$\lim_{x \to \infty} e^{-r_1 x}\left(c_1 y_1 + \cdots + c_m y_m\right) = 0$$

On the LHS, solutions corresponding to roots other than r_1 will give terms that have $\lim = 0$ because of the exponential. Removing those terms yields

$$\lim_{x \to \infty} e^{-r_1 x}\left(k_0 e^{r_1 x} + k_1 x e^{r_1 x} + \cdots + k_i x^i e^{r_1 x}\right) = 0$$

$$\lim_{x \to \infty} \left(k_0 + k_1 x + \cdots + k_i x^i\right) = 0$$

So all of these coefficients must be 0.

Repeating this process for each root (in decreasing order), we conclude that all $c_1, \ldots, c_m$ must be 0. So the solutions are independent. ■

Complex numbers

Recall, $i = \sqrt{-1}$. Multiples of i are called "imaginary".

Numbers of the form $a+bi$ (where $a,b \in \mathbb{R}$) are "complex".

Complex numbers add and multiply in the expected ways:

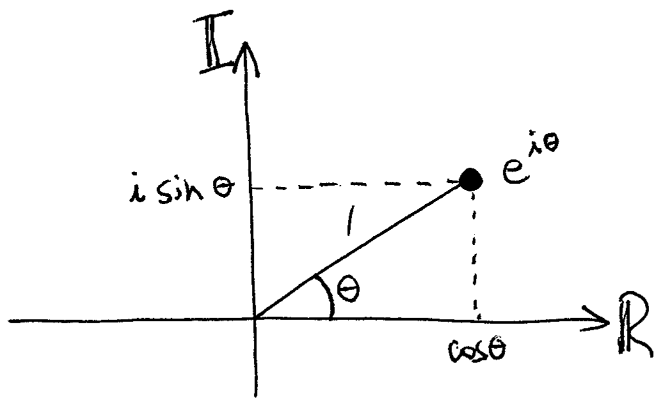
$$(a+bi) + (c+di) = (a+c) + (b+d)i$$

$$(a+bi)(c+di) = (ac-bd) + (ad+bc)i$$

A surprising result of Taylor series is:

$$e^{i\theta} = (\cos\theta) + (\sin\theta)\, i$$

The angle interpretation of θ and the trig in the above equation suggest thinking of an imaginary axis $\perp$ to the real axis:



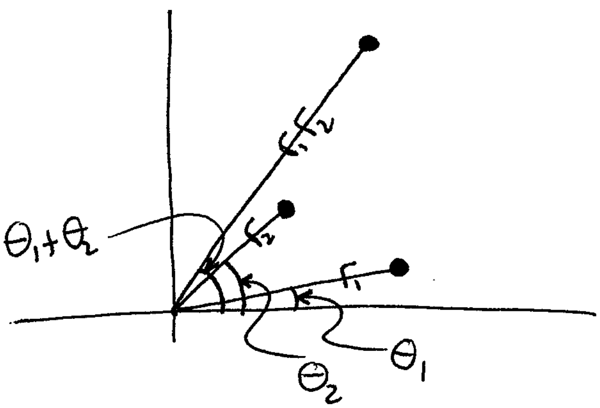
This is called the "complex plane". Note, every complex number can be written in the form $re^{i\theta}$.

Many algebraic facts about complex numbers and functions have appealing geometric interpretations on the complex plane.

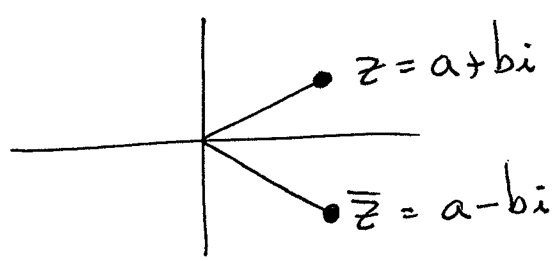
For example, complex multiplication :

$$\left(r_1 e^{i\theta_1}\right)\left(r_2 e^{i\theta_2}\right) = \left(r_1 r_2\right) e^{i(\theta_1+\theta_2)}$$

Geometrically, we note that lengths multiply and angles add.



The "complex conjugate" of a complex number $(a+bi)$ is the number $(a-bi)$. It is the reflection through the real axis.



Note: ① $z = re^{i\theta} \Rightarrow \bar{z} = re^{-i\theta}$

② $\left(\bar{z}\right)^k = \overline{\left(z^k\right)}$

Some useful facts about polynomials:

① Every polynomial with complex coefficients will factor <u>completely</u> (product of linear factors)

Ex: $x^2+1 = (x-i)(x+i)$

Ex: $x^3-x^2+2 = (x+1)(x-(1+i))(x-(1-i))$

② If $p(x)$ has <u>real</u> coefficients and if z is a root, then $\bar{z}$ is a root also.

(see above examples)

Pf: $p(\bar{z}) = \sum a_k(\bar{z})^k = \sum a_k\left(\overline{z^k}\right) = \sum \bar{a}_k\left(\overline{z^k}\right)$

$= \sum \overline{a_k z^k} = \overline{\sum a_k z^k} = \overline{p(z)} = \bar{0} = 0.$

Another useful fact: Every non zero complex number has exactly n nth roots.

(Note, this is a special case of fact ① on the previous page, using the polynomial $x^n - \alpha$.)

These roots are related by the "roots of unity". Let

$\rho = e^{2\pi i/n}$; note $\rho^n = 1$.

We also have $(\rho^k)^n = 1$. So,

$$1, \rho, \rho^2, \ldots, \rho^{n-1}$$

(these are called the "roots of unity")

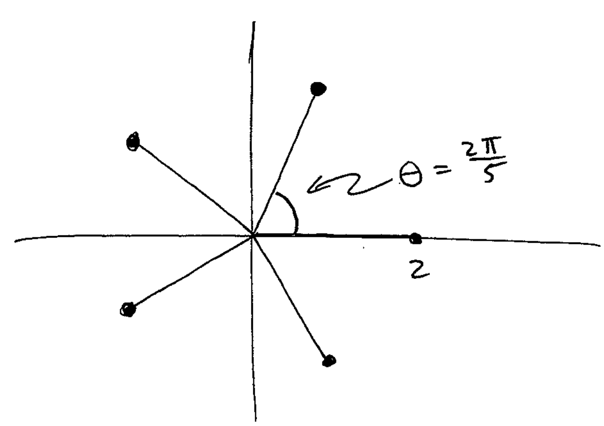
are all nth roots of 1. They are also all distinct.

Letting r be any nth root of α then, we have

$$r, r\rho, r\rho^2, \ldots, r\rho^{n-1}$$

are n distinct nth roots of α.

Ex: In the figure are the five fifth roots of 32.



Complex Roots

If $p(\lambda)$ has a complex root $r = a+bi$, then we have the solution

$$y = e^{rx} = e^{(a+bi)x}$$
$$= e^{(ax)} e^{(bx)i}$$
$$= e^{ax}\left(\cos(bx) + i\sin(bx)\right)$$

But this function has complex values...

Conveniently, we can note though that the real and imaginary parts do not interact in the differential equation, because:

$$\left(u(x) + i\,v(x)\right)^{[k]} = u^{[k]}(x) + i\,v^{[k]}(x)$$

$$a_k\left(u(x) + i\,v(x)\right) = \left(a_k u(x)\right) + i\left(a_k v(x)\right)$$

real

So,
$$L\left(u+iv\right) = 0 \iff \begin{array}{l} L(u) = 0 \\ L(v) = 0 \end{array}$$

That is, if $Y(x)$ is a complex-valued solution, then $\mathrm{Re}\left(Y(x)\right)$ and $\mathrm{Im}\left(Y(x)\right)$ are also real-valued solutions.

Applying this observation to the solution (with $\lambda = r = a+bi$)

$$y = e^{rx}$$
$$= \left(e^{ax}\cos(bx)\right) + i\left(e^{ax}\sin(bx)\right)$$

we can conclude two real solutions,

$$e^{ax}\cos(bx) \quad \text{and} \quad e^{ax}\sin(bx)$$

Does this give us "too many" solutions? No — because if $r = a+bi$ is a root, remember that the conjugate $\bar{r} = a-bi$ is a root also, and gives us the <u>same two</u> real solutions.

Thm: If $L(y) = 0$ is a real CCLDE and $r, \bar{r}$ are a pair of roots of the char. poly., then

$$e^{ax}\cos(bx) \quad \text{and} \quad e^{ax}\sin(bx)$$

are real solutions, independent, and with the same span as the solutions

$$e^{rx} \quad \text{and} \quad e^{\bar{r}x}$$

Ex:) Solve the CCLDE:

$$y''' - y'' + 2y = 0$$

The char. poly is

$$\begin{aligned} p(\lambda) &= \lambda^3 - \lambda^2 + 2 \\ &= (\lambda+1)(\lambda-(1+i))(\lambda-(1-i)) \end{aligned}$$

with roots $-1,\ 1+i,\ 1-i$

This gives us independent solutions

$$\{e^{-x},\ e^x\cos x,\ e^x\sin x\}$$

Ex:) Solve the CCLDE:

$$y'''' - 4y''' + 8y'' - 8y' + 4y = 0$$

The char. poly. is

$$\begin{aligned} p(\lambda) &= \lambda^4 - 4\lambda^3 + 8\lambda^2 - 8\lambda + 4 \\ &= (\lambda-(1+i))^2(\lambda-(1-i))^2 \end{aligned}$$

(NB, this is hard to factor...)

We have roots $r = (1+i)$ and $\bar{r} = (1-i)$, each with multiplicity 2. So solutions are

$$\{e^{rx},\ e^{\bar{r}x},\ xe^{rx},\ xe^{\bar{r}x}\}$$

or

$$\{e^x\cos x,\ e^x\sin x,\ xe^x\cos x,\ xe^x\sin x\}$$

Ex:) Solve the IVP:

$$y''' - y'' + 2y = 0 \,, \qquad \begin{aligned} y(3) &= 1 \\ y'(3) &= 5 \\ y''(3) &= 8 \end{aligned}$$

We already know the general solution from a previous example, so we need only solve for the constants.

We have

$$y = c_1 e^{-x} + c_2 e^{x} \cos x + c_3 e^{x} \sin x$$

$$\begin{aligned} y' &= -c_1 e^{-x} + c_2 e^{x} \cos x - c_2 e^{x} \sin x + c_3 e^{x} \sin x + c_3 e^{x} \cos x \\ &= (-c_1) e^{-x} + (c_2 + c_3) e^{x} \cos x + (c_3 - c_2) e^{x} \sin x \end{aligned}$$

$$y'' = c_1 e^{-x} + (2c_3) e^{x} \cos x + (-2c_2) e^{x} \sin x$$

So at $x=3$, we have

$$\begin{aligned} 1 &= (e^{-3}) c_1 + (e^{3} \cos(3)) c_2 + (e^{3} \sin(3)) c_3 \\ 5 &= (-e^{-3}) c_1 + (e^{3}(\cos(3) - \sin(3))) c_2 + (e^{3}(\cos(3) + \sin(3))) c_3 \\ 8 &= (e^{3}) c_1 + (-2e^{3} \sin(3)) c_2 + (2e^{3} \cos(3)) c_3 \end{aligned}$$

We could solve this system for c_1, c_2, c_3, but note that the coefficients are very inconvenient.

Instead, let's make the substitution $t = x-3$, and then rewrite the IVP. Note $\frac{dy}{dt} = \frac{dy}{dx}$, so the equation remains the same; only the initial conditions change.

$$y''' - y'' + 2y = 0 \qquad \begin{aligned} y(0) &= 1 \\ y'(0) &= 5 \\ y''(0) &= 8 \end{aligned}$$

derivs w.r.t. t

value of t

The general solution has the same form, and thus so do the derivatives.

$$y = c_1 e^{-t} + c_2 e^{t}\cos t + c_3 e^{t}\sin t$$
$$y' = (-c_1)e^{-t} + (c_2+c_3)e^{t}\cos t + (c_3-c_2)e^{t}\sin t$$
$$y'' = c_1 e^{-t} + (2c_3)e^{t}\cos t + (-2c_2)e^{t}\sin t$$

At $x=3$, which is $t=0$, we then have

$$1 = (1)\,c_1 + (1)\,c_2 + (0)\,c_3$$
$$5 = (-1)\,c_1 + (1)\,c_2 + (1)\,c_3$$
$$8 = (1)\,c_1 + (0)\,c_2 + (2)\,c_3$$

This can be much more easily solved for c_1, c_2, c_3 (not the same as the c_1, c_2, c_3 on the previous page!), and then we can back substitute $t = x-3$.

4.3 - Undetermined Coefficients

For nonhomogeneous equations, recall that we need only to find <u>any single particular solution</u>; we can then get the complete solution by adding the general solution to the assoc. hom. equation.

"Undetermined coefficients" is basically a guess-and-check method, advised by experience.

Ex i) Find a particular solution to

$$y'' + 2y' - 3y = 4e^{2x}$$

Note that if we chose $y_p = Ae^{2x}$, then y' and y'' would have similar forms, so the entire LHS would be a multiple of e^{2x}, as needed. We can then solve for (determine) the as yet "undetermined coefficient" A.

$$y_p = Ae^{2x}$$

$$(4Ae^{2x}) + 2(2Ae^{2x}) - 3(Ae^{2x}) = 4e^{2x}$$

$$5Ae^{2x} = 4e^{2x}$$

$$A = \frac{4}{5}$$

$$y_p = \frac{4}{5}e^{2x}$$

Ex:) Find a particular solution to

$$y'' + 2y' - 3y = x^2$$

We would have a shot at balancing this equation with

$$y = Ax^2$$

but, ... y' and y'' then take forms not on RHS... ☹

But if we put <u>those</u> forms in y too, then that could fix it:

$$y_p = Ax^2 + Bx + C$$

(Note, the derivs of these new terms are of the same forms, so no new terms are suggested.)

Plugging this in to the DE, we get

$$(2A) + 2(2Ax + B) - 3(Ax^2 + Bx + C) = x^2$$

$$(-3A)x^2 + (4A - 3B)x + (2A + 2B - 3C) = x^2$$

$$A = -1/3 \qquad B = -4/9 \qquad C = -14/27$$

So a particular solution is

$$y_p = -\frac{1}{3}x^2 - \frac{4}{9}x - \frac{14}{27}$$

Ex:) What form would be natural to try for a particular solution to

$$y'' + 2y' - 3y = 3\sin x$$

It appears we will need $\sin x$ on the LHS; but that will also create $\cos x$ on the LHS. But if we use both $\sin x$ and $\cos x$, we can hope to balance...

$$y_p = A\cos x + B\sin x$$

This form works, with $A = \frac{-3}{10}$, $B = \frac{-3}{5}$

Ex:) What form should we try for

$$y'' + 2y' - 3y = -2e^x \cos x$$

Well, using $Ae^x\cos x$ creates a need for $Be^x \sin x$; this term does not suggest any others. You can check that

$$y_p = Ae^x\cos x + Be^x\sin x$$

works, with $A = 2/17$, $B = -8/17$

Ex:) Find a particular solution for

$$y'' + 2y' - 3y = -2e^x$$

Tempting to try Ae^x... but, note Ae^x is a solution to the homogeneous equation, for all A!

Remember that xe^x has derivative terms with e^x...

So we might try $y_p = Axe^x$:

$$\left(A(x+2)e^x\right) + 2\left(A(x+1)e^x\right) - 3\left(Axe^x\right) = -2e^x$$

$$4Ae^x = -2e^x$$

$$A = -\tfrac{1}{2}$$

$$y_p = -\tfrac{1}{2}xe^x$$

Moral: If you know the homogeneous solutions, you know what forms not to bother trying.

The following theorem makes recommendations of forms to try:

Thm: Consider the CCLDE

$$L(y) = a_n y^{[n]} + \cdots + a_1 y' + a_0 y = g(x)$$

with char. poly. $p(\lambda)$.

If $g(x) = Ax^k e^{ax}\cos bx + Bx^k e^{ax}\sin bx \quad (k \in \mathbb{N})$

then ① If $r = a+bi$ is not a root of $p(\lambda)$, try

$$\boxed{\begin{aligned} y_p &= (c_k x^k + \cdots + c_0)e^{ax}\cos(bx) \\ &\quad + (d_k x^k + \cdots + d_0)e^{ax}\sin(bx) \end{aligned}}$$

② If $r = a+bi$ is a root of $p(\lambda)$ of mult. m, try

$$\boxed{\begin{aligned} y_p &= x^m(c_k x^k + \cdots + c_0)e^{ax}\cos(bx) \\ &\quad + x^m(d_k x^k + \cdots + d_0)e^{ax}\sin(bx) \end{aligned}}$$

Note ① $r = a+bi = 0$ gives $g(x) = Ax^k$

② $b = 0$ gives $g(x) = Ax^k e^{ax}$ with $r = a$

③ $a = 0$ gives $g(x) = Ax^k\cos(bx) + Bx^k\sin(bx)$ with $r = bi$

Ex i) What form should we try for a particular solution to

$$y'' - 2y' + 5y = x e^{x} \cos(2x)$$

The char. poly is

$$p(\lambda) = \lambda^2 - 2\lambda + 5$$
$$= \big(\lambda - (1+2i)\big)\big(\lambda - (1-2i)\big)$$

Our $g(x)$ is of the form

$$x e^{x} \cos(2x) = x^{k} e^{ax} \cos(bx)$$

with $k=1$, $r = a+bi = 1+2i$, and r <u>is</u> a root of $p(\lambda)$ with $m=1$.

So we try

$$y_p = x\left(c_1 x + c_0\right) e^{x} \cos(2x) + x\left(d_1 x + d_0\right) e^{x} \sin(2x)$$

$$= c_1 x^2 e^{x} \cos(2x) + c_0 x e^{x} \cos(2x) + d_1 x^2 e^{x} \sin(2x) + d_0 x e^{x} \sin(2x)$$

If there are multiple terms on the RHS that do not fit the same form, you can do "one at a time".

That is, if

$$L(y_{p_1}) = g_1$$

$$L(y_{p_2}) = g_2$$

then

$$L(y_{p_1} + y_{p_2}) = g_1 + g_2$$

Ex:) Find a particular solution to

$$L(y) = y'' + 2y' + y = x^2 + e^x$$

The equation $L(y) = x^2$ has a particular solution

$$y_{p_1} = Ax^2 + Bx + C$$

The equation $L(y) = e^x$ has a particular solution

$$y_{p_2} = De^x$$

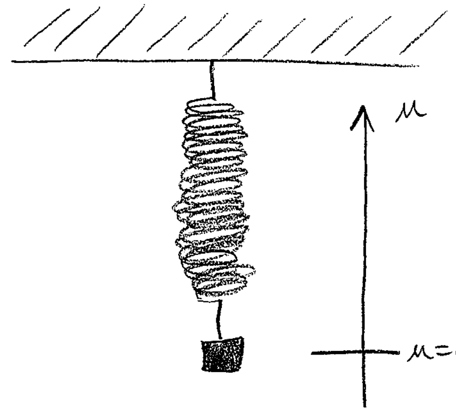
So, the particular solution we are looking for is of the form

$$y_p = Ax^2 + Bx + C + De^x$$

4.5 - Applications

Mass on a spring

Say a mass is attached to a fixed object by a spring, subject to forces from the spring, friction/damping, and external forces. What is the position u as a function of time t?



Forces: Hooke's law $F = -ku$

friction $F = -fu'$

external $F = h(t)$

Total forces:

$$F = -ku - fu' + h(t)$$

Recall that $F = ma = mu''$. So this becomes

$$mu'' = -ku - fu' + h(t)$$

$$mu'' + fu' + ku = h(t)$$

As previously claimed, note that this is a CCLDE.

Unforced examples

① No friction (undamped)

$$mu'' + ku = 0$$

Char. poly. is $\quad m\lambda^2 + k = 0$

$$\Rightarrow \lambda = i\sqrt{\tfrac{k}{m}} = i\omega_0$$

Solutions are $\left\{ \cos(\omega_0 t),\ \sin(\omega_0 t) \right\}$

So gen. sol. is $\quad c_1 \cos\omega_0 t + c_2 \sin\omega_0 t$

Trig. can rewrite this as

$$u(t) = A\cos(\omega_0 t - \phi)$$

If initial conditions are given, can solve for

c_1, c_2 (or A, ϕ).

② Friction (damped)

$$mu'' + fu' + ku = 0$$

Char. poly. is

$$m\lambda^2 + f\lambda + k = 0$$

$$\Rightarrow \lambda = \frac{-f \pm \sqrt{f^2 - 4mk}}{2m}$$

Observe that : — this simplifies to previous formula when $f=0$

— the $Re(\lambda)$ is <u>always</u> <0

ⓐ Under damped

If $f^2 - 4mk < 0$....

then we have two complex roots

$$\lambda = -a \pm bi \qquad \left(\begin{aligned} a &= f/2m \\ b &= \frac{\sqrt{f^2-4mk}}{2m} \end{aligned} \right)$$

and then solutions

$$u = C_1 e^{-at} \cos bt + C_2 e^{-at} \sin bt$$

or

$$u = e^{-at} \left(A \cos(bt - \phi) \right)$$

This is a decaying oscillation.

