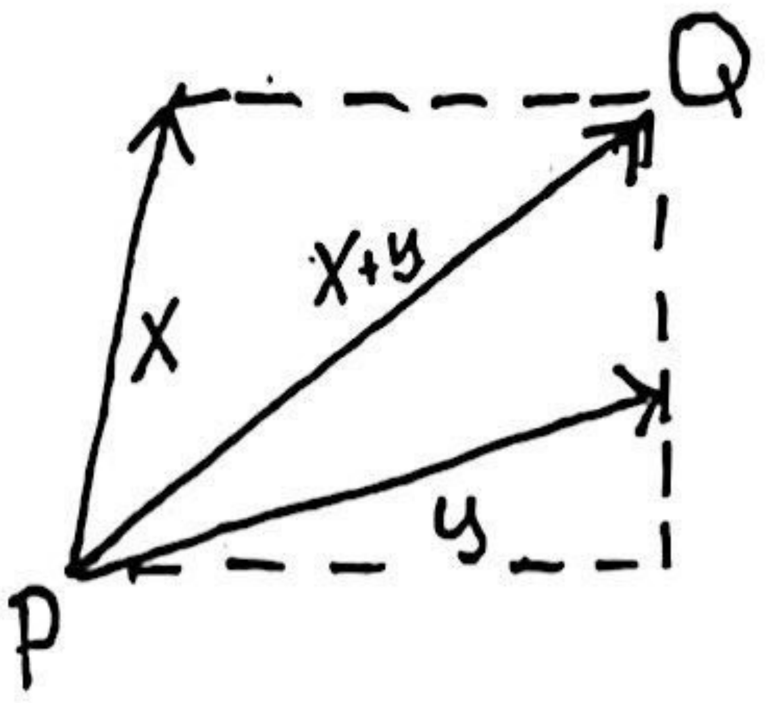
Chapter 1: Vector Spaces

1.1: Introduction

vector: an arrow with both magnitude & direction



Parallelogram Law for Vector Addition: the sum of two vectors (x & y) that begin at the same point (P) is the diagonal of the parallelogram with the vectors as sides

Properties of Vector Addition & Scalar Multiplication

For all vectors x & y:

1. $x+y = x+y$ → commutativity of addition
2. $(x+y)+z = x+(y+z)$ → associativity of addition
3. There exists a zero-vector st $x+0=x$
4. For each x there exists a y st $x+y=0$ → addative inverse
5. $1x=x$ → unit property
6. For each pair of real numbers a & b: $(ab)x = a(bx)$
7. $a(x+y) = ax+ay$
8. $(a+b)x = ax+bx$

Theorem 1.3: Subspaces

Let V be a vectorspace & let W be a subset of V

W is a subspace of V iff

1. $0 \in W$
2. $x+y \in W$ for $x, y \in V$
3. $cx \in W$ for $c \in F$ & $x \in V$

Vector Space: A set over (

addition & scalar multipli

Theorem 1.1: Cancellation La

If $x, y,$ & z are vectors in

$x+z = y+z$, then $x=y$

Theorem 1.2: In any vector s

a. zero vector: $0x = 0$ for a

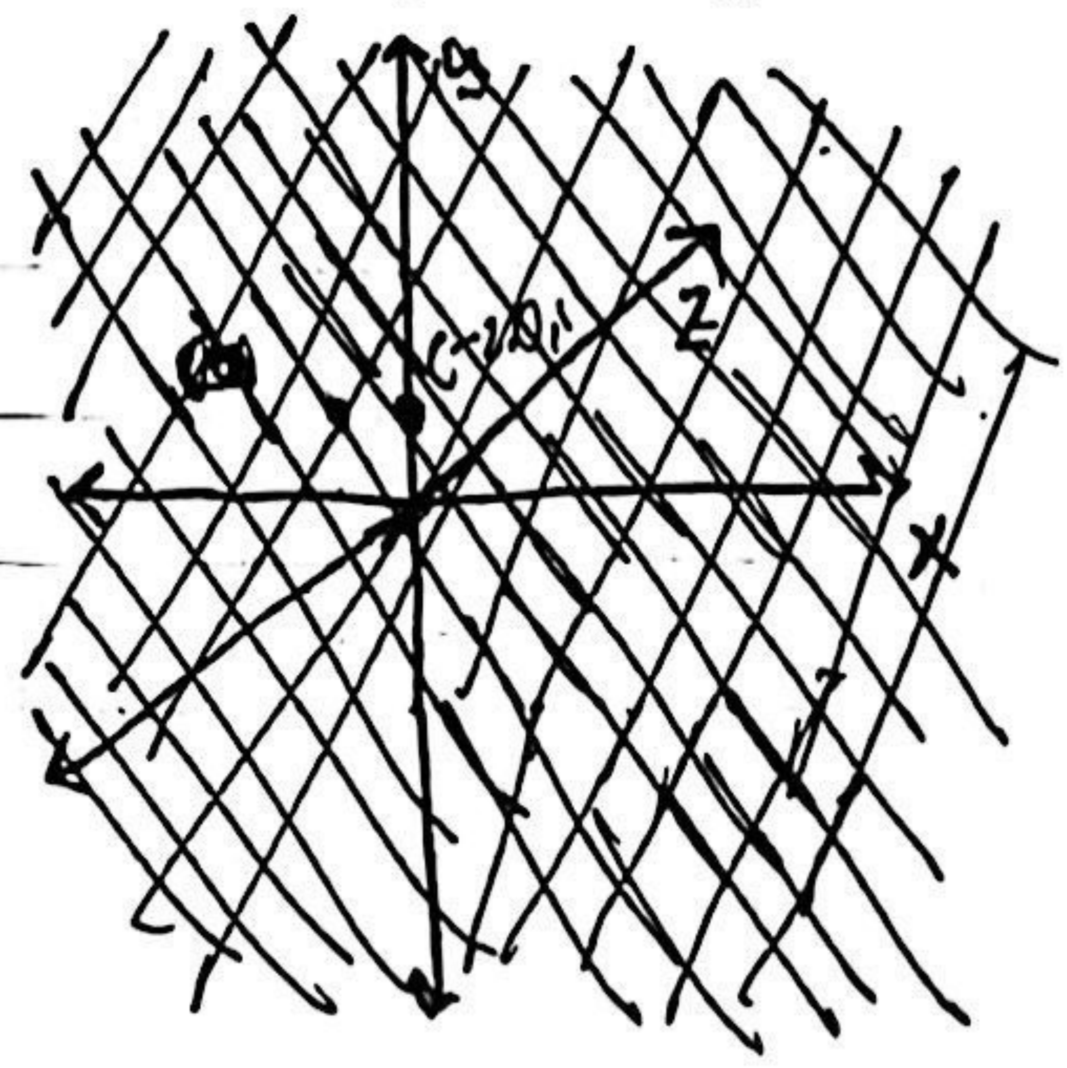
b. addative inverse: $(-a)x = -(ax) = a(-x)$

