MATB24

Prof. Michael Cavers

W1 (3.1)

Def. Vector Space

- A **non-empty set** V of objects called **vectors**, along with **rules** for
  - adding any 2 vectors v and w in V
  - multiplying any vector v in V by any scalar r in $\mathbb{R}$
- Must satisfy **closure** properties
  - vector addition: $\forall\, v, w \in V, v + w \in V$
  - scalar multiplication: $\forall\, r \in \mathbb{R}, v \in V, rv \in V$
- Must satisfy the following axioms for **vector addition**

  1. **Associative**

     $(u + v) + w = u + (v + w)$

  2. **Commutative**

     $u + v = v + u$

  3. **An additive identity exists**

     $\exists$ a zero vector $0 \in V$ s.t. **0 + v = v**, $\;\forall\, v \in V$

  4. **An additive inverse exists**

     $\exists$ an additive inverse $-v \in V$ s.t. **v + (-v) = 0**

- Must satisfy the following axioms for **scalar multiplication**

  1. **Distributes over vector addition**

$$r \cdot (v + w) = (r \cdot v) + (r \cdot w)$$

2. **Distributes over scalar addition**

$$(r + s) \cdot v = (r \cdot v) + (s \cdot v)$$

3. **Associative**

$$r \cdot (s \cdot v) = (r \cdot s) \cdot v$$

4. **A multiplicative identity exists**

$$1 \cdot v = v$$

Note: ***vectors*** could refer to matrices, polynomials, functions, power series, etc.

> **Q** Let $V = \left\{\vec{v} \in \mathbb{R}^n \middle| \text{components of } \vec{v} \text{ sum to } 0\right\}$ with usual operations. Is V a vector space?
>
> **A** Let $\vec{u}, \vec{v} \in V$. Suppose $\vec{u} = \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}, \vec{v} = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$
>
> Check closure under addition:
>
> $$\vec{u} + \vec{v} = \sum_{i=1}^{n} (u_i + v_i) = \sum_{i=1}^{n} u_i + \sum_{i=1}^{n} v_i = 0 + 0 = 0 \in V$$
>
> Check closure under scalar mult:
>
> $$r\vec{u} = \sum_{i=1}^{n} (ru_i) = r \sum_{i=1}^{n} u_i = r \cdot 0 = 0 \in V$$
>
> Usual ops $\implies$ A1-A4, S1-S4 all hold $\implies$ V is a vector space.

> **Q** Find the zero vector in $V = \mathbb{R}^2$ where addition is defined by $\begin{bmatrix} a \\ b \end{bmatrix} + \begin{bmatrix} c \\ d \end{bmatrix} = \begin{bmatrix} a + c \\ b + d + 1 \end{bmatrix}$
>
> **A** $\begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} \implies \begin{bmatrix} x + a \\ y + b + 1 \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix} \implies x = 0, y = 1 \implies \vec{0} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$

To define a vector space, need to specify 4 parts:

1. Objects that make up the set V

2. The field where scalars come from (if the field of scalars is the set of real #s, then V is a real vector space)
3. How to add vectors
4. How to multiply scalars with objects in V

The below sets with given rules for $\oplus, \odot$ are vector spaces.

Notation	**Definition**
$\mathbb{R}^n$	set of vectors of length n with real entries
$\mathbb{C}^n$	set of vectors of length n with complex entries
P	set of all polynomials in x with real coefficients
P_n	P with degree $\leq$ n
M_{mn}	set of all real matrices of size m(rows) by n(columns)
F	set of real-valued functions with domain $\mathbb{R}$

Thm 3.1 (Elementary Properties of Vector Spaces)

Every vector space V has the following properties:

1. $\vec{0}$ **is unique** for all vectors in V
2. $-\vec{v}$ **is unique** for each vector in V
3. $u+v=u+w \implies v=w$
4. $0\vec{v}=\vec{0},\ \forall\, \vec{v}\in V$
5. $r\vec{0}=\vec{0},\ \forall\, r\in\mathbb{R}$
6. $(-r)v=r(-v)=-(rv),\ \forall\, r\in\mathbb{R}, v\in V$

Proof

1. Suppose $\exists$ **2** zero vectors, x and y. Then for all v in V, $x+v=v$ and $y+v=v$.

   Since $y\in V$, sub $v=y$ into $x+v=v \implies x+y=y$

   Since $x\in V$, sub $v=x$ into $y+v=v \implies y+x=x$

   $x+y=y+x \implies y=x \implies \vec{0}$ is unique

2. By definition, $v+(-v)=0$. Suppose $v+u=0$.

   $u=u+0=u+[v+(-v)]=(u+v)+(-v)=0+(-v)=(-v)$

Any $u = -v \implies$ additive inverse is unique

3. If $u + v = u + w$

   Then $v = 0 + v = (-u + u) + v = -u + (u + v) = -u + (u + w) = (-u + u) + w = 0 + w = w$

   Equate LHS & RHS $\implies v = w$

4. $\vec{0} + 0\vec{v} = 0\vec{v} = (0 + 0)\vec{v} = 0\vec{v} + 0\vec{v}$.

   Equate LHS & RHS $\implies \vec{0} = 0\vec{v}$

5. $r\vec{v} + r\vec{0} = r(\vec{v} + \vec{0}) = r\vec{v} = r\vec{v} + \vec{0}$.

   Equate LHS & RHS $\implies r\vec{0} = \vec{0}$

6. $r\vec{v} + (-r\vec{v}) = (r + (-r))\vec{v} = 0\vec{v} = \vec{0} \implies -(r\vec{v}) = (-r)\vec{v}$

   $\vec{v} + (-\vec{v}) = \vec{0} \implies r\vec{v} + r(-\vec{v}) = \vec{0} \implies -(r\vec{v}) = r(-\vec{v})$

   $\implies -(r\vec{v}) = (-r)\vec{v} = r(-\vec{v})$

W2 (3.2)

Vector Space vs Real Space

- Concepts for real space extend to vector spaces
- Except that vector spaces could be spanned by an **infinite** set of vectors

> **E.g.** The vector space P of all polynomials is spanned by the set of monomials $M = \{1, x, x^2, \dots, x^n, \dots\}$
>
> - M is a basis for P $\implies$ M spans P $\iff$ P = span(M)
> - M generates P and so P is NOT finitely generated

Def. Linear Combination

$\vec{x}$ is a linear combination of $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_k \in V$ if there $\exists$ scalars $c_1, c_2, \dots c_k \in \mathbb{R}$ s.t. $\vec{x} = c_1\vec{x}_1 + \dots + c_k\vec{x}_k$

Note linear combinations must be of a finite number of vectors (otherwise called Schauder basis)

Def. Span

$sp(X)$ is the set of all linear combinations of elements in X, where X is a subset of V.

If $V = sp(X)$, then the vectors in X span or generate V.

> **Q** Is $1+x^2 \in sp(1+x, x+x^2, 2+x^2)$? Assume standard polynomial addition and scalar mult.
>
> **A** Check if $1+x^2$ can be written as $a(1+x)+b(x+x^2)+c(2+x^2)$
>
> $$1+x^2 = a+ax+bx+bx^2+2c+cx^2$$
>
> $$\implies \begin{cases} 1 = a+2c \\ 0 = a+b \\ 1 = b+c \end{cases} \text{ (derived using the coefficients)}$$
>
> $$\implies \begin{cases} a = -1/3 \\ b = 1/3 \\ c = 2/3 \end{cases} \implies \text{Unique solution, so yes, } 1+x^2 \text{ is in the span of the set.}$$

Thm (Span of a subset is a subspace)

If X is a non-empty subset of vector space V, then $sp(X)$ is a subspace of V.

Proof

1. Since X is non-empty, sp(X) must be non-empty
2. Let $r_1, r_2 \in \mathbb{R}$, let $v_1, v_2 \in sp(X)$

   Since $v_1, v_2 \in sp(X), \exists\, a_1, \ldots, a_n \in \mathbb{R}, b_1, \ldots b_n \in \mathbb{R}, x_1, \ldots, x_n \in X$ s.t.

   $$v_1 = a_1x_1 + \ldots + a_nx_n$$

   $$v_2 = b_1x_1 + \ldots + b_nx_n$$

   Then $r_1v_1 + r_2v_2 = r_1(a_1x_1 + \ldots + a_nx_n) + r_2(b_1x_1 + \ldots + b_nx_n)$

   $$= (r_1a_1 + r_2b_1)x_1 + \ldots + (r_1a_n + r_2b_n)x_n \in sp(X)$$

Non-empty, closed under additional and scalar mult -> span of X is a subspace of V.

> **E.g.** Let $X = \{1, x, x^2\} \in P$

$(1+x)+(1+x^2)=x^2+x+2$, which is still a linear combination of vectors in X, so we have closure under addition.

$r(1+x)=r+rx$, which is a multiple of - again - linear combinations of vectors in X, so we have closure under mult.

Corollary (Not covered in class)

$sp(X)$ is the smallest subspace of V containing all the vectors in X.

Proof

Let M be the smallest subspace of V containing $v_1, \ldots, v_n$

Claim: $M = span(v_1, \ldots, v_n)$

Since $v_i = 0 \cdot v_1 + \ldots + 1 \cdot v_i + \ldots + 0 \cdot v_n$ for all $1 \leq i \leq n$,

we have $v_i \in span(v_1, \ldots, v_n)$ by definition of span, so $M \subseteq span(v_1, \ldots v_n)$

Since $v_i \in M$ and every subspace is a vector space with closure properties,

we have $a_1 v_1 + \ldots + a_n v_n \in M$, so $span(v_1, \ldots v_n) \subseteq M$

Thus $M = span(v_1, \ldots, v_n)$

Def. Subspace

Let V be a vector space and U a subset of V. Then U is a subspace of V if U is a vector space using the defined ops for V.

E.g. Counter examples

Subsets of $\mathbb{R}^2$ that are closed under addition but not scalar multiplication:

$S = \left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x, y \geq 0 \right\}$ if multiply by r=-1, the result $\ngeq 0$ $so \notin S$

$S = \{(n,0) \in \mathbb{R}^2 : n \in \mathbb{Z}\}$ if multiply by r = 0.5, the result $\notin \mathbb{Z}$

$S = \{(r,s) \in \mathbb{R}^2 : r, s \in \mathbb{Q}\}$ if multiply by r = π, the result $\notin \mathbb{Q}$

> Subset of $\mathbb{R}^2$ which is closed under scalar multiplication but not addition:
>
> $$\left\{ \begin{bmatrix} x \\ y \end{bmatrix} : x = 0 \; or \; y = 0 \right\}$$

Ways to prove U is a subspace

1. Show U is a vector space using the definition (too much work)
2. Use the span theorem (probably easiest)
3. Use the subspace test

Thm 3.2 (The Subspace Test)

A subset U of V is a subspace of V if

1. The zero vector $\vec{0} \in U$ (Sometimes omitted as it is implied by 3)
2. U is closed under vector addition
3. U is closed under scalar multiplication

Shortcut for 2 & 3: check both at the same time by using $a\vec{x} + b\vec{y} \in U$

> **Q** Is the set of nxn invertible matrices a subspace of M_n (the vector space of all nxn matrices)?
>
> **A** No. The sum of 2 invertible matrices may not be invertible (violates 2).
>
> Also, the 0 matrix is not invertible (violates 1).

Def. Linear (In)dependence

Let X be a set of vectors in a vector space V. (X can be an infinite set.)

A dependence relation in X is an equation of the form $r_1\vec{v}_1 + \ldots + r_k\vec{v}_k = \vec{0}$ with some $r_i \neq 0$ where $\vec{v}_i \in X$ for i = 1, ... , k.

If such a dependence relation exists, then X is linearly dependent. Otherwise, X is linearly independent.

If X is finite, then the following definition applies:

The set $\{\vec{x}_1, \ldots, \vec{x}_k\}$ with $\vec{x}_1, \ldots, \vec{x}_k \in V$ is linearly independent if $c_1\vec{x}_1 + \ldots + c_k\vec{x}_k = \vec{0}$ implies $c_1 = \ldots = c_k = 0$

Note: { } is independent (vacuous statement)

- There is no collection of vectors from { } satisfying a dependence relation because there is nothing in the set at all
- sp({ }) = $\{\vec{x} : \underbrace{\vec{x} \text{ is the linear combination of vectors in } \{\ \}}_{\text{an empty sum}}\}$
- Convention: the empty summation is taken to be its additive identity ($\vec{0}$).

Q Suppose $\{\vec{x}, \vec{y}, \vec{z}\}$ is linearly independent. Prove $\{x+y, 2x+z, y-5z\}$ is also linearly independent.

A Assume $c_1(x+y)+c_2(2x+z)+c_3(y-5z)=\vec{0}$

WTS this implies $c_1=c_2=c_3=0$

$c_1x+c_1y+2c_2x+c_2z+c_3y-5c_3z=\vec{0}$

$(c_1+c_2)x+(c_1+c_3)y+(c_2-5c_3)z=\vec{0}$

Since $\{\vec{x}, \vec{y}, \vec{z}\}$ is linearly independent, $\begin{cases} c_1+2c_2=0 \\ c_1+c_3=0 \\ c_2-5c_3=0 \end{cases} \implies c_1, c_2, c_3 = 0 \implies \textit{independent}$

Q Consider F, the vector space of all functions from $\mathbb{R} \to \mathbb{R}$ (with usual operations).

Is $X = \{\sin^2 x, \cos^2 x, 1\}$ linearly independent?

A X is linearly dependent since $(1)sin^2x+(1)cos^2x+(-1)(1)=0$

Q Is $X = \{\sin x, \cos x\}$ linearly independent?

A X is linearly independent.

Assume $a\sin x + b\cos x = 0$ holds $\forall x$

(The equation must hold for $x=0, \pi/2$ since it holds for all x.)

$a\sin 0 + b\cos 0 = 0 \implies b=0$

$a\sin(\pi/2)+b\cos(\pi/2)=0 \implies a=0$

$\because a, b = 0 \therefore$ X is linearly independent

If we find a non-trivial solution (a or b or both != 0), NO conclusion can be made.

> To generate more equations, we can take derivatives first then plug in points.

To check if a set is linearly independent:

If it cannot be observed, set up the dependence relation equation as the first step.

If the set spans a space with a dimension lower than the cardinality of the set (i.e. dim(span of the set) < |set|), then the set cannot be linearly independent.

If given a set of polynomials, put the coefficients as column vectors in a matrix, reduce to REF, and solve the system of equations.

Other ways:

- create a system of equations by plugging in dif values of x (# values = # unknowns) and solve
- differentiate the dependence relation, plug in the same value of x each iteration
- expand and rearrange in terms of a linear independent set, then the coefficients must be 0, solve
- see if the set contains the zero vector

Def. Basis

A subset X of V is a basis for V if:

1. X spans V
2. X is linearly independent

Def. Dimension

Let V be a finitely generated vector space. The number of elements in a basis for V is the dimension of V, dim(V).

Note: When giving definitions, "a" and "the" imply different things - "the" implies that the object described is unique.

Q Is dimension well defined? I.e. Can 2 bases have different sizes?

A It is well defined since any 2 bases of V must have the same number of elements. This is true by the following theorem.

Thm (Dimension and Linear Independence)

Suppose that V is a vector space of dimension n. Then any subset X with size greater than n is linearly dependent.

Proof

Since V has dimension n, there must exist a basis of size n: $B = \{v_1, \ldots, v_n\}$

Let $X = \{x_1, \ldots, x_m\}$ with $m > n$

Since B is a basis, each x_j can be written as a linear combination of elements of B: $x_i = \sum_{j=1}^{n} a_{ji}v_j, i \in \{1, \ldots, m\}$.

i.e. $x_1 = a_{11}v_1 + \ldots + a_{n1}v_n$

$$\vdots$$

$$x_m = a_{1m}v_1 + \ldots + a_{nm}v_n$$

Assume $\exists$ a vector $(\lambda_1, \ldots, \lambda_m)$ such that $\begin{bmatrix} a_{11} & \ldots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \ldots & a_{nm} \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_m \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$

Since m > n, we have more unknowns than equations, so there must be a non-trivial (non-zero) solution, i.e. the lambda vector must be non-zero.

Setup dependence relation: $\sum_{i=1}^{m} \lambda_i x_i = 0$

$$\sum_{i=1}^{m} \lambda_i x_i = \sum_{i=1}^{m} \lambda_i (\sum_{j=1}^{n} a_{ji}v_j) = \sum_{j=1}^{n} (\sum_{i=1}^{m} \lambda_i a_{ji}) v_j \overset{optional}{=} \sum_{j=1}^{n} (\lambda_1 a_{j1} + \ldots + \lambda_m a_{jm}) v_j = \sum_{j=1}^{n} \lambda_1 a_{j1} v_j + \ldots + \lambda_m a_{jm} v_j$$

a_{ji} cannot all be 0, and λ_i cannot all be 0, so there exists at least one non-zero coefficient in the dependence relation
so X is not linearly independent.

■

E.g. $\mathbb{R}^n$ (with usual operations) has a standard basis of $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ldots, \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right\}$ and $\dim(\mathbb{R}^n) = n$

E.g. $M_{m,n}$ (with usual operations) has $\dim(M_{m,n}) = m \cdot n$

E.g. Diagonal matrices M_n (with usual operations) has $\dim(M_n) = n$

E.g. P_n (with usual operations) has a standard basis of $\{1, x, \ldots, x^n\}$ of $\dim(P_n) = n + 1$

Note The degree of the zero polynomial is conventionally -1 or -∞ (not defined in the textbook).

Note The subset $\left\{\vec{0}\right\}$ is a subspace of dimension 0, since its basis is { } or ∅ which has dimension 0.

- Any set containing $\vec{0}$ is dependent because $c_1 \cdot \vec{0} = \vec{0}$ has a solution $c_1 = 1$
- So its basis cannot be $\left\{\vec{0}\right\}$ since a basis cannot be a linearly dependent set

Thm (Lin Ind Set ⊆ Basis ⊆ Spanning Set)

- Every linearly independent set of vectors in U can be enlarged to a basis of U by adding vectors from **any** fixed basis of U. (Be careful which ones you add - to preserve independence.)
- Any spanning set for U can be cut down (by deleting vectors) to a basis of U. (Be careful which ones you delete to preserve span.)

Thm 3.3 (Unique Representation)

Let B be a set of non-zero vectors in a vector space V.

Then B is a basis for V iff each vector $\vec{v}$ in V can be **uniquely** expressed as $\vec{v} = r_1\vec{b}_1 + \ldots + r_k\vec{b}_k$ where $r_i \in \mathbb{R}$ and $\vec{b}_i \in B$.

Proof for Thm 3.3

When B is finite:

(⇒) Suppose B is a basis (so it is lin ind) and $\exists\, \vec{v}$ that can be expressed the following 2 ways:

$$\begin{matrix} \vec{v} = r_1\vec{b}_1 + \ldots + r_k\vec{b}_k \\ \vec{v} = s_1\vec{b}_1 + \ldots + s_k\vec{b}_k \end{matrix} \implies \vec{0} = (r_1 - s_1)\vec{b}_1 + \ldots + (r_k - s_k)\vec{b}_k$$

B is linearly independent, so $r_1 - s_1 = 0, \ldots, r_k - s_k = 0 \implies r_1 = s_1, \ldots, r_k = s_k \implies \vec{v}$ has unique representation.

(⇐) Suppose every vector can be uniquely expressed in terms of B. This implies V = span(B).

$\vec{0}$ can be uniquely expressed in terms of B, so $r_1\vec{b}_1 + \ldots + r_k\vec{b}_k = \vec{0} \implies r_1 = \ldots = r_k = 0 \implies$ B is lin ind.

B spans V and is linearly independent, so B must be a basis.

When B is infinite:

- pad two different expressions of a vector (in terms of a finite number of elements in B) with zeros
- so that the two expressions use the same finite subset of vectors from B
- As a result, B now represents this finite subset (of the infinite set) used in our representations of the vector v.

Def. Finitely Generated

V is finitely generated if it is the span of a finite set, i.e. if $V = sp(X)$ where X is finite.

> **Q** Which of these vector spaces is finitely generated?
>
> 1. The vector space of all functions $\mathbb{R} \to \mathbb{R}$?
> 2. The vector space of all polynomials.
> 3. The vector space of all polynomials of degree less than or equal to n.
> 4. The vector space of all 2×2 matrices.
>
> **A** 3 is finitely generated as it is the span of the finite set $\{1, x^2, \ldots, x^n\}$
>
> 4 is finitely generated as it is the span of the finite set of 2x2 matrices E_{ij} where the ij-th entry = 1 and all other = 0

Thm (Shortcut for proving basis)

Let B be a set of m vectors from V. If |B| = dim(V), then showing one of (B is lin ind, B spans V) will suffice.

Proof (If $\dim(V) = m$, then any lin ind B is a basis of V.)

B is a basis for V if it spans V and is lin ind. Lin ind is given, so just need to show it spans V.

Suppose that B does not span V. Then $\exists\, v \in V, v \notin sp(B) \implies \dim(B \cup \{v\}) = m + 1$.

This contradicts the fact that any set with >m elements cannot be lin ind in a vector space of dimension m.

Proof (If $\dim(V) = m$, then any B spanning V is a basis of V.)

B is a basis for V if it spans V and is lin ind. We know B spans V, so just need to show lin ind.

Suppose that B is not lin ind. Then at least 1 of its elements can be written as a lin combo of the other elements.

Delete that element to get B', which would have the same span as B. B' would have m-1 elements.

Continue deleting until lin ind. Since the set has the same span as B, it must be a basis for V.

However, it would have <m elements, which contradicts $\dim(V) = m$.

Q Let $U = \{p(x) \in P_2 : p(1) = 0\}$. Show U is a subspace of P_2 and find a basis for U.

A Any $p(x) \in U$ satisfies $p(1) = 0$

Consider $p(x) = ax^2 + bx + c$, then $a(1)^2 + b(1) + c = 0 \implies a + b + c = 0 \implies \begin{cases} a = -s - t \\ b = s \\ c = t \end{cases}$ where $s, t \in \mathbb{R}$

Plugging in a, b, c, we have $p(x) = (-s-t)x^2 + sx + t$

Then

$$U = \left\{(-s-t)x^2 + sx + t : s, t \in \mathbb{R}\right\} = \left\{s(-x^2 + x) + t(-x^2 + 1) : s, t \in \mathbb{R}\right\} = sp(\left\{-x^2 + x, -x^2 + 1\right\})$$

By Span Theorem, U is a subspace of P_2.

$\left\{-x^2 + x, -x^2 + 1\right\}$ is linearly independent, so it is a basis of U and $\dim(U) = 2$.

W3 (3.3, 3.4 pg. 213-216)

Recall: If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then T is induced by a matrix $A_{mn} = \left[T(\vec{e}_1)|\dots|T(\vec{e}_n)\right]$ where $\left\{\vec{e}_1, \dots, \vec{e}_n\right\}$ is the standard basis of $\mathbb{R}^n$. Then for $\vec{x} \in \mathbb{R}^n$ we have $T(\vec{x}) = A\vec{x}$.

Ordered Bases

We can give an order to every basis. **Note:** To emphasize the basis is ordered, we use ( ) instead of { }.