(b) Overdamped

If $f^2 - 4mk > 0$ ....

then we have two real roots $r_1, r_2 < 0$

and solutions

$$u = c_1 e^{r_1 t} + c_2 e^{r_2 t}$$

This is decaying with no oscillations

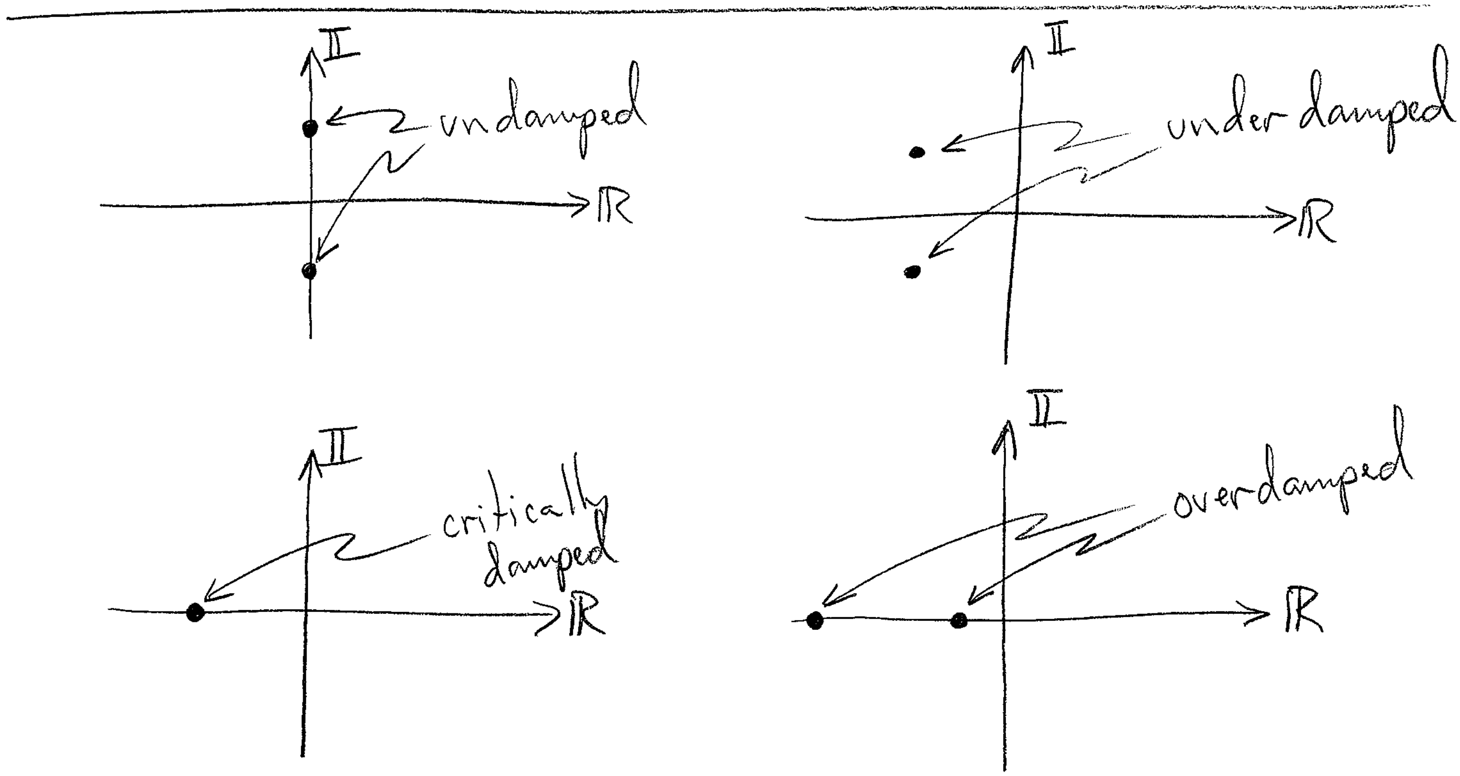
(c) Critically damped

If $f^2 - 4mk = 0$ ....

then we have one real root $r = \frac{-f}{2m}$

and solutions

$$u = c_1 e^{rt} + c_2 t e^{rt}$$



Forced examples

① No friction (undamped)

$$m u'' + k u = h(t)$$

Suppose the forcing function is sinusoidal; then we have

$$u'' + \omega_0^2 u = a \cos \omega t \quad \left(\text{freq} = \frac{\omega}{2\pi}\right)$$

(recall $\omega_0^2 = k/m$)

We already know the homogeneous (unforced) solutions.

$$u_H = c_1 \cos \omega_0 t + c_2 \sin \omega_0 t$$

Undet. coeffs. gives us a particular solution

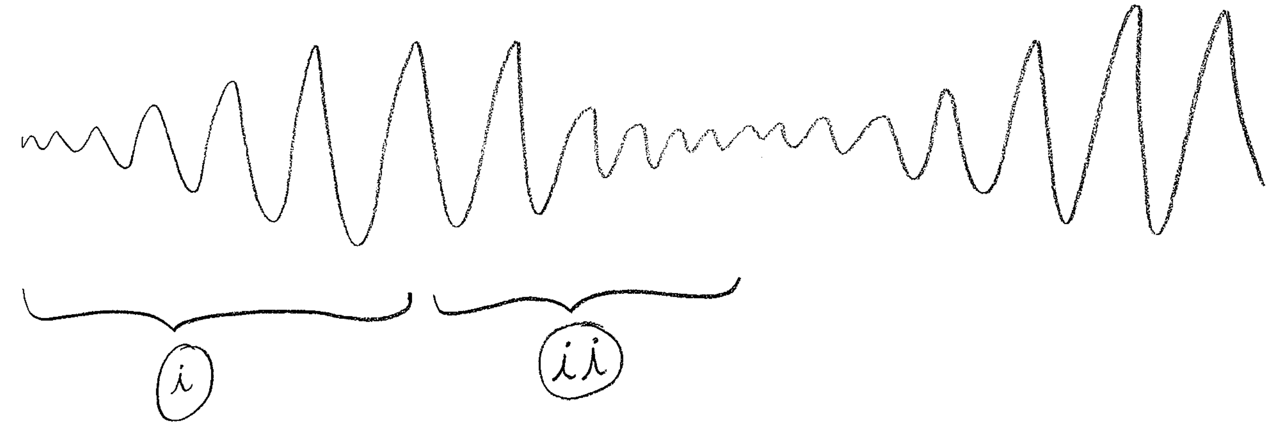
$$u_p = \frac{a}{\omega_0^2 - \omega^2} \cos \omega t$$

Using initial conditions $u(0) = 0$, $u'(0) = 0$, we can find the I.V.P. solution

$$u = \frac{a}{\omega_0^2 - \omega^2} \left(\cos \omega t - \cos \omega_0 t \right)$$

(a) $\omega_0 \neq \omega$

In this case the two cosines have a beat:



The natural frequency ω_0 and the forcing frequency are different. So sometimes (i above) the forcing is pushing with the oscillation, and sometimes (ii above) the forcing is working against the oscillation.

<u>Algebraically:</u>

$$\cos(a-b) = \cos a \cos b + \sin a \sin b$$

$$\cos(a+b) = \cos a \cos b - \sin a \sin b$$

$$\Rightarrow \boxed{\cos(a-b) - \cos(a+b) = 2 \sin a \sin b}$$

Let $a-b = \omega t$ and solve for a, b
$a+b = \omega_0 t$

$$\Rightarrow a = \left(\frac{\omega_0 + \omega}{2} t\right), \quad b = \left(\underbrace{\frac{\omega_0 - \omega}{2}}\, t\right)$$

"beat frequency" $= |\omega_0 - \omega|$

(b) $\omega = \omega_0$

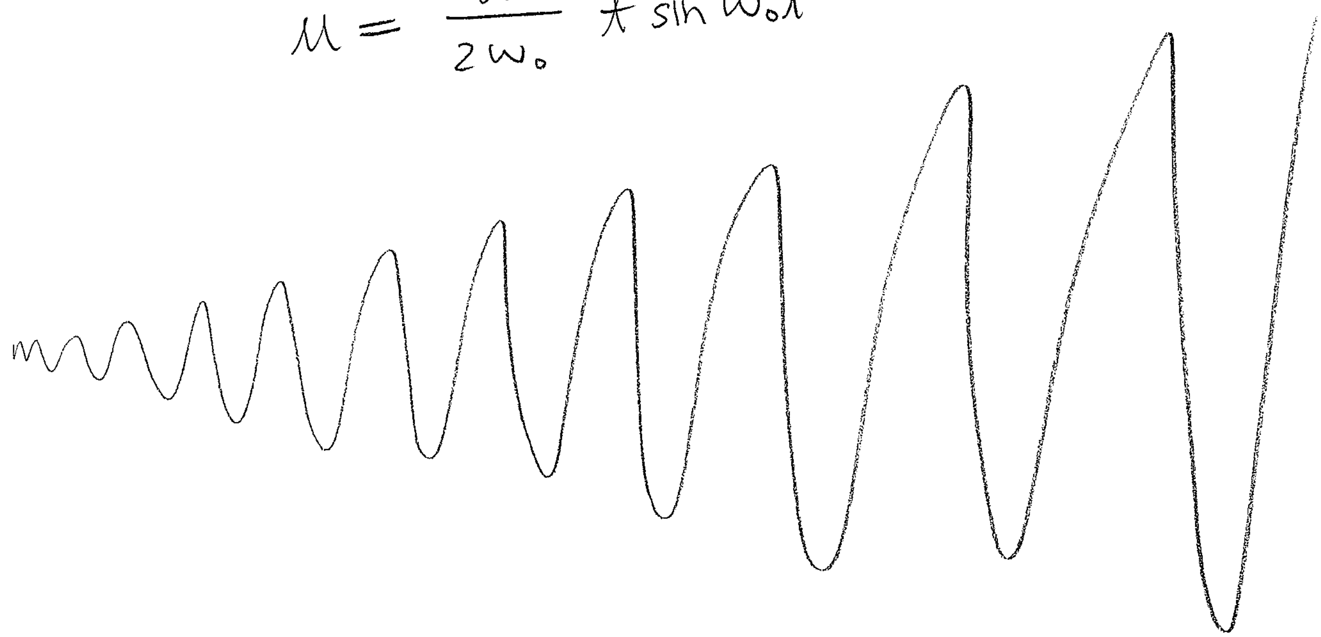
Then the previous solution no longer makes sense...

Undet. coeffs. then gives us a different solution because

$\omega i = \omega_0 i = i\sqrt{k/m}$ <u>is</u> a root of $p(\lambda)$...

We get an extra factor of t:

$$u = \frac{a}{2\omega_0} t \sin \omega_0 t$$



The forcing freq and natural oscillating frequency are the same, so they never go out of phase, so it is always growing.

This is called resonance. It can knock down or damage bridges (forced by wind, marching soldiers, or even just people walking!)

② Friction (damped)

Whether over, under, or critically damped, all of the homogeneous solutions decay exponentially; so they are irrelevant to the long-term behavior of the solution (they might come up from initial conditions though)

They are called transients.

(In book example, p. 226, transient freq. is 6...)

Let's consider the particular solution, again with a sinusoidal forcing function.

$$u'' + 2cu' + w_0^2 u = a \cos wt$$

Note that $a \cos wt = \mathrm{Re}\left(a e^{iwt}\right)$

So we can consider

$$z'' + 2cz' + w_0^2 z = a e^{iwt}$$

and then choose

$$u = \mathrm{Re}(z)$$

This equation

$$z'' + 2cz' + \omega_0^2 z = ae^{i\omega t}$$

has a solution that is a multiple of the RHS,

$$z = T\, a e^{i\omega t}$$

This number T has a magnitude & a direction, so we write it as

$$T = G e^{-i\phi}$$

(Both G and ϕ depend on ω, ω_0, c)

So we have

$$z = G\, a e^{i(\omega t - \phi)}$$

and so

$$u = G\, a \cos(\omega t - \phi)$$

So — the response is like the input, but with a change in magnitude and phase

G is called the "gain".

ϕ is called the "phase shift".

$$G = \frac{1}{\sqrt{(\omega_0^2 - \omega^2)^2 + 4c^2\omega^2}}$$

(Note G large when $\omega_0 \approx \omega$ and c small...)

5.1 - Linear Transformations

Given $f: X \to Y$, we consider also the set $B \subset Y$ defined by $B = \{\vec{y} \in Y \mid \vec{y} = f(\vec{x}), \vec{x} \in X\}$

Terminology:

X = "domain"

B = "image", "~~range~~"

$\cap$

Y = "~~range~~", "target", "codomain", "~~image~~"

<u>Def:</u> Given vector spaces V, W, a function $T: V \to W$ is called a <u>linear transformation</u> if

$$T(a\vec{x}_1 + b\vec{x}_2) = aT(\vec{x}_1) + bT(\vec{x}_2)$$

for all $\vec{x}_1, \vec{x}_2 \in X$, $a, b \in \mathbb{R}$.

<u>Equiv:</u> $T(c\vec{x}) = cT(\vec{x})$ and $T(\vec{x}_1 + \vec{x}_2) = T(\vec{x}_1) + T(\vec{x}_2)$

<u>Equiv:</u> $T(c_1\vec{x}_1 + \cdots + c_k\vec{x}_k) = c_1T(\vec{x}_1) + \cdots + c_kT(\vec{x}_k)$

<u>Interp:</u> T "commutes with linear combinations".

Ex: Is $T: \mathbb{R}^2 \to \mathbb{R}^2$, defined by

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x - y \\ 3x + 2y \end{pmatrix}$$

a l.t.?

Check:

$$T(a\vec{x}_1 + b\vec{x}_2) = T\left(a\begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + b\begin{pmatrix} x_2 \\ y_2 \end{pmatrix}\right)$$

$$= T\begin{pmatrix} ax_1 + bx_2 \\ ay_1 + by_2 \end{pmatrix}$$

$$= \begin{pmatrix} (ax_1 + bx_2) - (ay_1 + by_2) \\ 3(ax_1 + bx_2) + 2(ay_1 + by_2) \end{pmatrix}$$

$$= \begin{pmatrix} ax_1 - ay_1 \quad + \quad bx_2 - by_2 \\ 3ax_1 + 2ay_1 \quad + \quad 3bx_2 + 2by_2 \end{pmatrix}$$

$$= a\begin{pmatrix} x_1 - y_1 \\ 3x_1 + 2y_1 \end{pmatrix} + b\begin{pmatrix} x_2 - y_2 \\ 3x_2 + 2y_2 \end{pmatrix}$$

$$= aT(\vec{x}_1) \quad + \quad bT(\vec{x}_2)$$

✓

Ex: Is $T: \mathbb{R}^2 \to \mathbb{R}^1$, defined by

$$T\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y$$

a l.t.?

No...

$$\left.\begin{aligned} T\begin{pmatrix} 2 \\ 0 \end{pmatrix} &= 4 \\ 2T\begin{pmatrix} 1 \\ 0 \end{pmatrix} &= 2 \end{aligned}\right\} \ast$$

Here are some interesting linear transformations:

① $D: D(a,b) \to F(a,b)$ defined by

$$D(f) = f'$$

This is a l.t. because

$$\begin{aligned} D(af_1 + bf_2) &= (af_1 + bf_2)' \\ &= af_1' + bf_2' \\ &= a\,D(f_1) + b\,D(f_2) \quad \checkmark \end{aligned}$$

② $T: \mathbb{R}^n \to \mathbb{R}^m$ defined by

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

(for a given $m \times n$ matrix A)

③ $S : C^{\infty}[a,b] \to \mathbb{R}^1$, defined by

$$S(f) = \int_a^b f(x)\,dx$$

④ $S_g : C^{\infty}[a,b] \to \mathbb{R}^1$, defined by

$$S_g(f) = \int_a^b f(x)\,g(x)\,dx$$

This is useful in statistics/probability, physics, ...

⑤ $\delta_0 : C^{\infty} \to \mathbb{R}^1$, defined by

$$\delta_0(f) = f(0)$$

This is the "Dirac distribution" or "Dirac measure".

(Note, it is possible to relate this to ④ above with limits... Then it is tempting to call this a function, but this is problematic.)

⑥ $\delta_a^{[n]} : C^{\infty} \to \mathbb{R}^1$, defined by

$$\delta_a^{[n]}(f) = f^{[n]}(a)$$

⑦ In a linear DE $L(y) = g(x)$, the linear diff. op.

L is a linear transformation

$$L : C^n \to C^0$$

Sidenote: Examples ④ and ⑤ can be related. Consider a bump function $g : \mathbb{R}^1 \to \mathbb{R}^1$, with $\int g \, dx = 1$, and $g = 0$ away from 0.

We have

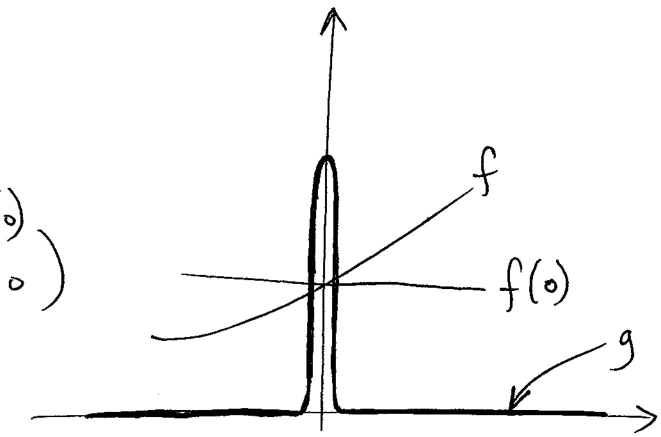
$$S_g(f) = \int f \, g \, dx$$

(b/c $f \approx f(0)$ where $g \neq 0$)

$$\approx \int f(0) \, g \, dx$$

$$\approx f(0) \int g \, dx$$

$$\approx f(0) \qquad = \delta_0(f)$$



One can define a sequence of functions $g_1, g_2, g_3, \ldots$ (increasingly "tall" and "thin") such that, for every continuous f,

$$\lim_{n \to \infty} S_{g_n}(f) = \delta_0(f)$$

It is tempting to think of δ_0 as being represented by a function like these g's that is "infinitely tall" and "infinitely thin" — but <u>no such function exists.</u>

(Sometimes people casually refer to the "Dirac delta function", though technically it is not actually a function.)

Thm: The images of a basis for the domain completely determine all of the values of a l.t.

Pf: We consider a l.t. $T: V \to W$, and suppose $\{\vec{v}_1, \ldots, \vec{v}_k\}$ are a basis for V.

Then, for any $\vec{v} \in V$, we can write

$$\vec{v} = c_1\vec{v}_1 + \cdots + c_k\vec{v}_k$$

and then compute

$$T(\vec{v}) = T(c_1\vec{v}_1 + \cdots + c_k\vec{v}_k)$$

$$= c_1 T(\vec{v}_1) + \cdots + c_k T(\vec{v}_k)$$

Ex: Say we know that $T: \mathbb{R}^2 \to \mathbb{R}^3$ has

$$T\begin{pmatrix}1\\1\end{pmatrix} = \begin{pmatrix}1\\2\\3\end{pmatrix} \quad \text{and} \quad T\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}2\\5\\7\end{pmatrix}$$

What is $T\begin{pmatrix}3\\4\end{pmatrix}$?

$$T\begin{pmatrix}3\\4\end{pmatrix} = T\left(2\begin{pmatrix}1\\1\end{pmatrix} + \begin{pmatrix}1\\2\end{pmatrix}\right)$$

$$= 2T\begin{pmatrix}1\\1\end{pmatrix} + T\begin{pmatrix}1\\2\end{pmatrix}$$

$$= 2\begin{pmatrix}1\\2\\3\end{pmatrix} + \begin{pmatrix}2\\5\\7\end{pmatrix} = \begin{pmatrix}4\\9\\13\end{pmatrix}$$

Def: For a l.t. $T: V \to W$, the <u>kernel</u> of T is the subset of V defined by

$$\ker(T) = \{\vec{v} \in V \mid T(\vec{v}) = \vec{0}\}$$

Ex: Say $T: \mathbb{R}^n \to \mathbb{R}^m$ is defined by

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

Then $\ker(T) = NS(A)$ = set of homog. solutions

Ex: Say $L: C^n \to C^0$ is an nth order lin. diff. op.
Then $\ker(L)$ is the set of homogeneous solutions to $L(y) = 0$.

Def: For a l.t. $T: V \to W$, the <u>image</u> (range) of T is the subset of W defined by

$$\operatorname{im}(T) = \{\vec{w} \in W \mid \vec{w} = T(\vec{v}),\ \vec{v} \in V\}$$

Ex: Say $T: \mathbb{R}^n \to \mathbb{R}^m$ is defined by

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

Let

$$A = \begin{pmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{pmatrix}$$

Then $A\vec{x} = x_1\vec{v}_1 + \cdots + x_n\vec{v}_n$

Then we have

$$\begin{aligned} \text{im}(T) &= \{\text{l.c.'s of cols of } A\} \\ &= CS(A) \end{aligned}$$

Thm: Given a l.t. $T: V \to W$, we have that

$\ker(T)$ is a subspace of V and

$\text{im}(T)$ is a subspace of W

Pf: You can check that each of these subsets is closed under addition and scalar multiplication.

Thm: If $T: V \to W$, and V is a finite-dim v.s., then

$$\dim\left(\ker(T)\right) + \dim\left(\mathrm{im}(T)\right) = \dim(V)$$

The proof is in the book. We will prove this in 5.3, using the already established fact

$$\dim\left(NS(A)\right) + \underbrace{\dim\left(CS(A)\right)}_{\substack{\| \\ \dim\left(RS(A)\right)}} = n$$

Note the following Corollary, which will be useful in Chapter 6 (and also Chapter 4!):

Cor: If the linear transformation $T: V \to W$ is both one-to-one and onto (bijective), then

$$\dim(V) = \dim(W)$$

(In Ch. 4, showing the dim. of the sol. set is n, we can now observe that $T: (\text{I.C.'s}) \to (\text{sols})$ is linear, one-to-one, and onto, and the space of I.C.'s is n-dim — so, the space of solutions is n-dim.)

5.2 - Algebra of Linear Transformations

Linear transformations are functions. But, they can also be viewed as objects with operations — like numbers, or matrices.

<u>Def:</u> Say T, S are l.t.'s from $V \to W$. Then

$$(T+S)(\vec{v}) = T(\vec{v}) + S(\vec{v})$$

$$(kT)(\vec{v}) = kT(\vec{v})$$

Note that these are <u>also</u> <u>l.t.'s</u> from $V \to W$.

We can't conveniently multiply like this though (on outputs...)

But a product-like operation can be defined with <u>compositions</u>.

<u>Def:</u> Say $T: V \to W$ and $S: W \to U$ are l.t.'s.
Then $ST: V \to U$ is defined by

$$ST(\vec{v}) = (S \circ T)(\vec{v}) = S(T(\vec{v}))$$

Note that this is <u>also</u> a <u>l.t.</u>.

(Note: <u>target</u> of T must be the <u>domain</u> of S.)

With these operations (addition, scalar prod, composition), we have an algebra on l.t.'s similar to matrix algebra:

Thm:

① $S+T = T+S$

② $R+(S+T) = (R+S)+T$

③ $c(dT) = (cd)T$

④ $c(S+T) = cS + cT$

⑤ $(c+d)T = cT + dT$

⑥ $R(ST) = (RS)T$

⑦ $R(S+T) = RS + RT$

⑧ $(R+S)T = RT + ST$

⑨ $c(ST) = (cS)T = S(cT)$

Note of course that, just as with matrices, we do <u>not</u> have composition commutativity:

$$ST \neq TS$$

To prove the above 9, note each side of each equation is a l.t.; prove they are the same by evaluating both on an arbitrary vector $\vec{v} \in V$.

Ex: To prove (8), we compute:

$$\Big((R+S)T\Big)(\vec{v}) = (R+S)\big(T(\vec{v})\big) = R\big(T(\vec{v})\big) + S\big(T(\vec{v})\big)$$

$$\Big(RT+ST\Big)(\vec{v}) = RT(\vec{v}) + ST(\vec{v}) = R\big(T(\vec{v})\big) + S\big(T(\vec{v})\big)$$

These are equal, as needed.

We can use this algebra to motivate some facts about solving DE's.

Setup (1) Define $D : C^\infty \to C^\infty$ by

$$D(f) = f'$$

(2) Define $T_g : C^\infty \to C^\infty$ by

$$T_g(f) = gf$$

(3) Observe that every LDE

$$g_n(x)\, y^{[n]}(x) + \dots + g_0(x)\, y(x) = g(x)$$

has LHS that is a l.t. on y:

$$L(y) = \left(T_{g_n} D^n + \dots + T_{g_0}\right)(y) = g(x)$$

We can write this l.t. as

$$L = T_{g_n} D^n + \cdots + T_{g_0}$$

$$= g_n D^n + \cdots + g_0$$

(4) Observe that <u>solving a hom. LDE</u> is equivalent to finding the <u>kernel of the l.t. L</u>.

(5) If L has <u>constant coefficients</u>, then

$$L = p(D)$$

where p is the characteristic polynomial.

(6) Note that LDO's do not commute:

$$DT_g(f) = D(gf) = f'g + fg'$$

$$T_g D(f) = T_g(f') = f'g$$

not equal!

But — <u>CC</u> LDO's <u>do</u> commute:

$$DT_{a_k}(f) = D(a_k f) = a_k f'$$

$$T_{a_k} D(f) = T_{a_k} f' = a_k f'$$

equal!

⑦ Note, CCLDO's factor just like their char. poly. :

$$p(\lambda) = (\lambda - r_1)^{m_1} \cdots (\lambda - r_k)^{m_k}$$

$$L = (D - r_1)^{m_1} \cdots (D - r_k)^{m_k}$$

Given these observations, we reconsider the question of solving a homogeneous CCLDE.

$$L(y) = 0$$

$$(D - r_1)^{m_1} \cdots (D - r_k)^{m_k} y = 0$$

Note that if r_i is one of the roots, the factors can be reordered as

$$\Big(\quad Q(D) \quad \Big)(D - r_i)\, y = 0$$

So, any y with $(D - r_1)y = 0$ would solve this.