in general, the union of two subspaces will be a VS if one contains the other

1.4: Linear Combinations & Systems of Linear Equations

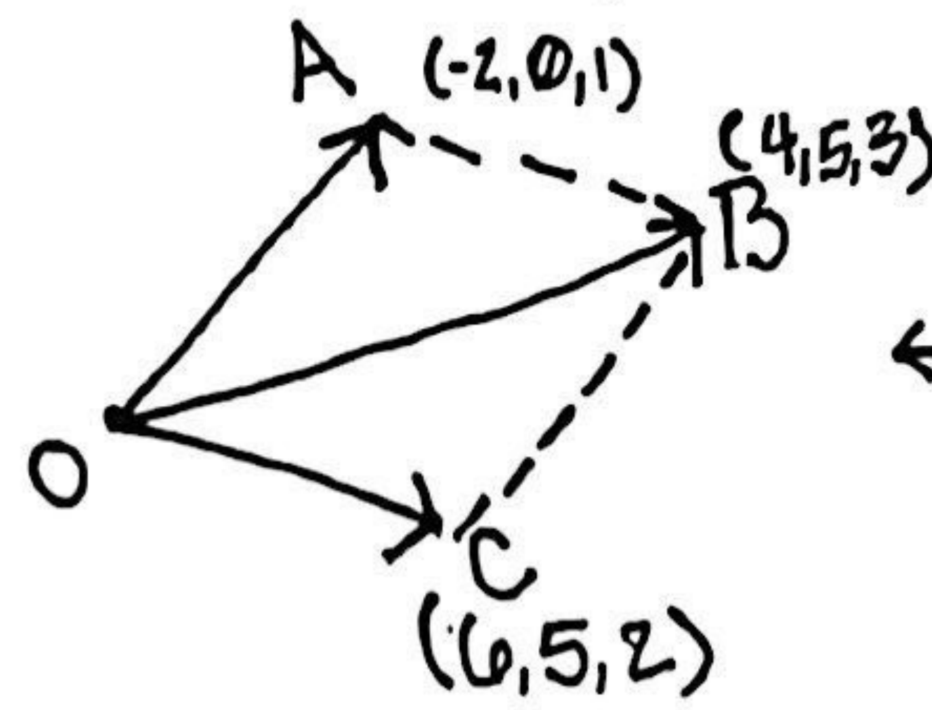
Let A & B points having coordinates $(-2, 0, 1)$ & $(4, 5, 3)$

The endpoint C of the vector emanating from the origin with the same direction as the line connecting A & B has the coordinates $(6, 5, 2)$



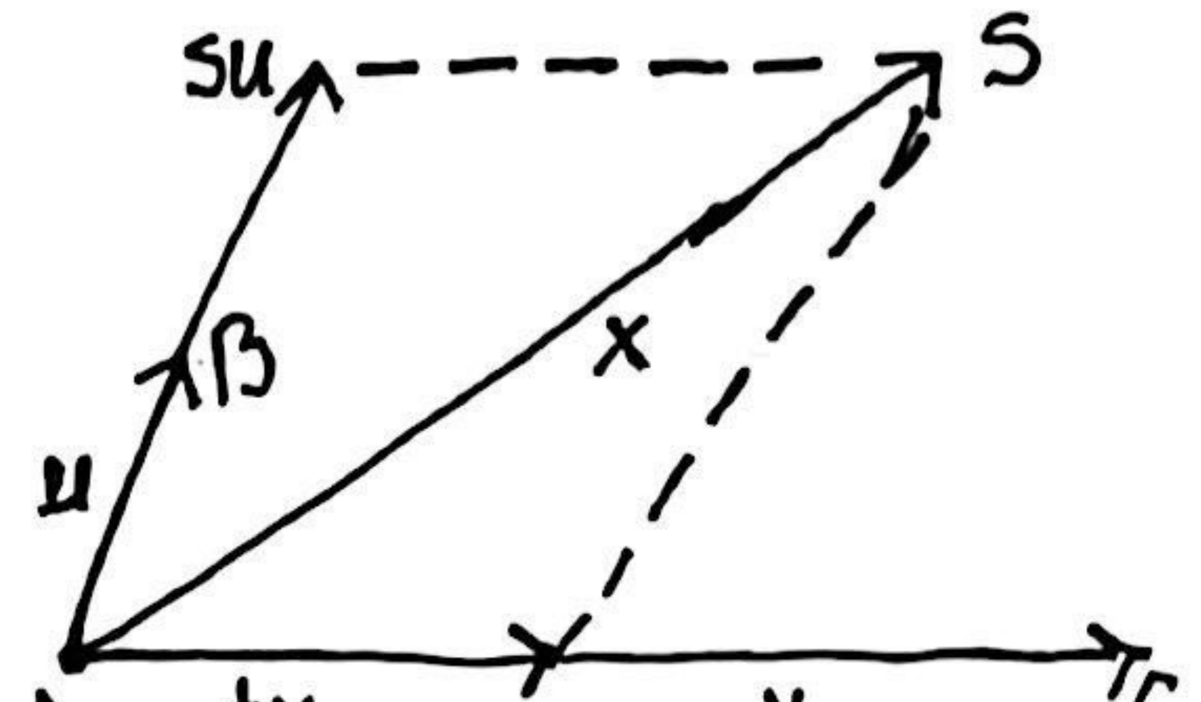
Thus, the equation of a line through A & B is given by:

$X = (-2, 0, 1) + t(6, 5, 2)$



← simplified depiction

A, B, & C are not colinear but they are coplanar & the plane containing all 3 is unique



→ Any point in this plane is an endpoint S of a vector x begining at A of the form:

$su + tv$

for some real numbers s & t

↳ By treating A as our origin & the vectors to B & C as our axises we can create a 2D plane the equation of which is given by

$x = A + su + tv$

↑ any point in the plane

In the standard Cartesian sense

s & t are similar to x & y except they originate at A instead of the zero vector

1.4: Linear Combos & Systems of Linear Combos Cont...

Recall: For any vectors A, B, & C the equation of the plane containing all 3 points is given by

$X = A + su + tv$ → When A is the origin this equation simplifies to: $X = su + tv$

↳ Where s & t are scalars & u & v are vectors

Let V be a vector space
& S be a nonempty subset of V

Linear Combination: any vector $v \in V$ that can be expressed in the form

$$v = a_1u_1 + a_2u_2 + \dots + a_nu_n$$

Where $a_1, a_2, \dots, a_n$ are scalars in F & $u_1, u_2, \dots, u_n$ are vectors in S

↑ coefficients of the linear combination

↳ since $0v = 0$ the zero vector is a linear combonation of any nonempty subset of V

Example: Prove that $2x^3 - 2x^2 + 12x - 6$ is a linear combo of $x^3 - 2x^2 - 5x - 3$ & $3x^3 - 5x^2 - 4x - 9$

↳ find any values of a & b such that

$$2x^3 - 2x^2 + 12x - 6 = a(x^3 - 2x^2 - 5x - 3) + b(3x^3 - 5x^2 - 4x - 9)$$
$$= (a + 3b)x^3 + (-2a - 5b)x^2 + (-5a - 4b)x + (-3a - 9b)$$

Rewriting as a system yields: ↵

$a + 3b = 2$ → add multiples of the first equation to the others to eliminate a
$-2a - 5b = -2$
$-5a - 4b = 12$
$-3a - 9b = -6$

$$\left.\begin{aligned} a + 3b &= 2 \\ b &= 2 \\ 11b &= 22 \\ 0b &= 0 \end{aligned}\right\} \rightarrow \begin{aligned} a &= -4 \\ b &= 2 \\ 0 &= 0 \\ 0 &= 0 \end{aligned}$$

$$2x^3 - 2x^2 + 12x - 6 = -4(x^3 - 2x^2 - 5x - 3) + 2(3x^3 - 5x^2 - 4x - 9)$$

↳ Linear Combonation!

1.5: Linear Dependence & Independence Cont...

Properties of Linearly Independent Sets:

→ The empty set is linearly Independent

→ A singlton set is linearly Independent

Theorem 1.6: Linear dependence of subsets

Let V be a vector space & let $S_1 \subseteq S_2 \subseteq V$. If S_1 is linearly dependent then S_2 is also linearly dependent

conversely...

If S_2 is linearily Independent, then S_1 is also linearly Independent

Theorem 1.7: Span & Linear Independence

Let S be a linearly independent subset of a vectorspace V, & let v be a vector within V that is not in S. Then, $S \cup \{v\}$ is linearly dependent iff $v \in span(S)$

If S is a generating set for a subspace W & no proper subset of S is a generating set of W, then S must be linearly independent meaning that every vector in W has a unique representation as a linear combination of the vectors in S

1.6: Bases & Dimension:

Basis: A linearly Independent, spanning set of a vectorspace → the vectors in this set "form a basis"

Example: In F^n let $e_1 = (1,0,0,\cdots,0)$

$\{e_1, e_2, \cdots, e_n\}$ is a basis for F^n

$e_2 = (0,1,0,\cdots,0) \cdots$

$e_n = (0,0,\cdots,0,1)$

↳ this is called the standard basis for F^n

The standard basis for $P_n(F)$ is $\{1, x, x^2, \cdots, x^n\}$

Theorem 1.8: Let V be a vectorspace & $u_1, u_2, \cdots, u_n$ be distinct vectors in V. Then $\beta = \{u_1, u_2, \cdots, u_n\}$ is a basis for V iff $v \in V$ can be uniquely expressed as a linear combination of vectors in β in the form:

$$v = a_1u_1 + a_2u_2 + \cdots + a_nu_n$$

↳ where $a_1, a_2, \cdots, a_n$ are unique scalars

Theorem 1.9: Finite Bases

Let V be a vectorspace generated by a finite set S, then some subset of S is a basis for V. Hence, V has a finite basis.