E.g. $\mathbb{R}^3$ has 6 ordered bases using $\{\vec{e}_1, \vec{e}_2, \vec{e}_3\}$ = $\left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$

They are: $(e_1, e_2, e_3), (e_1, e_3, e_2), (e_2, e_1, e_3), (e_2, e_3, e_1), (e_3, e_1, e_2), (e_3, e_2, e_2)$

Def. Coordinate Vector

Let $B = (\vec{b}_1, \ldots, \vec{b}_n)$ be an ordered basis for a vector space V.

Suppose $\vec{v} = c_1\vec{b}_1 + \ldots + c_n\vec{b}_n$, then the **coordinate vector of $\vec{v}$ relative to the basis B** is defined as $\vec{v}_B = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$

By the unique representation theorem, this vector is well defined.

> **E.g.** Let $\vec{v} = 2x^3 - 3x + 1 \in P_2$
>
> If $B = (1, x, x^2)$, then $\vec{v}_B = \begin{bmatrix} 1 \\ -3 \\ 2 \end{bmatrix}$
>
> If $B = (x, 2x^2, 1)$, then $\vec{v}_B = \begin{bmatrix} -3 \\ 2 \\ 1 \end{bmatrix}$

Thm (Coordinatization of Vector Spaces)

Any finite dimensional vector space V with dimension n can be coordinatized to look like $\mathbb{R}^n$.

We choose an ordered basis $(\vec{b}_1, \ldots, \vec{b}_n)$.

Since each vector $\vec{v} \in V$ has a *unique* coordinate vector relative to B, there is a 1-1 correspondence between objects in V and vectors in $\mathbb{R}^n$

We may now work over $\mathbb{R}^n$ instead of V

To see this, we verify that vector space operations are mirrored by operations on coordinate vectors in $\mathbb{R}^n$

Proof

We show that for $\vec{v}, \vec{w} \in V, t \in \mathbb{R}$, vector space ops = $\mathbb{R}^n$ ops:

(1) Show that $(\vec{v} + \vec{w})_B = \vec{v}_B + \vec{w}_B$

Suppose (by unique representation theorem) we have vectors v and w:

$\vec{v} = r_1\vec{b}_1 + \ldots + r_n\vec{b}_n \implies \vec{v}_B = (r_1, \ldots, r_n)$

$\vec{w} = s_1\vec{b}_1 + \ldots + s_n\vec{b}_n \implies \vec{w}_B = (s_1, \ldots, s_n)$

Then $(\vec{v}+\vec{w})_B = (r_1\vec{b}_1 + \ldots + r_n\vec{b}_n + s_1\vec{b}_1 + \ldots + s_n\vec{b}_n)_B$

$$= ((r_1+s_1)\vec{b}_1 + \ldots + (r_n+s_n)\vec{b}_n)_B$$

$$= \begin{bmatrix} r_1+s_1 \\ \vdots \\ r_n+s_n \end{bmatrix} = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix} + \begin{bmatrix} s_1 \\ \vdots \\ s_n \end{bmatrix} = \vec{v}_B + \vec{w}_B$$

(2) Show that $(t\vec{v})_B = t\vec{v}_B$

$$(t\vec{v})_B = (t(r_1\vec{b}_1 + \ldots + r_n\vec{b}_n))_B = (tr_1\vec{b}_1 + \ldots + tr_n\vec{b}_n)_B = [tr_1, \ldots, tr_n] = t[r_1, \ldots, r_n] = t\vec{v}_B$$

Def. Isomorphic

When we rename vectors in V using coordinates, the vector space of coordinates (i.e. $\mathbb{R}^n$) has the same structure as V.

If V appears structurally identical to W, we say V and W are isomorphic, or $V \cong W$.

> **E.g.** $P_n \cong \mathbb{R}^{n+1}$, $M_{m,n} \cong P_{mn-1} \cong \mathbb{R}^{mn}$

Thm (Dimension Determines Structure)

If V is a vector space over $\mathbb{R}$ and $\dim(V) = n$, then V is isomorphic to $\mathbb{R}^n$

Hence, to know the structure of a vector space (no matter how complicated $\oplus, \odot$ are), we only need to know its dimension - i.e. V and W are isomorphic $\iff$ dim(V) = dim(W)

This is contrary to other mathematical structures like groups, rings, etc.

Def. Linear Transformation

Let V and W be vector spaces, then the function $T: V \to W$ is a linear transformation if

1. T preserves addition

   For $\vec{v}_1, \vec{v}_2 \in V$, we need $\underbrace{T(\vec{v}_1+\vec{v}_2)}_{in\,V} = \underbrace{T(\vec{v}_1)+T(\vec{v}_2)}_{in\,W}$

2. T preserves scalar multiplication

   For all $\vec{V} \in V, k \in \mathbb{R}$, we need $T(k\cdot\vec{v}) = k\cdot T(\vec{v})$

It is clear to see that T preserves linear combinations: $T(a_1\vec{v}_1 + ... + a_k\vec{v}_k) = a_1T(\vec{v}_1) + ... + a_kT(\vec{v}_k)$

Fact 2 lin transf are the same if they have the same value at each basis vector $\vec{b}_i$

Proof Let T and T' be 2 linear transformations such that $T(\vec{b}_i) = T'(\vec{b}_i)\ \forall\, \vec{b}_i \in B$

Let $\vec{v} \in V$, then $\exists\, \vec{b}_1, \ldots, \vec{b}_k \in B$ and $r_1, .., r_k \in \mathbb{R}$ such that $\vec{v} = r_1\vec{b}_1 + ... + r_k\vec{b}_k$

$$\begin{aligned} T(\vec{v}) &= T(r_1\vec{b}_1 + ... + r_k\vec{b}_k) \\ &= r_1T(\vec{b}_1) + ... + r_kT(\vec{b}_k) \\ &= r_1T'(\vec{b}_1) + ... + r_kT'(\vec{b}_k) \\ &= T'(r_1b_1 + ... + r_kb_k) \\ &= T'(\vec{v}) \end{aligned}$$

Thus T, T' are the same transformation

Other facts:

1. T preserves zero vector

   $T(\vec{0}_V) = \vec{0}_W$

   Proof $T(\vec{0}) \overset{r=0}{=} T(0\vec{0}) = 0T(\vec{0}) = \vec{0}$

2. T preserves additive inverse (and thus substraction)

   $T(-\vec{v}) = -T(\vec{v})$

> **E.g.** $T : M_{n,n} \to \mathbb{R}$ defined $T(A) = tr(A)$ is a linear transformation
>
> > **Recall** $tr(A) = \sum_{i=1}^{n} a_{ii}$
>
> 1. T preserves addition
>
> $T(A + B) = tr(A + B) = tr(A) + tr(B) = T(A) + T(B)$
>
> 2. T preserves scalar mult
>
> $T(kA) = tr(kA) = k\, tr(A) = k\, T(A)$

E.g. $T: M_{n,n} \to \mathbb{R}, n > 1$ defined $T(A) = \det A$ is not a linear transformation

1. T preserves addition

   Consider A = B = I

   $T(A) + T(B) = T(I) + T(I) = \det I + \det I = 1 + 1 = 2$

   $T(A + B) = T(I + I) = \det(I + I) = \det(2I) = 2^n$

   > **Recall** Determinant of a diagonal matrix is the product of the diagonal entries

   The above are not equal if n > 1

2. T preserves scalar mult - no need to check

E.g. Derivatives and integrals are linear transformations

$T(f) = f'$

- $T(f + g) = (f + g)' = f' + g' = T(f) + T(g)$
- $T(rf) = (rf)' = r(f') = rT(f)$

$T(f) = \int_a^b f(x)dx$

- $T(f + g) = \int_a^b [f(x) + g(x)]dx = \int_a^b f(x)dx + \int_a^b g(x)dx = T(f) + T(g)$
- $T(rf) = \int_a^b rf(x)dx = r\int_a^b f(x)dx = rT(f)$

Def. Domain, Codomain

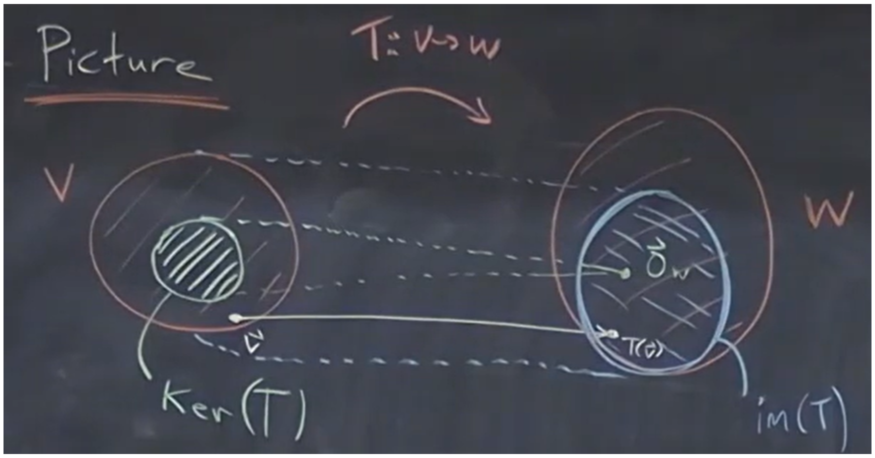
Let V and W be vector spaces and $T: V \to W$ be a linear transformation.

Then V is the **domain** of T, W is the **codomain** of T.

Def. Image, Inverse Image

The **image** of T (or image of V under T, in blue) is a subspace of W: $\mathfrak{I}(T) = \{T(\vec{v}) : \vec{v} \in V\} = T[V]$

Labelling $\mathfrak{I}(T)$ as W', then the **inverse image** of W' under T is a subspace of V: $\{\vec{v} : T(\vec{v}) \in W'\} = T^{-1}[W']$



Def. Range, Kernel / Null Space

$T[V]$ is the **range** of T (also image of V under T).

The **kernel** of T (or inverse image of $\vec{0}_W$ under T, in green) is a subspace of V:
$\ker(T) = \{\vec{v} : T(\vec{v}) = \vec{0}_W\} = T^{-1}[\{\vec{0}_W\}]$

It is clear to see that $\ker(T)$ is solution set of the **homogeneous transformation equation** $T(\vec{v}) = \vec{0}_W$

Thm (Rank-Nullity)

$\dim(V) = rank(T) + \text{nullity}(T) = \dim(\mathfrak{I}(T)) + \dim(\ker(T))$

Intuition: each dimension (or component) in V either gets preserved or compressed to $\vec{0}_W$

Proof

Let $\{\vec{v}_1, \ldots, \vec{v}_k\}$ be a basis for $\ker(T)$ so $\dim(\ker(T)) = k$

Extend it to a basis for $V : \{\vec{v}_1, \ldots, \vec{v}_k, \vec{w}_1, \ldots, \vec{w}_m\}$ so $\dim(V) = k + m$

We know $\dim(V) = k + m, \dim(\ker(T)) = k$, just need to prove $\dim(\mathfrak{I}(T)) = m$

Do so by finding a basis for the image:

By Unique Rep Thm, we can write any $\vec{v} \in V$ as $\vec{v} = a_1\vec{v}_1 + \ldots + a_k\vec{v}_k + b_1\vec{w}_1 + \ldots + \vec{b}_m\vec{w}_m$, so $\Im(T) = \{T(\vec{v}) : \vec{v} \in V\} = sp(\{\underbrace{T(\vec{v}_1), \ldots, T(\vec{v}_k)}_{T(\vec{v}_i)=\vec{0}_W}, T(\vec{w}_1), \ldots, T(\vec{w}_m)\}) = sp(\{T(\vec{w}_1), \ldots, T(\vec{w}_m)\})$

WTS $\{T(\vec{w}_1), \ldots, T(\vec{w}_m)\}$ is lin ind: set up the dependence relation $\lambda_1 \cdot T(\vec{w}_1) + \ldots + \lambda_m \cdot T(\vec{w}_m) = \vec{0}_W$

$\implies T(\lambda_1\vec{w}_1 + \ldots + \lambda_m\vec{w}_m) = \vec{0}_W$ since T is linear

$\implies \lambda_1\vec{w}_1 + \ldots + \lambda_m\vec{w}_m \in \ker(T)$

$\implies \lambda_1\vec{w}_1 + \ldots + \lambda_m\vec{w}_m = \mu_1\vec{v}_1 + \ldots + \mu_k\vec{v}_k$ since $\{\vec{v}_1, \ldots, \vec{v}_k\}$ is a basis for $\ker(T)$

$\implies \lambda_1\vec{w}_1 + \ldots + \lambda_m\vec{w}_m + (-\mu_1)\vec{v}_1 + \ldots + (-\mu_k)\vec{v}_k = \vec{0}_V$

$\implies \lambda_1 = \ldots = \lambda_m = \mu_1 = \ldots = \mu_k = 0$ since $\{\vec{v}_1, \ldots, \vec{v}_k, \vec{w}_1, \ldots, \vec{w}_m\}$ is lin ind (it is a basis)

So $\{T(\vec{w}_1), \ldots, T(\vec{w}_m)\}$ is lin ind, which means it is a basis for the image.

Thus $\dim(\Im(T)) = m, \dim(V) - \dim(\ker(T)) = k$

Technique to Find Bases for Image(T)

1. find a basis for $\ker(T)$
2. extend it to a basis for V
3. apply T to what's been added to get a basis for $\Im(T)$

E.g. Let $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ and $T : M_{2,2} \to M_{2,2}$ be a lin transf defined by $T(X) = XA - AX \; \forall \; X \in M_{2,2}$

Find a basis for $\Im(T)$ using the above technique.

1. Find a basis for $\ker(T)$

$$X = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \ker(T) \iff \begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \iff \begin{bmatrix} b-c & a-d \\ d-a & c-b \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\implies \begin{cases} b-c=0 \\ a-d=0 \\ d-a=0 \\ c-b=0 \end{cases} \implies \left[\begin{array}{cccc|c} 0 & 1 & -1 & 0 & 0 \\ 1 & 0 & 0 & -1 & 0 \\ -1 & 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 \end{array}\right] \implies \left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right] \implies \begin{cases} c=s \\ d=t \\ a=t \\ b=s \end{cases}$$

$$\implies \ker(T) = \left\{\begin{bmatrix} t & s \\ s & t \end{bmatrix} : s,t \in \mathbb{R}\right\} = \left\{s\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + t\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} : s,t \in \mathbb{R}\right\} = sp\left(\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}\right)$$

2. Extend it to a basis for V

   Extend by adding elements of the standard basis for $M_{2,2}$

$$\left\{\begin{bmatrix}0 & 1\\1 & 0\end{bmatrix},\begin{bmatrix}1 & 0\\0 & 1\end{bmatrix},\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix},\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}\right\}$$ (This is not the only way)

3. Apply T to what's been added to get a basis for $\mathfrak{I}(T)$

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

E.g. Let $T : P_4 \to P_4$ be the linear transformation defined by $T(p) = xp'(x) - p(x)$. Find a basis for the image of T and a basis for the kernel of T.

Method 1: finding a matrix rep rel. to ordered bases and observing the null/column space

Let $B = B' = \left(1, x, x^2, x^3, x^4\right)$ be an ordered basis for P_4.

$T(1) = x(0) - 1 = -1$

$T(x) = x(1) - x = 0$

$T(x^2) = x(2x) - x^2 = x^2$

$T(x^3) = x(3x^2) - x^3 = 2x^3$

$T(x^4) = x(4x^3) - x^4 = 3x^4$

Then the matrix A for T relative to B, B' is

$$A = \begin{bmatrix} | & | & | & | & | \\ T(1)_{B'} & T(x)_{B'} & T\left(x^2\right)_{B'} & T\left(x^3\right)_{B'} & T\left(x^4\right)_{B'} \\ | & | & | & | & | \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 3 \end{bmatrix}$$

A basis for the **null space** of A is $\{[0, 1, 0, 0, 0]\}$, thus, a basis for $\ker(T)$ is $\{x\}$.

Note: dim(ker(T)) = dim(null space(A))

A basis for the **column space** of A is $\{[-1, 0, 0, 0, 0], [0, 0, 1, 0, 0], [0, 0, 0, 2, 0], [0, 0, 0, 0, 3]\}$, thus, a basis for $\operatorname{im}(T)$ is $\left\{-1, x^2, 2x^3, 3x^4\right\}$

Method 2: finding a basis for the kernel of T and extending it to get a basis for the image of T

$\ker(T) = T^{-1}[\{\vec{0}_W\}]$ so to find the kernel, we need to find p such that $T(p) = 0$

> Using $p(x) = ax^4 + bx^3 + cx^2 + dx + e$,
> $T(p) = xp'(x) - p(x) = 0 \implies x\left(4ax^3 + 3bx^2 + 2cx + d\right) - \left(ax^4 + bx^3 + cx^2 + dx + e\right) = 0$
>
> Expand and simplify to obtain $3ax^4 + 2bx^3 + cx^2 - e = 0 \implies a = b = c = e = 0,\ d = k, k \in \mathbb{R}$
>
> (Or don't expand and obtain the corresponding system $[4a = a, 3b = b, 2c = c, d = d, 0 = e]$)
>
> Thus, $p = kx \implies \ker(T) = \{kx : k \in \mathbb{R}\} = \operatorname{sp}(x) \implies \{x\}$ is a basis for $\ker(T)$.
>
> Now extend this to a basis for P_4 — add $1, x^2, x^3, x^4$ to the set.
>
> This implies that a basis for $\operatorname{im}(T)$ is $\left\{T(1), T\left(x^2\right), T\left(x^3\right), T\left(x^4\right)\right\} = \left\{-1, x^2, 2x^3, 3x^4\right\}$

Def. Row Space

Motivation Suppose U is a subspace of $\mathbb{R}^n$ with a basis of $\left\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\right\}$.
What operations can be performed on this basis while preserving its span and linear independence?

- Swap 2 vectors
- Multiply a vector by a non-zero scalar
- Add a scalar multiple of 1 vector to another

If $\left\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_m\right\}$ are the rows of a matrix, these are simply elementary row operations!

> If A is a m x n matrix, then the **row space** of A is the subspace of $\mathbb{R}^n$ spanned by its rows, denoted **row(A)**

Facts

- Elementary row operations do not change the row space
- In general, the rows of A may be linearly dependent

> **Thm** The non-zero rows of any REF is a basis for row(A)

E.g. Find a basis for $U = sp\left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}1\\-1\end{bmatrix}, \begin{bmatrix}2\\-2\end{bmatrix}\right\}$ using row space

$$A = \begin{bmatrix}1 & 0\\1 & -1\\2 & -2\end{bmatrix} \implies row(A) = sp\left\{\begin{bmatrix}1\\0\end{bmatrix}^T, \begin{bmatrix}1\\-1\end{bmatrix}^T, \begin{bmatrix}2\\-2\end{bmatrix}^T\right\}$$

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

By the theorem, row(A) has a basis of $\{[1,0],[0,1]\} \implies U$ has a basis of $\left\{\begin{bmatrix}1\\0\end{bmatrix}, \begin{bmatrix}0\\1\end{bmatrix}\right\}$

E.g. Find dim(row(A)) if $A = \begin{bmatrix}1 & -1 & 1 & 3 & 2\\2 & -1 & 1 & 5 & 1\\3 & -1 & 1 & 7 & 0\\0 & 1 & -1 & -1 & -3\end{bmatrix}$

A row reduces to $\begin{bmatrix}1 & -1 & 1 & 3 & 2\\0 & 1 & -1 & -1 & -3\\0 & 0 & 0 & 0 & 0\\0 & 0 & 0 & 0 & 0\end{bmatrix} \implies \{[1,-1,1,3,2],[0,1,-1,-1,-3]\}$ is a basis

So dim(row(A)) = 2

Def. Column Space

> If A is a m x n matrix, then the **column space** of A is the subspace of $\mathbb{R}^m$ spanned by its columns, denoted **col(A)**;
>
> col(A) = row(A^T), so a basis of $col(A)$ can be computed by reducing A^T to REF

Note $\quad \Im(A) = col(A)$

E.g. Find a basis for col(A) if $A = \begin{bmatrix}1 & 2 & -1 & -2 & 0\\2 & 4 & -1 & 1 & 0\\3 & 6 & -1 & 4 & 1\\0 & 0 & 1 & 5 & 0\end{bmatrix}$

A row-reduces to $\begin{bmatrix}1 & 2 & 0 & 3 & 0\\0 & 0 & 1 & 5 & 0\\0 & 0 & 0 & 0 & 1\\0 & 0 & 0 & 0 & 0\end{bmatrix}$

Since (col 2) = 2(col 1), and (col 4) = 3(col 1) + 5(col 3),

a basis for col(A) is the other 3 columns: $\left\{\begin{bmatrix}1\\2\\3\\0\end{bmatrix}, \begin{bmatrix}-1\\-1\\-1\\1\end{bmatrix}, \begin{bmatrix}0\\0\\1\\0\end{bmatrix}\right\}$

Note We can't use columns of the REF for a basis, since row operations mess up the vectors!

Strategy for finding basis of row(A) and col(A)

1. reduce A to REF = R
2. A basis for row(A) is formed by rows of **R** containing leading 1's
3. A basis for col(A) is formed by cols of **A** *corresponding to cols of R* with leading 1's

Def. Rank

> The rank of a m x n matrix A is the **number of leading 1's** in its **REF**, and $rank(A) \leq \min(m, n)$
>
> $rank(A) = \dim(row(A)) = \dim(col(A)) = \dim(\Im(A))$

Both of the below have $rank(A) = 2 = \dim(row(A)) = \dim(col(A))$

(1) $m = 2, n = 3 : \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$

$$rank(A) = m \iff col(A) = \mathbb{R}^m \iff \text{columns of } A \text{ span } \mathbb{R}^m\text{; rows of } A \text{ are independent in } \mathbb{R}^n$$

(2) $m = 3, n = 2 : \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$

$$rank(A) = n \iff row(A) = \mathbb{R}^n \iff \text{rows of } A \text{ span } \mathbb{R}^n\text{; columns of } A \text{ are independent in } \mathbb{R}^m$$

E.g. Can a 5 x 6 matrix have linearly independent rows/cols?

The matrix would reduce to $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$

Cols cannot be linearly independent since each column has 5 entries (so $\mathbb{R}^5$) but there are 6 columns. Only 5 are needed to span $\mathbb{R}^5$. The 6th column can be written as a lin combo of the first 5. The 6th variable can be anything.

Rows can be linearly dependent or independent.

Note The converse: if # rows > # columns, the rows must be lin dep, the cols may be lin dep.

E.g. Let A be m x n with $rank(A) = m$, prove $m \leq n$

$rank(A) \leq n$ by definition, and $rank(A) = m$, so $m \leq n$

E.g. Let A be a 5 x 9 matrix, prove $\dim(null(A)) \geq 4$

$rank(A) \leq 5$ by definition, and $rank(A) = n - \dim(null(A)) \implies 9 - \dim(null(A)) \leq 5 \implies 4 \leq \dim(null(A))$

Note rank-nullity theorem was used here: for $A_{m,n}$, $rank(A) + \text{nullity}(A) = n$

W4 (remainder of 3.4)

Thm 3.7 (Lin transf preserves subspaces)

Let V and V' be vector spaces and let $T : V \to V'$ be a linear transformation.