That is : $\quad y' = r_1 y$

$$\Longrightarrow y = e^{r_1 x}$$

Similarly, note

$$(D-r)\left(x^k e^{rx}\right)$$

$$= k x^{k-1} e^{rx} + x^k r e^{rx} - r x^k e^{rx}$$

$$= k x^{k-1} e^{rx}$$

So we can conclude that, whenever $k < l$,

$$(D-r)^l \left(x^k e^{rx}\right) = 0$$

So, if $L = p(D)$ has a repeated root r with multiplicity m, then

$$e^{rx}, x e^{rx}, \ldots, x^{m-1} e^{rx}$$

are all solutions to

$$p(D) y = L(y) = 0$$

We can also use this to motivate part of the big theorem about undetermined coefficients. Consider this related theorem:

Thm:) Consider $L(y) = g(x)$

with $(p(D))(y) = L(y)$, $g(x) = Ax^k e^{rx}$ $\left(\begin{array}{c} r=a+bi \\ k \in \mathbb{N} \end{array}\right)$

① If $r=a+bi$ is not a root of $p(\lambda)$, a particular sol'n is

$$y_p = \left(c_k x^k + \dots + c_0\right) e^{rx}$$

② If $r=a+bi$ is a root of $p(\lambda)$ of mult. m, a particular sol. is

$$y_p = x^m \left(c_k x^k + \dots + c_0\right) e^{rx}$$

Outline of Pf:)

① Note $(D - r_i)(x^j e^{rx}) = \underbrace{a_1}_{=(r-r_i) \neq 0} x^j e^{rx} + a_2 x^{j-1} e^{rx}$

so

$$(p(D))(x^j e^{rx}) = \underbrace{b_j}_{=p(r) \neq 0} x^j e^{rx} + \dots + b_0 e^{rx}$$

So we can start by adding a term $c_k x^k e^{rx}$ to y to get the correct $x^k e^{rx}$ term on the right.

Then add an appropriate $c_{k-1} x^{k-1} e^{rx}$ to y to get the correct $x^{k-1} e^{rx}$ term on right.

Repeat until all terms are correct on right.

(2) If r is a root of mult. m, then

$$p(\lambda) = Q(\lambda)(\lambda - r)^m \qquad \left(\text{where } r \text{ is \underline{not} a root of } Q\right)$$

$$L = p(D) = Q(D)(D-r)^m$$

Choosing the form

$$y_p = x^m\left(c_k x^k + \ldots + c_0\right)e^{rx}$$

and recalling

$$(D-r)\left(x^j e^{rx}\right) = \underbrace{j}_{\neq 0} x^{j-1}e^{rx}$$

we get

$$L(y_p) = \left(p(D)\right)(y_p) = \left(Q(D)(D-r)^m\right)(y_p)$$

$$= \left(Q(D)\right)\left((D-r)^m\left(x^m(c_k x^k + \ldots + c_0)e^{rx}\right)\right)$$

$$= \left(Q(D)\right)\left((d_k x^k + \ldots + d_0)e^{rx}\right) \qquad = g(x)$$

Since r is <u>not</u> a root of Q, we know (from (1)) we can find $d_k, \ldots, d_0$.

We can then back-solve for $c_k, \ldots, c_0$.

5.3 - Matrices and Change of Basis

We have already seen that:

<u>Thm:</u> $\left(T(x) = A\vec{x}\right) \implies \left(T \text{ is a l.t. from } \mathbb{R}^n \to \mathbb{R}^m\right)$

In fact this works in both directions:

<u>Thm:</u> If T is a l.t. from $\mathbb{R}^n \to \mathbb{R}^m$, then there exists an $m \times n$ matrix A with

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

<u>Pf:</u> Let $\vec{a}_i = T(\vec{e}_i)$, and

$$A = \begin{pmatrix} | & & | \\ \vec{a}_1 & \cdots & \vec{a}_n \\ | & & | \end{pmatrix}$$

Then we simply check that

$$\begin{aligned} A\vec{x} &= x_1\vec{a}_1 + \cdots + x_n\vec{a}_n \\ &= x_1 T(\vec{e}_1) + \cdots + x_n T(\vec{e}_n) \\ &= T(x_1\vec{e}_1 + \cdots + x_n\vec{e}_n) \\ &= T(\vec{x}) \end{aligned}$$

Ex:) Note that rotations in $\mathbb{R}^2$ are l.t.'s ; now we see that we can compute these rotations with matrices <u>and</u> that we can find this matrix by looking at the $\vec{e}_i$.

Let R_θ be rotation ccwise around $\vec{0}$ by angle θ.

Note $R_\theta(\vec{e}_1) = \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix}$ $\quad R_\theta(\vec{e}_2) = \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$

So $R_\theta(\vec{x}) = A\vec{x}$

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

<u>Note</u>: The <u>columns</u> of A are the images by T of the <u>standard basis vectors</u>

Ex:) We know that for linear transformations $\mathbb{R}^n \to \mathbb{R}^m$

- composing l.t.'s gives l.t.'s
- we view this as a product on l.t.'s
- for each such l.t. there is a corresponding matrix

But, we do <u>not</u> yet know if the matrix for the product is the product of those matrices...

Q: What is the matrix for the composition of 2 l.t.'s?

$$\mathbb{R}^n \xrightarrow{T} \mathbb{R}^m \xrightarrow{S} \mathbb{R}^k$$

$$T(\vec{x}) = A\vec{x}, \quad S(\vec{y}) = B\vec{y} \qquad ST(\vec{x}) = C\vec{x}$$

We compute C one column at a time:

$$\vec{c}_i = ST(\vec{e}_i) = S(T(\vec{e}_i)) = B(A(e_i))$$

cannot use associativity — we have not proved ~~that~~ yet!

$$= B\vec{a}_i$$

By matrix multiplication, this <u>is</u> the ith column of the product BA.

So the product we defined for l.t.'s corresponds nicely to matrix multiplication.

$$S(\vec{y}) = B\vec{y}$$

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

$$\begin{tikzcd} \mathbb{R}^n \arrow[r, "T"] \arrow[rr, bend right, "ST"'] & \mathbb{R}^m \arrow[r, "S"] & \mathbb{R}^k \end{tikzcd}$$

$$ST(\vec{x}) = (BA)\vec{x}$$

<u>Comments</u>: ① Proving associativity of matrix multiplication is now an immediate consequence of associativity of compositions.

$$R(ST) = (RS)T$$

$$\Updownarrow$$

$$C(BA) = (CB)A$$

② One can use this relationship between matrices and l.t.'s as the <u>definition</u> of matrices and operations ...

Motivates matrix addition, scalar mult., and matrix mult. ...

Bases and coordinates

Recall that if $\mathcal{V} = \{\vec{v}_1, \ldots, \vec{v}_n\}$ is a basis for V, then there is a <u>unique</u> way to write any $\vec{v} \in V$ as

$$\vec{v} = c_1 \vec{v}_1 + \cdots + c_n \vec{v}_n$$

These unique coefficients are the "coordinates of $\vec{v}$ relative to this basis", and we write

$$[\vec{v}]_{\mathcal{V}} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

We can thus think of the basis $\mathcal{V}$ as giving a bijection between V (made up of vectors) and $\mathbb{R}^n$ (made up of coordinates relative to the basis).

Notation:

$$V \overset{\mathcal{V}}{\longleftrightarrow} \mathbb{R}^n$$

$$\vec{v} \longrightarrow [\vec{v}]_{\mathcal{V}}$$

$$\begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}_{\mathcal{V}} \longleftarrow \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$$

You can check that this is itself a l. t..

Ex:) Consider P_2, the v.s. of polynomials of degree ≤ 2, and the basis $\mathcal{V} = \{\underbrace{x^2-1}_{\vec{v}_1}, \underbrace{2x-1}_{\vec{v}_2}, \underbrace{1}_{\vec{v}_3}\}$

We can write the polynomial $f = 3x^2+6x+5$

as $f = 3\vec{v}_1 + 3\vec{v}_2 + 11\vec{v}_3$

Using the notation on the previous page then:

$$[f]_{\mathcal{V}} = \begin{pmatrix} 3 \\ 3 \\ 11 \end{pmatrix}$$

and

$$f = {}_{\mathcal{V}}\begin{pmatrix} 3 \\ 3 \\ 11 \end{pmatrix}$$

Ex:) $\mathbb{R}^3$ has a basis $\mathcal{V} = \left\{ \underbrace{\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}}_{\vec{v}_1}, \underbrace{\begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix}}_{\vec{v}_2}, \underbrace{\begin{pmatrix} 7 \\ 2 \\ 1 \end{pmatrix}}_{\vec{v}_3} \right\}$

We have
$$\begin{pmatrix} 5 \\ 6 \\ -3 \end{pmatrix} = 4\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} - 2\begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} + 1\begin{pmatrix} 7 \\ 2 \\ 1 \end{pmatrix}$$

So
$$\begin{pmatrix} 5 \\ 6 \\ -3 \end{pmatrix}_{\mathcal{V}} = \begin{pmatrix} 4 \\ -2 \\ 1 \end{pmatrix}$$

and
$$\begin{pmatrix} 5 \\ 6 \\ -3 \end{pmatrix} = {}_{\mathcal{V}}\begin{pmatrix} 4 \\ -2 \\ 1 \end{pmatrix}$$

Linear transformations

Using these ideas, we can use bases as a language for communicating about linear transformations.

Consider $T: V \to W$, with $\mathcal{V} = \{\vec{v}_1, \dots, \vec{v}_n\}$ a basis for V and $\mathcal{W} = \{\vec{w}_1, \dots, \vec{w}_m\}$ a basis for W:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, leftrightarrow, "\mathcal{V}"] & W \arrow[d, leftrightarrow, "\mathcal{W}"] \\ \mathbb{R}^n & \mathbb{R}^m \end{tikzcd}$$

All of the arrows are l.t.'s, so the composition from $\mathbb{R}^n \to \mathbb{R}^m$ is a l.t. — and thus is rep'd by a matrix:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, leftrightarrow, "\mathcal{V}"] & W \arrow[d, leftrightarrow, "\mathcal{W}"] \\ \mathbb{R}^n \arrow[r, "A"] & \mathbb{R}^m \end{tikzcd}$$

We call this matrix A the <u>matrix of T w.r.t. the bases $\mathcal{V}$ and $\mathcal{W}$</u>, written $A = [T]_{\mathcal{V}}^{\mathcal{W}}$

That is,

$$\left[T(\vec{v})\right]_{\mathcal{W}} = \left[T\right]_{\mathcal{V}}^{\mathcal{W}} \left[\vec{v}\right]_{\mathcal{V}}$$

Ex 1) Consider the linear transformation $D : P^3 \to P^2$ defined by $D(f) = f'$; and bases

$$\mathcal{V} = \{ \underbrace{x^3}_{\vec{v}_1}, \underbrace{x^2}_{\vec{v}_2}, \underbrace{x}_{\vec{v}_3}, \underbrace{1}_{\vec{v}_4} \} \quad \text{for } P^3 \text{ and}$$

$$\mathcal{W} = \{ \underbrace{x^2}_{\vec{w}_1}, \underbrace{x}_{\vec{w}_2}, \underbrace{1}_{\vec{w}_3} \} \quad \text{for } P^2$$

We can check that $[D]_{\mathcal{V}}^{\mathcal{W}}$ is the matrix

$$[D]_{\mathcal{V}}^{\mathcal{W}} = \begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

For example, $[x^2+x]_{\mathcal{V}} = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix}$, and

$\begin{bmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$. And, $D(x^2+x) = 2x+1$,

and $[2x+1]_{\mathcal{W}} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix}$.

How do we find the matrix $[T]_{\mathcal{V}}^{\mathcal{W}}$? We go "up and over" on our diagram:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, leftrightarrow, "\mathcal{V}"] & W \arrow[d, leftrightarrow, "\mathcal{W}"] \\ \mathbb{R}^n \arrow[r, "{[T]_{\mathcal{V}}^{\mathcal{W}}}"] & \mathbb{R}^m \end{tikzcd}$$

And recall that each column of $[T]_{\mathcal{V}}^{\mathcal{W}}$ is the image of the appropriate standard basis vector:

$$\begin{aligned} i\text{th col. of } [T]_{\mathcal{V}}^{\mathcal{W}} &= [T]_{\mathcal{V}}^{\mathcal{W}}\, \vec{e}_i \\ &= \Big[T\big({}_{\mathcal{V}}[\vec{e}_i]\big)\Big]_{\mathcal{W}} \quad \leftarrow \text{(up and over!)} \\ &= \big[T(\vec{v}_i)\big]_{\mathcal{W}} \end{aligned}$$

That is — the cols of $[T]_{\mathcal{V}}^{\mathcal{W}}$ are the $\mathcal{W}$ coordinates of the images of the corresponding $\mathcal{V}$ vectors.

Ex i) As in previous example:

$$D(\vec{v}_1) = D(x^3) = 3x^2 = 3\vec{w}_1$$

$$D(\vec{v}_2) = D(x^2) = 2x = 2\vec{w}_2$$

$$D(\vec{v}_3) = D(x) = 1 = 1\vec{w}_3$$

$$D(\vec{v}_4) = D(1) = 0 = 0$$

These give us columns:

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

Ex:) Suppose we have

$$[T]_{\mathcal{V}}^{\mathcal{W}} = \begin{pmatrix} 1 & 3 & 0 \\ 0 & 1 & 1 \\ 1 & 2 & 4 \end{pmatrix}$$

How do we interpret this matrix as a statement about T?

Note, the columns are the $\mathcal{W}$ coordinates of the images of the $\mathcal{V}$ vectors... So:

$$T(\vec{v}_1) = 1\vec{w}_1 + 0\vec{w}_2 + 1\vec{w}_3$$

$$T(\vec{v}_2) = 3\vec{w}_1 + 1\vec{w}_2 + 2\vec{w}_3$$

$$T(\vec{v}_3) = 0\vec{w}_1 + 1\vec{w}_2 + 4\vec{w}_3$$

If the bases are the same, we get a similar conclusion.

Ex:) Suppose

$$[T]_{\mathcal{V}}^{\mathcal{V}} = \begin{pmatrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{pmatrix}$$

Then

$$T(\vec{v}_1) = 4\vec{v}_1$$

$$T(\vec{v}_2) = 1\vec{v}_1 + 4\vec{v}_2$$

$$T(\vec{v}_3) = 1\vec{v}_2 + 4\vec{v}_3$$

Ex:) If $T: V \to V$, and if we use the <u>same</u> basis $\mathcal{V}$ for input and output, note that we get a familiar connection:

$$\text{ith col. of } [T]_{\mathcal{V}}^{\mathcal{V}} = [T(\vec{v}_i)]_{\mathcal{V}}$$

That is — <u>cols of matrix</u> are <u>images of basis vectors</u> (all in terms of the chosen basis.)

Ex:) Say we consider the l.t. that flips vectors over the line $Y=X$. Consider the basis $\mathcal{V} = \{\vec{v}_1, \vec{v}_2\}$ with

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \vec{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

Note $T(\vec{v}_1) = \vec{v}_1$, $T(\vec{v}_2) = -\vec{v}_2$. So, using the formula above, we have

$$[T]_{\mathcal{V}}^{\mathcal{V}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

Ex:) Say $T : \mathbb{R}^n \to \mathbb{R}^m$, where

$\mathbb{R}^n$ has bases $\mathcal{E}_1 = \{\vec{e}_1, \dots, \vec{e}_n\}$, $\mathcal{V} = \{\vec{v}_1, \dots, \vec{v}_n\}$

$\mathbb{R}^m$ has bases $\mathcal{E}_2 = \{\vec{e}_1, \dots, \vec{e}_m\}$, $\mathcal{W} = \{\vec{w}_1, \dots, \vec{w}_n\}$

Let $[\vec{v}_1]_{\mathcal{E}_1} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix}$, $[\vec{w}_1]_{\mathcal{E}_2} = \begin{pmatrix} d_1 \\ \vdots \\ d_m \end{pmatrix}$

$$[T]_{\mathcal{E}_1}^{\mathcal{E}_2} = A \quad , \quad [T]_{\mathcal{V}}^{\mathcal{W}} = M$$

Suppose that it happens that $T(\vec{v}_1) = \vec{w}_1$. Then the following all make that same statement:

$$A \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} d_1 \\ \vdots \\ d_m \end{pmatrix}$$

$$[T]_{\mathcal{E}_1}^{\mathcal{E}_2} [\vec{v}_1]_{\mathcal{E}_1} = [\vec{w}_1]_{\mathcal{E}_2}$$

$$T(\vec{v}_1) = \vec{w}_1$$

$$[T]_{\mathcal{V}}^{\mathcal{W}} [\vec{v}_1]_{\mathcal{V}} = [\vec{w}_1]_{\mathcal{W}}$$

$$M \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Change of Basis (for vectors)

Suppose we know $[\vec{v}]_{\mathcal{V}}$ and we want to find $[\vec{v}]_{\mathcal{W}}$.

We call this a <u>change of basis</u>.

Note that this is accomplished by the identity transformation, with these bases used for input, output:

$$\begin{tikzcd} V \arrow[r, "I"] \arrow[d, leftrightarrow, "\mathcal{V}"] & V \arrow[d, leftrightarrow, "\mathcal{W}"] \\ \mathbb{R}^n \arrow[r, "{[I]_{\mathcal{V}}^{\mathcal{W}}}"] & \mathbb{R}^n \end{tikzcd}$$

So we have (Change of basis matrix)

$$[\vec{v}]_{\mathcal{W}} = [I]_{\mathcal{V}}^{\mathcal{W}} [\vec{v}]_{\mathcal{V}}$$

As before, the <u>columns of $[I]$</u> are the <u>images of</u> the basis vectors:

$$\text{ith col. of } [I]_{\mathcal{V}}^{\mathcal{W}} = [I(\vec{v}_i)]_{\mathcal{W}}$$
$$= [\vec{v}_i]_{\mathcal{W}}$$

In this case then, we have that the columns are the $\mathcal{V}$ vectors written in the $\mathcal{W}$ basis.

Ex i) Let $\mathcal{V} = \{\vec{v}_1, \vec{v}_2\}$, $\mathcal{E} = \{\vec{e}_1, \vec{e}_2\}$, with

$$[\vec{v}_1]_{\mathcal{E}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix} , \quad [\vec{v}_2]_{\mathcal{E}} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

What is the change of basis matrix from $\mathcal{V}$ to $\mathcal{E}$?

From our previous discussion,

$$[I]_{\mathcal{V}}^{\mathcal{E}} = \begin{pmatrix} | & | \\ [\vec{v}_1]_{\mathcal{E}} & [\vec{v}_2]_{\mathcal{E}} \\ | & | \end{pmatrix} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

Note $[\vec{v}_1]_{\mathcal{V}} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $[\vec{v}_1]_{\mathcal{E}} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

and $\quad [\vec{v}_1]_{\mathcal{E}} = [I]_{\mathcal{V}}^{\mathcal{E}} [\vec{v}_1]_{\mathcal{V}}$

as it should be

<u>Note</u> : $\left([I]_{\mathcal{V}}^{\mathcal{W}}\right)^{-1} = [I]_{\mathcal{W}}^{\mathcal{V}}$

$$[I]_{\mathcal{V}}^{\mathcal{W}} [\vec{v}]_{\mathcal{V}} = [\vec{v}]_{\mathcal{W}}$$

~~Book: "C.o.B. matrix from $\mathcal{W}$ to $\mathcal{V}$"~~

Math 107: "C.O.B matrix from $\mathcal{V}$ to $\mathcal{W}$".

$$[I]_{\mathcal{V}}^{\mathcal{W}} = \begin{bmatrix} | & & | \\ [\vec{v}_1]_{\mathcal{W}} & \cdots & [\vec{v}_n]_{\mathcal{W}} \\ | & & | \end{bmatrix} \cancel{= P}$$

Change of Basis (for linear transformations)

Suppose we have a l.t. $T: V \to W$, bases $\mathcal{V}, \mathcal{V}'$ for V, and bases $\mathcal{W}, \mathcal{W}'$ for W.

Suppose we know $[T]_{\mathcal{V}}^{\mathcal{W}}$. How do we find $[T]_{\mathcal{V}'}^{\mathcal{W}'}$?

$$\begin{tikzcd}
V \arrow[rrr, "T"] \arrow[dr, leftrightarrow, "\mathcal{V}"] \arrow[dd, leftrightarrow, "\mathcal{V}'"'] & & & W \arrow[dr, leftrightarrow, "\mathcal{W}"] \arrow[dd, leftrightarrow, "\mathcal{W}'"'] & \\
& \mathbb{R}^n \arrow[rrr, "{[T]_{\mathcal{V}}^{\mathcal{W}}}"] & & & \mathbb{R}^m \arrow[dl, "{[I]_{\mathcal{W}}^{\mathcal{W}'}}"] \\
\mathbb{R}^n \arrow[ur, "{[I]_{\mathcal{V}'}^{\mathcal{V}}}"] \arrow[rrr, "{[T]_{\mathcal{V}'}^{\mathcal{W}'}}"] & & & \mathbb{R}^m &
\end{tikzcd}$$

We see from the above diagram that

$$[T]_{\mathcal{V}'}^{\mathcal{W}'} = [I]_{\mathcal{W}}^{\mathcal{W}'} [T]_{\mathcal{V}}^{\mathcal{W}} [I]_{\mathcal{V}'}^{\mathcal{V}}$$

Ex:) Suppose we want to find the matrix A with

$$A\begin{pmatrix}1\\3\end{pmatrix} = \begin{pmatrix}2\\5\end{pmatrix} \quad \text{and} \quad A\begin{pmatrix}2\\5\end{pmatrix} = \begin{pmatrix}1\\3\end{pmatrix}$$

Note, we can say $\quad A = [T]_{\mathcal{S}}^{\mathcal{S}}$

But consider instead the basis $\mathcal{V} = \{\vec{v}_1, \vec{v}_2\}$ with

$$\vec{v}_1 = \begin{pmatrix}1\\3\end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix}2\\5\end{pmatrix}$$

We can easily see that

$$[T]_{\mathcal{V}}^{\mathcal{V}} = \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}$$

and

$$[I]_{\mathcal{V}}^{\mathcal{S}} = \begin{pmatrix}1 & 2\\3 & 5\end{pmatrix}$$

Then $A = [T]_{\mathcal{S}}^{\mathcal{S}} = [I]_{\mathcal{V}}^{\mathcal{S}} [T]_{\mathcal{V}}^{\mathcal{V}} [I]_{\mathcal{S}}^{\mathcal{V}}$

$$= \underbrace{\begin{pmatrix}1 & 2\\3 & 5\end{pmatrix}\begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}} \underbrace{\begin{pmatrix}1 & 2\\3 & 5\end{pmatrix}^{-1}}$$

$$= \begin{pmatrix}2 & 1\\5 & 3\end{pmatrix} \quad \begin{pmatrix}-5 & 2\\3 & -1\end{pmatrix}$$

$$= \begin{pmatrix}-7 & 3\\-16 & 7\end{pmatrix} \qquad \checkmark$$

Ex:) Suppose we want to find the matrix A with

$$A\begin{pmatrix}1\\3\end{pmatrix} = \begin{pmatrix}3\\2\end{pmatrix} \quad \text{and} \quad A\begin{pmatrix}2\\5\end{pmatrix} = \begin{pmatrix}4\\3\end{pmatrix}$$

Note we can say $A = [T]_{\mathcal{S}}^{\mathcal{S}}$

Now consider bases $\mathcal{V} = \{\vec{v}_1, \vec{v}_2\}$, $\mathcal{W} = \{\vec{w}_1, \vec{w}_2\}$

with $\vec{v}_1 = \begin{pmatrix}1\\3\end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix}2\\5\end{pmatrix}$, $\vec{w}_1 = \begin{pmatrix}3\\2\end{pmatrix}$, $\vec{w}_2 = \begin{pmatrix}4\\3\end{pmatrix}$

We can easily see that

$$[T]_{\mathcal{V}}^{\mathcal{W}} = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$$

and $[I]_{\mathcal{V}}^{\mathcal{S}} = \begin{pmatrix}1 & 2\\3 & 5\end{pmatrix}$, $[I]_{\mathcal{W}}^{\mathcal{S}} = \begin{pmatrix}3 & 4\\2 & 3\end{pmatrix}$

Then

$$A = [T]_{\mathcal{S}}^{\mathcal{S}} = [I]_{\mathcal{W}}^{\mathcal{S}}[T]_{\mathcal{V}}^{\mathcal{W}}[I]_{\mathcal{S}}^{\mathcal{V}}$$

$$= \underbrace{\begin{pmatrix}3 & 4\\2 & 3\end{pmatrix}\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}}\begin{pmatrix}1 & 2\\3 & 5\end{pmatrix}^{-1}$$

$$= \begin{pmatrix}3 & 4\\2 & 3\end{pmatrix}\begin{pmatrix}-5 & 2\\3 & -1\end{pmatrix}$$

$$= \begin{pmatrix}-3 & 2\\-1 & 1\end{pmatrix} \quad \checkmark$$

5.4 – Eigenvalues & Eigenvectors

Often, a matrix A will have special vectors whose images are scalar multiples of themselves. That is,

$$A\vec{v} = \lambda \vec{v}$$

These vectors are called <u>eigenvectors</u>, and the corresponding scalar λ is called the <u>eigenvalue</u>.