↳ A finite spanning set for V can be reduced to a basis for V

1.6: Bases & Dimension Cont...

Example: Let $S = \{(2,-3,5), (8,-12,20), (1,0,-2), (0,2,-1), (7,2,0)\}$
Show that S generates R^3

↳ Start by selecting any nonzero vector in S $(2,-3,5)$ to be in the basis, & since $4(2,-3,5) = (8,-12,20)$ the two vectors are linearly dependent & thus $(8,-12,20)$ can exclude $(8,-12,20)$ from our basis

↳ Next, comparing $\{(2,-3,5), (1,0,-2)\}$ it is clear the two vectors are linearly Independent

Theorem 1.10 Replacement Theorem

Let V be a vector space that is generated by a set G containing exactly n vectors, & let L be a linearly Independent subset of V containing exactly m vectors Then, $m \leq n$ & there exists a subset H of G that Contains exactly $n-m$ vectors such that $L \cup H$ generates the vector space V

↳ If V is a vectorspace having a finite basis, Then, all bases for V are finite & they all contain the same number of Vectors

Dimension: the number of vectors in the basis of a given vector space

In terms of dimension the first conclusion of the replacement theorem states that if V is a finitely dimensional vector space then no linearly Independent subset of V can contain more than dim(V) vectors

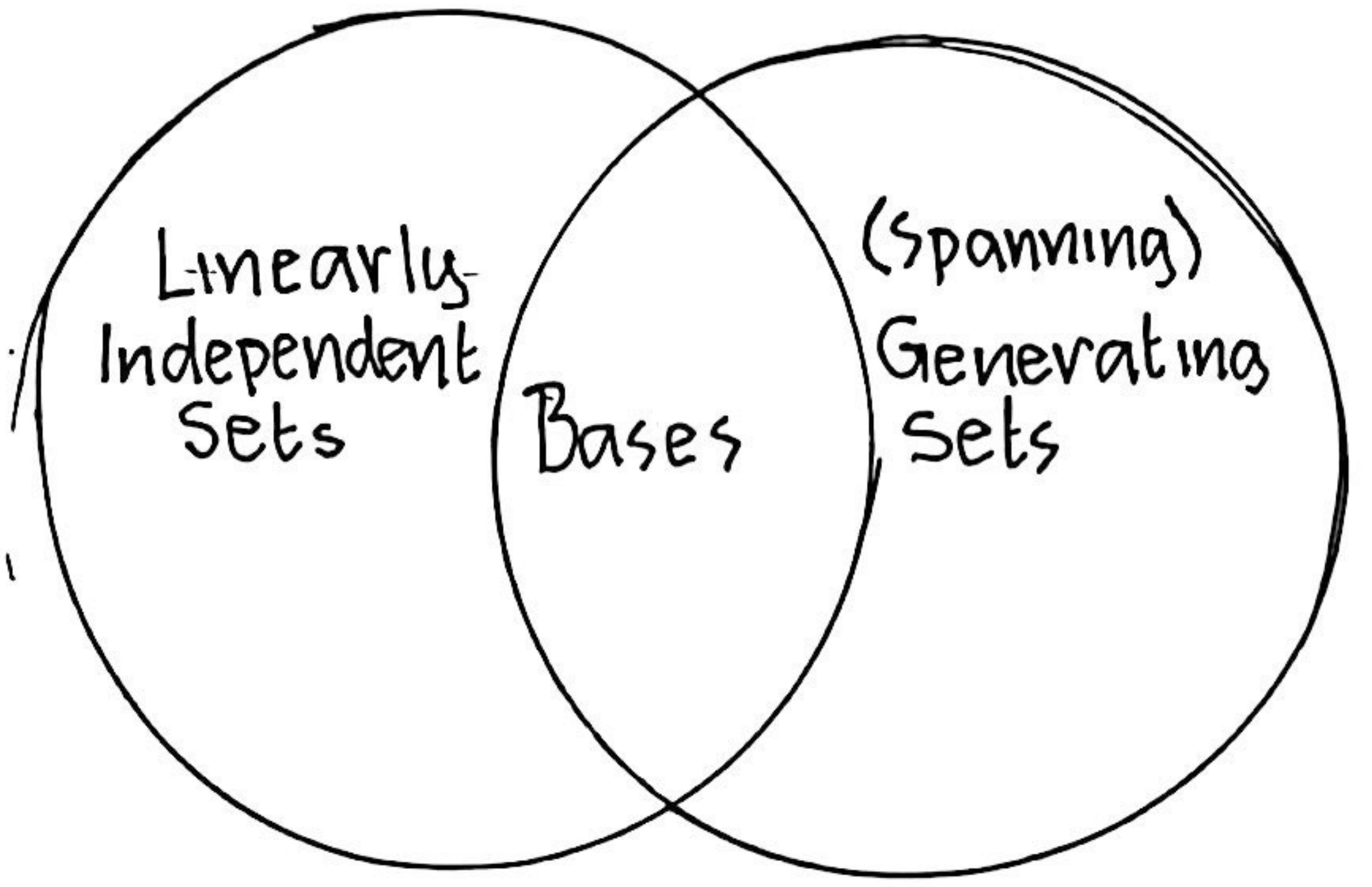
1.6: Bases & Dimension cont...

Properties of Dimension: Let V be a vectorspace with dimension n:

a. Any generating set for V contains at least n vectors & any generating set that contains exactly n vectors is a basis for V

b. Any linearly independent subset of V that contains exactly n vectors is a basis for V

c. Every linearly independent subset of V can be expanded to a basis of V, that is if L is a linearly independent subset of V, then there is a basis β of V st $L \subseteq \beta$



Summary:

1. A basis for a vector space V is a linearly independent spanning subset
2. The number of vectors in a basis is called the dimension
3. Let $\dim(V)=n$, there is no linearly independent subset of V containing more than n vectors

Theorem 1.11: Dimension of Subspaces:

Let W be a subspace of a finitely dimensonal vectorspace V. Then, W is finitely dimensional & $\dim(W) \leq \dim(V)$, & if $\dim(W) = \dim(V)$, then $W = V$

Example: Let $W = \{(a_1, a_2, a_3, a_4, a_5) \in F^5 : a_1 + a_3 + a_5 = 0,\ a_2 = a_4\}$

↳ W is a subspace of F^5 with a basis of:

$\{(-1,0,1,0,0), (-1,0,0,0,1), (0,1,0,1,0)\} \rightarrow \dim(W) = 3$

1.6: Example of Subspace Dimension

Consider the Vector space of all polynomials with degree less than or equal to 18, W. of the form $a_{18}x^{18} + a_{16}x^{16} + \cdots + a_2x^2 + a_0$

↳ A basis for W is given by: $\{1, x^2, \ldots, x^{16}, x^{18}\}$

note: this basis is a subset of the standard basis of $P_{18}(F)$

Lagrange Interpolation Formula:

Interpolation → the process of estimating intermediate values known is called interpolation

From the Replacement Theorem we get the following formula:

$$f_i(x) = \frac{(x-c_0)\cdots(x-c_{i-1})(x-c_{i+1})\cdots(x-c_n)}{(c_i-c_0)\cdots(c_i-c_{i-1})(c_i-c_{i+1})\cdots(c_i-c_n)} = \prod_{k=0}^{n} \frac{x-c_i}{c_i-c_k}$$