Proof (If S is a subspace of V, then $T[S]$ is a subspace of V')

Subspace test:

1. WTS $T[S]$ contains the zero vector

   Since S is a subspace, $\vec{0}_V \in S \implies T(\vec{0}_V) \in T[S] \overset{\text{linearity of T}}{\implies} \vec{0}_{V'} \in T[S]$

2. WTS $T(\vec{s}_1) + T(\vec{s}_2) \in T[S]$, where $T(\vec{s}_1), T(\vec{s}_2) \in T[S]$ and so $\vec{s}_1, \vec{s}_2 \in S$

   Since S is a subspace, $\vec{s}_1 + \vec{s}_2 \in S \implies T(\vec{s}_1) + T(\vec{s}_2) \in T[S] \overset{\text{linearity of T}}{\implies} T(\vec{s}_1 + \vec{s}_2) \in T[S]$

3. WTS $rT(\vec{s}) \in T[S]$, where $T(\vec{s}) \in T[S]$ and so $\vec{s} \in S$

   Since S is a subspace, $r\vec{s} \in S \implies T(r\vec{s}) \in T[S] \overset{\text{linearity of T}}{\implies} rT(\vec{s}) \in T[S]$

Proof (If S' is a subspace of V', then $T^{-1}[S']$ is a subspace of V)

1. WTS $T^{-1}[S']$ contains the zero vector

   Since S' is a subspace, $\vec{0}_{V'} \in S' \implies T^{-1}(\vec{0}_{V'}) \in T^{-1}[S'] \overset{\text{linearity of } T^{-1}}{\implies} \vec{0}_V \in T^{-1}[S']$

2. WTS $\vec{v}_1 + \vec{v}_2 \in T^{-1}[S']$, where $\vec{v}_1, \vec{v}_2 \in T^{-1}[S']$, and so $T(\vec{v}_1), T(\vec{v}_2) \in S'$

   Since S' is a subspace,
   $T(\vec{v}_1) + T(\vec{v}_2) \in S' \implies T^{-1}(T(\vec{v}_1) + T(\vec{v}_2)) \in T^{-1}[S'] \overset{\text{linearity of T}}{\implies} \implies \vec{v}_1 + \vec{v}_2 \in T^{-1}[S']$

3. WTS $r\vec{v} \in T^{-1}[S']$, where $\vec{v} \in T^{-1}[S']$ and so $T(\vec{v}) \in S'$

   Since S' is a subspace, $rT(\vec{v}) \in S' \implies T^{-1}(rT(\vec{v})) \in T^{-1}[S'] \overset{\text{linearity of T}}{\implies} r\vec{v} \in T^{-1}[S']$

Def. Onto & 1-1

Let V and W be vector spaces and $T: V \to W$ is a linear transformation

Onto (surjective)	**1-1 (injective)**
for any $\vec{w} \in W, \exists\, \vec{v} \in V$ such that $T(\vec{v}) = \vec{w}$	for any $\vec{v}_1, \vec{v}_2 \in V, T(\vec{v}_1) = T(\vec{v}_2) \implies \vec{v}_1 = \vec{v}_2$ or $\vec{v}_1 \neq \vec{v}_2 \implies T(\vec{v}_1) \neq T(\vec{v}_2)$
all elements in W have been mapped to **at least** once	all elements in W have been mapped to **at most** once
$\forall\, \vec{w} \in W, T(\vec{v}) = \vec{w}$ has **at least** 1 solution	$\forall\, \vec{w} \in W, T(\vec{v}) = \vec{w}$ has **at most** 1 solution
$\forall\, \vec{w} \in W, A\vec{v} = \vec{w}$ is consistent	$\forall\, \vec{w} \in W, A\vec{v} = \vec{w}$ has a unique sol'n or is inconsistent
A has a pivot in every **row**	A has a pivot in every **column**
The columns of A **span W**	The columns of A are **lin ind** i.e. $A\vec{x} = \vec{0}$ has only the trivial solution i.e. $\ker(T) = \{\vec{0}_V\}$
T preserves **spanning sets**: Suppose $V = sp\left\{\vec{v}_1, \ldots, \vec{v}_k\right\}$ then $W = sp\left\{T(\vec{v}_1), \ldots, T(\vec{v}_k)\right\}$	T preserves **lin ind**: Suppose $\left\{\vec{v}_1, \ldots, \vec{v}_k\right\}$ is a lin ind subset of V, then $\left\{T(\vec{v}_1), \ldots, T(\vec{v}_k)\right\}$ is a lin ind subset of W.
dim(range of T) = dim(**W**)	dim(range of T) = dim(**V**)

Note If T is both onto and 1-1, then it is bijective (and it must be an isomorphism, so it's invertible).

For bijective $T: V \to W, \dim(V) = \dim(W), T(\vec{v}) = A\vec{v} = \vec{w}$ has a unique solution (i.e. A has a pivot in every row and column)

Thm (Extra TB notes)

Let $T(\vec{p}) = \vec{b}$ for some particular vector in V.

Then the solution set of $T(\vec{x}) = \vec{b}$ is the set $\{\vec{p} + \vec{h} | \vec{h} \in \ker(T)\}$ (h is the solution to the homogeneous transf eq)

This means:

- If $\ker(T) = \{\vec{0}_V\}$, then $\vec{h} = \vec{0} \implies \vec{p} + \vec{h} = \vec{p} \implies T(\vec{x}) = \vec{b}$ has at most 1 solution which is $\vec{p}$, so T is 1-1
- If T is 1-1, then $T(\vec{x}) = \vec{0}_W$ has only the trivial solution by definition, so $\ker(T) = \{\vec{0}_V\}$

So 1-1 $\iff \ker(T) = \vec{0}_V$ follows from this.

E.g. Show that the solution of $y'' - 4y = x^2$ is the set $\{p + h | h \in \ker(T)\}$

(So $T(y) = y'' - 4y$, and $\vec{b} = x^2$)

- Want to find $\vec{p}$ such that $T(\vec{p}) = \vec{b}$, i.e. find y such that $y'' - 4y = x^2$

  By inspection, try $y = \frac{-x^2}{4}$. Then we have $-\frac{1}{2} - 4y = -4y \implies -\frac{1}{2} = 0$ which can't be true

  So kill off the $-\frac{1}{2}$ with $y = \frac{-x^2 - \frac{1}{2}}{4}$, then we have $-\frac{1}{2} - 4(\frac{-x^2 - \frac{1}{2}}{4}) = x^2$ which holds

  So we've found the solution: $p(x) = y = \frac{-x^2 - \frac{1}{2}}{4}$

- The homogeneous transformation eq $T(\vec{x}) = \vec{0}$ in this case is $y'' - 4y = 0$

  Its general solution is $h(x) = c_1 e^{2x} + c_2 e^{-2x} \iff \{e^{2x}, e^{-2x}\}$ is a basis for $\ker(T) \iff h \in \ker(T)$

- The general solution for $y'' - 4y = x^2$ is thus $h(x) + p(x) = c_1 e^{2x} + c_2 e^{-2x} + (\frac{-x^2 - \frac{1}{2}}{4})$

Thm (Compositions of lin transf is a lin transf)

Let $T : U \to V$ and $S : V \to W$ be linear transformations (U, V, W are vector spaces).

Then $S \circ T : U \to W$ is also a linear transformation

Proof Let $\vec{x}, \vec{y} \in V, a, b \in \mathbb{R}$

WTS $S \circ T(ax + by) = aS \circ T(x) + bS \circ T(y)$

$S \circ T(ax + by) = S(T(ax + by)) = S(aT(x) + bT(y)) = aS(T(x)) + bS(T(y)) = aS \circ T(x) + bS \circ T(y)$

Def. Invertible Transformation

A linear transformation $T : V \to W$ is invertible if there is a linear transformation $T^{-1} : W \to V$ such that $T^{-1} \circ T = I_V, T \circ T^{-1} = I_W$, where I_W is the identity transformation mapping vectors from W to W as $I_W(\vec{x}) = \vec{x}$

T^{-1} is the inverse for T (using "the" because it is unique).

Fact T is invertible $\iff$ T is an isomorphism

Thm (T is invertible if A is invertible)

Suppose that $B = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$ is a basis for the vector space V, $C = \{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ is a basis for the vector space W, and $T : V \to W$ is a linear transformation. Show that T is an invertible transformation iff the matrix representation $[T]_B^C$ (or denoted by $R_{B,C}$) of T wr.t. the bases B and C is an invertible matrix.

> Suppose T is invertible, then $\exists\, T^{-1} : W \to V$ such that $T^{-1} \circ T = I_V, T \circ T^{-1} = I_W$.
>
> Let $\left[T^{-1}\right]_C^B$ be the matrix representation of T^{-1} with respect to the bases C and B.
>
> Since $[T^{-1}]_C^B [T]_B^C = I_n$ and $[T]_B^C [T^{-1}]_C^B = I_n$, $[T]_B^C$ is an invertible matrix.
>
> Assuming $[T]_B^C$ is an invertible matrix, there must be a matrix A such that $A[T]_B^C = I_n$, $[T]_B^C A = I_n$
>
> Let A be the matrix rep of the linear transf. $S : W \to V$, i.e. $[S]_C^B = A$.
>
> Then $S \circ T = I_V$, and $T \circ S = I_W$ (reads identity transf. on W), so S is an inverse for T, and T is invertible.

Thm 3.8 (T is invertible iff 1-1 and onto)

Proof Suppose $T : V \to V'$

($\Rightarrow$): invertibility implies 1-1 & onto

Since T is invertible, $\exists\, \vec{v} \in V$ such that $T^{-1}(\vec{v}') = \vec{v}$ for all $v' \in V'$. This means T must be onto V'.

$v' = (T \circ T^{-1})(\vec{v}') = T(T^{-1}(\vec{v}')) = T(\vec{v})$, so $\vec{v}_1 \neq \vec{v}_2 \implies T(\vec{v}_1) \neq T(\vec{v}_2)$. Hence T must be 1-1.

($\Leftarrow$): 1-1 & onto implies invertibility

Since T is onto V', $\exists\, \vec{v} \in V$ such that $T^{-1}(\vec{v}') = \vec{v}$ for all $v' \in V'$. Since T is 1-1, this $\vec{v}$ must be unique.

Define $T^{-1} : V' \to V$ as $T^{-1}(\vec{v}') = \vec{v}$, where $\vec{v}$ is that unique vector satisfying $T(\vec{v}) = \vec{v}'$

Then $(T \circ T^{-1})(\vec{v}') = T(T^{-1}(\vec{v}')) = T(\vec{v}) = \vec{v} \implies T \circ T^{-1} = I_{V'}$

and $(T^{-1} \circ T)(\vec{v}) = T^{-1}(T(\vec{v}')) = T^{-1}(\vec{v}') = \vec{v} \implies T^{-1} \circ T = I_V$

Hence T is invertible.

Remains to show that T^{-1} is indeed a linear transformation:

1. Preserves addition

$$T^{-1}(v_1' + v_2') = T^{-1}(T(v_1) + T(v_2)) = T^{-1}(T(v_1 + v_2)) = v_1 + v_2 = T^{-1}(v_1) + T^{-1}(v_2)$$

2. Preserves scalar mult

$$T^{-1}(rv_1') = T^{-1}(rT(v_1)) = T^{-1}(T(rv_1)) = rv_1 = rT^{-1}(v_1')$$

Def. Isomorphism

A linear transformation $T : V \to W$ is an isomorphism if it is 1-1 and onto.

> **E.g.** Prove $M_{2,2} \cong P_3$
>
> Construct a linear transformation $T : M_{2,2} \to P_3$ that is 1-1 and onto:
>
> Consider $T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a + bx + cx^2 + dx^3$
>
> 1. Show linearity (T preserves addition and scalar mult)
>
> $$\begin{aligned} &T\left(r\begin{bmatrix} a & b \\ c & d \end{bmatrix} + s\begin{bmatrix} e & f \\ g & h \end{bmatrix}\right) \\ &= T\left(\begin{bmatrix} ra+se & rb+sf \\ rc+sg & rd+sh \end{bmatrix}\right) \\ &= (ra+se) + (rb+sf)x + (rc+sg)x^2 + (rd+sh)x^3 \\ &= r(a+bx+cx^2+dx^3) + s(e+fx+gx^2+hx^3) \\ &= rT\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) + sT\left(\begin{bmatrix} e & f \\ g & h \end{bmatrix}\right) \end{aligned}$$
>
> 2. Show 1-1 (by showing $\ker(T) = \vec{0}$)
>
> $$\begin{aligned} T(A) = 0 \iff & T\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = 0 + 0x + 0x^2 + 0x^3 \\ \iff & a + bx + cx^2 + dx^3 = 0 + 0x + 0x^2 + 0x^3 \\ \iff & a = b = c = d = 0 \\ \iff & \ker(T) = \left\{\begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}\right\} \end{aligned}$$
>
> 3. Show onto
>
> For any polynomial $p(x) = q + rx + sx^2 + tx^3 \in P_3$, we can find a matrix in $M_{2,2}$ that maps to it, namely
>
> $$A = \begin{bmatrix} q & r \\ s & t \end{bmatrix}$$

Thm (Iso proof shortcut if same dim)

Let V, W be finite dimensional vector spaces and $T : V \to W$ a linear transformation.

If $\dim V = \dim W$, then T is an isomorphism iff T is OR(onto, 1-1).

Prove using the rank nullity thm: $\dim(\Im(T)) + \dim(\ker(T)) = \dim(V)$

> same dim & 1-1 $\implies$ invertible
>
> If $\dim(W) = \dim(V) = n$, and T is 1-1, then $\ker(T) = \left\{\vec{0}\right\} \implies \dim(\ker(T)) = 0$
>
> By the rank nullity thm, $\dim(\Im(T)) + 0 = n$.
>
> And since $\Im(T)$ is a subspace of W which also has dimension n, we have that $\Im(T) = W$. Hence, T is onto.
>
> Since T is both 1-1 and onto, it must be invertible.
>
> same dim & onto $\implies$ invertible
>
> If $\dim(W) = \dim(V) = n$, and T is onto, then $\Im(T) = W \implies \dim(\Im(T)) = n$
>
> By the rank nullity thm, $n + \dim(\ker(T)) = n \implies \dim(\ker(T)) = 0 \implies \ker(T) = \left\{\vec{0}\right\} \implies T$ is 1-1.
>
> Since T is onto and 1-1, it must be invertible.

Thm 3.9 (Coordinatization is an isomorphism)

Let V be a finite dimensional vector space with ordered basis $B = (\vec{b}_1, \ldots, \vec{b}_n)$

Then $T : V \to \mathbb{R}^n$ defined by $T(\vec{v}) = \vec{v}_B$ is an isomorphism.

Proof

T is a lin transf since it preserves addition and scalar mult. as proved in W3: Coordinatization of Vector Spaces

T is an isomorphism:

1. T is 1-1 since the coordinate vector $\vec{v}_B$ uniquely determines $\vec{v}$
2. T is onto since the range of T is all of $\mathbb{R}^n$

Note The isomorphism of V with $\mathbb{R}^n$ is not unique, since there is a corresponding isomorphism for each choice of an ordered basis of V.

W5 (7.1, 7.2)

Def. Standard Matrix Rep of Lin Transf

Recall concepts for $T:\mathbb{R}^n \to \mathbb{R}^m$

- A transformation $T:\mathbb{R}^n \to \mathbb{R}^m$ is linear iff it has the form $T(\vec{x}) = A\vec{x}$.
- The matrix $A_{m\times n} = [T(\vec{e}_1)|\ldots|T(\vec{e}_n)] = T\left(I_{m\times n}\right)$ is the **standard matrix representation of** T

  The action of T on unit vectors $\vec{e}_j$ determines the structure of A
- Properties of T (onto/1-1) are deeply related to existence & uniqueness of solutions to $A\vec{x} = \vec{b}$ or $T(\vec{x}) = \vec{b}$

> **Q** Find the standard matrix rep of $T([x_1, x_2, x_3]) = [x_1 - x_2 + 3x_3, \quad x_1 + x_2 + x_3, \quad x_1]$
>
> **A** For $T:\mathbb{R}^n \to \mathbb{R}^m$, the standard matrix rep is given by $A_{m\times n} = [T(\vec{e}_1)|\ldots|T(\vec{e}_n)] = T\left(I_{m\times n}\right)$
>
> Since $T:\mathbb{R}^3 \to \mathbb{R}^3$, $A = \begin{bmatrix} 1 & -1 & 3 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}$
>
> Shortcut for finding A: instead of calculating each col vector by transforming $\vec{e}_j$,
> just write out the coefficients of the transformed input as row vectors

Now, how do we find the matrix rep of $T:V\to W$, where V, W are non-Euclidean vector spaces?

> **E.g.** Define $\tau: M_{2,2} \to P_3$ by $\tau\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = a + (a+c)x + (b-c)x^2 + dx^3$
>
> To represent τ with a matrix, we start with an ordered basis for each of $M_{2,2}$ and P_3
>
> For $M_{2,2} : B = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right)$
>
> For $P_3 : C = (1, x, x^2, x^3)$
>
> Since it is easier to work over $\mathbb{R}^n$, we use the following coord. transformations:

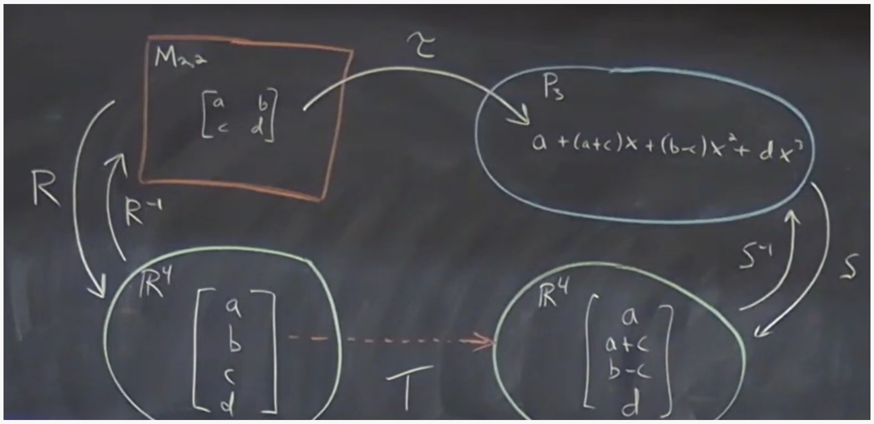
$R : M_{2,2} \to \mathbb{R}^4$ defined by $R\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$

$S : P_3 \to \mathbb{R}^4$ defined by $S(a + bx + cx^2 + dx^3) = \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}$

$T : \mathbb{R}^4 \to \mathbb{R}^4$ defined by $T\left(\begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix}\right) = \begin{bmatrix} a \\ a+c \\ b-c \\ d \end{bmatrix}$

This way, we can write T as a composition of linear transformations: $T = S \circ \tau \circ R^{-1}$

(since $\mathbb{R}^4 \to \mathbb{R}^4 \equiv \mathbb{R}^4 \to M_{2,2} \to P_3 \to \mathbb{R}^4$, illustrated below)



Since T is a composition of lin transf, it must be linear itself, so find the standard matrix for T :

$A = T\left(\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$ is said to be the matrix of τ with respect to B and C.

Also, A is invertible (it has a non-zero det) $\implies$ T is invertible $\implies$ T is an isomorphism.

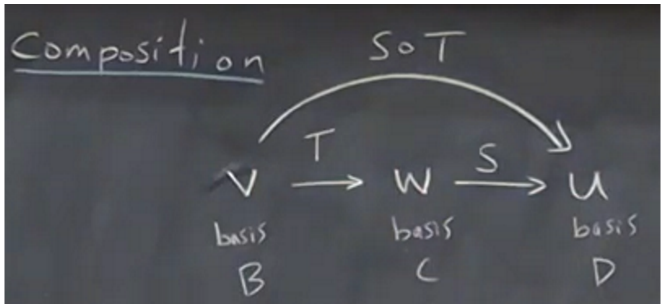
$\tau = S^{-1} \circ T \circ R \implies \tau$ is also an isomorphism.

Matrix Rep Notation

Matrix rep of T relative to B and B' can be denoted as $R_{B,B'}$, $R_{B,B'}(T)$, or $[T]_B^{B'}$

The following denote the matrix rep of **T inverse**

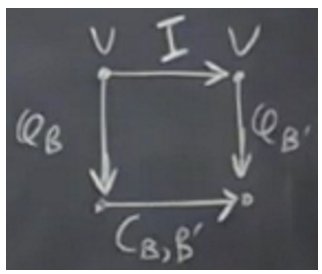
- $R_{C,B}(T^{-1}) = [R_{B,C}(T)]^{-1}$
- $[T^{-1}]_C^B = ([T]_B^C)^{-1}$



The following denote **S composed with T** illustrated above

- (matrix for $S \circ T$) = (matrix for S)·(matrix for T)
- $R_{B,D}(S \circ T) = R_{C,D}(S) \cdot R_{B,C}(T)$ $\quad \leftarrow$ outside must match
- $[S \circ T]_B^D = [S]_C^D \cdot [T]_B^C$ $\quad \leftarrow$ bottom left and top right must match

The following denote the **change of basis** matrix $C_{B,B'}$ illustrated below



- $R_{B,B'}(I)$ where I is the identity transformation
- $[I]_B^{B'}$

Def. Matrix of T Rel. to Basis B

Let V be a finite dimensional vector space with ordered bases $B = (\vec{b}_1, \ldots, \vec{b}_n)$

Let $T : V \to V$ be a lin transf, then $R_B = [T(\vec{b}_1)_B \,|\ldots|\, T(\vec{b}_n)_B]$ is the matrix of T relative to basis B

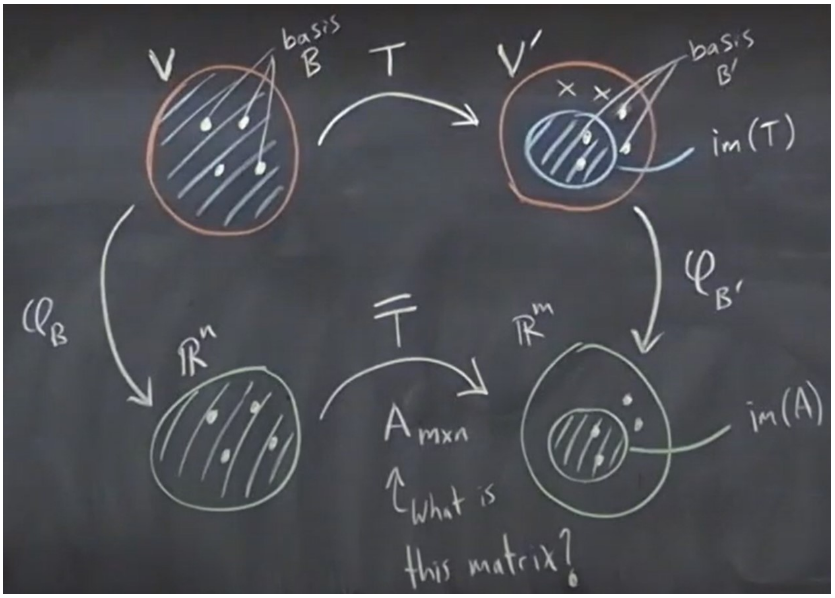
$T(\vec{v})_B = R_B \vec{v}_B$

Def. Matrix of T Rel. to Bases B and B'

Let V and V' be finite dimensional vector spaces with ordered bases $B = (\vec{b}_1, \ldots, \vec{b}_n)$ and $B' = (\vec{b}'_1, \ldots, \vec{b}'_m)$.