(Sometimes the vectors represent other things — functions, quantum states, ... similarly then one uses terms such as "eigenfunction", "eigenstate", ...)

How do we find them? Note

$$A\vec{v} = \lambda\vec{v} \iff (A - \lambda I)\vec{v} = \vec{0}$$

So, a nontrivial eigenvector will have an eigenvalue for which $(A - \lambda I)$ is <u>singular</u>.

Often this is convenient to note algebraically by

$$\det(A - \lambda I) = 0$$

LHS is a polynomial in λ; called the <u>"characteristic poly"</u>.

Note, the eigenvalues are simply the roots of the characteristic polynomial

Ex!) What are the eigenvalues & eigenvectors of

$$A = \begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix}$$

The char. poly. is

$$\det(A-\lambda I) = \det\begin{pmatrix} 1-\lambda & -3 \\ -2 & 2-\lambda \end{pmatrix}$$

$$= (1-\lambda)(2-\lambda) - (-2)(-3)$$

$$= \lambda^2 - 3\lambda - 4$$

$$= (\lambda+1)(\lambda-4)$$

The roots, and thus the eigenvalues, are $-1, 4$.

To find the eigenvectors, we look for solutions to $(A-\lambda I)\vec{v} = \vec{0}$:

For $\lambda = -1$:

$$(A-\lambda I)\vec{v} = \begin{pmatrix} 2 & -3 \\ -2 & 3 \end{pmatrix}\vec{v} = \vec{0}$$

$$\ldots \Rightarrow \vec{v} = k\begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

For $\lambda = 4$:

$$(A-\lambda I)\vec{v} = \begin{pmatrix} -3 & -3 \\ -2 & -2 \end{pmatrix} \vec{v} = \vec{0}$$

$$\ldots \Rightarrow \quad \vec{v} = k\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

In this example, we had a $\underline{2} \times \underline{2}$ matrix, a degree $\underline{2}$ char. poly., $\underline{2}$ distinct roots/eigenvalues, and $\underline{2}$ eigenvectors (l.i.).

It's great when this happens... unfortunately this is not always the case.

Ex!) Consider

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

Char. poly is

$$(1-\lambda)(1-\lambda) - (1)(0) = (1-\lambda)^2$$

$$\Rightarrow \lambda = 1$$

Eigenvectors are solutions to

$$(A-\lambda I)\vec{v} = \vec{0}$$

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \vec{v} = \vec{0}$$

$$\ldots \Rightarrow \vec{v} = k \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

There were 2 roots if you count multiplicity... but, in this case there is only 1 eigenvector.

Ex:) Let

$$A = \begin{pmatrix} 2 & -1 & 3 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

Char. poly is $\det(\lambda I - A) = \ldots = (\lambda - 2)(\lambda + 1)^2$

Eigen values are $2, -1$.

For $\lambda = 2$:

$$(A - \lambda I)\vec{v} = \begin{pmatrix} 0 & -1 & 3 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix} \vec{v} = \vec{0}$$

$$\ldots \Rightarrow \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \vec{v} = \vec{0}$$

$$\Rightarrow \vec{v} = k \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

For $\lambda = -1$:

$$(A-\lambda I)\vec{v} = \begin{pmatrix} 3 & -1 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{v} = \vec{0}$$

$$\ldots \Rightarrow \vec{v} = k_1 \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix} + k_2 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

Here we get 2 (l.i.) eigenvectors for the eigenvalue of multiplicity 2. In total we have 3 eigenvalues (with multiplicity) and 3 eigenvectors (l.i.).

<u>Note</u>: 1) eigenvalues are roots of $\det(A-\lambda I)$

2) eigenvectors are solutions to $(A-\lambda I)\vec{v} = \vec{0}$

.... in $NS(A-\lambda I)$

We define the <u>eigenspace</u> for the eigenvalue λ by

$$E_\lambda = NS(A-\lambda I)$$

Ex: In previous example:

E_2 has basis $\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\}$

E_{-1} has basis $\left\{ \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$

Thm: If λ is a root of multiplicity m, then

$$1 \leq \dim(E_\lambda) \leq m$$

(We will not prove this)

So the most (l.i.) eigenvectors possible is when one has equality in the above for all λ, in which case

$$\begin{array}{ccccc} \dim(E_{\lambda_1}) & + \cdots + & \dim(E_{\lambda_k}) & & \\ \| & & \| & & \\ m_1 & + \cdots + & m_k & = & n \end{array}$$

Ex: In previous example,

$\lambda=2$ had multiplicity $\underline{1}$, and $\dim(E_2) = \underline{1}$

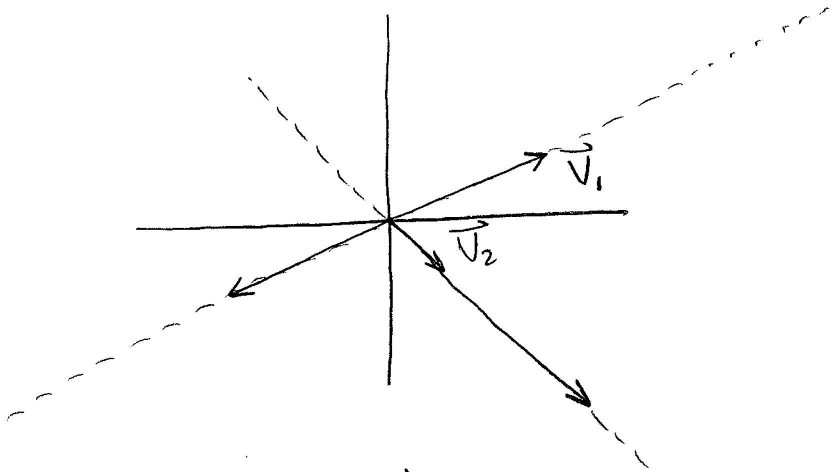
$\lambda=-1$ had multiplicity $\underline{2}$, and $\dim(E_{-1}) = \underline{2}$

In these cases, the l.t. has a simple interpretation w.r.t. the eigenvectors.

Geometrically:

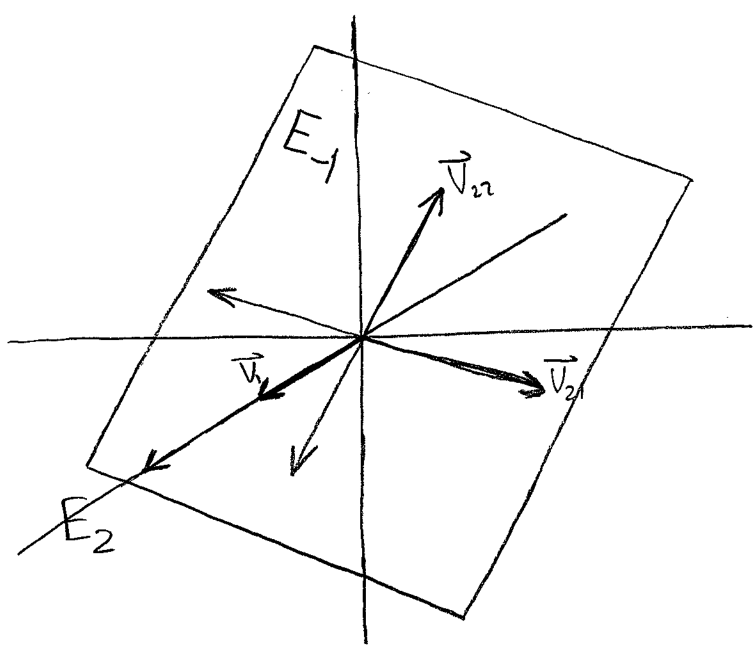
Ex: $A = \begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix}$, $\lambda_1 = -1$, $\vec{v}_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$

$\lambda_2 = 4$, $\vec{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$



Ex: $A = \begin{pmatrix} 2 & -1 & 3 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ $\lambda_1 = 2$ $\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

$\lambda_2 = -1$ $\vec{v}_{21} = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}$ $\vec{v}_{22} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$



Note, these matrices will be simple w.r.t. the basis of evects.

Note that our char. poly is a poly. with real coefficients. But, it might have complex roots...

Ex: Consider the rotation ($\pi/2$ ccwise) given by

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

Geometrically it would seem impossible for this to have eigenvectors... but:

$$\det(\lambda I - A) = \lambda^2 + 1 = 0$$

$$\Rightarrow \lambda = i, -i$$

For $\lambda = i$: $(A - \lambda I)\vec{v} = \vec{0}$

$$\begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \vec{v} = \vec{0}$$

$$\vdots$$

$$\begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix} \vec{v} = \vec{0} \Rightarrow \vec{v}_1 = \begin{pmatrix} 1 \\ -i \end{pmatrix}$$

For $\lambda = -i$:

$$\begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} \vec{v} = \vec{0}$$

$$\vdots$$

$$\begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix} \vec{v} = 0 \quad \Rightarrow \quad \vec{v}_2 = \begin{pmatrix} 1 \\ i \end{pmatrix}$$

So the rotation <u>matrix</u> has <u>complex</u> eigenvectors.

Thm: If A is real with complex eigenvalue λ, corresponding eigenspace E_λ with basis $\{\vec{v}_1, \dots, \vec{v}_k\}$, then $\bar{\lambda}$ is an eigenvalue, and $\{\overline{\vec{v}_1}, \dots, \overline{\vec{v}_k}\}$ is a basis for $E_{\bar{\lambda}}$.

(Proved in book.)

5.5 – Diagonalization and Jordan Canonical Form

In this section we will see strong connections between change of basis and eigenvectors.

Def:) A, B are <u>similar</u> if there is an invertible P with

$$B = P^{-1} A P$$

(here we say we are "conjugating" A by P; and A, B are "conjugates" of each other.)

Note, we have seen this sort of equation before...

$$[T]^{\mathcal{V}'}_{\mathcal{V}'} = \left([I]^{\mathcal{V}}_{\mathcal{V}'}\right)^{-1} [T]^{\mathcal{V}}_{\mathcal{V}} [I]^{\mathcal{V}}_{\mathcal{V}'}$$

In fact, all similar matrices can be viewed this way, because every invertible matrix P is a change of basis matrix.

Thm:) If P is invertible, then there exists a basis $\mathcal{V}$ for which $P = [I]^{\mathcal{S}}_{\mathcal{V}}$.

Pf:) The columns of P are this basis. (Note, because P is invertible, its columns are independent and thus form a basis.)

This theorem allows us to make a nice interpretation of this next definition.

Def: A matrix is <u>diagonalizable</u> if it is similar to a diagonal matrix.

In light of the previous theorem, we can rephrase this as

Thm: Say $A = [T]_{\mathcal{S}}^{\mathcal{S}}$. Then A is diagonalizable iff there exists a basis $\mathcal{V}$ for which

$$[T]_{\mathcal{V}}^{\mathcal{V}}$$

is a diagonal matrix.

Diagonal matrices are useful for many reasons. One of these is that the eigenvalues and eigenvectors are obvious.

Thm: If D is diagonal with ith diagonal entry equal to d_i, then the eigenvalues are $\{d_i\}$ and the eigenvectors are $\vec{e}_i$.

Similarly then, if A is diagonalizable, then the basis vectors in $\mathcal{V}$ are eigenvectors.

Ex:) Recall that

$$A = \begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix}$$

has eigenvalues and eigenvectors

$$\lambda_1 = -1 \qquad \vec{v}_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix}$$

$$\lambda_2 = 4 \qquad \vec{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

What happens if we form a basis $\mathcal{V} = \{\vec{v}_1, \vec{v}_2\}$ and write $A = [T]_{\mathcal{E}}^{\mathcal{E}}$ in this basis? Well,

$[T]_{\mathcal{V}}^{\mathcal{V}}$ has columns: (1) $[T(\vec{v}_1)]_{\mathcal{V}} = [-\vec{v}_1]_{\mathcal{V}} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$

(2) $[T(\vec{v}_2)]_{\mathcal{V}} = [4\vec{v}_2]_{\mathcal{V}} = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$

So $[T]_{\mathcal{V}}^{\mathcal{V}} = \begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}$

The change of basis connection is that

$$[T]_{\mathcal{V}}^{\mathcal{V}} = \left([I]_{\mathcal{V}}^{\mathcal{E}}\right)^{-1} [T]_{\mathcal{E}}^{\mathcal{E}} [I]_{\mathcal{V}}^{\mathcal{E}}$$

$$\begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 2 & -1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 2 & -1 \end{pmatrix}$$

eigenvalues

eigenvectors

All diagonalizable matrices can be viewed this way.

Thm: A is diagonalizable iff there is a basis for $\mathbb{R}^n$ consisting of eigenvectors.

Pf:

$$\underbrace{\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}}_{\text{eigenvalues}} = \begin{pmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{pmatrix}^{-1} \begin{pmatrix} & & \\ & A & \\ & & \end{pmatrix} \underbrace{\begin{pmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{pmatrix}}_{\text{change of basis matrix}}$$

eigenvectors

$$\Updownarrow$$

$$\begin{pmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{pmatrix} \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = \begin{pmatrix} & & \\ & A & \\ & & \end{pmatrix} \begin{pmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{pmatrix}$$

$$\Updownarrow$$

$$\begin{pmatrix} | & & | \\ \lambda_1\vec{v}_1 & \cdots & \lambda_n\vec{v}_n \\ | & & | \end{pmatrix} = \begin{pmatrix} & & \\ & A & \\ & & \end{pmatrix} \begin{pmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{pmatrix}$$

Alt:) $A = [T]_{\mathcal{E}}^{\mathcal{E}}$, and if A is diagonalizable we can also write

$$D = [T]_{\mathcal{V}}^{\mathcal{V}} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

for some basis $\mathcal{V} = \{\vec{v}_1, \dots, \vec{v}_n\}$. D has evident eigenvalues and eigenvectors

$$D \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} = \lambda_i \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

and this can be rewritten as

$$[T]_{\mathcal{V}}^{\mathcal{V}} [\vec{v}_i]_{\mathcal{V}} = \lambda_i [\vec{v}_i]_{\mathcal{V}}$$

$$T(\vec{v}_i) = \lambda_i \vec{v}_i$$

$$[T]_{\mathcal{E}}^{\mathcal{E}} [\vec{v}_i]_{\mathcal{E}} = \lambda_i [\vec{v}_i]_{\mathcal{E}}$$

$$A \vec{v}_i = \lambda_i \vec{v}_i$$

So $\mathcal{V}$ is a basis of eigenvectors.

(And this argument can be reversed.)

To determine if we can find such a basis, we use the following.

Thm: If we take bases for all of the eigenspaces and form a single set of vectors, that set is independent.

(Proved in book.)

The question then becomes simply whether there are a total of n vectors in this set...

Recalling

$$\begin{array}{ccccc} \dim(E_{\lambda_1}) & & & \dim(E_{\lambda_k}) & \\ \geq & & & \geq & \\ m_1 & + & \cdots & + \; m_k & = n \end{array}$$

we see that this will only happen if

$$\dim(E_{\lambda_i}) = m_i$$

for all eigenspaces.

Ex:) From our first example, the char. poly was

$$(\lambda+1)(\lambda-4)$$

So we had

$$\lambda_1 = -1 \qquad m_1 = 1$$

$$\lambda_2 = 4 \qquad m_2 = 1$$

Each of these eigenvalues had an eigenvector, so

$$\dim(E_{\lambda_1}) = 1 = m_1, \quad \dim(E_{\lambda_2}) = 1 = m_2$$

Ex:) Recall that for

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

we found the char. poly was $(\lambda-1)^2$, so there is only one eigenvalue:

$$\lambda_1 = 1 \qquad m_1 = 2$$

But there was only 1 eigenvector... so

$$\dim(E_{\lambda_1}) = 1 \neq m_1$$

So A is <u>not</u> diagonalizable.

Applications

Suppose you want to raise a matrix to a high power.

Diagonalizability helps!

Ex:)
$$\underbrace{\begin{pmatrix} -1 & 0 \\ 0 & 4 \end{pmatrix}}_{D} = \underbrace{\begin{pmatrix} 3 & 1 \\ 2 & -1 \end{pmatrix}^{-1}}_{P^{-1}} \underbrace{\begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 3 & 1 \\ 2 & -1 \end{pmatrix}}_{P}$$

So also $\quad A = PDP^{-1}$

Then
$$\begin{aligned} A^2 &= (PDP^{-1})(PDP^{-1}) \\ &= PD\cancel{(P^{-1}P)}DP^{-1} \\ &= PD^2P^{-1} \end{aligned}$$

Similarly,
$$A^n = PD^nP^{-1}$$

So
$$\begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix}^n = \begin{pmatrix} 3 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} (-1)^n & 0 \\ 0 & 4^n \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 2 & -1 \end{pmatrix}^{-1}$$

Ex: What is the nth Fibonacci number?

$$0, 1, 1, 2, 3, 5, \dots$$

$$f_0, f_1, f_2, f_3, f_4, f_5, \dots \qquad f_n = f_{n-1} + f_{n-2}$$

Note we can write

$$\begin{pmatrix} f_{n+1} \\ f_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} f_n \\ f_{n-1} \end{pmatrix}$$

And thus

$$\begin{pmatrix} f_{n+1} \\ f_n \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n \begin{pmatrix} 1 \\ 0 \end{pmatrix}$$

So, f_n is the bottom-left element of

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n = A^n$$

A has eigenvalues $g, \ -1/g$ where $g = \frac{1+\sqrt{5}}{2}$

(golden ratio)

Eigenvectors are $\begin{pmatrix} g \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -g \end{pmatrix}$

So

$$A^n = \begin{pmatrix} g & 1 \\ 1 & -g \end{pmatrix} \begin{pmatrix} g^n & 0 \\ 0 & (-1/g)^n \end{pmatrix} \begin{pmatrix} g & 1 \\ 1 & -g \end{pmatrix}^{-1}$$

Ex:) Suppose two quantities a and b take values in stages, and each value depends on the values in the previous stage.

$$\begin{pmatrix} a_{k+1} \\ b_{k+1} \end{pmatrix} = \begin{pmatrix} & M & \end{pmatrix} \begin{pmatrix} a_k \\ b_k \end{pmatrix}$$

Then
$$\begin{pmatrix} a_k \\ b_k \end{pmatrix} = M^k \begin{pmatrix} a_0 \\ b_0 \end{pmatrix}$$

Again, diagonalization allows us to compute this.

$$M = PDP^{-1}$$
$$M^k = PD^kP^{-1}$$

$$\begin{pmatrix} a_k \\ b_k \end{pmatrix} = PD^kP^{-1} \begin{pmatrix} a_0 \\ b_0 \end{pmatrix}$$

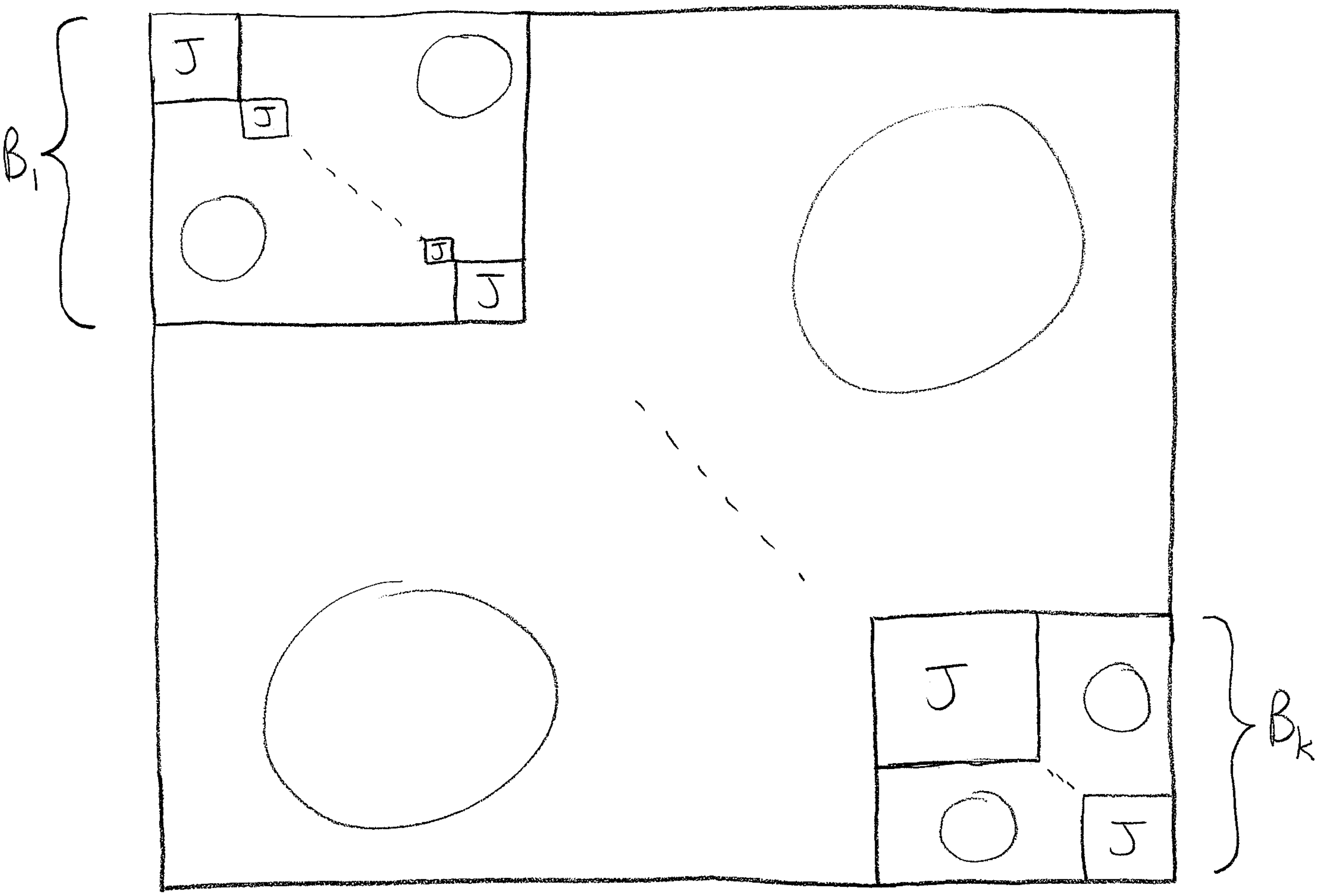
Jordan Canonical Form

What if A is not diagonalizable?