↳ The polynomials $f_0(x), f_1(x), \ldots, f_n(x)$ are called the **Lagrange Polynomials**

each $f_i(x)$ is a polynomial with degree n meaning it is in $P_n(F)$

↳ Expressing $f_i(x)$ as a polynomial function $f_i: F \to F$ yields: $f_i(c_j) = \begin{cases} 0 & \text{if } i \neq j \\ 1 & \text{if } i = j \end{cases}$

Using the polynomial function above, one can show that $\beta = \{f_0, f_1, \ldots, f_n\}$ is a linearly independent subset of $P_n(F)$ → Not only is β linearly independent, but also a basis of $P_n(F)$

↳ this means every polynomial g in $P_n(F)$ is a linear combination of polynomials in β, say

$g = \sum_{i=0}^{n} b_i f_i$ Then, $g = \sum_{i=0}^{n} g(c_i) f_i$

Is the unique representation of g as a linear combination of β

1.6: Example of Lagrange Interpolation:
Construct the real polynomial of degree 2 whose graph contains the points (1, 8), (2, 5), & (3, -4)
↳ Begin by setting up the Lagrange Polynomial

$$f_0(x) = \frac{(x-2)(x-3)}{(1-2)(1-3)} = \frac{1}{2}(x^2 - 5x + 6)$$

$$f_1(x) = \frac{(x-1)(x-3)}{(2-1)(2-3)} = -1(x^2 - 4x + 3)$$

$$f_2(x) = \frac{(x-1)(x-3)}{(3-1)(3-2)} = \frac{1}{2}(x^2 - 3x + 2)$$

Next set up the Interpolation from these polynomials

$$g(x) = \sum_{i=0}^{2} b_i f_i(x) = 8f_0(x) + 5f_1(x) - 4f_2(x)$$

$$= 4(x^2 - 5x + 6) - 5(x^2 - 4x + 3) - 2(x^2 - 3x + 2)$$

$$= -3x^2 + 6x + 5$$

Note: If $f \in P_n(F)$ & $f(c_i) = 0$ for $n+1$ distinct scalars $c_0, c_1, \ldots, c_n$ in F, then f is the zero function

Chapter 2: Matricies & Linear Transformations

2.1: Linear Transformations, Null Spaces, & Ranges:

Recall (from 165):

Linear Transformation: A function T with domain V & codomain W st for all $x, y \in V$ & $c \in F$

1. $T(x+y) = T(x) + T(y)$ &
6. $T(cx) = cT(x)$

denoted $T: V \to W$

Properties of Linear Transformations

1. $T(0) = 0$
2. $T(cx+y) = cT(x) + T(y)$ for all $x, y \in V$ ← This property easily tests wether or not a given Transformation is linear
3. $T(x-y) = T(x) - T(y)$ for all $x, y \in V$
4. for $x_1, x_2, \cdots, x_n \in V$ & $a_1, a_2, \cdots, a_n \in F$

$$T\left(\sum_{i=1}^{n} a_i x_i\right) = \sum_{i=1}^{n} a_i T(x_i)$$

Example: Let

$T: R^2 \to R^2$ by $T(a_1, a_2) = (2a_1 + a_2, a_1)$

1. let $c \in R$ & $x, y \in R^2$

$x = (b_1, b_2)$

$y = (d_1, d_2)$

2. Since $cx + y = (cb_1 + d_1, cb_2 + d_2)$

$T(cx+y) = (2(cb_1 + d_1) + cb_2 + d_2, cb_1 + d_1)$

→ Shes linear!

Note: Most common geometric transformations are linear!

$T: A \to B$ → ~~[illegible] the image~~

T(x) is called the image

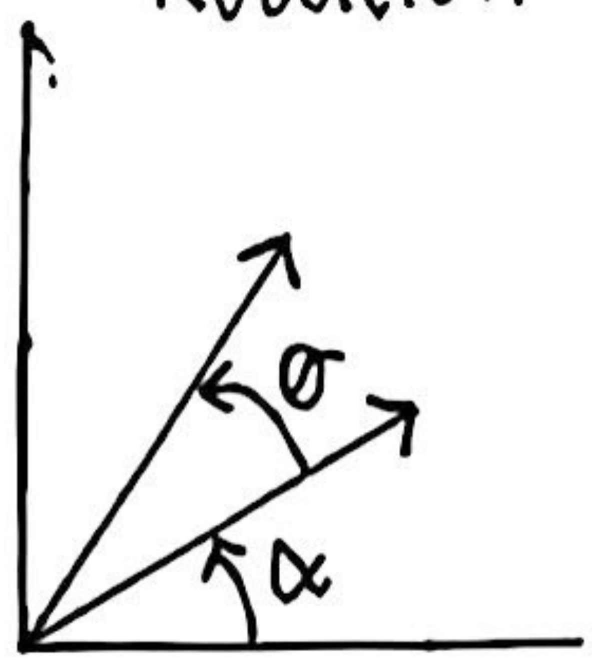
x is the preimage

~~elements of a are called preimages~~

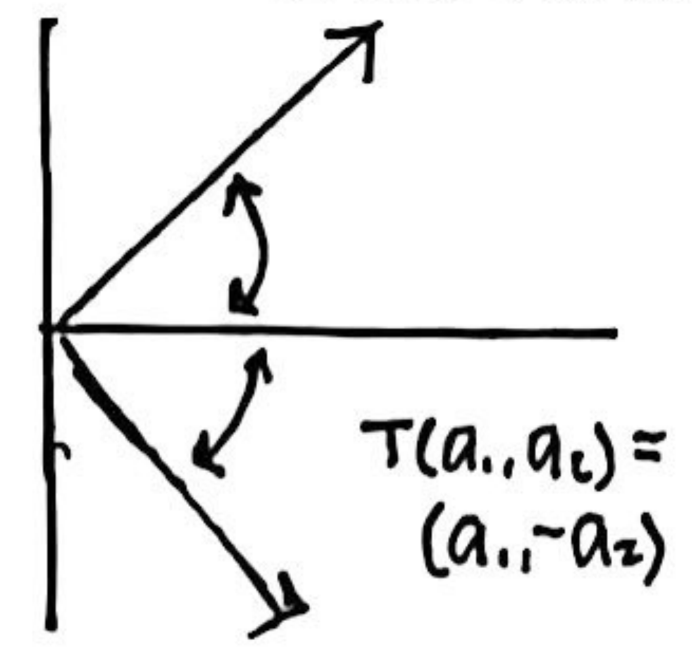
2.1: Linear Transformations, Nullspaces, & Ranges

More Examples of Linear Transformations:

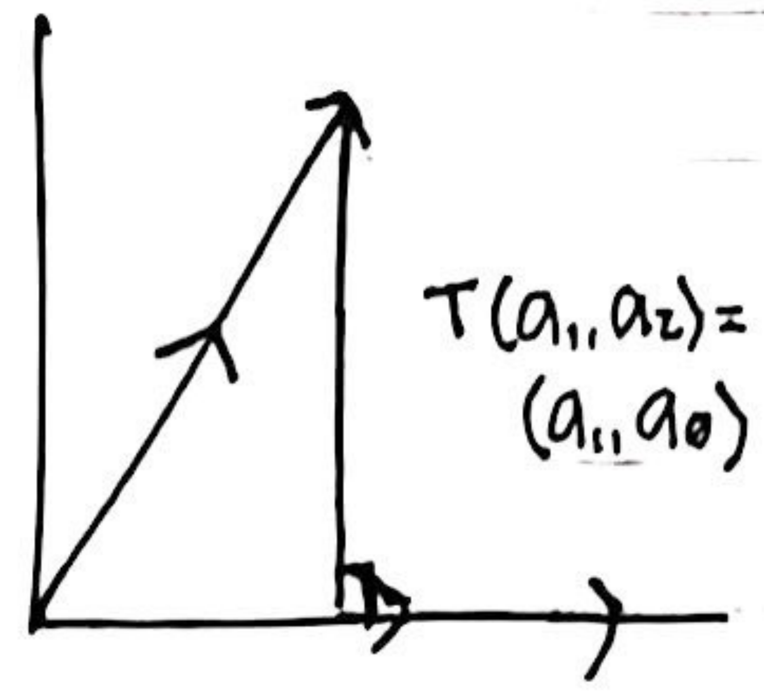
Rotation



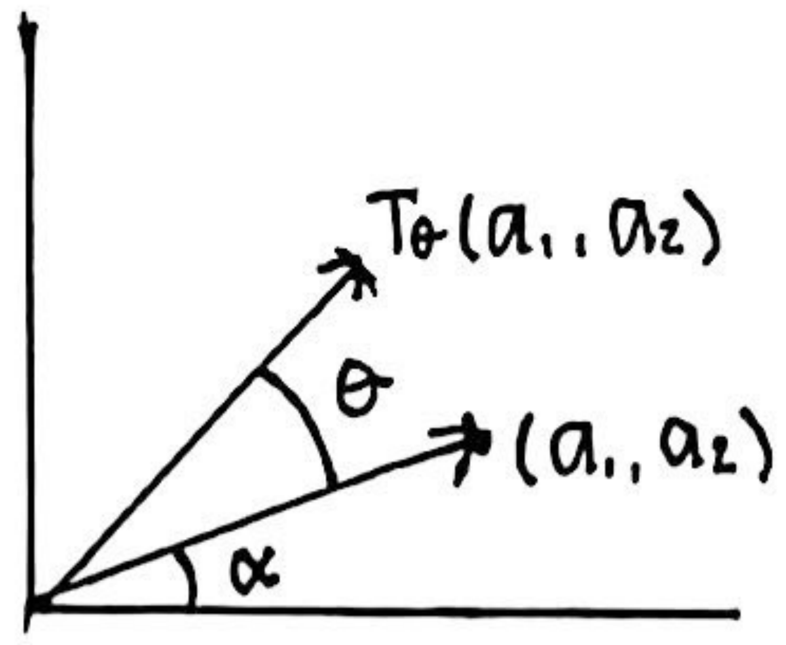
Reflection



Projection



Rotation: Let $T_\theta : R^2 \to R^2$ be defined by $T_\theta(a_1, a_2)$ is the vector rotating (a_1, a_2) counter clock wise by θ



Let $r = \sqrt{a_1^2 + a_2^2}$

↳ the magnitude of (a_1, a_2)

$T_\theta(a_1, a_2) = (r\cos(\alpha + \theta), r\sin(\alpha + \theta))$

$= (r\cos\alpha\cos\theta - r\sin\alpha\sin\theta, r\cos\alpha\sin\theta + r\sin\ldots$

$= (a_1\cos\theta - a_2\sin\theta, a_1\sin\theta + a_2\cos\theta)$

Reflection involves flipping a vector around a point or line

Projection involves changing the direction of a vector by using a different vector

Once we know that a given transformation is linear we can apply some vector space concepts to analyze them further

2.1: Nullspaces & Ranges

Let V & W be vector spaces, & let $T: V \to W$ be linear

Nullspace (Kernel): The set of all vectors in V st $T(x)=0$

Range (Image): The subset of W consisting of all images under T of vectors in V

Theorem 2.1: Nullspace & Range subspaces

Let V & W be vector spaces & $T: V \to W$ be linear.
Then, $N(T)$ & $R(T)$ are subspaces of V & W respectively

Theorem 2.2: Spanning Sets & Linear Transformations

Let V & W be vector spaces & let $T: V \to W$ be linear
If $\beta = \{v_1, v_2, \dots, v_n\}$ is a basis for V, then

$$R(T) = \text{span}(T(\beta)) = \text{span}(\{T(v_1), T(v_2), \dots, T(v_n)\})$$

Example: define $T: P_2(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ by

$$T(f(x)) = \begin{pmatrix} f(1)-f(2) & 0 \\ 0 & f(0) \end{pmatrix}$$ $\to \beta = \{1, x, x^2\}$ is a basis for $P_2(\mathbb{R})$, by the above theorem

$$R(T) = \text{span}(T(\beta)) = \text{span}(\{T(1), T(x), T(x^2)\})$$
$$= \text{span}\left(\left\{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} -3 & 0 \\ 0 & 0 \end{bmatrix}\right\}\right)$$
$$= \text{span}\left(\left\{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}\right\}\right)$$ $\leftarrow$ basis for ~~R(T)~~ $R(T)$

$\dim(R(T)) = 2$ (range)

To find a basis of the nullspace of T

$$\begin{bmatrix} f(1)-f(2) & 0 \\ 0 & f(0) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$ $\to$ let $f(x) = a + bx + cx^2$

$0 = f(1) - f(2)$
$= (a+b+c) - (a+2b+4c) = -b-3c$
& $0 = f(0) = a$

$f(x) = a + bx + cx^2 = -3cx + cx^2$
$= c(-3x + x^2)$

Thus, $-3x + x^2$ is a basis for the nullspace $N(T)$

2.1: Rank & Nullity

Let $T: V \to W$ be a linear transformation

Nullity: the dimension of the basis of the nullspace

Rank: the dimension of the basis of the range

Theorem 2.3: Dimension Theorem:

Let V & W be vector spaces & let $T: V \to W$ be linear
If V is finitely dimensional, then

$$\text{nullity}(T) + \text{rank}(T) = \dim(V)$$
$$\dim(N(T)) + \dim(R(T)) = \dim(V)$$

Theorem 2.4: One-to-One functions & Nullity

Let V & W be vector spaces & let $T: V \to W$ be linear.
Then, T is one-to-one if & only if $N(T) = \{0\}$

Theorem 2.5: Equal Dimensional Linear Transformation

Let V & W be vector spaces of equal dimension & let $T: V \to W$ be linear. Then the following are equivaelent

a. T is one-to-one
b. T is onto
c. $\text{rank}(T) = \dim(V)$

Theorem 2.6: Uniqueness of Linear Transformations

Let V & W be vector spaces over F, & suppose that $\{v_1, v_2, \ldots, v_n\}$ is a basis for V. For $w_1, w_2, \ldots, w_n$ in W there exists one linear transformation $T: V \to W$ such that $T(v_i) = w_i$ for $i = 1, 2, \ldots n$

2.2: Matrix Representation of a Linear Transformations

Ordered Basis: a finitie sequence of linearly independent vectors in a vectorspace V that generates V

↳ the standard ordered basis for F^n is:
$\{e_1, e_2, \ldots, e_n\}$

Coordinate Vectors:

Let $\beta = \{u_1, u_2, \ldots, u_n\}$ be an ordered basis for a finite dimensional vectorspace V. For $x \in V$ let $a_1, a_2, \ldots, a_n$ be the unique scalars such that:

$$x = \sum_{i=1}^{n} a_i u_i$$

Then...
The coordinate vector of x relative to β is given by:

$$[x]_\beta = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

Note: $[u_i]_\beta = e_i$

Example: Let $V = P_2(\mathbb{R})$, & let $\beta = \{1, x, x^2\}$ be the standard ordered basis of V.
If $f(x) = 4 + 6x - 7x^2$, then

$$[f]_\beta = \begin{bmatrix} 4 \\ 6 \\ -7 \end{bmatrix}$$

The moral of the story: making vectors of the coeffrents of a linear Equation

2.2: Matrix Representation of Linear Transformations

Let V & W be finitely dimensional vectorspaces with orded bases $\beta = \{v_1, v_2, \dots, v_n\}$ & $\gamma = \{w_1, w_2, \dots, w_n\}$ respectively. Let $T: V \to W$ be linear.
Then...
for each $j = 1, 2, \dots, n$ there exists unique scalars $a_{ij} \in F$, $i = 1, 2, \dots, m$ such that

$$T(v_j) = \sum_{i=1}^{m} a_{ij} w_i \quad \text{for } j = 1, 2, \dots, n$$

Matrix representation of T in the ordered bases β & γ: the $m \times n$ matrix A, defined by $A_{ij} = a_{ij}$ denoted $A = [T]_\beta^\gamma$