Let $T : V \to V'$ be a lin transf and $\bar{T} : \mathbb{R}^n \to \mathbb{R}^m$ be the lin transf such that $\bar{T}(\vec{v}_B) = T(\vec{v})_{B'}$.

I.e. $\bar{T}$ transforms $\vec{v}_B$ (coord. of $\vec{v}$ rel. to B) into the coord. of transformed $\vec{v}$ rel. to B'



Then the matrix rep of T relative to B and B' is the m x n matrix $A = \begin{bmatrix} T(\vec{b}_1)_{B'} & \dots & T(\vec{b}_n)_{B'} \end{bmatrix}$

I.e. A is the coordinates of the transformed basis vectors (of B) relative to B'

And so $T(\vec{v})_{B'} = A(\vec{v}_B)\ \forall\, \vec{v} \in V$

I.e. the coord. of any transformed vector can be obtained by:
multiplying A (coord. of transformed basis vectors rel to B') with its coord. relative to B

Note A^{-1} is the matrix rep of T^{-1} relative to B', B.

Prove: j-th column of A = $T(\vec{b}_j)_{B'}$

Since the coord vector of $\vec{v} = a_1\vec{b}_1 + \ldots + a_n\vec{b}_n$ rel. to B is given by $\vec{v}_B = \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$

the coord vector of $\vec{b}_j = 0\vec{b}_1 + \ldots + 1\vec{b}_j + \ldots + 0\vec{b}_n$ rel. to B is given by $[\vec{b}_j]_B = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \vec{e}_j \quad$ so $[\vec{b}_j]_B = \vec{e}_j$

$$T(\vec{v})_{B'} = A(\vec{v}_B) \implies T(\vec{b}_j)_{B'} = A[\vec{b}_j]_B = A\vec{e}_j = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix} = \text{j-th column of A}$$

E.g. Find $R_{B,B'}$ where $B = B' = (\sin x \cos x, \sin^2 x, \cos^2 x)$, and use it to compute $f'(x)$, given

$$f(x) = 3\sin x\cos x - 5\sin^2 x + 7\cos^2 x$$

$$\begin{aligned} R_{B,B'} &= \left[T(\sin x\cos x)_{B'} \quad T(\sin^2 x)_{B'} \quad T(\cos^2 x)_{B'}\right] \\ &= \left[(-\sin^2 x + \cos^2 x)_{B'} \quad (2\sin x\cos x)_{B'} \quad (-2\sin x\cos x)_{B'}\right] \\ &= \begin{bmatrix} 0 & 2 & -2 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \end{aligned}$$

$$T(\vec{v})_{B'} = R_{B,B'}\vec{v}_B$$

$$[f'(x)]_{B'} = R_{B,B'}[f(x)]_B = \begin{bmatrix} 0 & 2 & -2 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ -5 \\ 7 \end{bmatrix} = \begin{bmatrix} -24 \\ -3 \\ 3 \end{bmatrix}$$

$$f'(x) = -24\sin x\cos x - 3\sin^2 x + 3\cos^2 x$$

E.g. Consider $T : P_3 \to M_{2,2}$ defined by $T(ax^3 + bx^2 + cx + d) = \begin{bmatrix} 2a - 2b & -2b - 3c \\ 2a - 2b + d & -3c + d \end{bmatrix}$

Let $B = (x^3, x^2, x, 1)$ be an ordered basis for P_3,

and $B' = \left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}\right)$ an ordered basis for $M_{2,2}$

Find $R_{B,B'}(T), R_{B',B}(T^{-1}), T^{-1}\left(\begin{bmatrix} p & q \\ r & s \end{bmatrix}\right)$

$$\begin{aligned} R_{B,B'}(T) &= [T(x^3)_{B'} \quad T(x^2)_{B'} \quad T(x)_{B'} \quad T(1)_{B'}] \\ &= \left[\begin{bmatrix} 2 & 0 \\ 2 & 0 \end{bmatrix}_{B'} \quad \begin{bmatrix} -2 & -2 \\ -2 & 0 \end{bmatrix}_{B'} \quad \begin{bmatrix} 0 & -3 \\ 0 & -3 \end{bmatrix}_{B'} \quad \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}_{B'}\right] \\ &= \begin{bmatrix} 2 & -2 & 0 & 0 \\ 0 & -2 & -3 & 0 \\ 2 & -2 & 0 & 1 \\ 0 & 0 & -3 & 1 \end{bmatrix} \end{aligned}$$

$$R_{B',B}(T^{-1}) = \begin{bmatrix} 2 & -2 & 0 & 0 \\ 0 & -2 & -3 & 0 \\ 2 & -2 & 0 & 1 \\ 0 & 0 & -3 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & -1/2 & -1/2 & 1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \\ -1/3 & 0 & 1/3 & -1/3 \\ -1 & 0 & 1 & 0 \end{bmatrix}$$

$$T^{-1}(v)_B = R_{B',B}(T^{-1})(v_{B'})$$

$$\implies T^{-1}\left(\begin{bmatrix} p & q \\ r & s \end{bmatrix}\right)_B = R_{B',B}(T^{-1})\begin{bmatrix} p \\ q \\ r \\ s \end{bmatrix} = \begin{bmatrix} p - \frac{1}{2}q - \frac{1}{2}r - \frac{1}{2}s \\ \frac{1}{2}p - \frac{1}{2}q - \frac{1}{2}r + \frac{1}{2}s \\ -\frac{1}{3}p + \frac{1}{3}r - \frac{1}{3}s \\ -p + r \end{bmatrix}$$

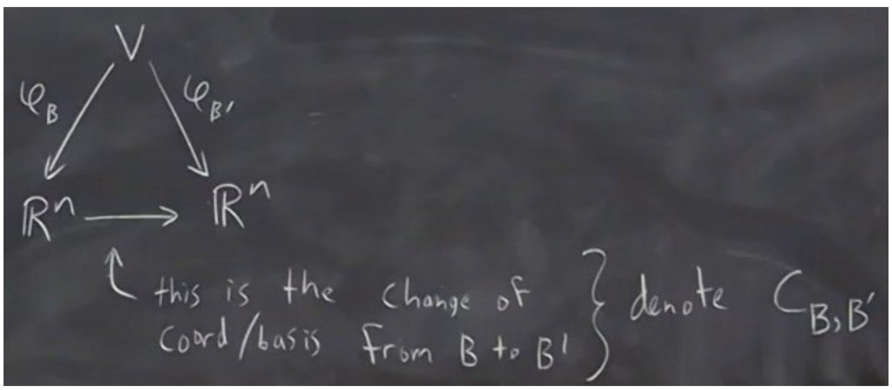
$$\implies T^{-1}\left(\begin{bmatrix} p & q \\ r & s \end{bmatrix}\right) = (p - \frac{1}{2}q - \frac{1}{2}r - \frac{1}{2}s)x^3 + (\frac{1}{2}p - \frac{1}{2}q - \frac{1}{2}r + \frac{1}{2}s)x^2 + (-\frac{1}{3}p + \frac{1}{3}r - \frac{1}{3}s)x + (-p + r)$$

Def. Change of basis matrix

If we have a vector space **V with basis B**, we can **find a new basis B'** using the matrix $C_{B,B'} = \begin{bmatrix} [\vec{b}_1]_{B'} & \dots & [\vec{b}_n]_{B'} \end{bmatrix}$ and the **change of basis formula** $\vec{v}_{B'} = C_{B,B'} \cdot \vec{v}_B$

I.e. (coord vec of $\vec{v}$ in new basis) = (change of basis matrix from B to B') · (coord vec of $\vec{v}$ in old basis)

Note Recall $T(\vec{v})_{B'} = A \cdot \vec{v}_B$ from thm 3.10. Note its similarity to the change of basis formula.



Note $C_{B',B} = C_{B,B'}^{-1}$

E.g. Find the change of basis matrix for $B = (x^2 - 1, x^2 + 1, x^2 + 2x + 1), B' = (x^2, x, 1)$

$$C_{B,B'} = [(x^2-1)_{B'} \quad (x^2+1)_{B'} \quad (x^2+2x+1)_{B'}] = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 2 \\ -1 & 1 & 1 \end{bmatrix}$$

If $V = \mathbb{R}^n$, then row reduce the matrix **[new basis vectors | old basis vectors]** to find $C_{B,B'}$:

$$\begin{bmatrix} \vec{b}'_1 & \dots & \vec{b}'_n \mid \vec{b}_1 & \dots & \vec{b}_n \end{bmatrix} \sim [I \mid C_{B,B'}]$$

E.g. Let $V = \mathbb{R}^3$. Find $C_{B,B'}$ where $B = (\vec{e}_1, \vec{e}_2, \vec{e}_3)$ and $B' = ([1,1,0],[2,0,1],[1,-1,0])$.

Also find $\vec{v}_{B'}$ where $\vec{v} = \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}$

$$\left[\begin{array}{ccc|ccc} 1 & 2 & 1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \end{array}\right] \sim \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 1/2 & 1/2 & -1 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1/2 & -1/2 & -1 \end{array}\right]$$

Note: $C_{B,B}$ is essentially $[\vec{e}_i]_{B'}$ since for i = 1, $[1 \quad 0 \quad 0] = \frac{1}{2}[1 \quad 1 \quad 0] + 0\,[2 \quad 0 \quad 1] + \frac{1}{2}[1 \quad -1 \quad 0]$

Then $\vec{v}_{B'} = C_{B,B'}\vec{v}_B = \begin{bmatrix} 1/2 & 1/2 & -1 \\ 0 & 0 & 1 \\ 1/2 & -1/2 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} -3/2 \\ 4 \\ -4/3 \end{bmatrix}$

Motivation for next topic

Consider $T : P_2 \to P_2$ where $T(P) = p'$ is the derivative

If $B = (1, x, x^2)$, then differentiate to get $R_B = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$

If $B' = (x+1, x-1, 2x^2)$, then differentiate (and decompose):

$$\begin{cases} T(x+1) = 1 = 1/2(x+1) - 1/2(x-1) + 0(2x^2) \\ T(x-1) = 1 = above \\ T(2x^2) = 4x = 2(x+1) + 2(x-1) + 0(2x^2) \end{cases}$$

so $R_{B'} = \begin{bmatrix} 1/2 & 1/2 & 2 \\ -1/2 & -1/2 & 2 \\ 0 & 0 & 0 \end{bmatrix}$

$R_B, R_{B'}$ use different bases, but represent the same transformation T. Their similarities are :

- both not onto with rank 2
- both have det = 0, trace = 0
- both have the eigen values {0, 0, 0}

These similarities are not a coincidence. Similar matrices have similar properties.

Def. Similar Matrices

Let A, B be square matrices, then A is similar to B if there is an invertible C such that $B = C^{-1}AC$

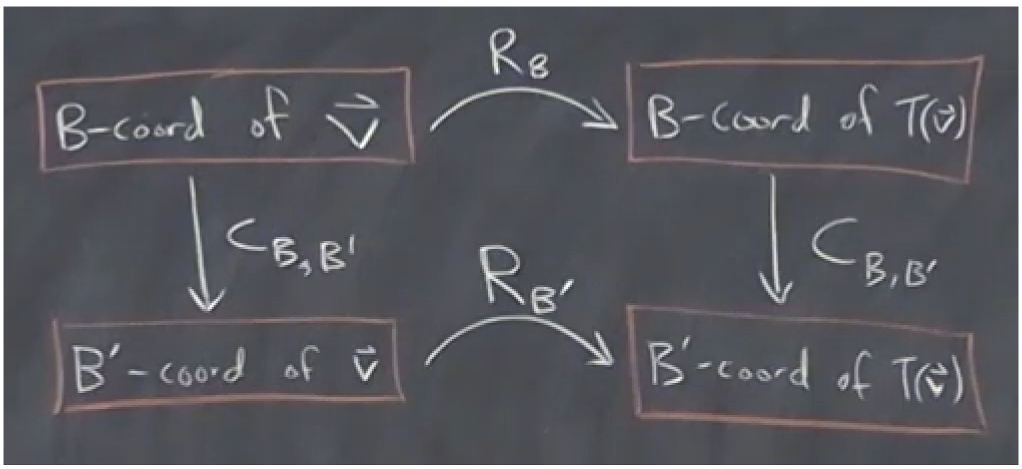
A and B would have the same rank, determinant, eigenvalues, trace, and change of basis matrix.

Note A and B are similar iff they are the matrix reps of the **same** linear transformation relative to different bases. (thm 7.1)

Thm 7.1 (Basis Change in A Vector Space)

If we have a finite dimensional vector space V with bases B, B', and a linear transformation $T : V \to V$, then the matrix reps for T, R_B and $R_{B'}$ must be similar, so $R_{B'} = C^{-1}R_BC$ where $C = C_{B',B}$

Rearranging, we have $C_{B,B'}R_B = R_{B'}C_{B,B'}$ which is illustrated below (right + down = down + right)



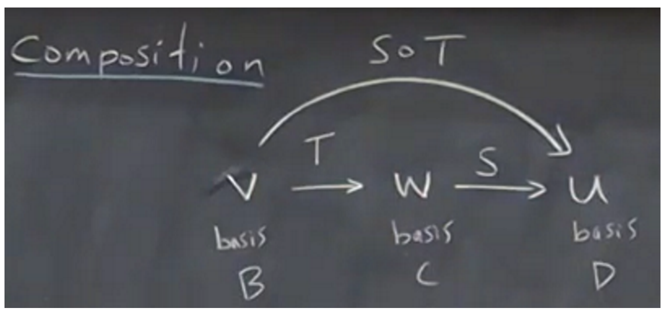
Since $C_{B,B'} = (C_{B',B})^{-1}$, we have $R_{B'} = C_{B,B'}R_B(C_{B,B'})^{-1} = (C_{B',B})^{-1}R_BC_{B',B}$

> **E.g.** $T : P_2 \to P_2, T(p(x)) = p(x-1)$
>
> Let $B = (x^2, x, 1), B' = (x, x+1, x^2-1)$
>
> Find $R_B, R_{B'}, C$ such that $R_{B'} = C^{-1}R_BC$
>
> $$C_{B',B} = \begin{bmatrix} [\vec{b}'_1]_B & [\vec{b}'_2]_B & [\vec{b}'_3]_B \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & -1 \end{bmatrix}$$

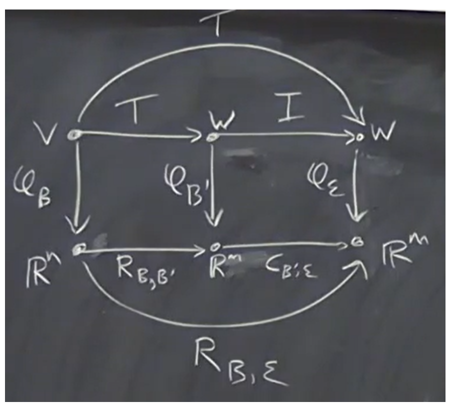
$$
\begin{aligned}
R_B &= [T(x^2)_B \quad T(x)_B \quad T(1)_B] \\
&= \left[(x-1)^2_B \quad (x-1)_B \quad 1_B\right] \qquad \leftarrow \text{shift x by 1} \\
&= [(x^2-2x+1)_B \quad (x-1)_B \quad 1_B] \\
&= \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & -1 & 1 \end{bmatrix}
\end{aligned}
$$

$$
\begin{aligned}
R_{B'} &= [T(x)_{B'} \quad T(x+1)_{B'} \quad T(x^2-1)_{B'}] \\
&= [(x-1)_{B'} \quad (x)_{B'} \quad ((x-1)^2-1)_{B'}] \qquad \leftarrow \text{shift x by 1} \\
&= [(x-1)_{B'} \quad (x)_{B'} \quad (x^2-2x)_{B'}] \\
&= \begin{bmatrix} 2 & 1 & -3 \\ -1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \qquad \leftarrow \begin{cases} x-1 & = 2(x) - 1(x+1) + 0(x^2-1) \\ x & = 1(x) + 0(x+1) + 0(x^2-1) \\ x^2-2x & = -3(x) + 1(x+1) + (x^2+1) \end{cases} \\
&= C^{-1} R_B C
\end{aligned}
$$

Shortcut to find Matrix Rep (Add basis E)



In the above, if we let $S = I$, and let W have bases B' and E. Then $U = W$ and we have the following



$\implies R_{B,E} = C_{B',E} R_{B,B'}$

Rearrange the above to get:

- $R_{B,B'}(T) = (C_{B',E})^{-1} R_{B,E}(T) = C_{E,B'} R_{B,E}(T)$
- or $[T]_B^{B'} = ([I]_{B'}^{E})^{-1} \cdot [T]_B^E = [I]_E^{B'} \cdot [T]_B^E$

E.g. Let V be P_3 polynomials, $T : P_3 \to P_3$ be a linear transf. defined by $T(p(x)) = x^2 p''(x)$

> Find R_B if $B = (1, 1+x, 1-x^2, 1+x^3)$

Let E be the standard basis $(1, x, x^2, x^3)$, then $R_{B,B} = (C_{B,E})^{-1} R_{B,E}$

$$C_{B,E} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Apply T to basis and decompose/express as lin combo:

$$\begin{aligned} T(1) &= 0 &&= 0(1) + 0(1+x) + 0(1-x^2) + 0(1+x^3) \\ T(1+x) &= 0 &&= 0(1) + 0(1+x) + 0(1-x^2) + 0(1+x^3) \\ T(1-x^2) &= -2x^2 &&= -2(1) + 0(1+x) + 2(1-x^2) + 0(1+x^3) \\ T(1+x^3) &= 6x^3 &&= -6(1) + 0(1+x) + 0(1-x^2) + 6(1+x^3) \end{aligned}$$

$$\text{so } R_{B,E} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}$$

Recall $[A|I] \to [I|A^{-1}] \quad \leftarrow \text{the row ops essentially multiply both sides by } A^{-1}$

$$[A|C] \to [I|A^{-1}C]$$

So to find $R_{B,B} = (C_{B,E})^{-1} R_{B,E}$, row reduce $[C_{B,E}|R_{B,E}] \to \left[I|(C_{B,E})^{-1} \cdot R_{B,E}\right] = [I|R_{B,B}]$

$$\left[\begin{array}{cccc|cccc} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 6 \end{array}\right] \to \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & 0 & 0 & 0 & -2 & -6 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 6 \end{array}\right]$$

We have found R_B (and to find $T(\vec{v})$, use $R_B \cdot \vec{v}_B = T(\vec{v})_B$)

Thm (Basis Change between 2 Vector Spaces)

For a linear $T : V \to W$, where V has ordered bases A, A', and W has ordered bases B, B', we can change coord from $A, B \to A', B'$ using

- $R_{A',B'} = C_{B,B'} R_{A,B} C_{A',A}$
- or $[T]_{A'}^{B'} = [I_W]_B^{B'} [T]_A^B [I_V]_{A'}^A$

Def. Eigen Values/Vectors

Given a transformation, an eigen vector is a non-zero vector that only gets stretched/scaled by a factor (not morphed)

$$
\begin{aligned}
A\vec{v} &= \lambda\vec{v} && \text{where } \lambda \text{ is the scaling factor, called eigen value} \\
A\vec{v} &= (\lambda I)\vec{v} && \text{Since LHS is matrix mult, turn RHS into matrix mult as well} \\
(A-\lambda I)\vec{v} &= \vec{0} && A-\lambda I \text{ transforms } \vec{v} \text{ to } \vec{0} \\
\det(A-\lambda I) &= 0 && \text{Since the determinant is the scaling factor for a volume under a transformation, it must} =
\end{aligned}
$$

To obtain the eigen value, solve the equation that arises from det = 0 (last line), called the characteristic polynomial

To obtain the eigen vectors, plug the eigen values back into the second to last line, and solve the system

Let $T: V \to V$ be linear, V be a vector space

- λ is an **eigen value** of T if $\exists$ a non-zero $\vec{v} \in V$ such that $T(\vec{v}) = \lambda\vec{v}$
- this $\vec{v}$ is the **eigen vector** of T corresponding to λ
  - direction of $\vec{v}$ is unchanged/invariant under T
- If λ is an eigen value of A corresponding to $\vec{v}$, then λ^k is an eigen value of A^k corresponding to $\vec{v}$

Fact If $T: \mathbb{R}^n \to \mathbb{R}^n$, then eigen values of T are eigen values of the standard matrix rep of T

E.g. Consider vector space P_2 and basis $E = (1, x, x^2)$.

Find the eigen values and eigen vectors of $T: P_2 \to P_2$ defined by $\begin{cases} T(1) = 3 + 2x + x^2 \\ T(x) = 2 \\ T(x^2) = 2x^2 \end{cases}$

Find matrix for T relative to E:

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

Find eigen values by performing cofactor expansion down the 3rd column:

$$\det(R_E - \lambda I) = \begin{vmatrix} 3-\lambda & 2 & 0 \\ 2 & -\lambda & 0 \\ 1 & 0 & 2-\lambda \end{vmatrix}$$
$$= (2-\lambda)\begin{vmatrix} 3-\lambda & 2 \\ 2 & -\lambda \end{vmatrix}$$
$$= (2-\lambda)((3-\lambda)(-\lambda)-4)$$
$$= (2-\lambda)(-3\lambda+\lambda^2-4)$$
$$= (2-\lambda)(\lambda+1)(\lambda-4)$$

Eigenvalues	Eigenvectors for P_2	Eigenvectors for T
$\lambda_1 = -1$	[-3, 6, 1]	$p_1(x) = -3+6x+x^2$
$\lambda_2 = 2$	[0, 0, 1]	$p_2(x) = x^2$
$\lambda_3 = 4$	[2, 1, 1]	$p_3(x) = 2+x+x^2$

Note If the basis consists of eigenvectors for T, i.e., $B = (-3+6x+x^2, x^2, 2+x+x^2)$, then the matrix rep of T relative to B is $R_B = \lambda I = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$

Def. Eigen Space

If λ is an eigen value of A, then the **eigen space** of A , denoted E_λ, is the set of eigen vectors corresponding to λ plus the zero vector (which isn't categorized as an eigenvector, even though it has the properties of one).

Note $E_\lambda = \left\{\vec{x} : (A-\lambda I)\vec{x} = \vec{0}\right\} = null(A-\lambda I) = \ker(A-\lambda I)$ so it is a subspace of n-space (either $\mathbb{R}^n$ or $\mathbb{C}^n$).