Can *always* put into Jordan form (by similarity / change of basis)

Thm: Say A has char. poly $= (\lambda - r_1)^{m_1} \cdots (\lambda - r_k)^{m_k}$

Then A is similar to :



(You can rearrange the blocks; this corresponds to permuting the basis vectors.)

We will refer to the blocks $B_1, \ldots, B_k$ as <u>eigenvalue blocks</u>. Each eigenvalue block B_i is associated to the corresponding eigenvalue λ_i. Good news — the dimensions of B_i are determined by λ_i's multiplicity m_i! (Remember this was not the case for eigenvectors.)

Each eigenvalue block is itself made up of blocks, called <u>basic Jordan blocks</u>. Each basic Jordan block looks like

$$\begin{pmatrix} \lambda_i & 1 & & 0 \\ & \lambda_i & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix}$$

Remember these are along the diagonal of A; and the <u>first</u> column of such a Jordan block is the <u>only</u> one without other stuff (other than the λ_i on the diagonal).

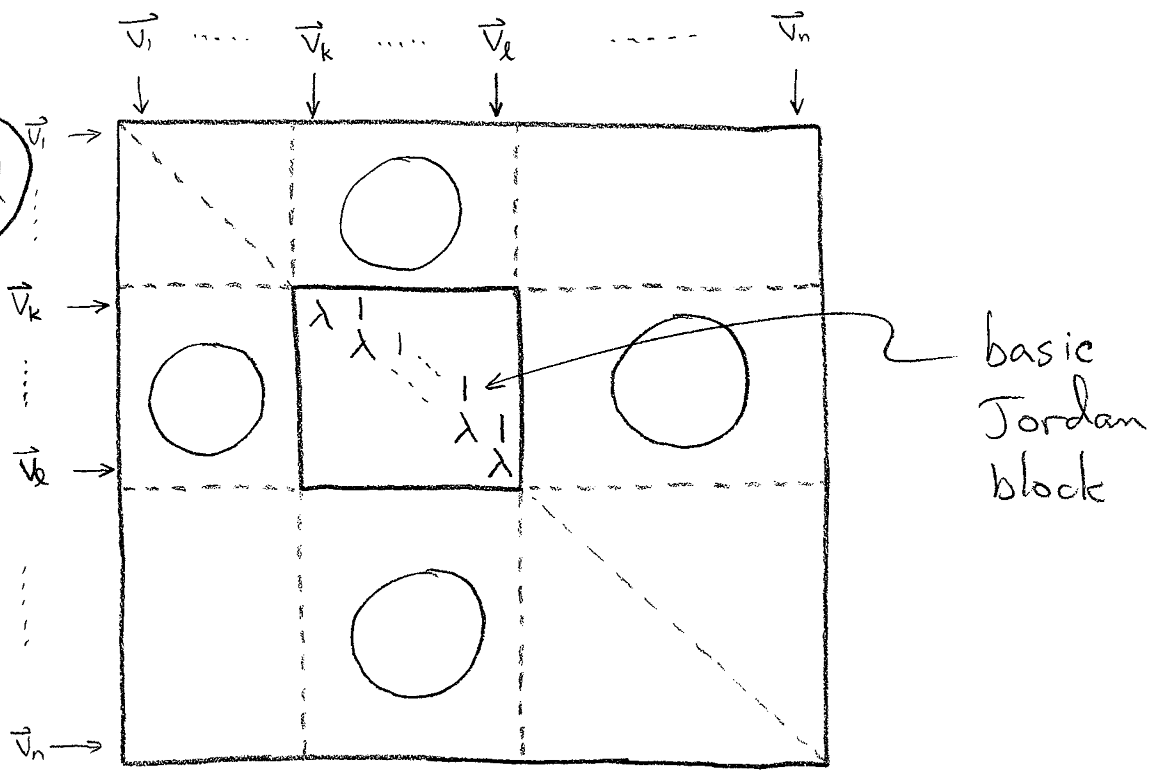
So — for <u>each</u> basic Jordan block there is <u>one</u> associated eigenvector.

So: $\begin{pmatrix} \text{\# of eigenvectors} \\ \text{for eigenvalue } \lambda_i \end{pmatrix} = \begin{pmatrix} \text{\# of basic Jordan} \\ \text{blocks in } B_i \end{pmatrix}$

$$J = [T]_{\mathcal{V}}^{\mathcal{V}} \quad , \quad \mathcal{V} = \{\vec{v}_1, \dots, \vec{v}_n\}$$

<u>Columns</u> give images of $\mathcal{V}$ basis vectors.

<u>Rows</u> correspond to coeffs of those images for $\mathcal{V}$ basis vectors.



Ex:) What does it mean about a matrix A if the Jordan form is

$$J = \begin{pmatrix} 4 & 1 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

A is the matrix for a l.t. T w.r.t. the standard basis $\mathcal{S}$; J is a change of basis on A, using the Jordan basis $\mathcal{V} = \{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$. That is,

$$J = [T]_{\mathcal{V}}^{\mathcal{V}}$$

So (writing in terms of the standard basis),

$$A\vec{v}_1 = 4\vec{v}_1$$

$$A\vec{v}_2 = 1\vec{v}_1 + 4\vec{v}_2$$

$$A\vec{v}_3 = \qquad\qquad 4\vec{v}_3$$

Notice — $\vec{v}_1, \vec{v}_3$ are eigenvectors, as we see from the fact that the corresponding <u>columns</u> have nonzero entries only on the <u>diagonal</u>.

The Jordan form is the matrix A with respect to a basis, which we will call the Jordan basis.

$$\left(J = [T]_{\mathcal{V}}^{\mathcal{V}}\right) = \left([I]_{\mathcal{V}}^{\mathcal{S}} \right)^{-1} \left(A = [T]_{\mathcal{S}}^{\mathcal{S}} \right) \left([\vec{v}_1]_{\mathcal{S}} \; \cdots \; [\vec{v}_n]_{\mathcal{S}} \right)$$

Jordan basis

Jordan form allows us to find something slightly weaker than eigenvectors, but which like eigenvectors and eigenspaces is preserved by the matrix.

Let $\{\vec{v}_{i1}, \ldots, v_{i\ell}\}$ be the vectors in the Jordan basis corresponding to the eigenvalue block B_i. Then, because of the block, we have

$$(A)\left(\operatorname{span}\{\vec{v}_{i1}, \ldots, \vec{v}_{i\ell}\}\right) = \operatorname{span}\{\vec{v}_{i1}, \ldots, \vec{v}_{i\ell}\}$$

This "Jordan subspace" is preserved by A, much like the corresponding eigenspace is:

$$A E_{\lambda_i} = E_{\lambda_i}$$

But the Jordan subspace contains more than the eigenvectors.

We will <u>not</u> prove the theorem about Jordan canonical form.

We will <u>not</u> discuss how to find the basis that puts a non-diagonalizable matrix into Jordan form.

Still, in some cases, we can still do this.

<u>Ex:</u> Note that if A is diagonalizable, then its diagonal form <u>is</u> Jordan form.
(Every basic Jordan block is 1×1, and every vector in the Jordan basis is an eigenvector.)

Ex:) Suppose A has characteristic polynomial

$$p(\lambda) = (\lambda-1)^2(\lambda-3)$$

We have $\lambda_1 = 1$, $m_1 = 2$

$\lambda_2 = 3$, $m_2 = 1$

So there are only 2 possible Jordan forms:

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

If λ_1 has 2 eigenvectors, then it is the former.

If λ_1 has 1 eigenvector, then it is the latter.

<u>Obs.</u> Within an eigenvalue block B_i with multiplicity m_i, the number of possibilities is the <u>partition function</u> (the number of different ways positive integers can add up to m_i.)

<u>Ex i)</u> How many Jordan forms are possible if the char. poly. is

$$(\lambda-7)^4(\lambda-5)^3 \quad ?$$

$$\begin{aligned} 4 &= 4 \\ &= 3+1 \\ &= 2+2 \\ &= 2+1+1 \\ &= 1+1+1+1 \end{aligned} \qquad \begin{aligned} 3 &= 3 \\ &= 2+1 \\ &= 1+1+1 \end{aligned}$$

There are 5 possible B_1 ; 3 possible B_2.

So there are 15 possible Jordan forms.

9.1 – Inner Product Spaces

Recall the following facts about the dot product in $\mathbb{R}^n$:

Def: $\vec{u}\cdot\vec{v} = u_1v_1 + \cdots + u_nv_n$

Facts:

$$\vec{u}\cdot\vec{v} = \vec{v}\cdot u$$

$$(\vec{u}+\vec{v})\cdot\vec{w} = \vec{u}\cdot\vec{w} + \vec{v}\cdot\vec{w}$$

$$(c\vec{u})\cdot\vec{v} = c(\vec{u}\cdot\vec{v}) \qquad \text{(note difference in book!)}$$

$$\vec{v}\cdot\vec{v} \geq 0 \text{ ; equality iff } \vec{v}=\vec{0}$$

More Facts:

$$\vec{0}\cdot\vec{v} = 0$$

$$\vec{v}\cdot(\vec{u}+\vec{w}) = \vec{v}\cdot\vec{u} + \vec{v}\cdot\vec{w}$$

$$\vec{v}\cdot(c\vec{w}) = c(\vec{v}\cdot\vec{w})$$

Still More Facts:

$$\|\vec{v}\| = \sqrt{\vec{v}\cdot\vec{v}} = \sqrt{\vec{v}^T\vec{v}}$$

$$|\vec{v}\cdot\vec{w}| \leq \|\vec{v}\|\,\|\vec{w}\|$$

$$\vec{v}\cdot\vec{w} = \|\vec{v}\|\,\|\vec{w}\|\cos\theta$$

$$\theta = \arccos\left(\frac{\vec{v}\cdot\vec{w}}{\|\vec{v}\|\,\|\vec{w}\|}\right)$$

We will use the first 4 facts above to define a generalization of this dot product for other kinds of vector spaces.

Def: V a vector space, suppose we have a function $\langle , \rangle$ on pairs of vectors, with real values, satisfying

1. $\langle u,v \rangle = \langle v,u \rangle$
2. $\langle u+v, w \rangle = \langle u,w \rangle + \langle v,w \rangle$
3. $\langle cu, v \rangle = c \langle u,v \rangle$
4. $\langle v,v \rangle \geq 0$; equality iff $v = 0$ (different in book!)

The function $\langle , \rangle$ is called an <u>inner product</u>, and V is called an <u>inner product space.</u>

Ex: $\mathbb{R}^n$, $\langle \vec{u}, \vec{v} \rangle = \vec{u} \cdot \vec{v}$ is an inner product space

Ex: On $C[a,b]$, define

$$\langle f,g \rangle = \int_a^b f(x) g(x) \, dx$$

Can check this is an inner product. (Called the <u>L^2 inner product</u>.)

Ex: Note dot product is <u>not</u> preserved by change of basis...

That is, $[\vec{v}]_{\mathcal{B}} \cdot [\vec{w}]_{\mathcal{B}} \neq [\vec{v}]_{\mathcal{G}} \cdot [\vec{w}]_{\mathcal{G}}$

But this <u>does</u> give us a different inner product;

$$\langle \vec{v}, \vec{w} \rangle = [\vec{v}]_{\mathcal{G}} \cdot [\vec{w}]_{\mathcal{G}}$$

(Think about this in Exercise 9.)

Note, the "More Facts" about dot products can be shown for all inner products, using the definition.

Note also that inner products are linear in each entry:

$$\langle c_1\vec{v}_1 + \cdots + c_n\vec{v}_n , \vec{w} \rangle = c_1\langle \vec{v}_1, \vec{w}\rangle + \cdots + c_n\langle \vec{v}_n, \vec{w}\rangle$$

$$\langle \vec{w}, c_1\vec{v}_1 + \cdots + c_n\vec{v}_n \rangle = c_1\langle \vec{w}, \vec{v}_1\rangle + \cdots + c_n\langle \vec{w}, \vec{v}_n\rangle$$

We have defined inner products for vector spaces that might not be Euclidean, and so might not have pre-existing notions of length and angle. But we can use the "Still More Facts" to motivate definitions for these.

Def:) For an inner product space V, define

$$\|\vec{v}\| = \sqrt{\langle \vec{v}, \vec{v}\rangle}$$

We call this the norm or magnitude of $\vec{v}$.
(Can think of this as a sort of "length".)

Ex:) $C^0[a,b]$ with the L^2 inner product gives us

$$\|f\| = \left(\int_a^b (f(x))^2\,dx\right)^{1/2}$$

This is called the L^2 norm (there are similar norms called L^p norms that do not come from inner products.)

Def: The <u>angle</u> between two vectors $\vec{v}, \vec{w}$ is defined as

$$\theta = \arccos\left(\frac{\langle \vec{v}, \vec{w} \rangle}{\|\vec{v}\| \, \|\vec{w}\|}\right)$$

Ex: What is the angle between $\sin(x)$, $\sin(x+\phi)$ in $C^0[0, 2\pi]$?

① $\|\sin x\| = \left(\int_0^{2\pi} \sin^2 x \, dx\right)^{1/2} = \sqrt{\pi}$

② $\|\sin(x+\phi)\| = \left(\int_0^{2\pi} \sin^2(x+\phi) \, dx\right)^{1/2} = \sqrt{\pi}$

③ $\langle \sin x, \sin(x+\phi) \rangle = \int_0^{2\pi} (\sin x)(\sin x \cos\phi + \cos x \sin\phi) \, dx$

$$= (\cos\phi)\left(\int_0^{2\pi} \sin^2 x \, dx\right) + (\sin\phi)\left(\int_0^{2\pi} \sin x \cos x \, dx\right)$$

$$= (\cos\phi)(\pi)$$

④ $\theta = \arccos\left(\frac{\pi \cos\phi}{\sqrt{\pi}\sqrt{\pi}}\right) = \phi$ ☺

How do we know the above arccos will always be defined?

The Cauchy-Schwarz inequality gives us this.

Thm: For <u>any</u> inner product space, we have

$$|\langle u, v \rangle| \leq \|u\| \, \|v\|$$

(proved in book)

Another example of angle involves statistics.

Ex:) Suppose X and Y are real random variables on the sample space S. The expression

$$\langle X,Y\rangle = E\big((X-\mu_X)(Y-\mu_Y)\big)$$

is an inner product on these random variables (called the covariance). Other statistical notions derive from this

$$\langle X,Y\rangle = \text{covariance} = \mathrm{cov}(X,Y)$$

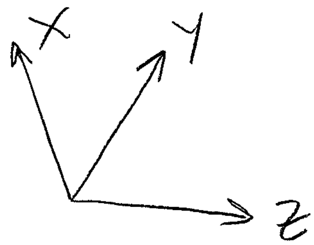
$$\langle X,X\rangle = \text{variance}$$

$$\|X\| = \sqrt{\langle X,X\rangle} = \text{standard deviation}$$

$$\cos\theta = \frac{\langle X,Y\rangle}{\|X\|\,\|Y\|} = \text{correlation}$$

This connection to angle leads naturally to a surprising fact about random variables — when X,Y are positively correlated, and Y,Z are positively correlated, X,Z might still be negatively correlated!

From a vector point of view, positive correlation relates to an acute angle. Geometrically then:



An example: Let $S = \{ \vec{v} \in \mathbb{R}^2 \mid \|\vec{v}\| \le 1 \}$, and

$\vec{x}, \vec{y}, \vec{z}$ three vectors in $\mathbb{R}^2$ arranged as above.

Then define
$$X = \vec{x} \cdot \vec{v}$$
$$Y = \vec{y} \cdot \vec{v}$$
$$Z = \vec{z} \cdot \vec{v}$$

These are random variables on S, with

$$\operatorname{cov}(X, Y) > 0$$
and
$$\operatorname{cov}(Y, Z) > 0$$

<u>but</u>
$$\operatorname{cov}(X, Z) < 0$$

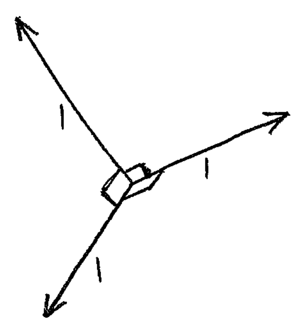
9.2 – Orthonormal Bases

Recall $\vec{v}, \vec{w} \in V$ are <u>orthogonal</u> if $\langle \vec{v}, \vec{w} \rangle = 0$.

Similarly:

<u>Def:</u> $\mathcal{V} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is an <u>orthogonal basis</u> for V iff it is a basis and the vectors are orthogonal.

<u>Def:</u> $\mathcal{V} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is an <u>orthonormal basis</u> for V iff it is orthogonal <u>and</u> $\|\vec{v}_i\| = 1$.



<u>Ex:</u> $\mathcal{E}$ is an orthonormal basis for $\mathbb{R}^n$

<u>Ex:</u> $\left\{ \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix}, \begin{pmatrix} 1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{pmatrix}, \begin{pmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ -1/\sqrt{3} \end{pmatrix} \right\}$ is an orthonormal basis for $\mathbb{R}^3$

Here are some properties of orthonormal bases.

Thm: If $\mathcal{V} = \{\vec{v}_1, \dots, \vec{v}_n\}$ is an orth. basis for V, then for any $\vec{v} \in V$, the coords. are the projections.

That is
$$[\vec{v}]_{\mathcal{V}} = \begin{pmatrix} \langle \vec{v}, \vec{v}_1 \rangle \\ \vdots \\ \langle \vec{v}, \vec{v}_n \rangle \end{pmatrix}$$

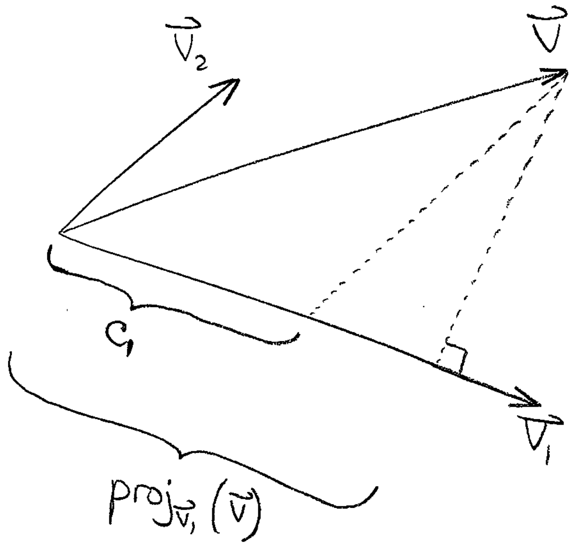
Pf: Say $\vec{v} = c_1\vec{v}_1 + \dots + c_n\vec{v}_n$. Then

$$\begin{aligned} \langle \vec{v}, \vec{v}_i \rangle &= \langle c_1\vec{v}_1 + \dots + c_n\vec{v}_n \;, \vec{v}_i \rangle \\ &= c_1 \underbrace{\langle \vec{v}_1, \vec{v}_i \rangle}_{0} + \dots + c_i \underbrace{\langle \vec{v}_i, \vec{v}_i \rangle}_{1} + \dots + c_n \underbrace{\langle \vec{v}_n, \vec{v}_i \rangle}_{0} \\ &= c_i \end{aligned}$$

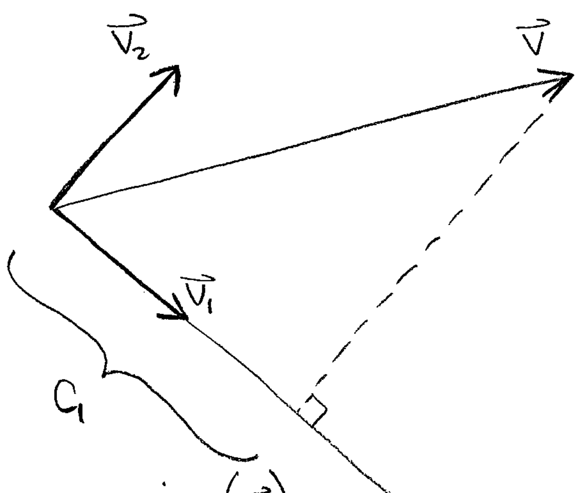
Ex: How do we write $\vec{v} = (1,2,3)$ as a l.c. of the orth. basis on prev. page?

$$\left.\begin{aligned} c_1 &= \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} = \frac{3}{\sqrt{2}} \\ c_2 &= \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{6} \\ -1/\sqrt{6} \\ 2/\sqrt{6} \end{pmatrix} = \frac{5}{\sqrt{6}} \\ c_3 &= \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{3} \\ -1/\sqrt{3} \\ -1/\sqrt{3} \end{pmatrix} = \frac{-4}{\sqrt{3}} \end{aligned}\right\} \Rightarrow [\vec{v}]_{\mathcal{V}} = \begin{pmatrix} 3/\sqrt{2} \\ 5/\sqrt{6} \\ -4/\sqrt{3} \end{pmatrix}$$

Geometrically:



When $\mathcal{V} = \{\vec{v}_1, \vec{v}_2\}$ is <u>not</u> orthonormal, the coordinates are <u>not</u> necessarily equal to the projections.



$= \text{proj}_{\vec{v}_1}(\vec{v})$

$$= \frac{\langle \vec{v}, v_1 \rangle}{\|\vec{v}_1\|} = \langle \vec{v}, \vec{v}_1 \rangle$$

Def: A matrix A is an <u>orthogonal matrix</u> if the column vectors form an orthonormal basis.

$$A = \begin{pmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{pmatrix},$$ $\{\vec{v}_1, \dots, \vec{v}_n\}$ is an orthonormal basis

(NB, the book reverses this def and thm)

Thm: A is orthogonal $\iff$ $A^TA = I$

Pf: $$A^TA = \begin{pmatrix} \text{---} \ \vec{v}_1 \ \text{---} \\ \vdots \\ \text{---} \ \vec{v}_n \ \text{---} \end{pmatrix} \begin{pmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{pmatrix}$$

The entries in this product are rows dot columns, which are

$$\vec{v}_i \cdot \vec{v}_j$$

which are 0 if $i \neq j$ and 1 if $i = j$.

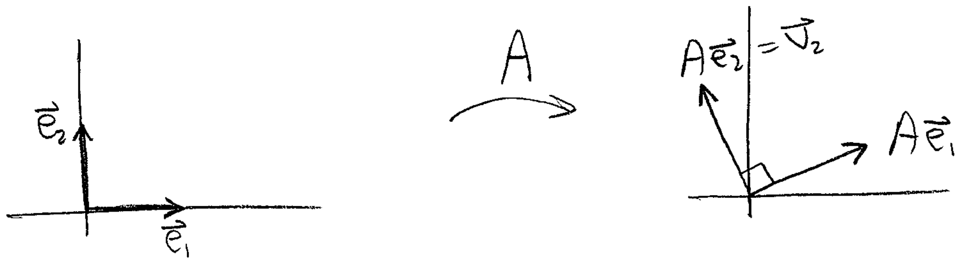
So $A^TA = I$.