↳ Note: the jth column of A is $[T(v_j)]_\gamma$

Example: Let $T: R^2 \to R^3$ be the linear transformation defined by $T(a_1, a_2) = (a_1 + 3a_2,\ 0,\ 2a_1 - 4a_2)$ &
Let β & γ be the standard ordered bases for R^2 & R^3

$T(1,0) = (1,0,2) = 1e_1 + 0e_2 + 2e_3$ &

$T(0,1) = (3,0,-4) = 3e_1 + 0e_2 - 4e_3$

Thus, the matrix representation of T is given by:

$$[T]_\beta^\gamma = \begin{bmatrix} 1 & 3 \\ 0 & 0 \\ 2 & -4 \end{bmatrix}$$

& if we let $\gamma' = \{e_3, e_2, e_1\}$

$$[T]_\beta^{\gamma'} = \begin{bmatrix} 2 & -4 \\ 0 & 0 \\ 1 & 3 \end{bmatrix}$$

2.2: Identity Matricies & The Kronecker Delta:

Let V & W be finitely dimensional vectorspaces with ordered bases $\beta = \{v_1, v_2, \dots, v_n\}$ & $\gamma = \{w_1, w_2, \dots, w_n\}$

Then:

$I_V(v_j) = v_j = 0v_1 + 0v_2 + \cdots + 0v_{j-1} + 1v_j + 0v_{j+1} + \cdots + 0v_n$

as a matrix:

Identity Matrix

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

& the jth column of I_V is e_j

Kronecker Delta: $\delta_{ij} = 1$ if $i = j$ & $\delta_{ij} = 0$ if $i \neq j$

↳ **Identity Matrix:** The nxn matrix I_n where $(I_n)_{ij} = \delta_{ij}$

Note: The matrix representation of the zero transformation is always the zero matrix

Theorem 2.7: Let V & W be vector spaces over a field F & let $T, U: V \to W$ be linear

a. For all $a \in F$, $aT + U$ is linear

b. The collection of all linear transformations from V to W is a vectorspace over F

↳ denoted: $\mathcal{L}(V, W)$

Theorem 2.8: Let V & W be finitely dimensional vectorspaces with ordered bases β & γ, respectively & let $T, U: V \to W$ be linear transformations

a. $[T+U]_\beta^\gamma = [T]_\beta^\gamma + [U]_\beta^\gamma$

b. $[aT]_\beta^\gamma = a[T]_\beta^\gamma$ for all scalars a

2.3: Composition of Linear Transformations & Matrix Multiplication:

Theorem 2.9:
Let $V, W, \& Z$ be vector spaces over the same Field F, & let $T: V \to W$ & $U: W \to Z$ be linear.
Then $UT: V \to Z$ is linear

Theorem 2.10: Let V be a vector space. Let $T, U_1, U_2 \in \mathcal{L}(V)$
a. $T(U_1 + U_2) = TU_1 + TU_2$ & $(U_1 + U_2)T = U_1T + U_2T$
b. $T(U_1U_2) = (TU_1)U_2$
c. $TI = IT = T$
d. $a(U_1U_2) = (aU_1)U_2 = U_1(aU_2)$ for all scalars a

Multiplication of Matricies:
Let $V, W, \& Z$ be finitely dimensional vector spaces & Let $T: V \to W$ & $U: W \to Z$ be linear transformations
↳ Suppose that $A = [U]_\beta^\gamma$ & $B = [T]_\alpha^\beta$, where
$\alpha = \{v_1, v_2, \dots, v_n\}$, $\beta = \{w_1, w_2, \dots, w_m\}$, &
$\gamma = \{z_1, z_2, \dots, z_p\}$
are ordered bases for $V, W, \& Z$ respectively
The product AB of two matricies is defined by
$AB = [UT]_\alpha^\gamma$

↳ $(UT)(v_j) = U(T(v_j)) = U\left(\sum_{k=1}^{m} B_{kj} w_k\right)$

$$= \sum_{k=1}^{m} B_{kj}\, U(w_k) = \sum_{k=1}^{m} B_{kj} \left(\sum_{i=1}^{p} A_{ik} z_i\right)$$

$$= \sum_{i=1}^{p} C_{ij} z_i \quad \text{where} \quad C_{ij} = \sum_{k=1}^{m} A_{ik} B_{kj}$$

2.3: Composition of Linear Transformations & Matrix Multiplication

Let A be an $m \times n$ matrix & B be an $n \times p$ matrix we define the product of A & B, denoted AB, to be the $m \times p$ such that:

$$(AB)_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj} \quad \text{for } 1 \leq i \leq m,\ 1 \leq j \leq p$$

meaning that $(AB)_{ij}$ is the sum of products of corresponding entries from the ith row of A & the jth column of B

However...

not all matricies can be crossmultiplied by one another

↳ $(m \times n) \cdot (n \times p) = (m \times p)$ in order for the cross product of two matricies to be defined the two "inner" ~~products~~ must be equal & the two "outer" dimensions give the size of the product

Example:

$$\begin{bmatrix} 1 & 2 & 1 \\ 0 & 4 & -1 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \cdot 4 + 2 \cdot 2 + 1 \cdot 5 \\ 0 \cdot 4 + 4 \cdot 2 + (-1) \cdot 5 \end{bmatrix} = \begin{bmatrix} 13 \\ 3 \end{bmatrix}$$

$$(2 \times 3) \cdot (3 \times 1) = 2 \times 1$$

Theorem 2.11: Let V, W, and Z be finite dimensional vectorspaces with ordered bases α, β, & γ respective. Let $T: V \to W$ & $U: W \to Z$ be linear transformations. Then, $[UT]_\alpha^\gamma = [U]_\beta^\gamma [T]_\alpha^\beta$

Example:

Let V be a finite-dimensional vector space & Let $U: P_3(\mathbb{R}) \to P_2(\mathbb{R})$ & $T: P_2(\mathbb{R}) \to P_3(\mathbb{R})$ be linear transformation defined by

$$U(f(x)) = f'(x) \quad \& \quad T(f(x)) = \int_0^x f(t)\,dt$$

Let α & β be the standard ordered bases of $P_3(\mathbb{R})$ & $P_2(\mathbb{R})$ then...

$$[UT]_\beta = [U]_\alpha^\beta [T]_\beta^\alpha = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{bmatrix}$$

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

↳ since it equals the identity matrix we know that it is true in all cases

Theorem 2.12: Let A be an $m \times n$ matrix, B & C be $n \times p$ matricies & D & E be $q \times m$ matricies then...

a. $A(B+C) = AB + AC$ & $(D+E)A = DA + EA$

b. $a(AB) = (aA)B = A(aB)$ for any scalar a

c. $I_m A = A = A I_n$

Theorem 2.13: Let A be an $m \times n$ matrix, & let B be an $n \times p$ matrix. For each j. $j = 1, 2, \ldots, p$ let u_j and v_j denote the jth columns of AB & B respectively. Then...

a. $u_j = Av_j$

b. $v_j = Be_j$ where e_j is the j^{th} standard vector F^p

The j^{th} column of AB is a linear combination of the columns of A with each coefficient being the corresponding entry of the j^{th} column of B

↳ the same is true when considering rows too

Theorem 2.14: Let V & W be finite-dimensional vector spaces having ordered bases β & γ respectively & let $T: V \to W$ be linear. Then, for each $u \in V$ we have

$$[T(u)]_\gamma = [T]_\beta^\gamma [u]_\beta$$

Let A be an $m \times n$ matrix with entries from a field F. The left multiplication transformat is defined as $L_A: F^n \to F^m$ where $L_A(x) = Ax$ for each column vector $x \in F^n$

2.2: Matrix Representation of a Linear Transformation

Example: Let $T: P_3(\mathbb{R}) \to P_2(\mathbb{R})$ be the linear transformation defined by $T(f(x)) = f'(x)$

Let $\beta = \{1, x, x^2, x^3\}$ & $\gamma = \{1, x, x^2\}$

$$\begin{bmatrix} T(1) = 0(1) + 0(x) + 0(x^2) \\ T(x) = 1 + 0x + 0x^2 \\ T(x^2) = 0 + 2x + 0x^2 \\ T(x^3) = 0 + 0 + 3x^2 \end{bmatrix} \rightarrow [T]_\beta^\gamma = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