Def. Alg & Geo Multiplicities

Algebraic multiplicity of λ = its multiplicity (as a root) in the characteristic polynomial

Geometric multiplicity of λ = $\dim(E_\lambda)$

Def. Diagonalizable

Def. A_n is diagonalizable if it is similar to a diagonal matrix $D = P^{-1}AP \iff A = PDP^{-1}$

P is said to diagonalize A_n

Note The diagonal entries of D are the eigenvalues λ_i, and the column vectors of P are the eigen vectors

Thm A_n is diagonalizable

$\iff$ A_n has n linearly independent eigen vectors (i.e. enough eigen vectors to form a basis for $\mathbb{R}^n$)

$\iff$ sum of (alg mult of eigen values of A) = n, and each eigen value has geo mult = alg mult

$\iff$ sum of (geo mult of eigen values of A) = n (so if $\exists\, r$ distinct eigen values, then $\mathbb{R}^n = E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_r}$

where $\oplus$ means direct sum, e.g. $\mathbb{R} \oplus \mathbb{R} = \mathbb{R}^2$)

E.g. Consider $T : P_2 \to P_2$ defined by $T(p) = p'$, e.g. $T(ax^2 + bx + c) = 2ax + b$. Is T diagonalizable?

Let $B = (1, x, x^2)$ be a standard basis of P_2

Then $[T]_B^B = R_{B,B} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$

Does there exist 3 lin ind. eigen vectors for the matrix rep of T?

The characteristic polynomial of $[T]_B^B$ is $p(\lambda) = \det([T]_B - \lambda I) = \begin{vmatrix} -\lambda & 1 & 0 \\ 0 & -\lambda & \lambda \\ 0 & 0 & -\lambda \end{vmatrix} = (-\lambda)^3 = -\lambda^3$

So $\lambda = 0$ is an eigenvalue for $[T]_B^B$ with algebraic multiplicity 3 ($\lambda_1 = \lambda_2 = \lambda_3 = 0$)

And $[T]_B^B - 0I = \vec{0} \implies \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix} \sim \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \implies$ eigenvectors are $r \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, r \neq 0$

So the eigenspace $E_\lambda = \left\{ r \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} : r \in \mathbb{R} \right\} = sp(\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix})$ and the geometric multiplicity of eigenvalue 0 is 1

Since there isn't enough eigenvectors, the sum of $\dim(E_\lambda) \neq$ dim(vector space), so T is not diagonalizable.

Note If T is diagonalizable, then it is easy to compute powers of T, e.g. T(T(V)) by finding $[T]_{B'}$ which is the (diagonal) matrix rep of T relative to the basis of eigen vectors (denoted B') and taking the appropriate power of the matrix .

W6 (6.1)

Def. Orthogonal

Let $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $\vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$ be vectors.

Their dot product is $\vec{x} \cdot \vec{y} = x_1 y_1 + \ldots + x_n y_n = \vec{y}^T \vec{x} = \|x\| \cdot \|y\| cos(\theta)$ where θ is the angle between x and y.

Since $\vec{x} \cdot \vec{y} = \vec{0}$ when $\cos\theta = 0 \implies \theta = \dfrac{\pi}{2} = 90^\circ$, they are said to be orthogonal (perpendicular) if $\vec{x} \cdot \vec{y} = \vec{0}$

The length of $\vec{x}$ is $\|\vec{x}\| = \sqrt{x_1^2 + \ldots + x_n^2}$ and $\|\vec{x}\|^2 = \vec{x} \cdot \vec{x}$

A unit vector in the direction of $\vec{v}$ is obtained by dividing by its length: $\vec{u} = \dfrac{1}{\|\vec{v}\|} \vec{v}$

Def. Orthogonal Set

The set $\{\vec{x}_1, \ldots, \vec{x}_k\} \subseteq \mathbb{R}^n$ is an orthogonal set if each $\vec{x}_i$ is not the zero vector and every pair of distinct vectors is orthogonal.

Note $\vec{0}$ not included so that certain properties hold (e.g. orthogonal sets are lin ind.)

Proof Orthogonal sets are linearly independent.

WTS $c_1 \vec{x}_1 + \ldots + c_k \vec{x}_k = \vec{0}$ only has the trivial solution $c_1 = \ldots = c_k = 0$

Take dot prod of eq'n with $\vec{x}_i, 1 \le i \le k$:

$$\begin{aligned} \vec{x}_i \cdot (c_1 \vec{x}_1 + \ldots + c_k \vec{x}_k) &= \vec{x}_i \cdot \vec{0} \\ c_1 \underbrace{(\vec{x}_i \cdot \vec{x}_1)}_{=0} + c_2 \underbrace{(\vec{x}_i \cdot \vec{x}_2)}_{=0} + \ldots + c_i \underbrace{(\vec{x}_i \cdot \vec{x}_i)}_{\neq 0} + \ldots + c_k \underbrace{(\vec{x}_i \cdot \vec{x}_k)}_{=0} &= 0 \\ c_i &= 0 \end{aligned}$$

Def. Orthonormal Set

A set is orthonormal if it is orthogonal and $\|\vec{x}_i\| = 1 \; \forall \; x_i$

E.g. $\left\{ \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}, \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix} \right\}$

Each pair has dot prod = 0 and each vector has length 1, so this set is orthonormal subset of $\mathbb{R}^4$

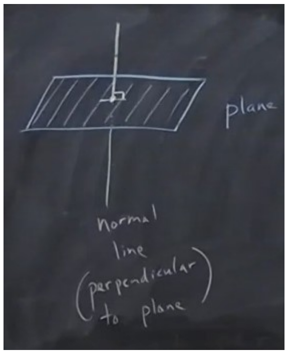
Def. Normalize

To normalize an orthogonal set, divide each element by its length to turn it into an orthonormal set

Def. Orthogonal Complement

Let W be a subspace of $\mathbb{R}^n$. The orthogonal component of W is $W^\perp = \left\{\vec{x} \in \mathbb{R}^n : \vec{x} \cdot \vec{w} = 0 \;\forall\, \vec{w} \in W\right\}$

E.g. If W is a plane through the origin in $\mathbb{R}^3$, then $W^\perp$ is the line through the origin perpendicular to the plane



- both line and plane are vector spaces assuming they pass through the origin
- every vector in the plane is orthogonal to every vector on the normal line

Facts Let W be a subspace of $\mathbb{R}^n$

- $W^\perp$ is a subspace of $\mathbb{R}^n$
- $\dim W + \dim W^\perp = n$
- $(W^\perp)^\perp = W$
- $(\mathbb{R}^n)^\perp = \left\{\vec{0}\right\}$ and $\left\{\vec{0}\right\}^\perp = \mathbb{R}^n$
- $W \cap W^\perp = \left\{\vec{0}\right\}$
- If $W = sp(\{\vec{w}_1, \ldots, \vec{w}_k\})$ then $\vec{x} \in W^\perp \iff \vec{x} \cdot \vec{w}_i = 0 \;\forall\, i = 1, \ldots, k$ (no need to check every vector in W, just check its generating set)
- Each vector $\vec{b}$ in $\mathbb{R}^n$ can be expressed uniquely in the form $\vec{b} = \vec{b}_W + \vec{b}_{W^\perp}$ where $\vec{b}_W \in W, \vec{b}_{W^\perp} \in W^\perp$

Thm (Ortho Complement of Row/Col Space)

Let A be an m x n matrix, then

1. $(row(A))^\perp = null(A)$
2. $(col(A))^\perp = null(A^T)$

Proof Let $\vec{x} \in \mathbb{R}^n$

1. $\vec{x}$ is in $(row(A))^\perp \iff \vec{x}$ is orthogonal to every row of $A \iff A\vec{x} = \vec{0} \iff \vec{x} \in null(A)$

2. $\vec{x}$ is in $\underbrace{(row(A^T))}_{col(A)}^{\perp} \iff \vec{x}$ is orthogonal to every row of $A^T \iff A^T\vec{x} = \vec{0} \iff \vec{x} \in null(A^T)$

Technique to Find Ortho Complement

1. Find a matrix A whose rows are vectors that span W
2. Find $null(A) = W^{\perp}$

E.g. Find $U^{\perp}$ if $U = sp\left(\left\{\begin{bmatrix}1\\0\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\1\\0\end{bmatrix}, \begin{bmatrix}0\\2\\0\\1\end{bmatrix}\right\}\right)$

$\vec{x} = \begin{bmatrix}a\\b\\c\\d\end{bmatrix} \in U^{\perp} \iff \begin{bmatrix}a\\b\\c\\d\end{bmatrix} \cdot \vec{u}_i = 0$ for i = 1, 2, 3. Solving the 3 equations, we have $\begin{cases}a = 0\\ b + c = 0\\ 2b + d = 0\end{cases}$

Solving the 3 equations is equivalent to just writing the u_i's in a matrix: $\left[\begin{array}{cccc|c}a & b & c & d & \\ 1 & 0 & 0 & 0 & 0\\ 0 & 1 & 1 & 0 & 0\\ 0 & 2 & 0 & 1 & 0\end{array}\right]$

which row reduces to $\left[\begin{array}{cccc|c}1 & 0 & 0 & 0 & 0\\ 0 & 1 & 1 & 0 & 0\\ 0 & 0 & 1 & -\frac{1}{2} & 0\end{array}\right] \implies \vec{x} = \begin{bmatrix}0\\-t/2\\t/2\\t\end{bmatrix}, t \in \mathbb{R}$

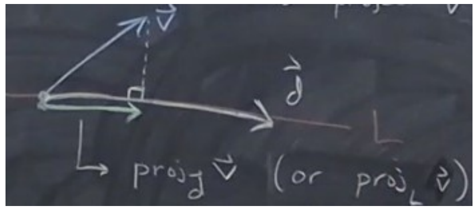
so $U^{\perp} = \left\{\vec{x} = \begin{bmatrix}0\\-t/2\\t/2\\t\end{bmatrix} : t \in \mathbb{R}\right\} = sp\left\{\begin{bmatrix}0\\-1\\1\\2\end{bmatrix} s : s \in \mathbb{R}\right\}$

Check that $\begin{bmatrix}a\\b\\c\\d\end{bmatrix} \cdot \vec{u}_i = 0$ and that $\dim(U^{\perp}) = \dim(V) - \dim(U)$

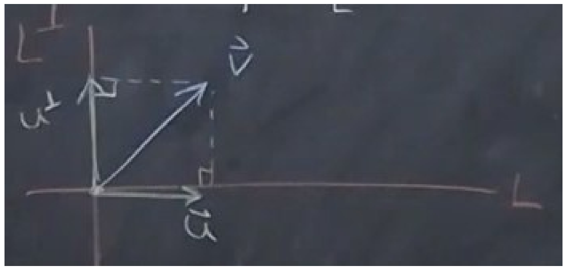
W8 (6.2 - 6.4)

Def. Projections (Textbook Def)

- Let L = span$\{\vec{d}\}$ be a line through the origin
- Take a vector $\vec{v}$ and project it onto L



- We can decompose $\vec{v}$ into the sum of 2 vectors: 1 parallel to L and 1 perpendicular to L



$$\vec{v} = \vec{u} + \vec{u}^{\perp} = \textbf{proj}_L\vec{v} + \textbf{proj}_{L^{\perp}}\vec{v}$$

Let $\vec{b}$ be a vector in $\mathbb{R}^n$ and W a subspace of $\mathbb{R}^n$. Let $\vec{b} = \underbrace{\vec{b}_W} + \vec{b}_{W^{\perp}}$, then $\vec{b}_W$ is the projection of $\vec{b}$ onto W

Steps to find projection of $\vec{b}$ onto W

1. Select a basis $\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k\}$ for W
2. Find a basis $\{\vec{v}_{k+1}, \vec{v}_{k+2}, \ldots, \vec{v}_n\}$ for $W^{\perp}$
3. Find coord. vector $\vec{r} = \begin{bmatrix} r_1 \\ \vdots \\ r_n \end{bmatrix}$ of $\vec{b}$ relative to basis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ (proof below) so that $\vec{b} = r_1\vec{v}_1 + \ldots + r_n\vec{v}_n$
4. Then $\vec{b}_w = r_1\vec{v}_1 + \ldots + r_k\vec{v}_k$

Proof If W is a subspace in $\mathbb{R}^n$, $\{w_1, \ldots, w_k\}$ is a basis for W, and $\{w_{k+1}, \ldots, w_n\}$ is a basis for $W^{\perp}$, show that $\{w_1, \ldots, w_k, w_{k+1}, \ldots, w_n\}$ is a basis for $\mathbb{R}^n$.

> NTS it spans $\mathbb{R}^n$ (which it clearly does), and is linearly independent.
>
> Set up the dependence relation $\sum_{i=1}^{n} a_i w_i = \vec{0}$
>
> Suppose $v = a_1w_1 + \ldots + a_kw_k \implies v \in W$
>
> By the dep. rel, we also have that $v = -(a_{k+1}w_{k+1} + \ldots + a_nw_n) \implies v \in W^{\perp}$
>
> Since $v \in W, W^{\perp}$, $v = \vec{0}$ (since $v \cdot v = \vec{0}$)
>
> Thus, $a_1w_1 + \ldots + a_kw_k = \vec{0}$ and $-(a_{k+1}w_{k+1} + \ldots + a_nw_n) = \vec{0}$
>
> Since $\{w_{k+1}, \ldots, w_n\}$ and $\{w_{k+1}, \ldots, w_n\}$ are linearly independent, $a_1 = \ldots = a_n = 0$

> $\implies \{w_1, \ldots, w_k, w_{k+1}, \ldots, w_n\}$ is lin ind.

Note If instead, $\{w_1, \ldots, w_k\}$ is an **orthonormal** basis for W, and $\{w_{k+1}, \ldots, w_n\}$ is an **orthonormal** basis for $W^\perp$, then $\{w_1, \ldots, w_k, w_{k+1}, \ldots, w_n\}$ is an **orthonormal** basis for $\mathbb{R}^n$. Thus, every $\vec{v} \in \mathbb{R}^n$ can be written uniquely in terms of $\vec{w}_i$ as $\vec{v} = \underbrace{(\vec{v}\cdot\vec{w}_1)\vec{w}_1 + \ldots + (\vec{v}\cdot\vec{w}_k)\vec{w}_k}_{\text{proj}_W\vec{v}} + \underbrace{(\vec{v}\cdot\vec{w}_{k+1})w_{k+1} + \ldots + (\vec{v}\cdot\vec{w}_n)\vec{w}_n}_{\text{proj}_{W^\perp}\vec{v}}$

Proof (that coefficients are $a_i = \vec{v}\cdot\vec{w}_i$)

If $\vec{v} = a_1\vec{w}_1 + \ldots + a_n\vec{w}_n$, then $\vec{v}\cdot\vec{w}_i = (a_1\vec{w}_1 + \ldots + a_n\vec{w}_n)\cdot\vec{w}_i = a_1\underbrace{\vec{w}_1\cdot\vec{w}_j}_{0} + \ldots + a_i\underbrace{\vec{w}_i\cdot\vec{w}_i}_{1} + \ldots + a_n\underbrace{\vec{w}_n\cdot\vec{w}_i}_{0} = a_i$

Def. Projections (Alternative Def)

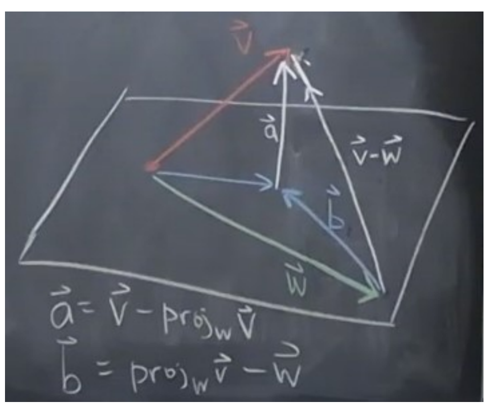
Let W be a subspace of $\mathbb{R}^n$ with an orthogonal basis $\{\vec{v}_1, \ldots, \vec{v}_k\}$. Then the projection of $\vec{b} \in \mathbb{R}^n$ on W is

$\text{proj}_W\vec{b} = \text{proj}_{\vec{v}_1}\vec{b} + \ldots + \text{proj}_{\vec{v}_k}\vec{b}$ or $\vec{b}_W = \dfrac{\vec{b}\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}\vec{v}_1 + \ldots + \dfrac{\vec{b}\cdot\vec{v}_k}{\vec{v}_k\cdot\vec{v}_k}\vec{v}_k$

Properties Let W be a subspace of $\mathbb{R}^n$ and $\vec{v} \in \mathbb{R}^n$. Then

- $\text{proj}_W\vec{v} \in W$
- $\vec{v} - \text{proj}_W\vec{v} \in W^\perp$
- $\text{proj}_W\vec{v}$ is the unique vector in W closest to $\vec{v}$

  Proof WTS $\|\vec{v} - \text{proj}_W\vec{v}\| \le \|\vec{v} - \vec{w}\|$ for all $\vec{w} \in W$, i.e. minimized when $\vec{w} = \text{proj}_W\vec{v}$



$$
\begin{aligned}
\|\vec{v}-\vec{w}\|^2 &= \left\|\vec{b}+\vec{a}\right\|^2 \\
&= (\vec{b}+\vec{a})\cdot(\vec{b}+\vec{a}) \\
&= \vec{b}\cdot\vec{b}+2\vec{a}\cdot\vec{b}+\vec{a}\cdot\vec{a} \\
&= \left\|\vec{b}\right\|^2+0+\|\vec{a}\|^2 \\
&\geq \|\vec{a}\|^2 = \|\vec{v}-\mathrm{proj}_W\vec{v}\|^2
\end{aligned}
$$

$$\|\vec{v}-\vec{w}\| \geq \|\vec{v}-\mathrm{proj}_W\vec{v}\| \implies$$

$$\|\vec{v}-\vec{w}\| = \|\vec{v}-\mathrm{proj}_W\vec{v}\| \text{ iff } \vec{w}=\mathrm{proj}_W\vec{v}$$

E.g. Let $W = sp(\begin{bmatrix}0\\-1\\1\end{bmatrix}, \begin{bmatrix}0\\-1\\-1\end{bmatrix})$. Find the vector in W closest to $\vec{x} = \begin{bmatrix}-2\\-3\\2\end{bmatrix}$

Since W is the span of an orthogonal set, the answer is $\mathrm{proj}_W\vec{x}$ by the projection theorem.

$$\vec{x}_W = \frac{\vec{x}\cdot\vec{w}_1}{\vec{w}_1\cdot\vec{w}_1}\vec{w}_1 + \frac{\vec{x}\cdot\vec{w}_2}{\vec{w}_2\cdot\vec{w}_2}\vec{w}_2 = \frac{5}{2}\begin{bmatrix}0\\-1\\1\end{bmatrix} + \frac{1}{2}\begin{bmatrix}0\\-1\\-1\end{bmatrix} = \begin{bmatrix}0\\-3\\2\end{bmatrix}$$

Observe that the $\mathrm{proj}_W\vec{x}$ is just $\vec{x}$ with its first value replaced with 0. This makes sense since W is the span of 2 vectors with x = 0, so it is like projecting onto the y-z plane. Hence another (simpler) way to approach this problem would be to choose the basis as the standard basis of the y-z plane $\left\{\begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\}$

Def. Gram Schmidt Algorithm

Every subspace W of $\mathbb{R}^n$ has an orthonormal basis.

Let $\{\vec{x}_1, \ldots, \vec{x}_k\}$ be a basis for W. The Gram Schmidt algo generates an orthogonal basis $\{\vec{v}_1, \ldots, \vec{v}_k\}$ for W.

$\vec{v}_1 = \vec{x}_1$

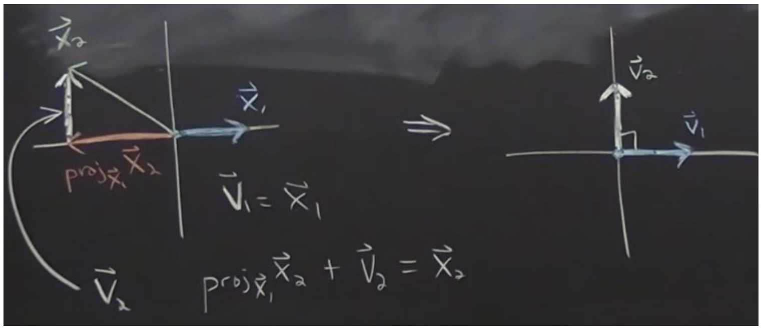
$$\vec{v}_2 = \vec{x}_2 - \frac{\vec{x}_2\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}\vec{v}_1 \quad \leftarrow \text{perp. to } \vec{v}_1$$

$$\vec{v}_3 = \vec{x}_3 - \frac{\vec{x}_3\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}\vec{v}_1 - \frac{\vec{x}_3\cdot\vec{v}_2}{\vec{v}_2\cdot\vec{v}_2}\vec{v}_2 \quad \leftarrow \text{perp. to } \vec{v}_1, \vec{v}_2$$

$\vdots$

$$\vec{v}_k = \vec{x}_k - \frac{\vec{x}_k\cdot\vec{v}_1}{\vec{v}_1\cdot\vec{v}_1}\vec{v}_1 - \ldots - \frac{\vec{x}_k\cdot\vec{v}_{k-1}}{\vec{v}_{k-1}\cdot\vec{v}_{k-1}}\vec{v}_{k-1} \quad \leftarrow \text{perp. to } \vec{v}_1, \ldots, \vec{v}_k$$

Illustration for $\mathbb{R}^2$:



For convenience, we can scale $\vec{v}_i$ to remove fractions during the algorithm.

To get an orthonormal basis, we divide each vector by its length at the end.

Def. Orthogonal Matrix

A square matrix Q is said to be orthogonal if any of the following is true

- its columns form an orthonormal set
- $Q^TQ = I$
- $Q^{-1} = Q^T$

Note If A is not square, then $A^TA = I \neq AA^T$. This chain of equality only holds if A is square. (They do have the same non-zero eigen values however.)

Properties

- both Q^{-1} and Q^T are orthogonal

  Proof $(Q^{-1})^T = (Q^T)^{-1} = (Q^{-1})^{-1} \implies Q^{-1}$ is orthogonal

  Proof $(Q^T)^{-1} = (Q^{-1})^{-1} = Q = (Q^T)^T \implies Q^T$ is orthogonal

- Since Q^T is orthogonal, rows of W also form an orthonormal set

- det(Q) = ± 1 (so orthogonal matrices are rotations or reflections)

  Proof $(\det(Q))^2 = \det(Q)\det(Q) = \det(Q^T)\det(Q) = \det(Q^TQ) = \det(I) = 1$

- $(Q\vec{x}) \cdot (Q\vec{y}) = \vec{x} \cdot \vec{y}$

  Proof $Q\vec{x} \cdot Q\vec{y} = (Q\vec{y})^T(Q\vec{x}) = \vec{y}^T(Q^TQ)\vec{x} = \vec{y}^T\vec{x} = \vec{x} \cdot \vec{y}$

- $\|Q\vec{x}\| = \|\vec{x}\|$

  Proof $\|Q\vec{x}\|^2 = (Q\vec{x}) \cdot (Q\vec{x}) = \vec{x} \cdot \vec{x} = \|\vec{x}\|^2$

- The angle between non-zero vectors $\vec{x}$ and $\vec{y}$ = the angle between $(Q\vec{x})$ and $(Q\vec{y})$

  Proof $\cos\theta = \dfrac{Q\vec{x}\cdot Q\vec{y}}{\|Q\vec{x}\|\,\|Q\vec{y}\|} = \dfrac{\vec{x}\cdot\vec{y}}{\|\vec{x}\|\,\|\vec{y}\|}$

- If λ is an eigenvalue of Q, then $|\lambda| = 1$ (all eigs are on the unit circle in the complex plane)

  Proof $Q\vec{x} = \lambda\vec{x}$ so $\|Q\vec{x}\| = \|\lambda\vec{x}\|$ but also $\|Q\vec{x}\| = \|\vec{x}\|$ so $\|\lambda\vec{x}\| = \|\vec{x}\| \implies \|\lambda\| = 1$

- If Q and R are orthogonal, then so is QR

  Proof $(QR)^T(QR) = R^TQ^TQR = R^TR = I$

Thm (Relation to Change of Basis)

If $B = \{\vec{w}_1, \ldots, \vec{w}_n\}$ is an orthonormal basis of $\mathbb{R}^n$,

then the change of basis matrix rel. to E is $[I]_B^E = [\vec{w}_1 \mid \ldots \mid \vec{w}_n]$, which is an orthogonal matrix.