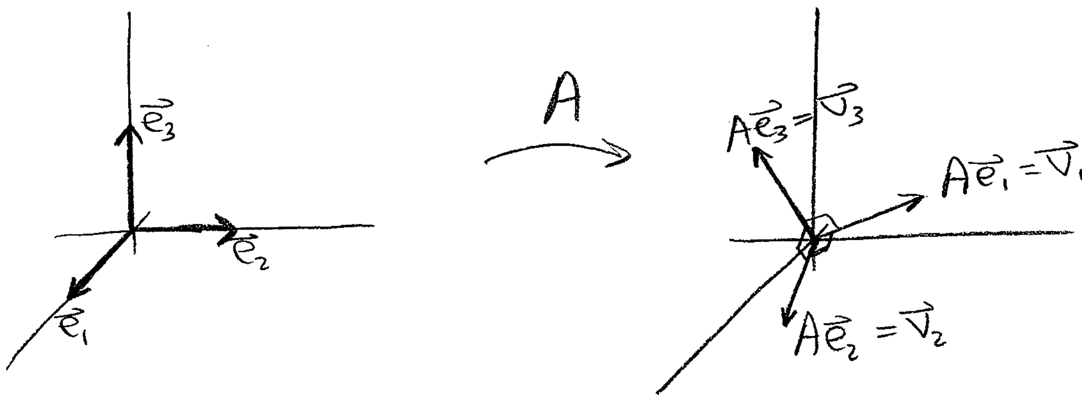
A nice application of this involves dot products.

Thm: If A is orthogonal, then it preserves dot products.

Pf: $$(A\vec{v}) \cdot (A\vec{w}) = (A\vec{v})^T (A\vec{w})$$
$$= \vec{v}^T A^T A \vec{w}$$
$$= \vec{v}^T \vec{w}$$
$$= \vec{v} \cdot \vec{w}$$

Note, since lengths and angles are written in terms of dot products, these are preserved too! Such matrices are rigid motions on $\mathbb{R}^n$.





They are <u>all</u> rotations with a possible reflection.

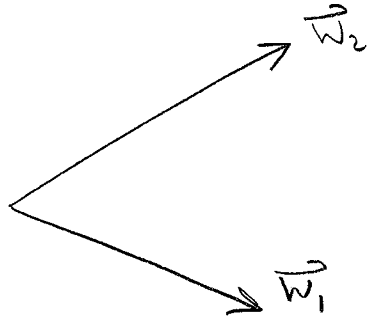
("SO_3" is the set of all rotations (no reflection) in $\mathbb{R}^3$.
There is a proof of a special property of SO_3 that <u>literally</u> involves handwaving... (demo in class).)

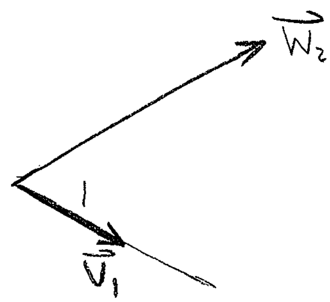
How can we convert a basis that is <u>not</u> orthonormal into another one that is?

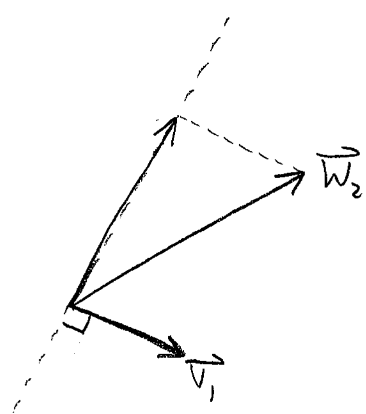
Gram-Schmidt Orthonormalization

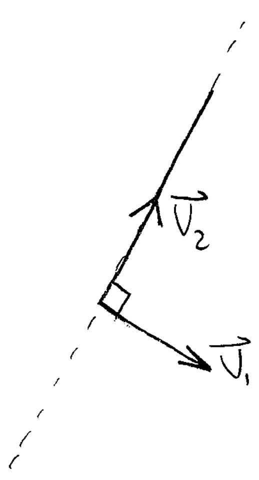
Geometrically:

Given





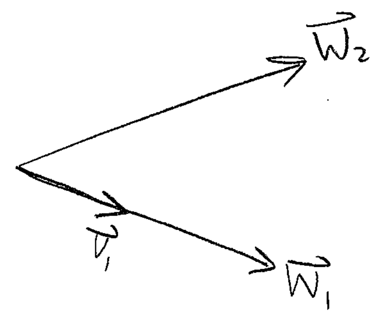




⟵ orthonormal!

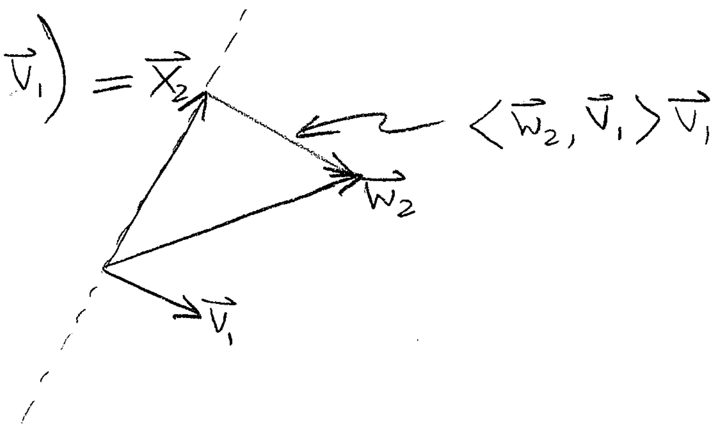
Basically, we use $\vec{v}_1$ to adjust $\vec{w}_2$ to something orthogonal to $\vec{v}_1$, and then divide by length to "normalize".

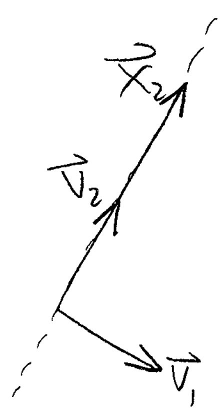
Algebraically:



$$\vec{v}_1 = \frac{\vec{w}_1}{\|\vec{w}_1\|}$$

$$\vec{w}_2 - \left(\langle \vec{w}_2, \vec{v}_1 \rangle \vec{v}_1\right) = \vec{x}_2$$

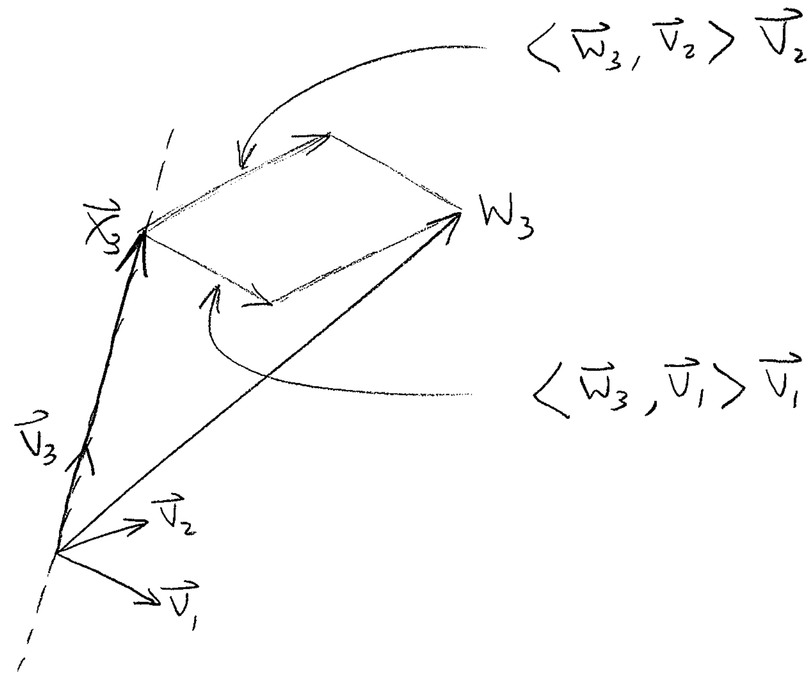




$$\vec{v}_2 = \frac{\vec{x}_2}{\|\vec{x}_2\|}$$

What if we have a basis with 3 vectors?

Same process for 1st two... Then:



$$\vec{X}_3 = \vec{W}_3 - \langle \vec{W}_3, \vec{V}_1 \rangle \vec{V}_1 - \langle \vec{W}_3, \vec{V}_2 \rangle \vec{V}_2$$

$$\vec{V}_3 = \frac{\vec{X}_3}{\| \vec{X}_3 \|}$$

Similarly for bases with even more vectors...

Ex:) Apply the Gram-Schmidt process to the basis

$$\mathcal{W} = \left\{ \begin{pmatrix} 2 \\ 3 \\ 6 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} \right\} = \{ \vec{w}_1, \vec{w}_2, \vec{w}_3 \}$$

Compute $\vec{v}_1$: $\quad \vec{v}_1 = \dfrac{\vec{w}_1}{\|\vec{w}_1\|} = \begin{pmatrix} 2 \\ 3 \\ 6 \end{pmatrix} / 7 = \begin{pmatrix} 2/7 \\ 3/7 \\ 6/7 \end{pmatrix}$

Compute $\vec{v}_2$: First, $\quad \vec{x}_2 = \vec{w}_2 - \langle \vec{w}_2, \vec{v}_1 \rangle \vec{v}_1$

$$= \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} - \left(\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 2/7 \\ 3/7 \\ 6/7 \end{pmatrix} \right) \begin{pmatrix} 2/7 \\ 3/7 \\ 6/7 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} - \left(\frac{17}{7} \right) \begin{pmatrix} 2/7 \\ 3/7 \\ 6/7 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} - \begin{pmatrix} 34/49 \\ 51/49 \\ 102/49 \end{pmatrix} = \begin{pmatrix} 15/49 \\ -2/49 \\ -4/49 \end{pmatrix}$$

Then $\quad \vec{v}_2 = \dfrac{\vec{x}_2}{\|\vec{x}_2\|} = \dfrac{\vec{x}_2}{\sqrt{245}/49} = \begin{pmatrix} 15/7\sqrt{5} \\ -2/7\sqrt{5} \\ -4/7\sqrt{5} \end{pmatrix}$

Compute $\vec{V}_3$:

$$\vec{X}_3 = \vec{W}_3 - \langle \vec{W}_3, \vec{V}_1 \rangle \vec{V}_1 - \langle \vec{W}_3, \vec{V}_2 \rangle \vec{V}_2$$

$$= \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} - \left(\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} 2 \\ 3 \\ 6 \end{pmatrix} / 7 \right) \begin{pmatrix} 2 \\ 3 \\ 6 \end{pmatrix} / 7$$

$$- \left(\begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} \cdot \begin{pmatrix} 15 \\ -2 \\ -4 \end{pmatrix} / 7\sqrt{5} \right) \begin{pmatrix} 15 \\ -2 \\ -4 \end{pmatrix} / 7\sqrt{5}$$

$$= \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} - \frac{26}{49} \begin{pmatrix} 2 \\ 3 \\ 6 \end{pmatrix} - \frac{-1}{245} \begin{pmatrix} 15 \\ -2 \\ -4 \end{pmatrix}$$

$$= \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} - \begin{pmatrix} 260 \\ 390 \\ 780 \end{pmatrix} / 245 + \begin{pmatrix} 15 \\ -2 \\ -4 \end{pmatrix} / 245$$

$$= \begin{pmatrix} 0 \\ -392 \\ 196 \end{pmatrix} / 245$$

$$= \begin{pmatrix} 0 \\ -8 \\ 4 \end{pmatrix} / 5 = \begin{pmatrix} 0 \\ -8/5 \\ 4/5 \end{pmatrix}$$

$$\vec{V}_3 = \frac{\vec{X}_3}{\|\vec{X}_3\|} = \frac{\vec{X}_3}{\sqrt{80}/5} = \frac{\vec{X}_3}{4/\sqrt{5}} = \begin{pmatrix} 0 \\ -2 \\ 1 \end{pmatrix} / \sqrt{5}$$

9.3 — Complex Inner Product Spaces, Symmetric Matrices

Recall that the vector spaces we have talked about up to now were "vector spaces <u>over $\mathbb{R}$</u>"; that is, the scalar multiplication was only by <u>real</u> numbers.

A similar (but different) thing is a "vector space over $\mathbb{C}$", in which we allow scalar multiplication by <u>complex</u> numbers.

<u>Ex:</u> $\mathbb{C}$ itself is a vector space over $\mathbb{C}$.

It "has 1 complex dimension" ($\{1\}$ is a basis).

On the other hand, we could have viewed $\mathbb{C}$ as a vector space over $\mathbb{R}$; as such it "has 2 real dimensions" ($\{1, i\}$ is a basis).

<u>Ex:</u> $\mathbb{C}^n = \{ (z_1, \ldots, z_n) \mid z_i \in \mathbb{C} \}$ is a vector space over $\mathbb{C}$ of dimension n.

It is a vector space over $\mathbb{R}$ of dimension $2n$.

We have to treat vector spaces over $\mathbb{C}$ totally differently. Specifically, we must reformulate the idea of an inner product for this new category of objects.

So, thinking of $\mathbb{C}^n$ as a vector space over $\mathbb{C}$, ...

How do we define an inner product on $\mathbb{C}^n$?

Can <u>not</u> use $\sum_i v_i w_i$!

<u>Ex:</u> $\vec{v} = \begin{pmatrix} 1 \\ i \end{pmatrix}$, $\sum_i v_i v_i = 1 \cdot 1 + i \cdot i = 0$

This defies one of the desirable properties of an inner prod!

<u>Def:</u> The <u>Hermitian dot product</u> on $\mathbb{C}^n$ is

$$\langle \vec{v}, \vec{w} \rangle_H = \sum v_i \overline{w_i} = \vec{v}^T \overline{\vec{w}}$$

<u>Properties:</u> ① $\langle \vec{v}, \vec{w} \rangle_H = \overline{\langle \vec{w}, \vec{v} \rangle_H}$

② $\langle \vec{u} + \vec{v}, \vec{w} \rangle_H = \langle \vec{u}, \vec{w} \rangle_H + \langle \vec{v}, \vec{w} \rangle_H$

③ $\langle c\vec{v}, \vec{w} \rangle_H = c \langle \vec{v}, \vec{w} \rangle_H$

$\langle \vec{v}, c\vec{w} \rangle_H = \overline{c} \langle \vec{v}, \vec{w} \rangle_H \Longleftarrow$ <u>careful</u>!!

④ $\langle \vec{v}, \vec{v} \rangle_H \geq 0$, equality iff $\vec{v} = \vec{0}$

$$\langle \vec{v}, \vec{v} \rangle_H = \sum v_k \overline{v_k} = \sum (a_k + ib_k)(a_k - ib_k)$$

$$= \sum a_k^2 + b_k^2$$

<u>Obs:</u> With the "forgetful function" $\mathbb{C}^n \to \mathbb{R}^{2n}$, the "Hermitian norm" corresponds to the usual "length".

Comment: Any function satisfying ①-④ is called a "Hermitian inner product" on $\mathbb{C}^n$.

Recall that real dot products relate to transposes by:

$$\langle A\vec{v}, \vec{w} \rangle = (A\vec{v})^T \vec{w} = \vec{v}^T A^T \vec{w} = \langle \vec{v}, A^T \vec{w} \rangle$$

So

$$\langle A\vec{v}, \vec{w} \rangle = \langle \vec{v}, A^T \vec{w} \rangle$$

With the Hermitian inner product there is a similar connection to a similar construction.

Def:) The Hermitian transpose of A is

$$A^* = \overline{A}^T$$

Thm:) $\langle A\vec{v}, \vec{w} \rangle_H = \langle \vec{v}, A^* \vec{w} \rangle_H$

Terminology: "Hermitian transpose" = "Hermitian conjugate"
= "conjugate transpose"
= "adjoint"

Recall that transposes relate to symmetry by definition.

We make an analogous definition for Hermitian transpose;

Def! A is Hermitian if $A^* = A$.

(One might think of this as being "Hermitian symmetric".)

So, for symmetric matrices : $\langle A\vec{v}, \vec{w} \rangle = \langle \vec{v}, A\vec{w} \rangle$

and for Hermitian matrices : $\langle A\vec{v}, \vec{w} \rangle_H = \langle \vec{v}, A\vec{w} \rangle_H$

Symmetric Matrices

Note first that real symmetric matrices are Hermitian.

From this we will get some remarkable facts about real symmetric matrices.

Thm:) If A is real and symmetric, then all of its eigenvalues are real, and all of its eigenspaces have real bases.

Pf:) Let λ be an eigenvalue, and $\vec{v}$ a corresponding eigenvector. Consider the Hermitian inner product:

$$\langle A\vec{v}, \vec{v} \rangle_H$$

(Remember, $\vec{v}$ might be complex...)

Since A is real and symmetric, it is Hermitian.

So

$$\langle A\vec{v}, \vec{v} \rangle_H = \langle \vec{v}, A\vec{v} \rangle_H$$

$$\langle \lambda\vec{v}, \vec{v} \rangle_H = \langle \vec{v}, \lambda\vec{v} \rangle_H$$

$$\lambda \langle \vec{v}, \vec{v} \rangle_H = \bar{\lambda} \langle \vec{v}, \vec{v} \rangle_H$$

$$\lambda = \bar{\lambda}$$

So λ is real.

We find a basis for the eigenspace by finding a basis for $NS(\lambda I - A)$, which is real, so we get a real basis.

Thm) If A is <u>real</u> and <u>symmetric</u>, with eigenvalues $\lambda_1 \neq \lambda_2$ and associated eigenvectors $\vec{v}_1, \vec{v}_2$, then $\vec{v}_1, \vec{v}_2$ are <u>orthogonal</u>.

Pf) Again, A is Hermitian, so

$$\langle A\vec{v}_1, \vec{v}_2 \rangle_H = \langle \vec{v}_1, A\vec{v}_2 \rangle_H$$

$$\langle \lambda_1\vec{v}_1, \vec{v}_2 \rangle_H = \langle \vec{v}_1, \lambda_2\vec{v}_2 \rangle_H$$

$$\lambda_1 \langle \vec{v}_1, \vec{v}_2 \rangle_H = \lambda_2 \langle \vec{v}_1, \vec{v}_2 \rangle_H$$

(because we know λ_2 is real)

$$(\lambda_1 - \lambda_2) \langle \vec{v}_1, \vec{v}_2 \rangle_H = 0$$

$(\neq 0)$

$$\langle \vec{v}_1, \vec{v}_2 \rangle_H = 0$$

So $\vec{v}_1, \vec{v}_2$ are orthogonal.

Thm: If A is <u>real</u> and <u>symmetric</u>, then it has an <u>orthonormal eigenbasis</u>, and is <u>diagonalizable</u>.

Pf: Choose an eigenvalue λ_1, and a unit eigenvector $\vec{v}_1$.
Define W_1 as the subspace of $\mathbb{R}^n$ orthogonal to $\vec{v}_1$:

$$W_1 = \left\{ \vec{w} \;\middle|\; \langle \vec{w}, \vec{v}_1 \rangle = 0 \right\}$$

Note $\langle A\vec{w}, \vec{v}_1 \rangle = \langle \vec{w}, A\vec{v}_1 \rangle = \langle \vec{w}, \lambda_1 \vec{v}_1 \rangle$

$$= \lambda_1 \langle \vec{w}, \vec{v}_1 \rangle$$

$$= 0$$

So: $\left(\vec{w} \in W_1\right) \Longrightarrow \left(A\vec{w} \in W_1\right)$

Since A then is a l.t. on W_1, it must also have an eigenvector in W_1 (find a basis for W_1, write $A|_{W_1}$ w.r.t. that basis, and of course all matrices have at least one eigenvalue and eigenvector...).

Call that eigenvector λ_2, and choose a unit eigenvector $\vec{v}_2$. Note, since $\vec{v}_2 \in W_1$, we have $\vec{v}_1 \perp \vec{v}_2$.

Now define W_2 as the subspace orthogonal to both $\vec{v}_1, \vec{v}_2$:

$$W_2 = \left\{ \vec{w} \;\middle|\; \langle \vec{w}, \vec{v}_1 \rangle = 0,\ \langle \vec{w}, \vec{v}_2 \rangle = 0 \right\}$$

Again,
$$\left(\vec{w} \in W_2 \right) \Longrightarrow \left(A\vec{w} \in W_2 \right)$$

So, again we can view A as a l.t. on W_2, and again find an eigenvalue and unit eigenvector, with $\vec{v}_3 \perp \vec{v}_1, \vec{v}_2$.

Repeating, we eventually get $\{\vec{v}_1, \dots, \vec{v}_n\}$, all orthogonal and all unit vectors.

So, this is an orthonormal eigenbasis of $\mathbb{R}^n$, and so A is diagonalizable.

Def: If A, B are similar by an <u>orthogonal</u> matrix P,

$$B = P^{-1}AP = P^TAP$$

then we say they are <u>orthogonally similar</u>.

Thm: (Schur) If A has all real eigenvalues, then A is orthogonally similar to an upper triangular matrix.

There is another version of Schur's Thm involving complex matrices, for which we need another definition.

Recall: $\quad (P \text{ orthogonal}) \iff (P^TP = I)$

As we already have a complex notion analogous to transpose, we can define:

Def: P is <u>unitary</u> iff $\quad P^*P = I$

Note, this means that the columns of P are orthogonal.

And

$$\langle P\vec{v}, P\vec{w}\rangle_H = \langle \vec{v}, \vec{w}\rangle_H$$

so it preserves dot products

Def: If A, B are similar by a <u>unitary</u> matrix P,

$$B = P^{-1}AP = P^*AP$$

then we say that they are <u>unitarily similar</u>.

Thm: Every square matrix is unitarily similar to an upper triangular matrix.

6.1 – Systems of First Order Linear DE's

We will be interested in systems of DE's of the form:

$$y_1' = a_{11}(x)\, y_1 + \cdots + a_{1n}(x)\, y_n + g_1(x)$$

$$\vdots$$

$$y_n' = a_{n1}(x)\, y_1 + \cdots + a_{nn}(x)\, y_n + g_n(x)$$

If we write $\vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, $\vec{g} = \begin{pmatrix} g_1 \\ \vdots \\ g_n \end{pmatrix}$, $A = \begin{pmatrix} a_{11}(x) & \cdots & a_{1n}(x) \\ \vdots & & \vdots \\ a_{n1}(x) & \cdots & a_{nn}(x) \end{pmatrix}$

then this becomes

$$\vec{y}\,'(x) = A\,\vec{y}(x) + \vec{g}$$

Variations ① If $\vec{g}(x) = \vec{0}$, we call this a homogeneous system

② If it is required that $\vec{y}(x_0) = \vec{b}$, this is called an initial value problem.

There is an analogy between the theory of first order linear systems and linear DE's.

Thm: If $a_{ij}(x)$ and $g_i(x)$ are all continuous on (a,b) containing x_0, then the initial value problem

$$\vec{y}' = A\vec{y} + \vec{g}, \quad \vec{y}(x_0) = \vec{b}$$

has a unique solution on (a,b).

We will not prove this in this course.

Note that if $\vec{y}(x)$ is a solution to

$$\vec{y}' = A\vec{y}$$

then $\vec{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ where each $y_i \in C^1(a,b)$. This is a vector space:

$$(C^1)^n = \left\{ \vec{f} = \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} \,\middle|\, f_i \in C^1(a,b) \right\}$$

And the set of solutions to the homogeneous system is a subspace, because it is the kernel of the linear operator

$$L(\vec{f}) = \vec{f}' - A\vec{f}$$

(Check that this is linear.)

Thm: The solutions to $\vec{y}\,' = A\vec{y}$ form a subspace of dimension n.