So $([I]_B^E)^{-1} = ([I]_B^E)^T \quad \leftarrow (Q^{-1} = Q^T)$

and $\|\vec{v}_B\| = \left\| \underbrace{[I]_B^E}_{Q}\ \underbrace{\vec{v}_E}_{\vec{v}} \right\| = \|\vec{v}_E\| \quad \leftarrow (\text{use Qv = v, proves Q preserves length})$

E.g. Let $(\vec{a}_1, \vec{a}_2, \vec{a}_3, \vec{a}_4)$ be an ordered orthonormal basis for $\mathbb{R}^4$, and let $[2, 1, 4, -3]$ be the coordinate vector of a vector $\vec{b}$ in $\mathbb{R}^4$ relative to this basis. Find $\|\vec{b}\|$.

> Let ϵ be the standard basis for $\mathbb{R}^4$ and $B = (\vec{a}_1, \vec{a}_2, \vec{a}_3, \vec{a}_4)$, then $[2, 1, 4, 3] = [\mathbf{b}]_B$
>
> $\|\mathrm{b}\| = \|[\mathrm{b}]_\varepsilon\| = \|[I]_B^\varepsilon[\mathrm{b}]_B\| = \|[\mathrm{b}]_B\| = \|[2, 1, 4, -3]\| = \sqrt{30}$

Def. Orthogonal Transformation

A linear transformation $T : \mathbb{R}^n \to \mathbb{R}^n$ is orthogonal if $T(\vec{v}) \cdot T(\vec{w}) = \vec{v} \cdot \vec{w}\ \forall\ \vec{v}, \vec{w} \in \mathbb{R}^n$

(T is orthogonal $\iff$ its matrix is orthogonal)

Def. Projection Matrix

Let W be a subspace of $\mathbb{R}^n$, then $T : \mathbb{R}^n \to \mathbb{R}^n$ defined by $T(\vec{x}) = \text{proj}_W \vec{x}$ is a linear transformation. The standard matrix for T is called the **projection matrix** for the subspace W.

> **Steps** (to find a projection matrix for subspace W)
>
> 1. form a matrix using vectors from W's basis: $A = [\vec{w}_1 \,|\dots|\, \vec{w}_k]_{n\times k}$ so that W is the column space of A
> 2. then $P = A(A^TA)^{-1}A^T$ is the projection matrix for the subspace W and $\text{proj}_W \vec{x} = P\vec{x}$
>
> If $\{\vec{w}_1, \dots, \vec{w}_k\}$ is an orthonormal basis, $P = AA^T$

E.g. Let $W = sp(\begin{bmatrix} \frac{1}{\sqrt{3}} \\ -\frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix})$. Find the projection matrix for W and $\text{proj}_W \vec{x}$

Let $A = \begin{bmatrix} \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{3}} & 0 \end{bmatrix}$. Note that the columns are orthonormal, so A is orthogonal, i.e. $A^TA = I$

Then $P = A(\underbrace{A^TA}_{I})^{-1}A^T = AA^T = \begin{bmatrix} \frac{5}{6} & \frac{1}{6} & \frac{1}{3} \\ \frac{1}{6} & \frac{5}{6} & -\frac{1}{3} \\ \frac{1}{3} & -\frac{1}{3} & \frac{1}{3} \end{bmatrix}$ and $\text{proj}_W \vec{x} = [P]\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \frac{1}{6}\begin{bmatrix} 5x_1 + x_2 + 2x_3 \\ x_1 + 5x_2 - 2x_3 \\ 2x_1 - 2x_2 + 2x_3 \end{bmatrix}$

Proof (for $P = A(A^TA)^{-1}A^T$)

Note that W = col(A), so all vectors in W can be written as $A\vec{x} = [w_1 \,|\dots|\, w_k]\begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} = \sum_i x_i \vec{w}_i$

$\implies \text{proj}_W \vec{b} \in W \implies \text{proj}_W \vec{b} \in col(A)$

$\implies \exists\, \hat{x} \in \mathbb{R}^k$ such that $\text{proj}_W \vec{b} = A\hat{x}$

$\implies b - \text{proj}_W \vec{b} = \vec{b} - A\hat{x}$ is perpendicular to vectors in W, which all have form $A\vec{x}$

$\implies (b - A\hat{x}) \cdot (A\vec{x}) = 0$

$\implies (A\vec{x})^T(\vec{b} - A\hat{x}) = 0 \quad \leftarrow \text{since } x \cdot y = y^Tx$

$\implies (\vec{x}^T)(A^T\vec{b} - A^TA\hat{x}) = 0 \quad \leftarrow \text{since } (A\vec{x})^T = \vec{x}^TA^T$

$\implies (A^T\vec{b} - A^TA\hat{x}) \cdot \vec{x} = 0 \quad \leftarrow \text{write as dot product}$

$\implies A^T\vec{b} - A^TA\hat{x} = \vec{0} \quad \leftarrow \text{can now solve for } \hat{x}$

$\implies \hat{x} = (A^TA)^{-1}A^T\vec{b} \quad \leftarrow \text{since } rank(A^TA) = rank(A) = k\text{, } (A^TA)_{k\times k} \text{ has full rank so must be invertible}$

$\implies \text{proj}_W\vec{b} = A\hat{x} = \underbrace{A(A^TA)^{-1}A^T}_{P}\vec{b}$

Properties

Every n x n matrix that is both **symmetric** and **idempotent** is a projection matrix.

- P is symmetric if $P^T = P$
- P is idempotent if $P^2 = P$

E.g. Find all invertible projection matrices.

Since projection matrices are idempotent, $P^2 = P$.

If P is invertible, then multiplying both sides of $P^2 = P$ by P^{-1} gives $P = I$.

W9 (9.1, 9.2)

Def. Least Squares

Let A be an m x n matrix. $A\vec{x} = \vec{b}$ could have no solution, i.e. if $\vec{b}$ is not in col(A).

Then the best "approximate solution" to $A\vec{x} = \vec{b}$ is the least squares solution $\hat{x}$, where
$\|b - A\hat{x}\| \leq \|b - A\vec{x}\| \; \forall \; \vec{x} \in \mathbb{R}^n$

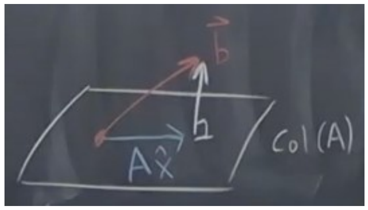
$\hat{x}$ minimizes error $e = \|A\hat{x} - \vec{b}\|^2$, which is the sum of squares of the errors in the m equations (when $m \geq n$)

> If no solution, then multiply $A\vec{x} = \vec{b}$ by A^T to get $A^TA\hat{x} = A^T\vec{b}$, and solve for $\hat{x} = (A^TA)^{-1}A^T\vec{b}$.

Justification Look for $\hat{x}$ for which $A\hat{x} \in W = col(A)$.

$(A\hat{x} - \vec{b}) \perp col(A) \implies (A\hat{x} - \vec{b}) \in (col(A))^{\perp} = null(A^T) = \{\vec{x} \in \mathbb{R}^n : A^T\vec{x} = 0\}$

$\implies A^T(A\hat{x} - \vec{b}) = \vec{0} \implies A^TA\hat{x} = A^T\vec{b}$



E.g. Find the line closest to points (0, 6) (1, 0) and (2, 0)

$$\text{y = mx + b} \implies \begin{cases} 0m + b = 6 \\ 1m + b = 0 \\ 2m + b = 0 \end{cases} \text{ so } A\vec{x} = \vec{b} \implies \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} m \\ b \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix}$$

$$A^T A\hat{x} = A^T\vec{b} \implies \hat{x} = (A^T A)^{-1} A^T \vec{b} = \left(\begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 6 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -3 \\ 5 \end{bmatrix}$$

So y = -3x + 5 is the line of best fit

Note $A^T A$ is invertible by the following thm.

Thm (Invertibility, Lin Ind, and Uniqueness)

Let A be an m x n matrix and $\vec{b} \in \mathbb{R}^m$. Then the following are equivalent:

- $A\vec{x} = \vec{b}$ has a unique least squares solution $\hat{x}$
- columns of A are linearly independent
- $A^T A$ is invertible

Def. Symmetric Matrices

A symmetric matrix $A = A^T$ always has real eigen values and is always diagonalizable (by a real orthogonal matrix).

Properties

- Eigenspaces of a symmetric matrix are mutually orthogonal.
- Thus, symmetric matrices have n mutually perpendicular eigen vectors.

Proof Let $\lambda \in \mathbb{C}$ be an eigen value and A be a real matrix.

Then $A\vec{x} = \lambda\vec{x}$ for some $\vec{x} \neq \vec{0}$

Since A is real, $A = \overline{A} = \overline{A^T}$

LHS: $\overline{\vec{x}^T} A\vec{x} = \overline{\vec{x}^T}\lambda\vec{x} = \lambda\left(\overline{\vec{x}^T}\vec{x}\right)$

RHS: $\overline{\vec{x}^T A^T \vec{x}} = \overline{(A\vec{x})^T \vec{x}} = \overline{(\lambda\vec{x})^T \vec{x}} = \overline{\lambda}\left(\overline{\vec{x}^T \vec{x}}\right)$

Equating LHS and RHS, we have that $\lambda = \overline{\lambda} \implies \lambda$ is real

Def. Orthogonally Diagonalizable

A square matrix A is **orthogonally diagonalizable** if there is an orthogonal matrix Q and a diagonal matrix D such that $A = QDQ^T \iff Q^T A Q = D$

Note For a real n x n matrix A, symmetric $\iff$ orthogonally diagonalizable

E.g. $\begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$ is diagonalizable but not orthogonally diagonalizable since it is not symmetric.

> Steps to orthogonally diagonalize a matrix A:
>
> 1. Find the eigen values and eigen vectors
> 2. Form a basis with the eigen vectors and normalize
> 3. Form Q with the normalized eigen vectors as columns

E.g. Diagonalize $A = \begin{bmatrix} 1 & 2 \\ 2 & -2 \end{bmatrix}$ using an orthogonal matrix.

The eigen values are $\lambda_1 = -3$ with eigen vector $\vec{v}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix}$ and $\lambda_2 = 2$ with eigen vector $\vec{v}_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$A = PDP^{-1}$ with $P = [\vec{v}_1 | \vec{v}_2] = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$ and $D = \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix}$ $\leftarrow$ regular diagonalization

Note that $\left\{\vec{v}_1, \vec{v}_2\right\}$ is an orthogonal set, so if we normalize it to $\left\{ \begin{bmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix} \right\}$, then we can use

$A = QDQ^{-1}$ with $Q = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{bmatrix}$ and D defined above. $\leftarrow$ orthogonal diagonalization

Since Q is orthogonal, $Q^{-1} = Q^T$, so $A = QDQ^T$.

Def. Complex Numbers

In **cartesian** form, $z = a + bi$

In **polar** form, $z = r(\cos\theta + i\sin\theta) = re^{i\theta}$

The **modulus** is its magnitude $|z| = \sqrt{a^2+b^2} = r$

The **principal argument**, denoted $\arg(z)$, is θ if $-\pi < \theta \le \pi$

The **complex conjugate** of $z = a + ib$ is $\bar{z} = a - ib$, and their product $z\overline{z} = |z|^2$ is a real number: $a^2 + b^2 > 0$

Thm z is a real number iff $z = \overline{z}$

Proof Let $z = a + bi$ for $a, b \in \mathbb{R}$. Then $\bar{z} = a - bi$.

$(\Rightarrow)$ Assume that z is a real number. Then $b = 0 \implies z = a = \bar{z}$.

$(\Leftarrow)$ Assume that $z = \overline{z}$. Then $a + bi = a - bi \implies b = -b \implies b = 0$

E.g. Find the modulus and principal argument for $z = \sqrt{3} - i$

$$r = |z| = \sqrt{(\sqrt{3})^2 + (-1)^2} = \sqrt{4} = 2$$

$$\arg(z) = \theta = -\pi/6$$

E.g. z = -1 + i. Express z^{-1} in the form a + bi where a and b are real numbers.

$$\frac{1}{-1+i} = \frac{1}{-1+i}\frac{-1-i}{-1-i} = \frac{-1-i}{1+1} = -\frac{1}{2} - \frac{1}{2}i$$

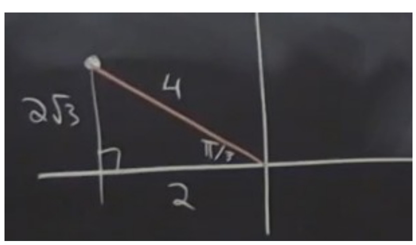
Def. De Moivre's Theorem

$(e^{i\theta})^n = e^{in\theta}$ and $(re^{i\theta})^n = r^n e^{in\theta}$ where n is a positive integer

Def. nth roots of z

E.g. Solve $z^4 = 2(\sqrt{3}i - 1)$

RHS is $-2 + 2\sqrt{3}i \implies r = \sqrt{4 + 4 \cdot 3} = 4, b_\triangle = 2, h_\triangle = 2\sqrt{3} \implies \theta = \dfrac{2\pi}{3}$



RHS in polar form is $4e^{i(\frac{2\pi}{3})}$

LHS in polar form is $(re^{i\theta})^4$

Eq'n in polar form is $(re^{i\theta})^4 = 4e^{i(\frac{2\pi}{3})} \implies r^4 e^{i\theta 4} = 4e^{i(\frac{2\pi}{3})}$

So $r^4 = 4 \implies r = \sqrt{2}$ and $4\theta = \dfrac{2\pi}{3} + 2k\pi \implies \theta = \dfrac{\pi}{6} + \dfrac{\pi}{2}k, k \in \mathbb{Z}$

k = 0 $\implies z = \sqrt{2}e^{(\pi/6)i} = \sqrt{2}(\dfrac{\sqrt{3}}{2} + \dfrac{1}{2}i)$

k = 1 $\implies z = \sqrt{2}e^{(2\pi/3)i} = \sqrt{2}(-\dfrac{1}{2} + \dfrac{\sqrt{3}}{2}i)$

k = 2 $\implies z = \sqrt{2}e^{(7\pi/6)i} = \sqrt{2}(-\dfrac{\sqrt{3}}{2} - \dfrac{1}{2}i)$

k = 4 $\implies z = \sqrt{2}e^{(5\pi/3)i} = \sqrt{2}(\dfrac{1}{2} - \dfrac{\sqrt{3}}{2}i)$

Formula for nth roots of $z = r(\cos\theta + i\sin\theta)$:

$$r^{\frac{1}{n}}\left(\cos\left(\frac{\theta}{n} + \frac{2k\pi}{n}\right) + i\sin\left(\frac{\theta}{n} + \frac{2k\pi}{n}\right)\right) \quad \text{for k = 0, 1, 2, ..., n-1}$$

E.g. Find the 4 fourth roots of 1.

> Method 1: Factor and solve.
>
> $z^4 - 1 = 0 \quad \rightarrow \quad \left(z^2 - 1\right)\left(z^2 + 1\right) = 0 \quad \rightarrow \quad z = \pm 1$ or $z = \pm i$.
>
> Method 2: Use the formula. $\theta = 0, r = 1, n = 4$
>
> $k = 0 \rightarrow \cos 0 + i\sin 0 = 1$
> $k = 1 \rightarrow \cos\dfrac{\pi}{2} + i\sin\dfrac{\pi}{2} = i$
> $k = 2 \rightarrow \cos\pi + i\sin\pi = -1$
> $k = 3 \rightarrow \cos\dfrac{3\pi}{2} + i\sin\dfrac{3\pi}{2} = -i$

E.g. Find eigen values/eigen vectors of $A = \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

> $p(\lambda) = (\lambda - 1)^2 + 1 = \lambda^2 - 2\lambda + 2 = 0$
>
> $\lambda = \frac{2\pm\sqrt{4-8}}{2} = 1 \pm i$
>
> $\lambda_1 = 1 + i$
>
> $$A - \lambda I = \begin{bmatrix} 1-(1+i) & -1 \\ 1 & 1-(1+i) \end{bmatrix} = \begin{bmatrix} -i & -1 \\ 1 & -i \end{bmatrix}$$

$$\begin{cases} R_1 = iR_1 = [1, -i] \\ R_2 = R_2 - i(R_1) = [1 + i^2 = 0, -i + i = 0] \end{cases} \implies \begin{bmatrix} 1 & -i \\ 0 & 0 \end{bmatrix} \implies a = ib \implies v_1 = \begin{bmatrix} i \\ 1 \end{bmatrix}$$

$\lambda_2 = 1 - i$

$$v_2 = \overline{v_1} = \begin{bmatrix} -i \\ 1 \end{bmatrix}$$

Def. Complex Vector Space

$\mathbb{C}^n =$ space of ordered n-tuples of complex numbers. Vectors in $\mathbb{C}^n$ are $\begin{bmatrix} a_1 + b_1 i \\ \vdots \\ a_n + b_n i \end{bmatrix}$. Assume scalars are in $\mathbb{C}$

Real scalars	**Complex scalars**
$\dim(\mathbb{C}^n) = 2n$	$\dim(\mathbb{C}^n) = n$
basis = $\left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} i \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ \vdots \\ 0 \\ i \end{bmatrix} \right\}$	basis = $\left\{ \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \dots, \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \right\}$

A complex vector space is one in which scalars are complex numbers.

E.g. $M_{2,2}(\mathbb{C}) = \left\{ \begin{bmatrix} a & b \\ c & d \end{bmatrix} : a, b, c, d \in \mathbb{C} \right\}$ is a complex vector space (with usual operations)

E.g. $\mathbb{R}^2$ is not a complex vector space since it is not closed under scalar mult ($v \in \mathbb{R}^2, i \cdot v \notin \mathbb{R}^2$)

E.g. If $\vec{v} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$, find $\vec{v}_B$ if $B = \left(\begin{bmatrix} i \\ i \end{bmatrix}, \begin{bmatrix} i \\ 1 \end{bmatrix} \right)$ is a basis of $\mathbb{C}^2$

$$\begin{bmatrix} -1 \\ 1 \end{bmatrix} = z_1 \begin{bmatrix} i \\ i \end{bmatrix} + z_2 \begin{bmatrix} i \\ 1 \end{bmatrix}$$

$$\left[\begin{array}{cc|c} i & i & -1 \\ i & 1 & 1 \end{array}\right] \to \left[\begin{array}{cc|c} i & i & -1 \\ 0 & 1-i & 2 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 1 & i \\ 0 & 1-i & 2 \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 1 & i \\ 0 & 1 & \frac{2}{1-i}\frac{1+i}{1+i} = 1+i \end{array}\right] \to \left[\begin{array}{cc|c} 1 & 0 & -1 \\ 0 & 1 & 1+i \end{array}\right] \implies$$

$$\begin{cases} z_1 = -1 \\ z_2 = 1 + i \end{cases} \implies v_b = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} -1 \\ 1+i \end{bmatrix}$$

Note $\left\{ \begin{bmatrix} i \\ i \end{bmatrix}, \begin{bmatrix} i \\ 1 \end{bmatrix} \right\}$ is linearly independent since the matrix has rank 2.

Def. Conjugate Transpose

The conjugate transpose of A is defined as $A^* = \overline{A}^T$

Properties

Let A, B be m x n matrices and $k \in \mathbb{C}$

1. $(A^*)^* = A$
2. $(A+B)^* = A^* + B^*$
3. $(kA)^* = \overline{k}A^*$
4. $(AB)^* = B^*A^*$ if A, B are square

Def. Euclidean Inner Product / Dot Product

In $\mathbb{R}^n : \vec{x} \cdot \vec{y} = x_1y_1 + \ldots + x_ny_n = \vec{y}^T\vec{x}$ and $\|\vec{x}\|^2 = \vec{x} \cdot \vec{x}$

In $\mathbb{C}^n : \langle \vec{x}, \vec{y} \rangle = x_1\overline{y_1} + \ldots + x_n\overline{y_n} = \vec{y}^*\vec{x}$

E.g. Find $\langle u, v \rangle$ where $u = \begin{bmatrix} i \\ 0 \\ 1+i \end{bmatrix}$ and $v = \begin{bmatrix} 3-i \\ 7 \\ 2i \end{bmatrix}$

> $\langle u, v \rangle = (i)(\overline{3-i}) + (1+i)(\overline{2i}) = (i)(3+i) + (1+i)(-2i) = 3i - 1 - 2i + 2 = 1 + i$

Complex Dot Product Properties

Let $u, v, w \in \mathbb{C}^n, k \in \mathbb{C}$

1. $\langle u, u \rangle \geq 0$ and $\langle u, u \rangle = 0$ iff $u = \vec{0}$
2. $\langle u, v \rangle = \overline{\langle v, u \rangle}$ $\quad \leftarrow$ not commutative! order matters!
3. $\langle u+v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$ and $\langle w, u+v \rangle = \langle w, u \rangle + \langle w, v \rangle$
4. $\langle ku, v \rangle = k\langle u, v \rangle$ and $\langle u, kv \rangle = \overline{k}\langle u, v \rangle$

Def. Euclidean Norm

The Euclidean norm (or length) of $u \in \mathbb{C}^n$ is $||u|| = \sqrt{\langle u, v \rangle}$

Def. Unitary Matrix

A complex matrix U is said to be unitary if any of the following is true (**analogous to orthogonal matrices** in $\mathbb{R}^n$)

- columns of U are orthonormal in $\mathbb{C}^n$
- $U^{-1} = U^*$
- $UU^* = I$

Property the product of unitary matrices is unitary. $AA^* = BB^* = I \implies AB(AB)^* = ABB^*A^* = AIA^* = I$

Def. Hermitian Matrix

> H is a Hermitian matrix if $H = H^*$ (analogous to symmetric matrices in $\mathbb{R}^n$).
>
> - It has real eigen values and is diagonalizable (by a unitary matrix).
> - Eigen vectors corresponding to distinct eigen values are orthogonal.
> - $A = A^* \iff \langle Az, w\rangle = \langle z, Aw\rangle \; \forall \; w, z, \in \mathbb{C}^n$

Proof (eigen values are real)

Let $\lambda \in \mathbb{C}$ be an eigen value and H be a Hermitian matrix.

Then $H\vec{x} = \lambda\vec{x}$ for some $\vec{x} \neq \vec{0}$

Since H is Hermitian, $H = H^*$

LHS: $\overline{\vec{x}^T} H\vec{x} = \overline{\vec{x}^T}\lambda\vec{x} = \lambda\left(\overline{\vec{x}^T}\vec{x}\right)$

RHS: $\overline{\vec{x}^T} H^*\vec{x} = \overline{(H\vec{x})^T}\vec{x} = \overline{(\lambda\vec{x})^T}\vec{x} = \overline{\lambda}\left(\overline{\vec{x}^T}\vec{x}\right)$

Equating LHS and RHS, we have that $\lambda = \overline{\lambda} \implies \lambda$ is real

Proof (eigen vectors are orthogonal)

Let λ_1, λ_2 be eigen values of A with (nonzero) eigen vectors v_1 and v_2, then $Av_1 = \lambda_1 v_1, Av_2 = \lambda v_2$

$\langle Av_1, v_2\rangle = \langle \lambda_1 v_1, v_2\rangle = \lambda_1\langle v_1, v_2\rangle$ and $\langle Av_1, v_2\rangle = \langle v_1, Av_2\rangle = \langle v_1, \lambda_2 v_2\rangle = \overline{\lambda}_2\langle v_1, v_2\rangle$, so $\lambda_1 = \overline{\lambda}_2$

(If we take $v_1 = v_2, \lambda_1 = \lambda_2$, then $\lambda_1 = \overline{\lambda}_1, \lambda_2 = \overline{\lambda}_2$ which also proves the eigs are real)

Since the eigs are real, $\overline{\lambda}_2 = \lambda_2 \implies \lambda_1\langle v_1, v_2\rangle = \lambda_2\langle v_1, v_2\rangle \implies (\lambda_1 - \lambda_2)\langle v_1, v_2\rangle = 0$

So if $\lambda_1 \neq \lambda_2$, then $\langle v_1, v_2 \rangle = 0 \implies v_1 \perp v_2$

Proof (Hermitian iff $\langle Az, w \rangle = \langle z, Aw \rangle$)

($\Rightarrow$) If $A = A^*$, then $\langle Az, w \rangle = w^*(Az) = W^* A^* z = (Aw)^* z = \langle z, Aw \rangle$

($\Leftarrow$) Let $A = [a_{ij}]$. Then $a_{ij} = e_i^T A e_j = e_i^* A e_j = \langle Ae_j, e_i \rangle = \langle e_j, Ae_i \rangle = (Ae_i)^* e_j = e_i^T A^* e_j = \overline{a_{ji}}$

Note if A is skew-Hermitian($A^* = -A$), then iA is Hermitian, and vice versa.

Fact Every square matrix can be decomposed into the sum of a Hermitian matrix and skew-Hermitian matrix

$$C = \frac{1}{2}\underbrace{(C + C^*)}_{Hermitian} + \frac{1}{2}\underbrace{(C - C^*)}_{skew-Hermitian}$$

Def. Normal Matrix

A square matrix A is normal if it commutes with its conjugate transpose: $AA^* = A^*A$

Property

- If A is normal, then $\|A\mathbf{z}\| = \|A^*\mathbf{z}\| \; \forall \; \mathbf{z} \in \mathbb{C}^n$

  Proof $\|A\mathbf{z}\|^2 = \;\; = (Az)^*(Az) = z^* A^* A z = z^* A A^* z = (A^* z)^*(A^* z) = \;\; = \|A^* z\|^2$

Shortcuts (to prove a matrix is normal)

- Every Hermitian matrix is normal.

  Proof $H = H^*$ so $HH^* = H^2$ and $H^*H = H^2 \implies HH^* = H^2 = H^*H$

- Every unitary matrix is normal.

  Proof $UU^* = I \iff U^* = U^{-1}$ so $UU^* = UU^{-1} = I = U^{-1}U = U^*U$

  Note for square matrices A and B, AB = I = BA always holds

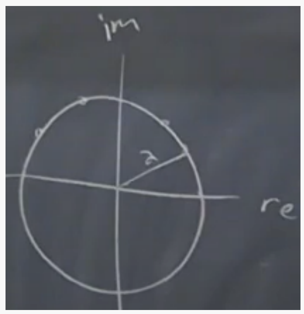
- Every skew-Hermitian matrix is normal.

  Proof $A^*A = -AA = -A^2$ and $AA^* = A(-A) = -A^2$

Def. Unitarily Diagonalizable

A square matrix A is unitarily diagonalizable iff A is a normal matrix.

E.g. Determine all values of $a \in \mathbb{C}$ such that $A = \begin{bmatrix} i & a \\ 2 & i \end{bmatrix}$ is unitarily diagonalizable.

> By thm, A must be normal, so $AA^* = A^*A$
>
> $$\begin{bmatrix} i & a \\ 2 & i \end{bmatrix} \begin{bmatrix} -i & 2 \\ \overline{a} & -i \end{bmatrix} = \begin{bmatrix} -i & 2 \\ \overline{a} & -i \end{bmatrix} \begin{bmatrix} i & a \\ 2 & i \end{bmatrix}$$
>
> Calculate top left entry on both sides $\implies i(-i) + a\overline{a} = i(-i) + 4 \implies 4 = a\overline{a} \implies 4 = |a|^2 \implies |a| = 2$
>
> Hence a is any number $x + iy$ such that $x^2 + y^2 = 4$
>
> 
>

E.g. Find all $a \in \mathbb{C}$ such that the matrix $\begin{bmatrix} i & 4 \\ a & i \end{bmatrix}$ is unitarily diagonalizable.

> By thm, A must be normal, so $AA^* = A^*A$
>
> $$\begin{bmatrix} i & 4 \\ a & i \end{bmatrix} \begin{bmatrix} -i & \overline{a} \\ 4 & -i \end{bmatrix} = \begin{bmatrix} -i & \overline{a} \\ 4 & -i \end{bmatrix} \begin{bmatrix} i & 4 \\ a & i \end{bmatrix}$$
>
> $-1 + 16 = -1 + a^2 \implies |a| = 4$
>
> Hence a is any number of the form $x + iy$ such that $x^2 + y^2 = 4$

W10 (9.3)

Comparison - Real vs Complex

Real (Symmetric matrix)	Complex (Hermitian matrix)
eigen values of A are real	eigen values of H are real
eigen vectors corresponding to distinct eigen values are orthogonal in $\mathbb{R}^n$	eigen vectors corresponding to distinct eigen values are orthogonal in $\mathbb{C}^n$
there exists an **orthogonal** matrix Q s.t. $Q^T AQ$ is diagonal	there exists a **unitary** matrix U s.t. U^*HU is diagonal

Notes (eigenvalues)

If a matrix has **odd order** (order = m x n), then the eigen values must be **real**.

For a real matrix, eigen values come in **conjugate pairs**, i.e. if $a+ib$ is an eigen value, then $a-ib$ is also an eigen value.

For an **orthogonal** matrix, the eigen value pairs all lie on the unit circle.

For a **Hermitian** matrix, eigen values lie on the **real** axis; for a **skew-Hermitian** matrix, eigen values lie on the **imaginary** axis.

Proof (eigenvalues of skew-Hermitian matrices have form $ib, b \in \mathbb{R}$)

Let $\vec{v}$ be an eigen vector for A with corresponding eigen value $\lambda = a+ib$

$\langle Av, v\rangle = v^*Av = v^*\lambda v = \lambda v^*v$

$\langle Av, v\rangle = v^*(-A^*)v = -(Av)^*v = -(\lambda v)^*v = -\overline{\lambda}v^*v$

So $\lambda = -\overline{\lambda} \implies a+ib = -(\overline{a+ib}) = -(a-ib) = -a+ib \implies a = -a \implies a = 0$

If we replace $\langle Av, v\rangle$ with $\langle Av, w\rangle$, then we can prove that $\langle v, w\rangle = 0$ (eigen vectors corresp to dif eigen values are orthogonal)

Def. Schur's Theorem

If A is an n x n **complex** matrix, then there must exist a **unitary** matrix U s.t. $U^{-1}AU = U^*AU = T$ is upper triangular. (The main diagonal of T are eigenvalues of A)

Proof.

- By induction on n ($n = 1$ is already upper triangular).
- Assume the theorem holds for $(n-1) \times (n-1)$ complex matrices.
- Let λ_1 be an eigenvalue of A with eigenvector w_1; assume $||w_1|| = 1$. Extend to a basis of $\mathbb{C}^n$.
- By Gram-Schmidt, construct an orthonormal basis $\{w_1, \ldots, w_n\}$ of $\mathbb{C}^n$.
- Let $U_1 = [w_1 \mid w_2 \mid \cdots \mid w_n]$ and note U is unitary.
- We analyze $U_1^* A U_1$:
  (1, 1)-entry: first row of U_1^* (i.e., w_1^*) times first column of AU_1 (i.e., Aw_1):
  $$w_1^*(Aw_1) = w_1^* \lambda_1 w_1 = \lambda_1 w_1^* w_1 = \lambda_1.$$
  $(i, 1)$-entry: ith row of U_1^* (i.e., w_i^*) times first column of AU_1 (i.e., Aw_1):
  $$w_i^*(Aw_1) = w_i^* \lambda_1 w_1 = \lambda_1 w_i^* w_1 = 0$$
  for $2 \leq i \leq n$.
- Thus, $U_1^* A U_1$ has the form $\begin{bmatrix} \lambda_1 & X_1 \\ 0 & M \end{bmatrix}$.
- By induction, there is a unitary $(n-1) \times (n-1)$ matrix W_1 such that $W_1^* M W_1 = T_1$ is upper triangular.
- Let $U_2 = \begin{bmatrix} 1 & 0 \\ 0 & W_1 \end{bmatrix}$.
- Then $U = U_1 U_2$ is unitary and
  $$U^* A U = U_2^*(U_1^* A U_1)U_2) = \begin{bmatrix} 1 & 0 \\ 0 & W_1^* \end{bmatrix} \begin{bmatrix} \lambda_1 & X_1 \\ 0 & M \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & W_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & X_1 W_1 \\ 0 & T_1 \end{bmatrix}$$
  is upper triangular.

Def. Spectral Theorem

Recall Let A be an n x n **real** matrix. A is **orthogonally** diagonalizable iff A is **symmetric**.

> Let A be an n x n **complex** matrix. A is **unitarily** diagonalizable iff A is **normal.**

Proof.

($\Longleftarrow$)

- Suppose A is unitarily diagonalizable.
- Then there exists a unitary matrix U and diagonal matrix D such that $U^* A U = D$.
- Observe that $DD^* = D^* D$ since D is a diagonal matrix.
- Since
  $$DD^* = (U^* A U)(U^* A U)^* = U^* A U U^* A^* U = U^*(AA^*)U$$
  and
  $$D^* D = (U^* A U)^*(U^* A U) = U^* A^* U U^* A U = U^*(A^* A)U$$
  we have $U^*(AA^*)U = U^*(A^* A)U$ implying $AA^* = A^* A$.
- Thus, A is a normal matrix.

(normal $\Longrightarrow$ unitarily diagonalizable)

- Suppose $AA^* = A^*A$.
- By Schur's theorem, there is a unitary matrix U such that $U^*AU = T$ is upper triangular.
- Note: T is also a normal matrix since

$$TT^* = (U^*AU)(U^*AU)^* = U^*AA^*U = U^*A^*AU = (U^*AU)^*(U^*AU) = T^*T.$$

- We show that every normal upper triangular $n \times n$ matrix T is a diagonal matrix by induction on n.
- It is trivially true for $n = 1$.
- Let $T = [t_{ij}]$. Equating the $(1,1)$-entries in TT^* and T^*T gives

$$|t_{11}|^2 + |t_{12}|^2 + \cdots + |t_{1n}|^2 = |t_{11}|^2.$$

- This implies that $t_{12} = t_{13} = \cdots = t_{1n} = 0$, thus,

$$T = \begin{bmatrix} t_{11} & 0 \\ 0 & T_1 \end{bmatrix}.$$

  Hence, $T^* = \begin{bmatrix} \bar{t}_{11} & 0 \\ 0 & T_1^* \end{bmatrix}$.
- Thus, $TT^* = T^*T$ implies that $T_1T_1^* = T_1T_1^*$.
- By the induction hypothesis, T_1 is diagonal, implying T is diagonal.

Lemma If A is **Hermitian**, then there exists a unitary matrix U such that U^*AU is diagonal.

Proof By Schur's thm, there exists an upper triangular matrix $T = U^*AU$.

Then $T^* = (U^*AU)^* = U^*A^*(U^*)^* = U^*AU = T \implies T^* = T$

Since T is both upper triangular and lower triangular, it must be diagonal

Note The converse is false.

E.g. $A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ is not Hermitian but it is unitarily diagonalizable.

W11 (9.4)

Def. Jordan Canonical Form

Every n x n complex matrix A is similar to a matrix of the following form (λ_i's are not necessarily distinct).

$$\begin{bmatrix} \lambda_1 & 1 & 0 & \cdots & 0 & & & & & & \\ 0 & \lambda_1 & 1 & \ddots & 0 & & & & & & \\ 0 & 0 & \lambda_1 & \ddots & 0 & & & & & & \\ \vdots & \ddots & \ddots & \ddots & 1 & & & & & & \\ 0 & 0 & 0 & \cdots & \lambda_1 & & & & & & \\ & & & & & \ddots & & & & & \\ & & & & & & \lambda_k & 1 & 0 & \cdots & 0 \\ & & & & & & 0 & \lambda_k & 1 & \ddots & 0 \\ & & & & & & 0 & 0 & \lambda_k & \ddots & 0 \\ & & & & & & \vdots & \ddots & \ddots & \ddots & 1 \\ & & & & & & 0 & 0 & 0 & \cdots & \lambda_k \end{bmatrix}$$

The matrix is composed of **Jordan blocks** along the diagonal.

Def. Block Matrix

A block matrix is one that has been partitioned into blocks.

- We can multiply block matrices (assuming sizes compatible):

  $\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} E & F \\ G & H \end{bmatrix} = \begin{bmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{bmatrix}$ Note that order matters!

- $\det \left(\begin{bmatrix} A & B \\ 0 & C \end{bmatrix} \right) = \det A \cdot \det C$

- A block diagonal matrix has square blocks down its diagonal. (JCF is block diagonal)

  Its eigen values are exactly those of the blocks, and $\det A = \det(A_1) \det(A_2) \dots \det(A_n)$

E.g. Find the eigen values of $A = \begin{bmatrix} 0 & 1 & * & * \\ -1 & 0 & * & * \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 4 \end{bmatrix}$

Observe that A can be partitioned into 2 x 2 blocks.

The top left block has eigs $\pm i$, and the bottom right block has 2 and 4.

Properties of Jordan Blocks

Let J be a k x k Jordan block $\begin{bmatrix} \lambda & 1 & 0 & \dots & 0 \\ 0 & \lambda & 1 & \ddots & \vdots \\ \vdots & \ddots & \lambda & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & \lambda \end{bmatrix}$

- J has eigen value λ with algebraic mult k

  Proof $p(x) = \det(J - xI) = (\lambda - x)^k \implies x = \lambda$ is a root with mult K

- J is not diagonalizable unless k = 1 since $\dim(E_\lambda) = 1$ (if k not = 1, the matrix is not diagonal)

  Proof $J - \lambda I = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & 0 & \ddots & 0 \\ \vdots & & \ddots & \ddots & 1 \\ 0 & \dots & \dots & 0 & 0 \end{bmatrix}$ so $E_\lambda = null(J - \lambda I) = sp\left(\begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}\right) = sp(e_1)$

- If P is the smallest positive integer such that $(J - \lambda I)^P = O$, then $P = k$

  I.e. $(J - \lambda I)^P = O$ but if $i < k$, then $(J - \lambda I)^i \neq O$

  Proof The 1's move up one line at a time when we take powers of $J - \lambda I$

  E.g. Let $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$ so k = 4

  $$A^2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad A^3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad A^4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Fact

$(J - \lambda I)e_i = e_{i-1}$ for $1 < i \leq k$

$(J - \lambda I)e_1 = 0 \implies e_1 \in null(J - \lambda I)$ note that e_1 is an eigen vector with rank 1

$(J - \lambda I)e_2 = e_1 \implies (J - \lambda I)^2 e_2 = (J - \lambda I)e_1 = \vec{0} \implies e_1, e_2 \in null(J - \lambda I)^2$

$\vdots$

$(J-\lambda I)^p e_p = \vec{0} \implies e_1, e_2, \ldots, e_p \in null(J-\lambda)^p$

Def. Generalized Eigenvector

For an n x n complex matrix A, a generalized eigenvector with rank p is a non-zero vector $\vec{x} \in \mathbb{C}^n$ corresponding to eigenvalue λ if $(A-\lambda I)^p \vec{x} = \vec{0}$ for some positive integer p.

Generalized eigenspace: set of all generalized eigenvectors for λ plus $\vec{0}$.
$\hat{E}_\lambda = \left\{ \vec{x} \in \mathbb{C}^n : (A-\lambda I)^p \vec{x} = \vec{0} \text{ for some } p \right\}$

Notes

- ordinary eigenvectors are generalized eigenvectors with rank 1.
- In general, if A is not diagonalizable, then $E_{\lambda_1} \oplus \ldots \oplus E_{\lambda_r} \subset \mathbb{C}^n$

  But the generalized eigenspaces form a basis: $\hat{E}_{\lambda_1} \oplus \ldots \oplus \hat{E}_{\lambda_r} = \mathbb{C}^n$
- $\dim(\hat{E}_\lambda) = \text{alg mult of } \lambda$

Why find generalized eigenvectors?

By forming P using generalized eigenvectors, we can turn A into JCF using $P^{-1}AP$

E.g. Find generalized eigenvectors for $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{bmatrix}$

Eigen values are 1, 1, 3

$\lambda = 3 \implies E_3 = sp\left(\begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} \right)$

$\lambda = 1 \implies E_1 = null\left(\begin{bmatrix} 0 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 3 \end{bmatrix} \right) = sp(e_1)$

When $p = 1$, $(A-\lambda I)\vec{x} = \vec{0} \implies \vec{x} = e_1$

When $p = 2$, $(A-\lambda I)^2 \vec{x} = \vec{0} \implies \vec{x} = null((A-\lambda I)^2) = null\left(\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 3 \end{bmatrix}^2 \right) = null\left(\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right)$

$$= \left\{ \begin{bmatrix} s \\ t \\ 0 \end{bmatrix} : s, t \in \mathbb{C} \right\} = sp(e_1, e_2) \text{ where } e_2 \text{ is a generalized eigen vector}$$

Thm (Similar to JCF)

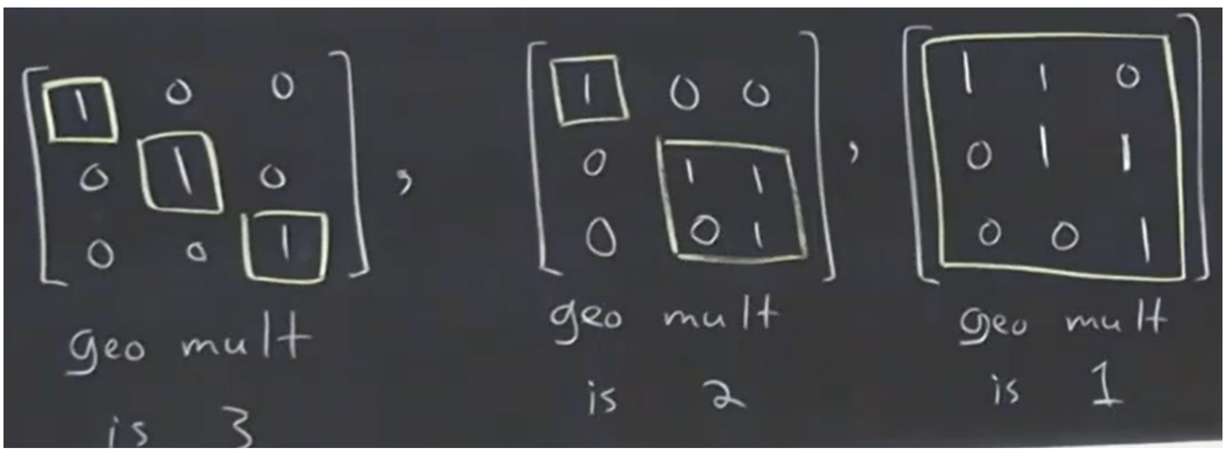
Every square complex matrix is similar to a JCF.

Properties

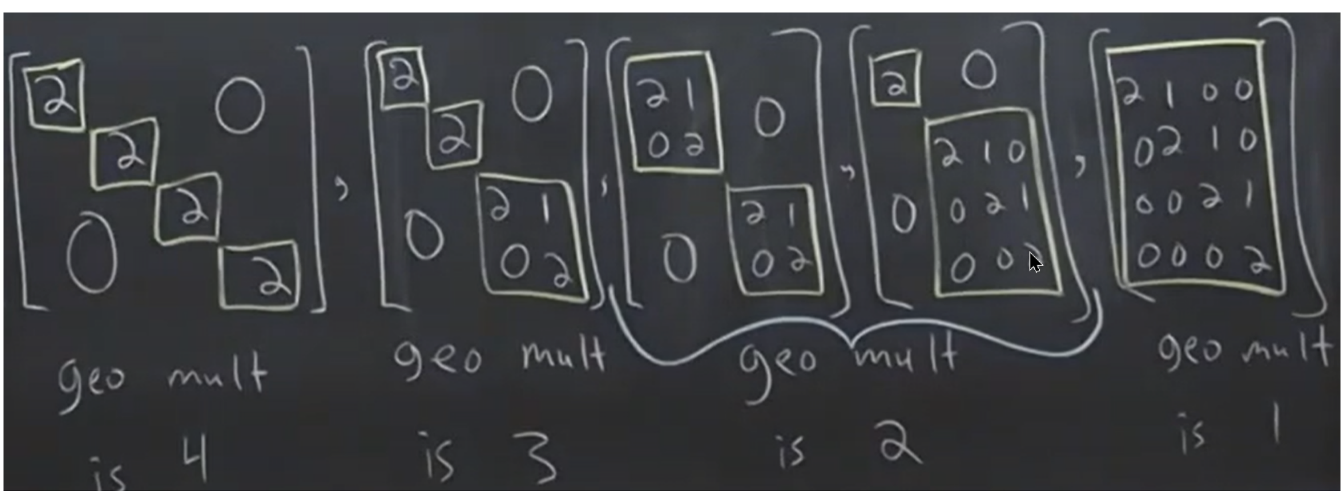
- JCF is unique (up to permutation of blocks)
- # blocks for λ is its geometric mult (each eigenvector of A generates a block)
- what happens for each λ is independent of other eigenvalues

E.g. If a 7 x 7 matrix A has eigenvalues {1, 1, 1, 2, 2, 2, 2}, what are the possible JCF's?