The proof of this is analogous to the similar result about nth order linear DE's. Roughly, the result again comes from the fact that the space of initial conditions is again n-dimensional.

A basis for this subspace is called a fundamental set of solutions. Such a basis, $\{\vec{y}_1, \ldots, \vec{y}_n\}$ can be used to form columns of a matrix called a matrix of fundamental solutions.

$$M = \begin{pmatrix} | & & | \\ \vec{y}_1 & \cdots & \vec{y}_n \\ | & & | \end{pmatrix} = \begin{pmatrix} y_{11} & \cdots & y_{1n} \\ \vdots & & \vdots \\ y_{n1} & \cdots & y_{nn} \end{pmatrix}$$

(Do not confuse these indices, distinguishing solution vectors, with those used to distinguish the variables making up those vectors!)

Ch. 4	Ch. 6.
nth order, 1 eq	1st order, n eqs.
Ex/uni. I.V.P.	Ex/uni I.V.P.

Ch. 4

$$\begin{tikzcd} \underbrace{(\text{I.C.})}_{n\ \text{dim}} \arrow[r, leftrightarrow] & (\text{sol.}) \arrow[d] \\ \text{basis } \underbrace{\{\vec{e}_1, \dots, \vec{e}_n\}} \arrow[u] & \text{basis? } \{y_1, \dots, y_n\} \end{tikzcd}$$

Show:
① l.i.
② span

Ch. 6

$$\begin{tikzcd} \underbrace{(\text{I.C.})}_{n\ \text{dim}} \arrow[r, leftrightarrow] & (\text{sol.}) \arrow[d] \\ \text{basis } \{\vec{e}_1, \dots, \vec{e}_n\} \arrow[u] & \text{basis? } \{\vec{y}_1, \dots, \vec{y}_n\} \end{tikzcd}$$

Show:
① l.i.
② span.

Ex:) Consider the system

$$\vec{y}' = \begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix} \vec{y}$$

This has a <u>2-dimensional</u> set of solutions.

Note the solutions below are independent:

$$\vec{y}_1 = \begin{pmatrix} -e^{4x} \\ e^{4x} \end{pmatrix}, \quad \vec{y}_2 = \begin{pmatrix} 3e^{-x} \\ 2e^{-x} \end{pmatrix}$$

So, the complete set of solutions is

$$\vec{y} = c_1 \begin{pmatrix} -e^{4x} \\ e^{4x} \end{pmatrix} + c_2 \begin{pmatrix} 3e^{-x} \\ 2e^{-x} \end{pmatrix}$$

and a fundamental set of solutions is

$$\left\{ \begin{pmatrix} -e^{4x} \\ e^{4x} \end{pmatrix}, \begin{pmatrix} 3e^{-x} \\ 2e^{-x} \end{pmatrix} \right\}$$

and the matrix of fundamental solutions is

$$M = \begin{pmatrix} -e^{4x} & 3e^{-x} \\ e^{4x} & 2e^{-x} \end{pmatrix}$$

Note the complete set of solutions is

$$\vec{y} = M\vec{c}, \quad \vec{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

As with linear DE's, non-homogeneous solutions relate to homogeneous solutions.

Thm: Say $\{\vec{y}_1, \dots, \vec{y}_n\}$ is a fund. set of sols. for

$$\vec{y}\,' = A\vec{y}$$

and suppose also that $\vec{y}_p$ is a solution to

$$\vec{y}\,' = A\vec{y} + \vec{g}$$

Then the complete set of solutions to the non-homogeneous system is

$$\vec{y} = c_1\vec{y}_1 + \dots + c_n\vec{y}_n + \vec{y}_p$$

(Proof is analogous to similar previous thms.)

Recall that for linear DE's, we defined the Wronskian of solutions $y_1, \ldots, y_n$ to be

$$w(x) = \det \begin{pmatrix} y_1(x) & & y_n(x) \\ \vdots & \cdots & \vdots \\ y^{[n-1]}(x) & & y_n^{[n-1]}(x) \end{pmatrix}$$

We showed that for solutions, $W(x)$ was either identically zero $(\Rightarrow$ l.d.$)$ or never zero $(\Rightarrow$ l.i.$)$.

Each column in the above matrix represent n pieces of information about the corresponding solution, relating to the given DE, which we know has an n-dimensional solution set.

We can do something similar for first order linear systems; in fact, since each solution has n components, we don't even need to take any derivatives to get the needed number of bits of information...

Consider the system

$$\vec{y}' = A\vec{y}$$

Given solutions $\vec{y}_1, \dots, \vec{y}_n$, we define

$$W(x) = \det \begin{pmatrix} | & & | \\ \vec{y}_1 & \cdots & \vec{y}_n \\ | & & | \end{pmatrix}$$

Thm: If $\vec{y}_1, \dots, \vec{y}_n$ are solutions to $\vec{y}' = A\vec{y}$, then either

(1) $W(x)$ is <u>identically zero</u> and $\{\vec{y}_1, \dots, \vec{y}_n\}$ <u>l.d.</u>

<u>or</u>

(2) $W(x)$ is <u>never zero</u> and $\{\vec{y}_1, \dots, \vec{y}_n\}$ l.i.

Ex: Consider the solutions found previously to

$$\vec{y}' = A\vec{y}, \quad A = \begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix}$$

The Wronskian is

$$W(x) = \det \begin{pmatrix} -e^{4x} & 3e^{-x} \\ e^{4x} & 2e^{-x} \end{pmatrix} = -5e^{3x}$$

This is <u>never zero</u>, so these solutions are l.i..

6.2 – Constant Coefficient Homogeneous Systems; Diagonalizable

In this section and the next, we will consider how to solve systems

$$\vec{y}\,' = A\vec{y}$$

where A is a <u>constant</u> matrix.

Note the similarity to the single variable equation

$$y' = ay$$

for which we know the solutions are

$$y = e^{ax} y_0$$

Naively, we would be tempted to try to write down solutions to our system then as

$$\vec{y} = e^{xA}\vec{y}_0$$

But — can we make sense out of a matrix exponent ??
And, if so, would this actually give solutions to the DE ??

We will see that the answer to each of these questions is

$$\boxed{\text{YES!}}$$

Recalling our Taylor series

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots$$

and that we used this to motivate our definition

$$e^{i\theta} = 1 + i\theta + \frac{(i\theta)^2}{2!} + \frac{(i\theta)^3}{3!} + \cdots$$

we are motivated to again use this to define:

Def:) $$e^M = I + M + \frac{M^2}{2!} + \frac{M^3}{3!} + \cdots$$

The following properties can be checked:

(1) $e^0 = I$ (where 0 is the zero matrix)

(2) $e^{rI} = e^r I$

(3) if $AB = BA$, then $e^{A+B} = e^A e^B$

(4) Considering e^{xA} as a function of x, we get

$$\left(e^{xA}\right)' = A e^{xA}$$

Check:
$$\begin{aligned}\left(e^{xA}\right)' &= \left(I + (xA) + \frac{(xA)^2}{2!} + \frac{(xA)^3}{3!} + \cdots\right)' \\ &= \left(A + xA^2 + \frac{x^2}{2!}A^3 + \cdots\right) \\ &= A\left(I + (xA) + \frac{(xA)^2}{2!} + \cdots\right) \\ &= A e^{xA}\end{aligned}$$

So indeed then, by (4), $\vec{y} = e^{xA}\vec{y}_0$ is a solution to

$$\vec{y}\,' = A\vec{y}$$

Specifically, it is a solution to the I.V.P.

$$\vec{y}\,' = A\vec{y} \quad , \quad \vec{y}(0) = y_0$$

And since we know I.V.P.'s have unique solutions, that means that <u>every</u> solution is of this form!

<u>Thm:</u> If $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is a basis for $\mathbb{R}^n$,

then $\left\{e^{xA}\vec{v}_1, \ldots, e^{xA}\vec{v}_n\right\}$ is a fundamental set of solutions to $\vec{y}\,' = A\vec{y}$.

Problem — How do we actually <u>compute</u> these in terms of functions we know how to evaluate?

(Note, at $x=0$, the Wronskian of the functions in this set is nonzero, because $\{\vec{v}_1, \ldots, \vec{v}_n\}$ is independent)

The following formula is very helpful in evaluating these functions.

$$e^{xA}\vec{v} = e^{x(\lambda I + A - \lambda I)}\vec{v}$$
$$= e^{x\lambda I}\, e^{x(A-\lambda I)}\vec{v}$$
$$= e^{\lambda x}\, e^{x(A-\lambda I)}\vec{v}$$

So:

$$\boxed{e^{xA}\vec{v} = e^{\lambda x}\left(I + x(A-\lambda I) + \frac{x^2}{2!}(A-\lambda I)^2 + \cdots\right)\vec{v}}$$

We have the following special case:

Thm: If λ is an eigenvalue and $\vec{v}$ a corresponding eigenvector of A, then

$$e^{xA}\vec{v} = e^{\lambda x}\vec{v}$$

Note, the RHS is easy to evaluate!

Ex:) Solve the system $\vec{y}' = D\vec{y}$, where

$$D = \begin{pmatrix} d_1 & & 0 \\ & \ddots & \\ 0 & & d_n \end{pmatrix}$$

Note that D has eigenvectors $\vec{e}_1, \dots, \vec{e}_n$, with corresponding eigenvalues $d_1, \dots, d_n$.

Using the eigenbasis $\mathcal{B} = \{\vec{e}_1, \dots, \vec{e}_n\}$, we get a fundamental set

$$\left\{ e^{d_1 x}\vec{e}_1 , \dots , e^{d_n x}\vec{e}_n \right\}$$

So the general solution is

$$\vec{y} = c_1 e^{d_1 x}\vec{e}_1 + \dots + c_n e^{d_n x}\vec{e}_n$$

$$= \begin{pmatrix} c_1 e^{d_1 x} \\ \vdots \\ c_n e^{d_n x} \end{pmatrix}$$

Of course the system in the previous example can be written out as

$$y_1' = d_1 y_1$$
$$\vdots$$
$$y_n' = d_n y_n$$

So, the general solution can also be derived by observing that these individual equations are independent of each other, and are natural growth equations.

Ex i) Solve the system $\vec{y}' = A\vec{y}$, where

$$A = \begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix}$$

Recall that the eigenvectors and eigenvalues are

$$\vec{v}_1 = \begin{pmatrix} 3 \\ 2 \end{pmatrix} \qquad \lambda_1 = -1$$

$$\vec{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \qquad \lambda_2 = 4$$

Using the eigenbasis $\mathcal{V} = \{\vec{v}_1, \vec{v}_2\}$, we get a fundamental set

$$\left\{ e^{-x}\begin{pmatrix} 3 \\ 2 \end{pmatrix}, \; e^{4x}\begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}$$

or

$$\left\{ \begin{pmatrix} 3e^{-x} \\ 2e^{-x} \end{pmatrix}, \; \begin{pmatrix} e^{4x} \\ -e^{4x} \end{pmatrix} \right\}$$

So the general solution is

$$\vec{y} = c_1 \begin{pmatrix} 3e^{-x} \\ 2e^{-x} \end{pmatrix} + c_2 \begin{pmatrix} e^{4x} \\ -e^{4x} \end{pmatrix}$$

$$= \begin{pmatrix} 3c_1e^{-x} + c_2e^{4x} \\ 2c_1e^{-x} - c_2e^{4x} \end{pmatrix}$$

Note, we can take a different motivation for these solutions:

Consider $$\vec{y}\,' = A\vec{y}$$

Exponentials and eigenvectors have something in common — they both give a <u>factor</u> after a particular operation.

$$e^{kx} \xrightarrow{\text{diff}} \text{factor of } k$$

$$\vec{v} \xrightarrow{A} \text{factor of } \lambda$$

Since one side of the equation above has a derivative and the other has a matrix mult., we can try to match these factors by looking at

$$\vec{y} = e^{\lambda x}\vec{v}$$

so the equation becomes

$$\left(e^{\lambda x}\vec{v}\right)' = A\left(e^{\lambda x}\vec{v}\right)$$

$$\left(e^{\lambda x}\right)'\vec{v} = e^{\lambda x}\left(A\vec{v}\right)$$

$$\lambda e^{\lambda x}\vec{v} = e^{\lambda x}\lambda\vec{v} \quad \checkmark$$

With complex eigenvalues and eigenvectors, we can apply the same procedure, and then look at Re and Im components.

Ex) Solve the system $\vec{y}' = A\vec{y}$, where

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

The eigenvectors and eigenvalues are

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \qquad \lambda_1 = 1$$

$$\vec{v}_2 = \begin{pmatrix} 1-i \\ i \\ 1 \end{pmatrix} \qquad \lambda_2 = 1+i$$

$$\vec{v}_3 = \begin{pmatrix} 1+i \\ -i \\ 1 \end{pmatrix} \qquad \lambda_3 = 1-i$$

The first gives us the solution $\vec{y}_1 = e^{1x}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} e^x \\ 0 \\ 0 \end{pmatrix}$

The second gives us $\vec{y} = e^{(1+i)x}\begin{pmatrix} 1-i \\ i \\ 1 \end{pmatrix}$

$$= e^x\left(\cos x + i\sin x\right)\begin{pmatrix} 1-i \\ i \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} (e^x\cos x + e^x\sin x) + i(e^x\sin x - e^x\cos x) \\ (-e^x\sin x) \quad + i\,(e^x\cos x) \\ (e^x\cos x) \quad + i\,(e^x\sin x) \end{pmatrix}$$

which has components

$$\vec{y}_2 = Re\left(\vec{y}\right) = \begin{pmatrix} e^x \cos x + e^x \sin x \\ -e^x \sin x \\ e^x \cos x \end{pmatrix}$$

$$\vec{y}_3 = Im\left(\vec{y}\right) = \begin{pmatrix} e^x \sin x - e^x \cos x \\ e^x \cos x \\ e^x \sin x \end{pmatrix}$$

This gives us a fundamental set $\left\{ \vec{y}_1, \vec{y}_2, \vec{y}_3 \right\}$

Note — we did not need to look at the third eigenvector and eigenvalue! These are conjugate to the second, and yield the same (up to sign) Re and Im components.

6.3 — Constant Coefficient Homogeneous Systems; Non-diagonalizable

Recall: (1) for any basis $\mathcal{V} = \{\vec{v}_1, \ldots, \vec{v}_n\}$,

$$\{e^{xA}\vec{v}_1, \ldots, e^{xA}\vec{v}_n\}$$

is a fund. sol. set for $\vec{y}' = A\vec{y}$.

(2) $e^{xA}\vec{v}$ is easy to evaluate when $\vec{v}$ is an eigenvector.

So, we have a good method for solving such systems when we have a basis of eigenvectors. That is, when A is diagonalizable.

What if A is not diagonalizable?

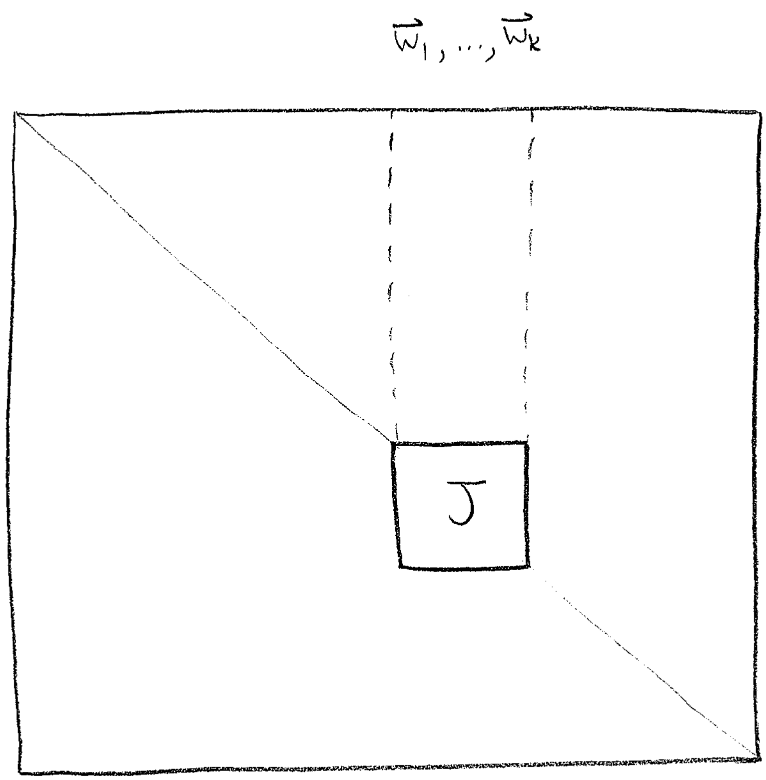
Recall that non-diagonalizable matrices can still be put into Jordan form. That is — there is a Jordan basis $\mathcal{V}$ such that $A = [T]_{\mathcal{S}}^{\mathcal{S}}$, changed to the $\mathcal{V}$ basis as $F = [T]_{\mathcal{V}}^{\mathcal{V}}$, is in Jordan form.

Above, we used the fact that $e^{xA}\vec{v}$ is easy to evaluate when $\vec{v}$ is in an eigenbasis.

We will see here that $e^{xA}\vec{v}$ is also not hard to evaluate when $\vec{v}$ is in a Jordan basis.

Consider a particular <u>basic</u> Jordan block J in the Jordan form, and the Jordan basis vectors $\vec{w}_1, \dots, \vec{w}_k$ that correspond to that block:

$[T]_{\mathcal{V}}^{\mathcal{V}} =$



We know that J has the form

$$J = \begin{bmatrix} \lambda & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{bmatrix}$$

Looking at the columns other than the first one, we have

$$A\vec{w}_i = \lambda \vec{w}_i + 1 w_{i-1}$$

which we can rewrite as

$$(A - \lambda I)\vec{w}_i = \vec{w}_{i-1}$$

And of course the first column corresponds to an eigenvector $\vec{w}_1$,

so

$$(A - \lambda I)\vec{w}_1 = \vec{0}$$

So — the elements of the Jordan basis (corresponding to this particular Jordan block) relate by the matrix $(A - \lambda I)$:

$$\vec{w}_k \xrightarrow{(A-\lambda I)} \vec{w}_{k-1} \xrightarrow{(A-\lambda I)} \cdots \xrightarrow{(A-\lambda I)} \vec{w}_2 \xrightarrow{(A-\lambda I)} \vec{w}_1 \xrightarrow{(A-\lambda I)} \vec{0}$$

This allows us to compute $e^{xA}\vec{w}_i$ by using the same formula we used for eigenvectors.

Recall:

$$e^{xA}\vec{v} = e^{\lambda x}\left(I + x(A-\lambda I) + \frac{x^2}{2}(A-\lambda I)^2 + \cdots\right)\vec{v}$$

So, using the above results, we have

$$e^{xA}\vec{w}_i = e^{\lambda x}\left(\vec{w}_i + x\vec{w}_{i-1} + \frac{x^2}{2}\vec{w}_{i-2} + \cdots + \frac{x^{(i-1)}}{(i-1)!}\vec{w}_1\right)$$

Ex:) Solve the system $\vec{y}' = A\vec{y}$, with

$$A = \begin{pmatrix} -1 & 3 & 5 \\ -4 & 6 & 4 \\ -3 & 2 & 7 \end{pmatrix}$$

This matrix is <u>not</u> diagonalizable — its Jordan form is

$$F = \begin{pmatrix} 4 & 1 & 0 \\ 0 & 4 & 1 \\ 0 & 0 & 4 \end{pmatrix}$$

using the Jordan basis with vectors

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \vec{v}_2 = \begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix} \quad \vec{v}_3 = \begin{pmatrix} 6 \\ 3 \\ 5 \end{pmatrix}$$

Choosing this Jordan basis, we get solutions

$$e^{xA}\vec{v}_1 = e^{4x}\left(\vec{v}_1\right)$$

$$e^{xA}\vec{v}_2 = e^{4x}\left(\vec{v}_2 + x\vec{v}_1\right)$$

$$e^{xA}\vec{v}_3 = e^{4x}\left(\vec{v}_3 + x\vec{v}_2 + \frac{x^2}{2}\vec{v}_1\right)$$

which we can write out explicitly as

$$e^{xA}\vec{v}_1 = e^{4x}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \vec{y}_1$$

$$e^{xA}\vec{v}_2 = e^{4x}\begin{pmatrix} 4+x \\ 2 \\ 3+x \end{pmatrix} = \vec{y}_2$$

$$e^{xA}\vec{v}_3 = e^{4x}\begin{pmatrix} 6+4x+x^2/2 \\ 3+2x \\ 5+3x+x^2/2 \end{pmatrix} = \vec{y}_3$$

These functions form a fundamental solution set $\{\vec{y}_1, \vec{y}_2, \vec{y}_3\}$

Note, each basic Jordan block is dealt with separately.

Ex:) Suppose our Jordan form is

$$F = \begin{bmatrix} \boxed{\begin{matrix} 3 & 1 \\ & 3 \end{matrix}} & & \\ & \boxed{\begin{matrix} 4 & 1 \\ & 4 \end{matrix}} & \\ & & \boxed{4} \end{bmatrix}$$

with Jordan basis $\{\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4, \vec{v}_5\}$.

Then

(1) $\vec{v}_1$ is an eigenvector with eigenvalue 3, and

$$\vec{v}_2 \xrightarrow{(A-3I)} \vec{v}_1 \xrightarrow{(A-3I)} \vec{0}$$

(2) $\vec{v}_3$ is an eigenvector with eigenvalue 4, and

$$\vec{v}_4 \xrightarrow{(A-4I)} \vec{v}_3 \xrightarrow{(A-4I)} \vec{0}$$

(3) $\vec{v}_5$ is an eigenvector with eigenvalue 4.

$$\vec{v}_5 \xrightarrow{(A-4I)} \vec{0}$$

This use of the Jordan form for solving systems of DE's motivates a strategy for finding a Jordan basis, and the Jordan form, in some cases

1. The characteristic polynomial tells you the eigenvalues, and the dimensions of the eigenvalue blocks.

2. The number of eigenvectors for an eigenvalue tells you how many basic Jordan blocks in that eigenvalue block.

3. For each eigenvector, you can try to solve

$$(A-\lambda I)\vec{w}_i = \vec{w}_{i-1}$$

to get the other Jordan basis vectors for that basic Jordan block.

<u>Disclaimer</u>: This method will not always work!

(But if there is only 1 eigenvector per eigenvalue, it will.)

Ex: Find a Jordan basis and the Jordan form for

$$A = \begin{pmatrix} -1 & 3 & 5 \\ -4 & 6 & 4 \\ -3 & 2 & 7 \end{pmatrix}$$

① The char. poly is

$$p(\lambda) = (\lambda - 4)^3$$

So, we know there is a single eigenvalue block, with $\lambda = 4$, and it is 3×3.

<u>Shortcut</u>: We will need to solve several equations involving $M = A - 4I$ ....

Row reducing $(M \mid I)$ to $(R \mid E)$ (rref) (prod. of elems)

we have

$$R = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \text{ and } E = \begin{pmatrix} 0 & -1 & 1 \\ 1 & -2 & 1 \\ -1/5 & 1/10 & 1/5 \end{pmatrix}$$

Given this, we can solve

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

instead with

$$R\vec{x} = E\vec{b}$$

(2) To find the eigenvectors, we solve

$$M\vec{v} = \vec{0}$$

or

$$R\vec{v} = E\vec{0}$$

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{v} = \vec{0}$$

Arbitrarily choosing 1 for the free (third) variable, we get the eigenvector

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

Note: there is only 1 free variable, so there is only 1 eigenvector, and thus only 1 basic Jordan block in this eigenvalue block. (So there are 3 basis vectors for this basic Jordan block.)

(3) To find the other Jordan vectors for this block,

we use $$M\vec{v}_i = \vec{v}_{i-1}$$

First then, we solve

$$M\vec{v}_2 = \vec{v}_1$$

or

$$R\vec{v}_2 = E\vec{v}_1$$

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{V}_2 = \begin{pmatrix} 0 & -1 & 1 \\ 1 & -2 & 1 \\ -1/5 & 1/10 & 1/5 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{V}_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$$

Arbitrarily choosing 3 for the free (third) variable, we get

$$\vec{V}_2 = \begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix}$$

Next, we solve

$$M \vec{V}_3 = \vec{V}_2$$

or

$$R \vec{V}_3 = E \vec{V}_2$$

$$\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \vec{V}_3 = \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}$$

Arbitrarily choosing 5 for the free (third) variable, we get

$$\vec{V}_3 = \begin{pmatrix} 6 \\ 3 \\ 5 \end{pmatrix}$$

So we get Jordan basis $\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 4 \\ 2 \\ 3 \end{pmatrix}, \begin{pmatrix} 6 \\ 3 \\ 5 \end{pmatrix} \right\}$

which we had assumed in a previous example.

The Jordan form then comes from a change of basis.

6.4 – Nonhomogeneous Linear Systems

Recall that the general solution to

$$\vec{y}\,' = A\vec{y} + \vec{g}$$

is obtained by finding <u>any</u> particular solution, and then adding the complete homogeneous solutions.

So, how do we find the particular solution?

Let's consider first the single variable case.

$$y' = a(x)\,y + g(x)$$

The homogeneous equation and solution are

$$y' = ay \quad \Longrightarrow \quad y = k\,e^{\int a\,dx}$$

So the fundamental solution is $m(x) = e^{\int a(x)\,dx}$

One might speculate that the particular solution could look similar. We could try:

$$y' = ay + g \quad \overset{?}{\leftarrow} \quad y_p = m(x)\,v(x)$$

and then try to solve for this factor $v(x)$ . . .

Plugging in to the equation and doing some algebra, we get

$$(mv)' = amv + g$$

$$m'v + mv' = amv + g$$

(These cancel because $m' = am$)

$$mv' = g$$

$$v = \int m^{-1} g \, dx$$

$$y_p = mv = m \int m^{-1} g \, dx$$

One can then plug this in to the equation to confirm that it actually is a solution.

(See 4.4, variation of parameters)

(Note this gives us the same solution as the method of integrating factors in this case.)

We can do the same in the case of systems.

For the homogeneous system

$$\vec{y}' = A\vec{y}$$

we know how to find a fundamental set of solutions.

Let M be the solution matrix

$$M = \begin{pmatrix} | & & | \\ \vec{y}_1 & \cdots & \vec{y}_n \\ | & & | \end{pmatrix}$$

We might speculate that the particular solution again will look similar to the $\vec{y}$'s ...

$$\vec{y}' = A\vec{y} + \vec{g} \quad \overset{?}{\leftarrow} \quad \vec{y}_p = M\vec{v}$$

Again, we can plug in and solve.

$$(M\vec{v})' = AM\vec{v} + \vec{g}$$

$$M'\vec{v} + M\vec{v}' = AM\vec{v} + \vec{g}$$

(These cancel because $M' = AM$, because every column of M is a solution to $\vec{y}' = A\vec{y}$.)

$$M\vec{v}' = g$$

$$\vec{v} = \int M^{-1}\vec{g}\,dx$$

$$\vec{y}_p = M\vec{v} = \boxed{M\int M^{-1}\vec{g}\,dx}$$

Ex i) Find a particular solution to

$$\vec{y}\,' = \begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix} \vec{y} + \begin{pmatrix} x \\ 3x \end{pmatrix}$$

Recall that the matrix of solutions is

$$M = \begin{pmatrix} -e^{4x} & 3e^{-x} \\ e^{4x} & 2e^{-x} \end{pmatrix}$$

Its inverse is

$$M^{-1} = \begin{pmatrix} 2e^{-x} & -3e^{-x} \\ -e^{4x} & -e^{4x} \end{pmatrix} \Big/ (-5e^{3x})$$

$$= \begin{pmatrix} -\frac{2}{5}e^{-4x} & \frac{3}{5}e^{-4x} \\ \frac{1}{5}e^{x} & \frac{1}{5}e^{x} \end{pmatrix}$$

Then $$M^{-1}g = \begin{pmatrix} \frac{7}{5}xe^{-4x} \\ \frac{4}{5}x\,e^{x} \end{pmatrix}$$

$$\int M^{-1}g\,dx = \begin{pmatrix} -\frac{7}{20}xe^{-4x} - \frac{7}{80}e^{-4x} \\ \frac{4}{5}x\,e^{x} - \frac{4}{5}e^{x} \end{pmatrix}$$

$$\vec{y}_p = M\int M^{-1}g\,dx = \begin{pmatrix} \left(\frac{7}{20}x + \frac{7}{80}\right) + \left(\frac{12}{5}x - \frac{12}{5}\right) \\ \left(\frac{-7}{20}x - \frac{7}{80}\right) + \left(\frac{8}{5}x - \frac{8}{5}\right) \end{pmatrix} = \begin{pmatrix} \frac{11}{4}x - \frac{37}{16} \\ \frac{5}{4}x - \frac{27}{16} \end{pmatrix}$$

Ex i) Find a particular solution to

$$\vec{y}\,' = \begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix} \vec{y} + \begin{pmatrix} \cos x \\ 0 \end{pmatrix}$$

We can use complex solutions to solve this real problem. We consider

$$\vec{z}\,' = \begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix} \vec{z} + \begin{pmatrix} e^{ix} \\ 0 \end{pmatrix}$$

M, M^{-1} are as in the previous example.

$$M^{-1} \begin{pmatrix} e^{ix} \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{2}{5} e^{(-4+i)x} \\ \frac{1}{5} e^{(1+i)x} \end{pmatrix}$$

$$\int M^{-1} \begin{pmatrix} e^{ix} \\ 0 \end{pmatrix} dx = \begin{pmatrix} -\frac{2}{5} \frac{1}{-4+i} e^{(-4+i)x} \\ \frac{1}{5} \frac{1}{1+i} e^{(1+i)x} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{8+2i}{85} e^{(-4+i)x} \\ \frac{1-i}{10} e^{(1+i)x} \end{pmatrix}$$

$$\vec{z}_p = M \int M^{-1} \begin{pmatrix} e^{ix} \\ 0 \end{pmatrix} dx = \begin{pmatrix} -\frac{8+2i}{85} e^{ix} + \frac{3-3i}{10} e^{ix} \\ \frac{8+2i}{85} e^{ix} + \frac{2-2i}{10} e^{ix} \end{pmatrix}$$

$$= \begin{pmatrix} \left(\frac{7}{34} - \frac{11}{34} i\right)\left(\cos x + i \sin x\right) \\ \left(\frac{5}{17} - \frac{3}{17} i\right)\left(\cos x + i \sin x\right) \end{pmatrix}$$

Since $\begin{pmatrix} \cos x \\ 0 \end{pmatrix} = \mathrm{Re}\begin{pmatrix} e^{ix} \\ 0 \end{pmatrix}$ we have $\vec{y}_p = \mathrm{Re}(\vec{z}_p)$:

$$\vec{y}_p = \begin{pmatrix} \frac{7}{34}\cos x + \frac{11}{34}\sin x \\ \frac{5}{17}\cos x + \frac{3}{17}\sin x \end{pmatrix}$$

The first example makes clear an awkward point about this technique ... that is, we used complicated algebra in a problem where both the question and the answer are fairly simple.

Instead then, we can try to <u>guess</u> the form of the solution, and use algebra to work out the details. (This is the idea behind "undetermined coefficients" in 4.3.)

Ex:) Find a particular solution to

$$\vec{y}' = \begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix} \vec{y} + \begin{pmatrix} x \\ 3x \end{pmatrix}$$

We guess $\vec{y}_p = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} x = \vec{a} + \vec{b}x$

Noting $\vec{y}_p' = \vec{b}$, the equation becomes

$$\vec{b} = A(\vec{a} + \vec{b}x) + \begin{pmatrix} 1 \\ 3 \end{pmatrix} x$$

Grouping like powers of x, we get

$$A\vec{b} = \begin{pmatrix} -1 \\ -3 \end{pmatrix} \quad \text{and} \quad A\vec{a} = \vec{b}$$

Since $A^{-1} = \begin{pmatrix} -1/2 & -3/4 \\ -1/2 & -1/4 \end{pmatrix}$ we have

$$\vec{b} = A^{-1} \begin{pmatrix} -1 \\ -3 \end{pmatrix} = \begin{pmatrix} 11/4 \\ 5/4 \end{pmatrix}$$

$$\vec{a} = A^{-1} \begin{pmatrix} 11/4 \\ 5/4 \end{pmatrix} = \begin{pmatrix} -37/16 \\ -27/16 \end{pmatrix}$$

So we get the same particular solution as previous:

$$\vec{y}_p = \begin{pmatrix} -37/16 + 11/4\, x \\ -27/16 + 5/4\, x \end{pmatrix}$$

Ex:) Find a particular solution to

$$\vec{y}\,' = \begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix} \vec{y} + e^{rx} \vec{v}$$

Try $\quad \vec{y}_p = e^{rx} \vec{a}$

$$r \cancel{e^{rx}} \vec{a} = A \cancel{e^{rx}} \vec{a} + \cancel{e^{rx}} \vec{v}$$

$$(A - rI) \vec{a} = -\vec{v}$$

You can solve for $\vec{a}$ and get a solution, ...
<u>unless</u> $\det(A - rI) = 0$ ... that is, unless
r is an eigenvalue of A.

Ex:) Find a particular solution to

$$\vec{y}\,' = \begin{pmatrix} 1 & -3 \\ -2 & 2 \end{pmatrix} \vec{y} + e^{4x} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

Note, 4 <u>is</u> an eigenvalue for A.

Try $\quad \vec{y}_p = e^{4x} (\vec{a} + \vec{b} x)$

The equation becomes

$$4\cancel{e^{4x}}(\vec{a}+\vec{b}x)+\cancel{e^{4x}}(\vec{b}) = \cancel{e^{4x}}\,A(\vec{a}+\vec{b}x)+\cancel{e^{4x}}\vec{v}$$

$$x\left(4\vec{b}-A\vec{b}\right)+\left(4\vec{a}+\vec{b}-A\vec{a}-\vec{v}\right)=0$$

$$(A-4I)\vec{b}=\vec{0} \quad \underline{\underline{\text{and}}} \quad (A-4I)\vec{a}=\vec{b}-\vec{v}$$

Since 4 is an eigenvalue with eigenvector $\begin{pmatrix}-1\\1\end{pmatrix}$, we conclude from the first equation that

$$\vec{b}=k\begin{pmatrix}-1\\1\end{pmatrix}$$

The second equation then becomes

$$(A-4I)\vec{a}=k\begin{pmatrix}-1\\1\end{pmatrix}-\vec{v}$$

$$\begin{pmatrix}-3 & -3\\-2 & -2\end{pmatrix}\vec{a}=\begin{pmatrix}-k-1\\k-2\end{pmatrix}$$

To have a solution we must have

$$\frac{-k-1}{k-2}=\frac{3}{2} \Rightarrow -2k-2=3k-6$$

$$\Rightarrow 4=5k$$

$$\Rightarrow k=\frac{4}{5}$$

The second equation then becomes

$$\begin{pmatrix} -3 & -3 \\ -2 & -2 \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = \begin{pmatrix} -9/5 \\ -6/5 \end{pmatrix}$$

So $\quad a_1 + a_2 = 3/5$

and we can choose

$$\vec{a} = \begin{pmatrix} 3/5 \\ 0 \end{pmatrix}$$

Our solution then is

$$\vec{y}_p = e^{4x}\left(\vec{a} + \vec{b}x\right)$$

$$= e^{4x}\begin{pmatrix} 3/5 - 4/5\,x \\ 0 + 4/5\,x \end{pmatrix}$$

6.5 – Converting Higher Order Equations and Systems

Consider the equation

$$a_2 y'' + a_1 y' + a_0 y = g$$

Note:

$$(y)' = y'$$
$$(y')' = y''$$

If we view y, y' as being separate functions, we can rewrite the equation as a system.

Let
$$y = u_0(x)$$
$$y' = u_1(x)$$

Then
$$u_0' = y' = u_1$$
$$u_1' = y'' = -\frac{a_1}{a_2} y' - \frac{a_0}{a_2} y + \frac{g}{a_2}$$
$$= -\frac{a_1}{a_2} u_1 - \frac{a_0}{a_2} u_0 + \frac{g}{a_2}$$

So we have the system
$$u_0' = 0\, u_0 + 1\, u_1$$
$$u_1' = \left(\frac{-a_0}{a_2}\right) u_0 + \left(\frac{-a_1}{a_2}\right) u_1 + \left(\frac{g}{a_2}\right)$$

This will work for any linear equation.

Thm: If $a_n y^{[n]} + \cdots + a_0 y = 0$
is a constant coefficient homogeneous linear DE, then the char. poly. of this equation is equal to the char. poly. of the equivalent system.

Ex: Consider $y'' + 5y' + 6y = 0$

The char. poly. is $p_1(\lambda) = \lambda^2 + 5\lambda + 6$

The equivalent system is

$$u_0' = 0u_0 + 1u_1$$
$$u_1' = -6u_0 - 5u_1$$

which can be rewritten as

$$\vec{u}' = \begin{pmatrix} 0 & 1 \\ -6 & -5 \end{pmatrix} \vec{u}$$

and the char. poly is

$$p_2(\lambda) = \det \begin{pmatrix} -\lambda & 1 \\ -6 & -5-\lambda \end{pmatrix} = \lambda(\lambda+5) + 6$$
$$= \lambda^2 + 5\lambda + 6$$

You can use this same idea to convert higher order systems to first order systems.

Ex:) Consider the system

$$y_1'' = 3y_1 + 2y_1' - y_2 + y_2'$$

$$y_2'' = 2y_1 - 4y_1' + 2y_2 - y_2'$$

We can choose

$$\begin{aligned} y_1 &= u_1 \\ y_1' &= u_2 \\ y_2 &= u_3 \\ y_2' &= u_4 \end{aligned}$$

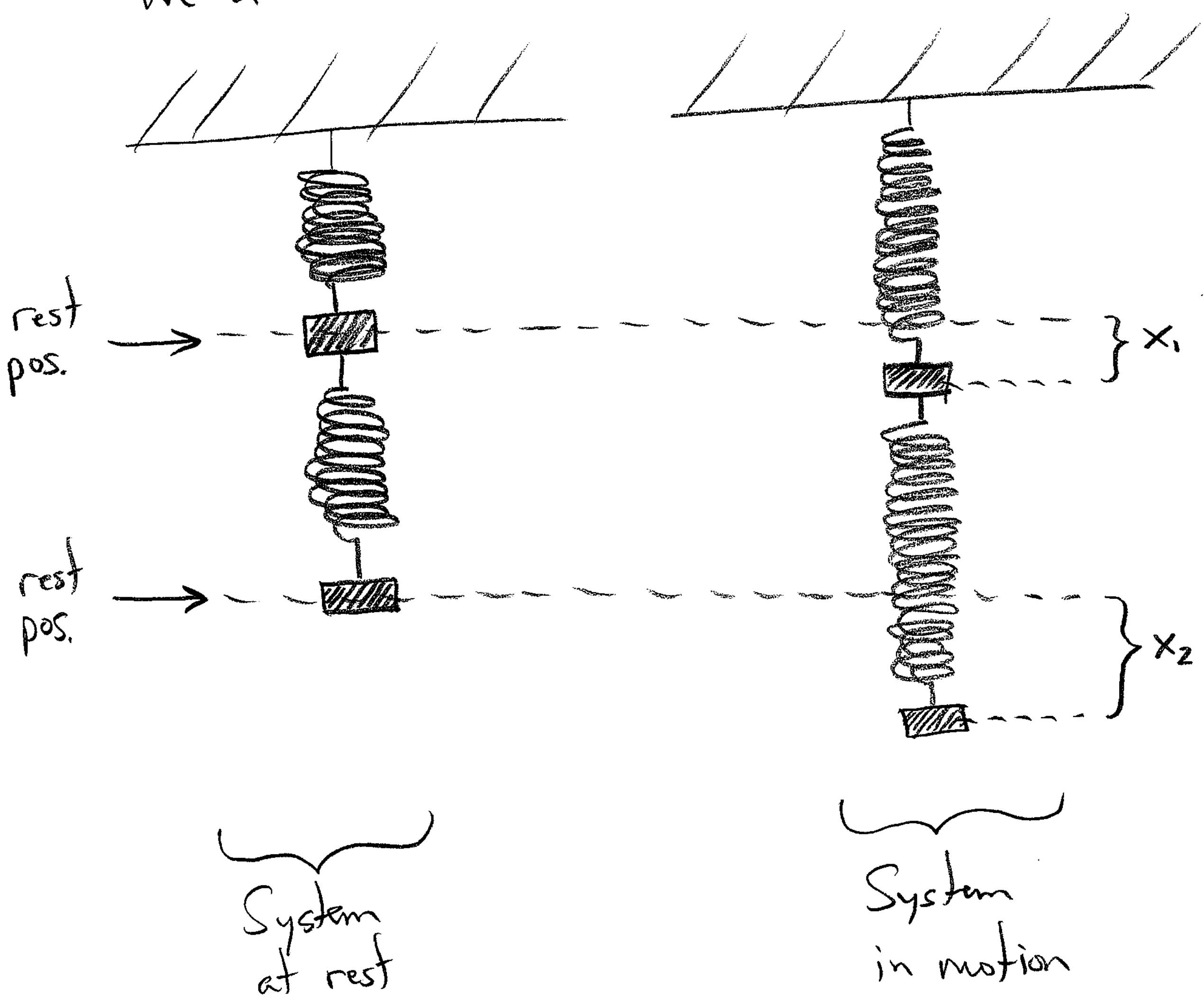
The system then becomes

$$\begin{aligned} u_1' &= 0u_1 + 1u_2 + 0u_3 + 0u_4 \\ u_2' &= 3u_1 + 2u_2 - 1u_3 + 1u_4 \\ u_3' &= 0u_1 + 0u_2 + 0u_3 + 1u_4 \\ u_4' &= 2u_1 - 4u_2 + 2u_3 - 1u_4 \end{aligned}$$

or $\vec{u}' = A\vec{u}$ with $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 3 & 2 & -1 & 1 \\ 0 & 0 & 0 & 1 \\ 2 & -4 & 2 & -1 \end{pmatrix}$

6.6 – Applications Involving Systems

Ex i) Consider a mass on a spring, with another spring and mass hanging from it. How can we describe the motion?



The forces on these masses again come from acceleration, friction, springs, and external.

The system is

$$
\begin{aligned}
m_1 x_1'' &= -k_1 x_1 + k_2(x_2 - x_1) - f_1 x_1' + h_1(t) \\
\underbrace{m_2 x_2''}_{\text{accel.}} &= \underbrace{-k_2(x_2 - x_1)}_{\text{spring forces}} \underbrace{- f_2 x_2'}_{\text{friction}} \underbrace{+ h_2(t)}_{\text{external}}
\end{aligned}
$$

or

$$
\begin{aligned}
x_1'' &= \left(\frac{-k_1 - k_2}{m_1}\right) x_1 + \left(\frac{-f_1}{m_1}\right) x_1' + \left(\frac{k_2}{m_1}\right) x_2 + (0) x_2' + \frac{h_1}{m_1} \\
x_2'' &= \left(\frac{k_2}{m_2}\right) x_1 + (0) x_1' + \left(-\frac{k_2}{m_2}\right) x_2 - \left(\frac{f_2}{m_2}\right) x_2' + \frac{h_2}{m_2}
\end{aligned}
$$

Writing

$$
\begin{aligned}
x_1 &= u_1 \\
x_1' &= u_2 \\
x_2 &= u_3 \\
x_2' &= u_4
\end{aligned}
$$

this becomes

$$
\begin{aligned}
u_1' &= 0\, u_1 + 1\, u_2 + 0\, u_3 + 0\, u_4 + 0 \\
u_2' &= \left(\frac{-k_1 - k_2}{m_1}\right) u_1 + \left(\frac{-f_1}{m_1}\right) u_2 + \left(\frac{k_2}{m_1}\right) u_3 + 0\, u_4 + \frac{h_1}{m_1} \\
u_3' &= 0\, u_1 + 0\, u_2 + 0\, u_3 + 1\, u_4 + 0 \\
u_4' &= \left(\frac{k_2}{m_2}\right) u_1 + 0\, u_2 + \left(\frac{-k_2}{m_2}\right) u_3 + \left(\frac{-f_2}{m_2}\right) u_4 + \frac{h_2}{m_2}
\end{aligned}
$$

We can write this in matrix form as

$$\vec{u}' = \begin{pmatrix} 0 & 1 & 0 & 0 \\ \left(\frac{-k_1-k_2}{m_1}\right) & \left(\frac{-f_1}{m_1}\right) & \left(\frac{k_2}{m_1}\right) & 0 \\ 0 & 0 & 0 & 1 \\ \left(\frac{k_2}{m_2}\right) & 0 & \left(\frac{-k_2}{m_2}\right) & \left(\frac{-f_2}{m_2}\right) \end{pmatrix} \vec{u} + \begin{pmatrix} 0 \\ h_1/m_1 \\ 0 \\ h_2/m_1 \end{pmatrix}$$

Ex: If there is no friction and no external forcing, we can solve the second order system <u>without converting</u>.

$$m_1 x_1'' = -k_1 x_1 + k_2 (x_2 - x_1)$$

$$\underbrace{m_2 x_2''}_{\text{accel.}} = \underbrace{-k_2 (x_2 - x_1)}_{\text{spring forces}}$$

$$\vec{x}'' = \begin{pmatrix} \frac{-k_1-k_2}{m_1} & \frac{k_2}{m_1} \\ \frac{k_2}{m_2} & \frac{-k_2}{m_2} \end{pmatrix} \vec{x}$$

This is homogeneous, and we might again hope to find exponential solutions of the form

$$\vec{x} = e^{rt} \vec{v}$$

The equation becomes

$$r^2 e^{rt} \vec{v} = A e^{rt} \vec{v}.$$

$$r^2 \vec{v} = A \vec{v}$$

This is an eigenvalue equation, with eigenvalue $\lambda = r^2$ and eigenvector $\vec{v}$.

That is — if we have an eigenvector $\vec{v}$ and choose $r = \pm\sqrt{\lambda}$, then

$$\vec{y} = e^{rt} \vec{v}$$

is a solution.

With some algebra it can be shown that for this particular matrix, λ is always negative. So for any eigenvalue, we choose

$$r = \pm\sqrt{\lambda} = \pm i \sqrt{|\lambda|} = \pm i \omega$$

If we have two eigenvalues λ_1, λ_2 with eigenvectors $\vec{v}_1, \vec{v}_2$, we then get solutions

$$\left\{ e^{i\omega_1 t}\vec{v}_1 , e^{-i\omega_1 t}\vec{v}_1 , e^{i\omega_2 t}\vec{v}_2 , e^{-i\omega_2 t}\vec{v}_2 \right\}$$

Looking at real and imaginary parts, we can choose a different (real) basis:

$$\left\{ \vec{v}_1 \cos(\omega_1 t), \vec{v}_1 \sin(\omega_1 t), \vec{v}_2 \cos(\omega_2 t), \vec{v}_2 \sin(\omega_2 t) \right\}$$