For $\lambda = 1$, the possible configurations are:



For $\lambda = 2$, the possible configurations are:

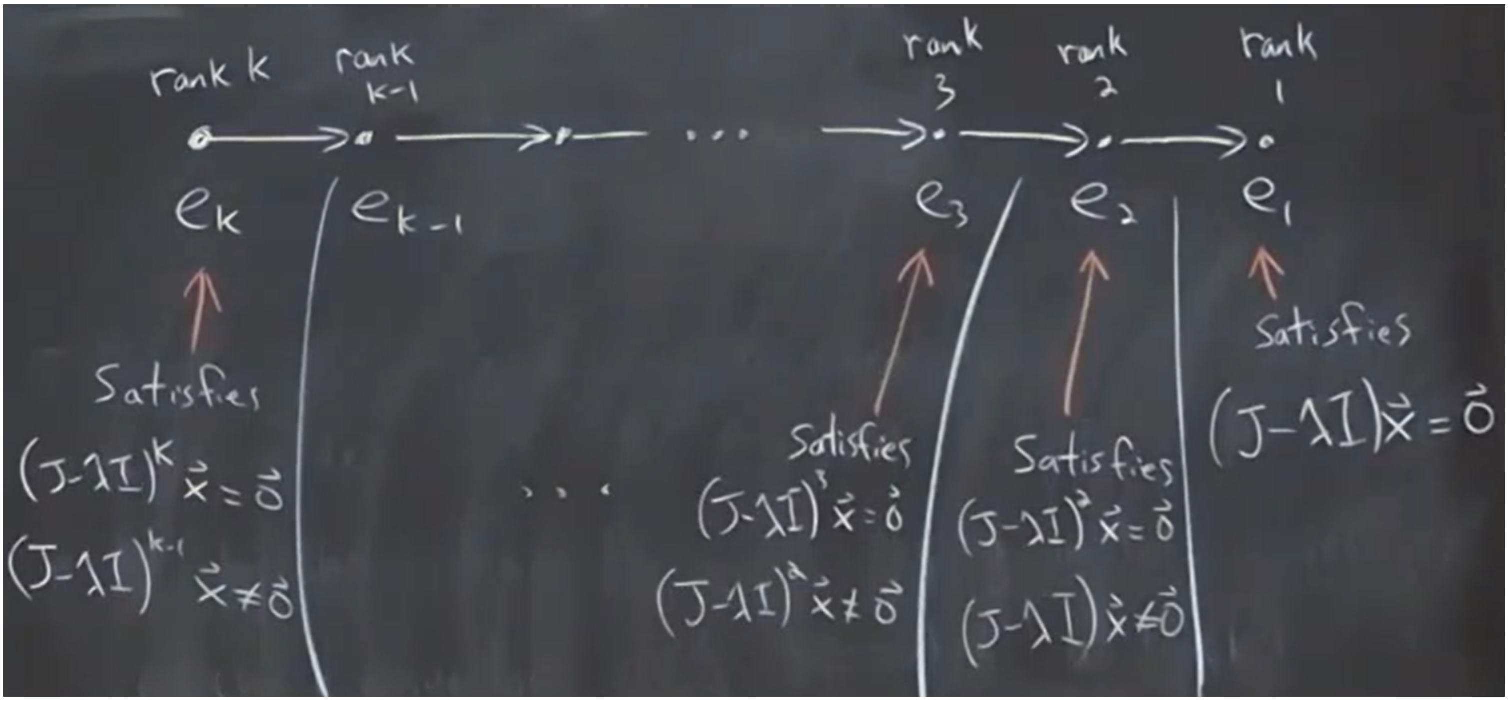


For small order, the geo mult gives the block structure, but it does not tell us the structure for alg mult = 4, geo mult = 2 (see above underbrace). How do we tell which structure to use?

Compute $\dim(null(A-\lambda I)^k)$ for $k \geq 1$, i.e. generalized eigenspaces.

Note Each k x k Jordan block has

- λ as eigen value with alg mult k and geo mult = $\dim(E_\lambda) = \dim(null(J-\lambda I)) = 1$
- a chain of k generalized eigen vectors $e_1, e_2, \ldots, e_k$

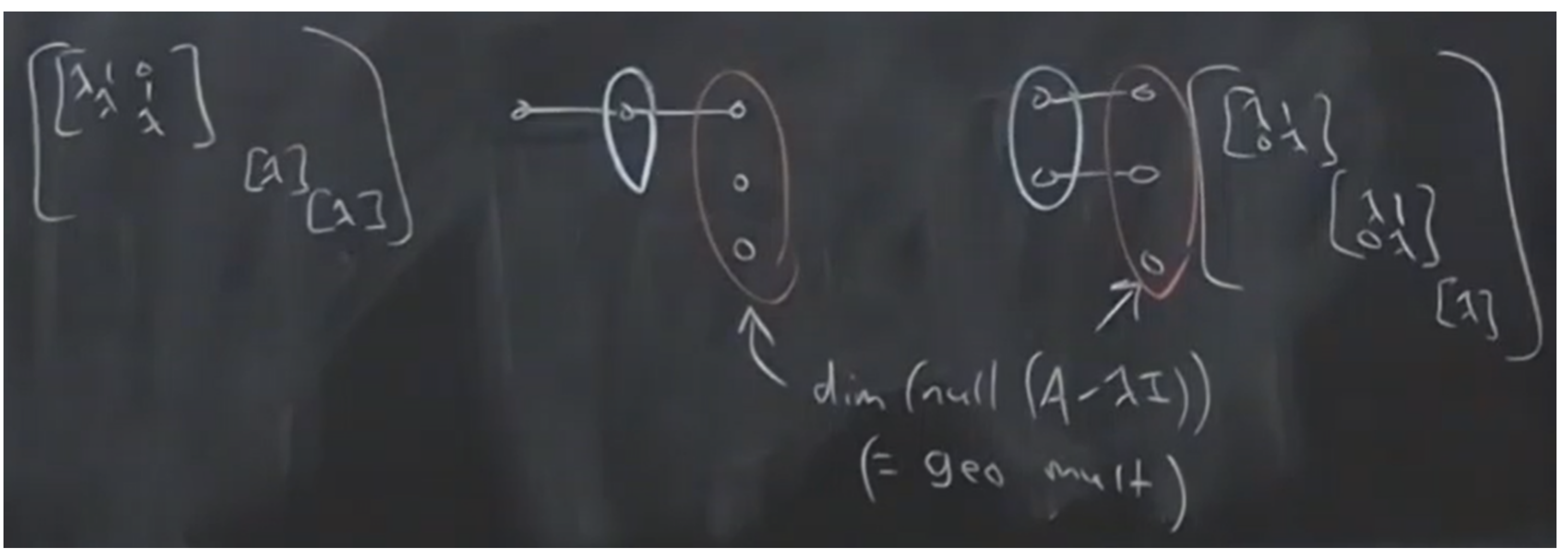


Note Solutions to $(A-\lambda I)^P \vec{x} = \vec{0}$ are also solutions to $(A-\lambda I)^{p+1} \vec{x} = \vec{0}$

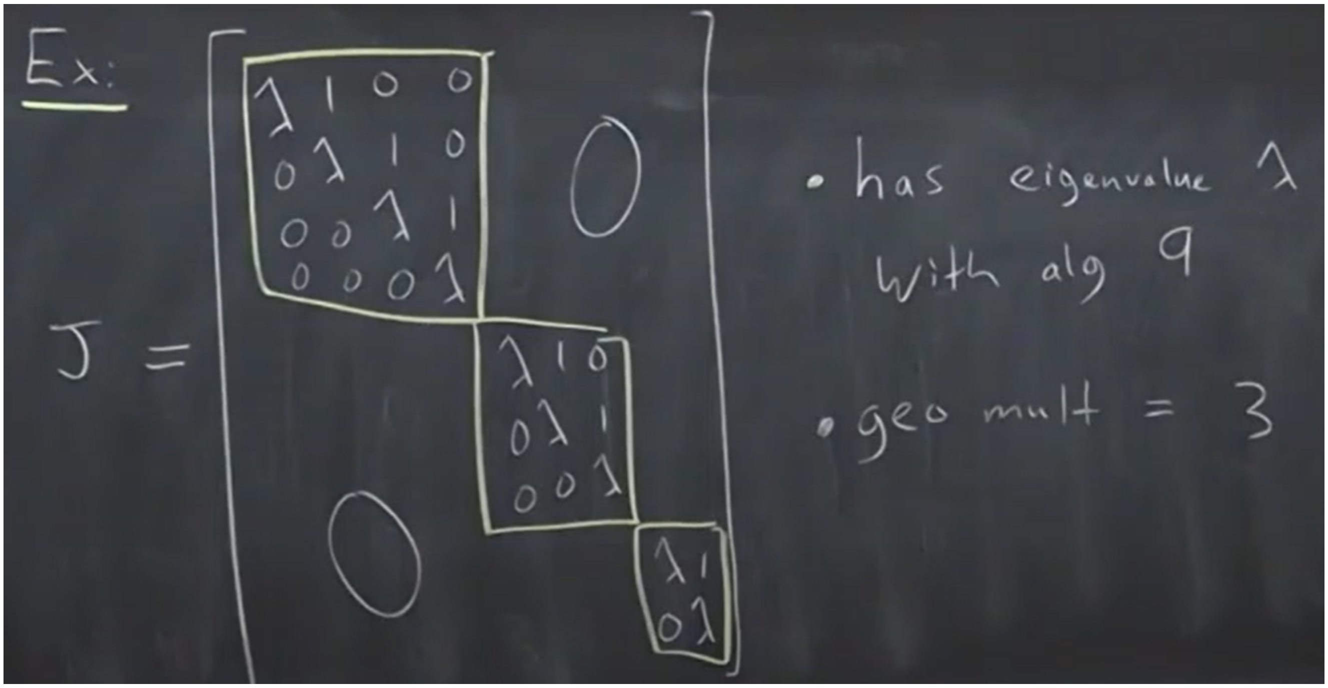
In general, an arbitrary matrix has a chain for each block in its JCF.

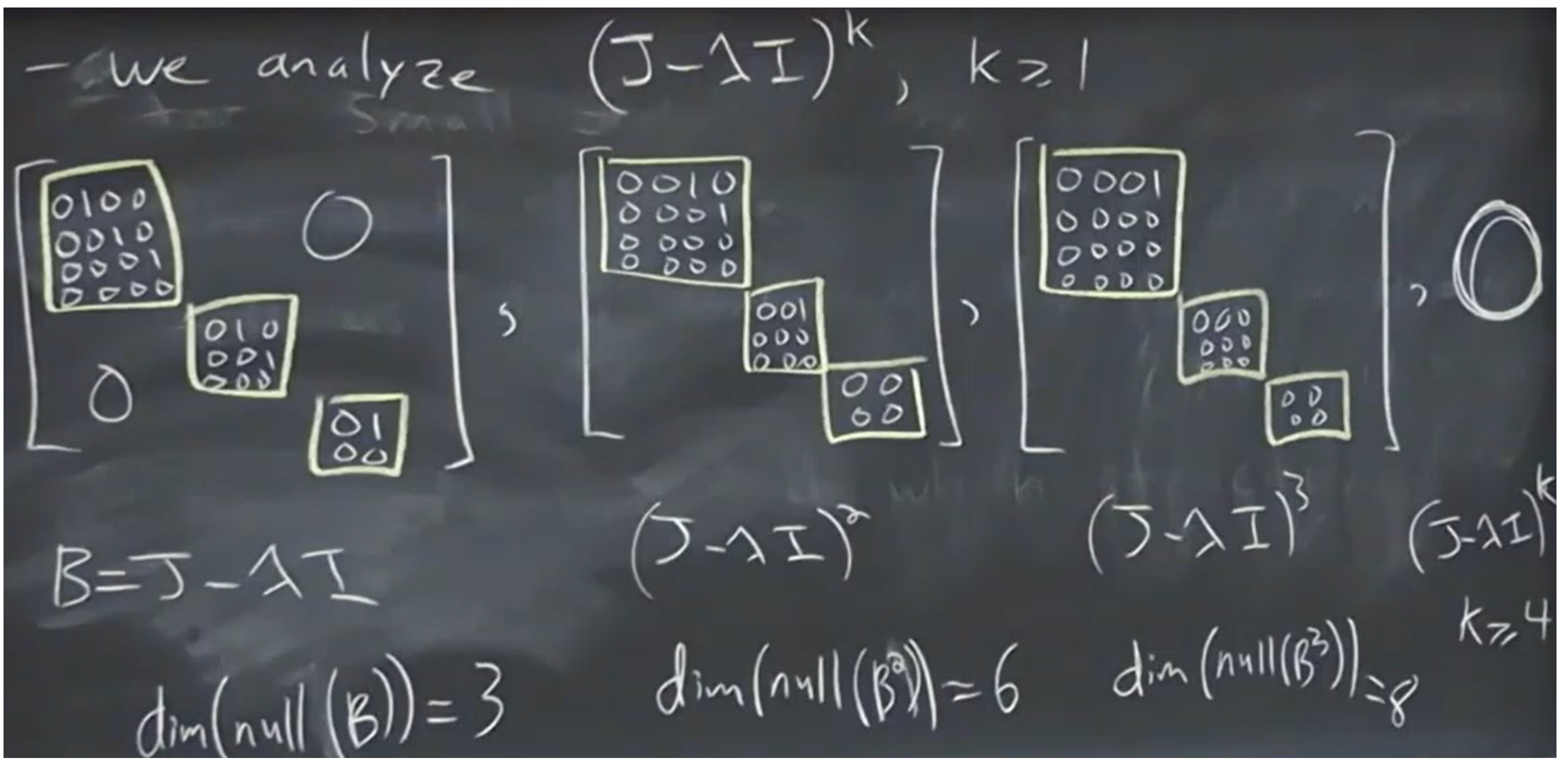
E.g. For a 5x5 matrix, an eigen value with alg mult = 5, geo mult = 3, we need 2 generalized eigenvectors:

- 3 chains – either 1 of them has 2 generalized eigenvectors, or 2 of them have 1 generalized eigenvector



E.g.





$\dim(null(B)) = 9$ for the last one

E.g.

$$A = \left(\begin{array}{ccc|cc} \lambda & 1 & 0 & & \\ 0 & \lambda & 1 & \multicolumn{2}{c}{\mathbf{0}} \\ 0 & 0 & \lambda & & \\ \hline & \mathbf{0} & & \lambda & 1 \\ & & & 0 & \lambda \end{array}\right)$$

For the 1st block, we have the chain $v_1 = \left[\begin{array}{c}1\\0\\0\\\hline 0\\0\end{array}\right] \to v_2 = \left[\begin{array}{c}0\\1\\0\\\hline 0\\0\end{array}\right] \to v_3 = \left[\begin{array}{c}0\\0\\1\\\hline 0\\0\end{array}\right]$

For the 2nd block, we have the chain $w_1 = \left[\begin{array}{c}0\\0\\0\\\hline 1\\0\end{array}\right] \to w_2 = \left[\begin{array}{c}0\\0\\0\\\hline 0\\1\end{array}\right]$

$v_1, w_1 \in null(A - \lambda I)$ are independent eigenvectors or rank 1

$v_2, w_2 \in null((A - \lambda I)^2), \notin null(A - \lambda I)$ are independent generalized eigenvectors of rank 2

$v_3 \in null((A - \lambda I)^3), \notin null((A - \lambda I)^2)$ is an independent generalized eigenvector of rank 3

Put them into a matrix $P = [v_1|v_2|v_3|w_1|w_2]$ and $J = P^{-1}AP$

> **Fact** $\dim(null(A - \lambda I)^k) =$ # ind. gen. eigenvectors of rank at most k

$\dim(null(A - \lambda I)) = 2$ since it has v_1, w_1

$\dim(null(A - \lambda I)^2) = 4$ since it has v_1, v_2, w_1, w_2

$\dim(null(A - \lambda I)^3) = 5$ since it has v_1, v_2, v_3, w_1, w_2

$\dim(null(A - \lambda I)^k) = 5 \quad k \geq 3$

Fact $\dim(null(A - \lambda I)^k) - \dim(null(A - \lambda I)^{k-1})$ = # ind. gen. eigenvectors of rank k = # Jordan blocks of size $\geq k$

# of Jordan blocks of size k = (# Jordan blocks of size $\geq k$) - (# Jordan blocks of size $\geq k+1$)

$= \dim(null((A - \lambda k)^k)) - \dim(null((A - \lambda k)^{k-1})) - \left(\dim(null((A - \lambda k)^{k+1})) - \dim(null((A - \lambda k)^k))\right)$

$= 2\dim(null(A - \lambda I)^k) - \dim(null(A - \lambda I)^{k-1}) - \dim(null(A - \lambda I)^{k+1})$

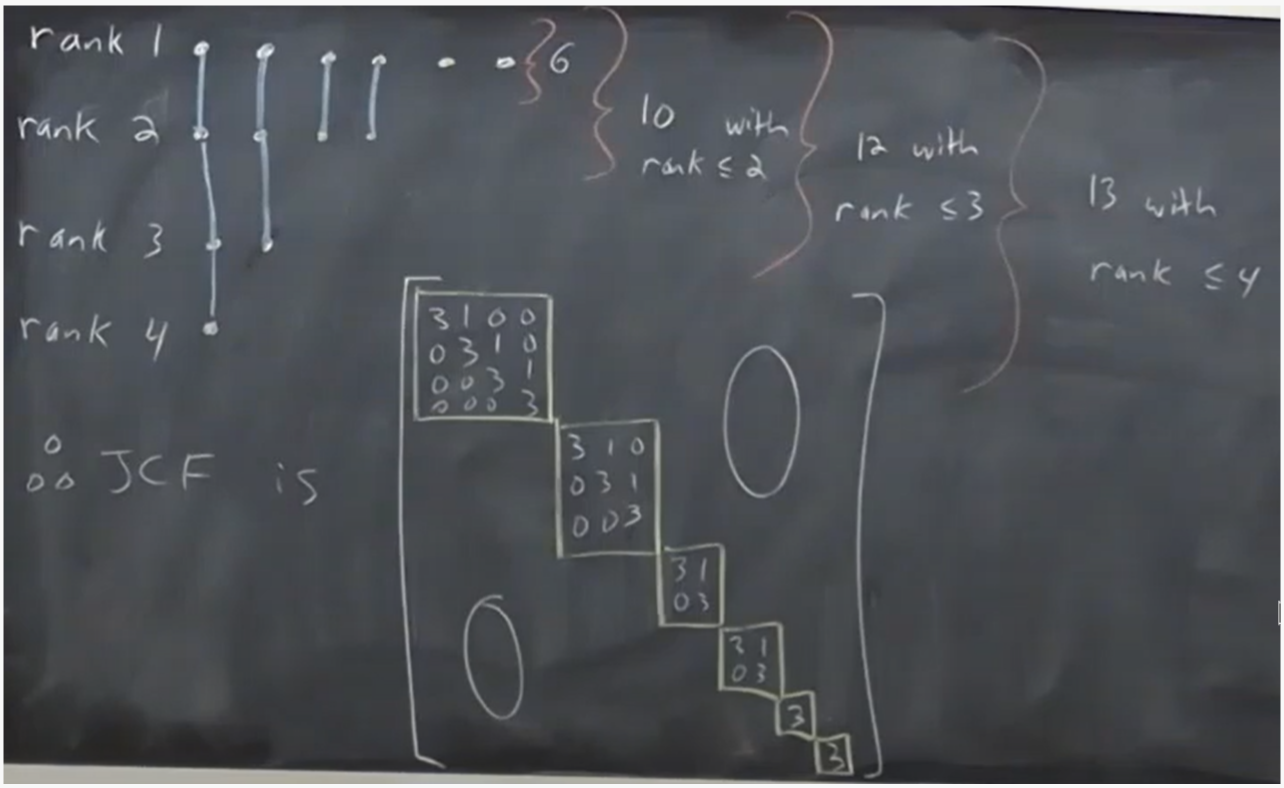
E.g. Find Jordan Form of A if $p(\lambda) = (\lambda - 3)^{13}$ and suppose:

$\dim(null((A-3I)^1)) = 6$ so # ind. gen. eigenvectors of rank at most 1 = 6

$\dim(null((A-3I)^2)) = 10$

$\dim(null((A-3I)^3)) = 12$

$\dim(null((A-3I)^4)) = 13$

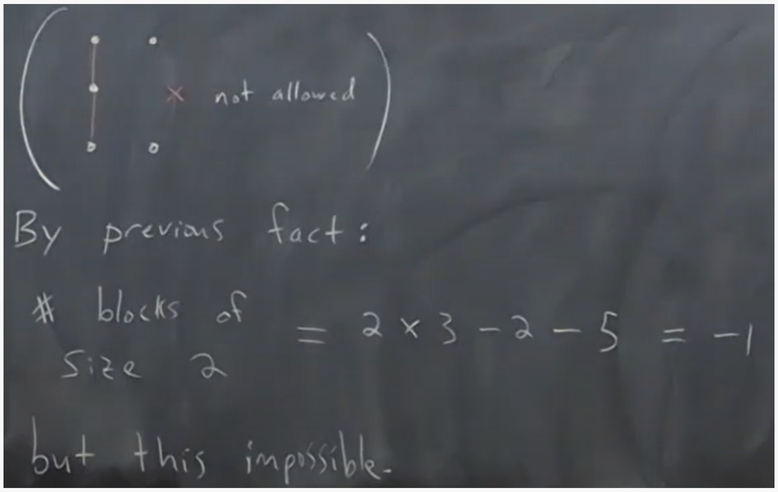
> So A is 13 x 13, $\lambda = 3$ with alg mult 13
>
> Obtain 6 chains, each corresponding to a Jordan block (each eigen vector produces a Jordan block), and the length of the chain = size of the block
>
> 
>
>
> Use formula: $2\dim(null(A-\lambda I)^k) - \dim(null(A-\lambda I)^{k-1}) - \dim(null(A-\lambda I)^{k+1})$
>
> $\dim(null((A-3I)^1)) = 6$ 2(6) - 0 - 10 = 2 blocks of size 1
>
> $\dim(null((A-3I)^2)) = 10$ 2(10) - 6 - 12 = 2 blocks of size 2
>
> $\dim(null((A-3I)^3)) = 12$ 2(12) - 10 - 13 = 1 block of size 3
>
> $\dim(null((A-3I)^4)) = 13$ 2(13) - 12 - 13 = 1 block of size 4
>
> $\dim(null((A-3I)^5)) = 13$ 2(13) - 13 - 13 = 0 block of size 5

E.g. Is it possible for a matrix to have $p(\lambda) = (\lambda - 7)^5$ and

$\dim(null((A - \lambda I)^2)) = 2$

$\dim(null((A - \lambda I)^3)) = 3$

$\dim(null((A - \lambda I)^4)) = 5$



Def. Cayley Hamilton Thm

> Every square matrix satisfies its own characteristic polynomial.

E.g. Let $A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$ where $p(\lambda) = \lambda^2 - 5\lambda - 2 = 0$. Find A^{-1} (in terms of powers of A).

Then by Cayley Hamilton, $p(A) = A^2 - 5A - 2I = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$

$$\implies 2I = A^2 - 5A \implies 2A^{-1} = A - 5I \implies A^{-1} = \frac{1}{2}(A - 5I) = \frac{1}{2}\begin{bmatrix} -4 & 2 \\ 3 & -1 \end{bmatrix}$$

E.g. Let $A = \begin{bmatrix} 2 & 1 \\ 5 & -2 \end{bmatrix}$. Find A^6.

$$p(\lambda) = \lambda^2 - 9 \implies p(A) = A^2 - 9I = 0 \implies A^2 = 9I \implies (A^2)^3 = (9I)^3 \implies A^6 = 729I = \begin{bmatrix} 729 & 0 \\ 0 & 729 \end{bmatrix}$$