Review of one-to-one & onto Transformations

One-to-one:

"A function is said to be one-to-one if each element in the of the range corresponds to a unique element in the domain"

$T: A \to B$ is one-to-one if $T(x) = T(y)$ implies $x = y$
(meaning if $x \neq y$ then $T(x) \neq T(y)$)

Onto:

"A function is said to be onto if & only if its range is equal to its codomain"

$T: A \to B$ is onto if & only if $T(A) = B$
(meaning every element in B can be expressed as a function of an element of A)

2.4: Invertibility & Isomorphisms

Let V & W be vector spaces & let $T: V \to W$ be linear.
The Inverse of T is any function $U: W \to V$ such that

$TU = I_W$ & $UT = I_V$

T is said to be Invertible if & only if it has an inverse

The following are true for invertible functions U & T

1. $(TU)^{-1} = U^{-1}T^{-1}$
2. $(T^{-1})^{-1} = T$

Theorem 2.17: Let V & W be vector spaces, & let $T: V \to W$ be linear & Invertible. Then $T^{-1}: W \to V$ is linear

Let $y_1, y_2 \in W$ & let $c \in F$. Since T is onto & one-to-one there exists vectors x_1 & x_2 such that $T(x_1) = y_1$ & $T(x_2) = y_2$
Thus, $x_1 = T^{-1}(y_1)$ & $T^{-1}(y_2) = x_2$; so

$$T^{-1}(cy_1 + y_2) = T^{-1}[cT(x_1) + T(x_2)] = T^{-1}[T(cx_1 + x_2)] = cx_1 + x_2 \quad ▨$$

An $n \times n$ matrix is said to be invertible if there exists an $n \times n$ B such that $AB = BA = I$

Theorem 2.18: Let V & W be vector spaces of finite dimensional with ordered bases β & γ respectively. Let $T: V \to W$ be linear Then T is invertible if & only if $[T]_\beta^\gamma$ is invertible. Furthermore, $[T^{-1}]_\gamma^\beta = ([T]_\beta^\gamma)^{-1}$

2.4: Isomorphisims

Let V & W be vector spaces we say V is **Isomorphic** to W if there exists some transformation $T: V \to W$ that is invertible

↳ such a linear transformation is called an **Isomorphism**

(Isomorphism is an equivalente relation)

Theorem 2.19: Let V & W be finite-dimensional vector spaces (over the same field). Then V is isomorphic to W if & only if $\dim(V) = \dim(W)$

Proof:

Suppose that V is isomorphic to W, & that $T: V \to W$ is an isomorphism from V to W, suppose that $\dim(V) = \dim(W)$

IIII hello
hell Hello

Review: Matrix Representations Review

Let $\beta = \{u_1, u_2, \cdots, u_n\}$ be an ordered basis for a finite-dimensional vector space V. For $x \in V$, let $a_1, a_2, \cdots,$ be the unique scalars such that $x = \sum_{i=1}^{n} a_i u_i$

The **Coordinate Vector** of x relative to β, denoted $[x]_\beta$ is defined as

$$[x]_\beta = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$$

Example:
Let $V = P_2(\mathbb{R})$ & let $\beta = \{1, x, x^2\}$
If $f(x) = 4 + 6x - 7x^2$, then

$$[f]_\beta = \begin{bmatrix} 4 \\ 6 \\ -7 \end{bmatrix}$$

Now...
Suppose V & W are finite-dimensional vector spaces with ordered bases $\beta = \{v_1, v_2, \cdots, v_n\}$ & $\gamma = \{w_1, w_2, \cdots, w_m\}$ respectively. Let $T: V \to W$ be linear. Then, for each $j = 1, 2, \cdots, n$ there exists unique scalars $a_{ij} \in F$, $i = 1, 2, \cdots, m$, such that

$$T(v_{ij}) = \sum_{i=1}^{m} a_{ij} w_i \quad \text{for } j = 1, 2, \cdots, n$$

Thus, the $m \times n$ matrix A defined by $A_{ij} = a_{ij}$ is called **The Matrix representation of T with respect of β & γ** denoted $A = [T]_\beta^\gamma$

Example:
Let $T: P_3(\mathbb{R}) \to P_2(\mathbb{R})$ be the linear transformation defined by $T(f(x)) = f'(x)$
Let β & γ be the standard ordered bases of $P_3(\mathbb{R})$ & $P_2(\mathbb{R})$ respectively, that is,

$\beta = \{1, x, x^2, x^3\}$
$\gamma = \{1, x, x^2\}$

Then,

$T(1) = 0(1) + 0(x) + 0(x^2)$
$T(x) = 1(1) + 0(x) + 0(x^2)$
$T(x^2) = 0(1) + 2(x) + 0(x^2)$
$T(x^3) = 0(1) + 0(x) + 3(x^2)$

$$[T]_\beta^\gamma = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$$

More Matrix Representation:

When representing a linear transformation as a matrix A the jth column is given by $[T(v_j)]_\gamma$
& if two transformations have the same matrix representation, then they are eqivalent.

Another Example:

Let $T: \mathbb{R}^2 \to \mathbb{R}^3$ be the linear transformation defined by:

$$T(a_1, a_2) = (a_1 + 3a_2, \ 0, \ 2a_1 - 4a_2)$$

Let β & γ be the standard orderd bases of $\mathbb{R}^2$ & $\mathbb{R}^3$

Then, $T(1,0) = (1,0,2) = 1e_1 + 0e_2 + 2e_3$

$T(0,1) = (3,0,-4) = 3e_1 + 0e_2 + -4e_3$

meaning,

$$[T]_\beta^\gamma = \begin{bmatrix} 1 & 3 \\ 0 & 0 \\ 2 & -4 \end{bmatrix}$$

Additionally, if we let $\gamma' = \{e_1, e_2, e_3\}$ then,

$$[T]_\beta^{\gamma'} = \begin{bmatrix} 2 & -4 \\ 0 & 0 \\ 1 & 3 \end{bmatrix}$$

Finally, these matrix representations are themselves, linear. That is...

Theorem 2.8: Let V & W be finite-dimensional vector spaces with ordered bases β & γ respectively, and let $T: V \to W$ & $U: V \to W$ be linear.
Then, $[T+U]_\beta^\gamma = [T]_\beta^\gamma + [U]_\beta^\gamma$ & $[aT]_\beta^\gamma = a[T]_\beta^\gamma$

for all scalars $a \in F$

Composition of Linear Transformations

Theorem 2.9: Let V, W, & Z be vector spaces over the same field F, & let $T: V \to W$ & $U: W \to Z$ be linear. Then, $UT: V \to Z$ is linear.

Proof:

Let $x, y \in V$ & $a \in F$. Then,

$$UT(ax+y) = U(T(ax+y)) =$$
$$= U(aT(x) + T(y)) = a(U(T(x)) + U(T(y)) =$$
$$= a(UT(x)) + UT(y)$$

Theorem 2.10:

Let V be a vector space. Let $T, U_1, U_2 \in \mathcal{L}(V)$, then,

a. $T(U_1 + U_2) = TU_1 + TU_2$ & $(U_1 + U_2)T = U_1 T + U_2 T$

b. $T(U_1 U_2) = (TU_1)U_2$

c. $TI = IT = T$

d. $a(U_1 U_2) = (aU_1)U_2 = U_1(aU_2)$ for all scalars $a \in F$

Theorem 2.11:

Let V, W, & Z be finite-dimensional vector spaces with ordered bases α, β, & γ respectively. Let $T: V \to W$ & $U: W \to Z$ be linear transformations. Then...

$$[UT]_\alpha^\gamma = [U]_\beta^\gamma [T]_\alpha^\beta$$

Matrix Multipication

Review:

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

$$= \begin{bmatrix} 1(4) + 2(2) + 1(5) \\ 0(4) + 4(2) + -1(5) \end{bmatrix} = \begin{bmatrix} 13 \\ 3 \end{bmatrix}$$

$(2 \times 3)(3 \times 1) = 2 \times 1$

$AB \neq BA$

Example:

Let $U: P_3(\mathbb{R}) \to P_2(\mathbb{R})$ & $T: P_2(\mathbb{R}) \to P_3(\mathbb{R})$ be the linear transformations defined by:

$U(f(x)) = f'(x)$ & $T(f(x)) = \int_0^x f(t)\,dt$

& let α & β be the standard bases of $P_3(\mathbb{R})$ & $P_2(\mathbb{R})$

$$[UT]_\beta = [U]_\alpha^\beta [T]_\beta^\alpha = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1/2 & 0 \\ 0 & 0 & 1/3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_\beta$$

Matrix Multiplication & Composition

Theorem 2.12: Let A be a $m \times n$ matrix, B & C be $n \times p$ matricies, & D & E be $q \times m$ matricies. Then...

a. $A(B+C) = AB + AC$ & $(D+E)A = DA + EA$

b. $a(AB) = (aA)B = A(aB)$

c. $I_m(A) = A = A(I_n)$

Theorem 2.13: Let A be an $m \times n$ matrix & B be an $n \times p$ matrix. For each $j = 1, 2, \ldots, p$ let u_j & v_j denote the jth column of AB & B respectively. Then...

a. $u_j = Av_j$

b. $v_j = Be_j$

where e_j is the jth standard vector of F^p

Theorem 2.14:
Let V & W be finite-dimensional vector spaces having ordered bases β & γ, respectively, & let $T: V \to W$ be linear. Then, for each $u \in V$

$$[T(u)]_\gamma = [T]_\beta^\gamma [u]_\beta$$

Proof: Fix $u \in V$ & define the linear transformation $f: F \to V$ by $f(a) = au$ & $g: F \to W$ by $g(a) = aT(u)$ for all $a \in F$. Note that $g = Tf$. Thus, by representing column vectors as matricies & composing trasformations (theorem 2.11) we get:

$$[T(u)]_\gamma = [g(1)]_\gamma = [g]_\alpha^\gamma = [Tf]_\alpha^\gamma = [T]_\beta^\gamma [f]_\alpha^\beta = [T]_\beta^\gamma [f(1)]_\beta$$
$$= [T]_\beta^\gamma [u]_\beta$$ ★

Composition & The Left Multiplication Transform.

Putting things together...

Let $T: P_3(\mathbb{R}) \to P_2(\mathbb{R})$ be the linear transformation defined by $T(f(x)) = f'(x)$, & let β & γ be the standard bases of $P_3(\mathbb{R})$ & $P_2(\mathbb{R})$ respectively. Two pages ago we showed that if

$A = [T]_\beta^\gamma$ Then... $A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix}$ Then, Theorem 2.14 states that:

$[T(p(x))]_\gamma = [T]_\beta^\gamma [p(x)]_\beta$ where $p(x) \in P_3(\mathbb{R})$ is the polynomial
$p(x) = 2 - 4x + x^2 + 3x^3$

Now...

Let $q(x) = T(p(x))$, then: $q(x) = p'(x) = -4 + 2x + 9x^2$
mening

$$[T(p(x))]_\gamma = [q(x)]_\gamma = \begin{bmatrix} -4 \\ 2 \\ 9 \end{bmatrix}$$

Additionally,

$$[T]_\beta^\gamma [p(x)]_\beta = A[p(x)]_\beta = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{bmatrix} \begin{bmatrix} 2 \\ -4 \\ 1 \\ 3 \end{bmatrix} = \begin{bmatrix} -4 \\ 2 \\ 9 \end{bmatrix}$$

To close this section...

The **Left-Multiplication Transformation**, denoted L_A, of a $m \times n$ matrix A is defined by $L_A: F^n \to F^m$ where $L_A(x) = Ax$ (the matrix-product of A & x) for each column vector $x \in F^n$

Example:

Let $A = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}$ Then, $A \in M_{2\times3}(\mathbb{R})$ & $L_A: \mathbb{R}^3 \to \mathbb{R}^2$ if

$x = \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix}$ Then,

$$L_A(x) = Ax = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ -1 \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \end{bmatrix}$$

Properties of The Left Multiplication Transform

Theorem 2.15: Let A be an $m \times n$ matrix with entries from F. Then, the left-multiplication transformation $L_A : F^n \to F^m$ is linear. Furthermore if B is any other $m \times n$ matrix & β & γ are the standard bases of F^n & F^m respectively, then:

a. $[L_A]_\beta^\gamma = A$
b. $L_A = L_B$ if & only if $A = B$
c. $L_{A+B} = L_A + L_B$ & $L_{aA} = aL_A$ for all $a \in F$
d. If $T : F^n \to F^m$ is linear then there exists C such that $C = [T]_\beta^\gamma$
e. If E is an $n \times p$ matrix, then $L_{AE} = L_A L_E$
f. If $m = n$, then $L_{I_n} = I_{F^n}$

Isomorphisms & Dual Spaces

We are well aquainted with representing linear transformations as matricies. The next step is to show that the collection of all linear transformations between two vector spaces is itself a vector space that can be expressed as a vector space of $m \times n$ matricies

Theorem 2.20: Let V & W be finite-dimensional vector spaces over F with dimensions n & m respectively, & let β & γ be ordered bases for V & W respectively. Then, the function $\Phi_\beta^\gamma : \mathcal{L}(V, W) \to M_{m \times n}(F)$, defined by:
$\Phi_\beta^\gamma(T) = [T]_\beta^\gamma$ for $T \in \mathcal{L}(V, W)$ is an isomorphism.

↳ *Proof:* First, we have to show that Φ_β^γ is one-to-one & onto. → Show that for every $m \times n$ matrix A, there exists a unique linear transformation $T: V \to W$ such that $\Phi_\beta^\gamma(T) = A$.
Let $\beta = \{v_1, v_2, \cdots, v_n\}$, $\gamma = \{w_1, w_2, \cdots, w_m\}$ & let A be a given $m \times n$ matrix. Then, there exists a transformation $T: V \to W$ such that $T(v_j) = \sum_{i=1}^{m} A_{ij} w_i$ for $1 \leq j \leq n$ meaning that
$[T]_\beta^\gamma = A$, or $\Phi_\beta^\gamma(T) = A$. Thus, Φ_β^γ is an isomorphism!

The Rank of a Matrix & Matrix Inverses

The Rank of a matrix A (denoted rank(A)) to be the rank of the linear transformation given by $L_A: F^n \to F^m$
↳(assuming $A \in M_{m\times n}(F)$)

Theorem 3.3:

Let $T: V \to W$ be a linear transformation between finite-dimensional vector spaces, & let β & γ be ordered bases of V & W respectively. Then $\text{rank}(T) = \text{rank}([T]_\beta^\gamma)$

Theorem 3.4:

Let A be an $m \times n$ matrix. If P & Q are invertible $m\times m$ & $n \times n$ matricies, respectively, then:

$\left.\begin{array}{l} \text{a. } \text{rank}(AQ) = \text{rank}(A) \\ \text{b. } \text{rank}(PA) = \text{rank}(A) \end{array}\right\}$ & therefore, c. $\text{rank}(PAQ) = \text{rank}(A)$

Proof: First, note that

↱ $R(L_{AQ}) = R(L_A L_Q) = L_A L_Q(F^n) = L_A(L_Q(F^n)) = L_A(F^n) = R(L_A)$

Since L_Q is onto. As a result

$\text{rank}(AQ) = \dim(R(L_{AQ})) = \dim(R(L_A)) = \text{rank}(A)$

Note that if B is obtained from A by an elementary row operation then there exists an elementary matrix E such that $B = AE$, E is invertible & therefore $\text{rank}(A) = \text{rank}(B)$
↳ Meaning EROs preserve rank

Theorem 3.5:

The rank of any matrix is equal to the maximum number of linearly independent columns it contains; that is, the rank of a matrix is the dimension of the subspace generated by its columns. → *Proof:* For any $A \in M_{m\times n}(F)$,

$\text{rank}(A) = \text{rank}(L_A) = \dim(R(L_A))$

let β be the standard basis of F^n. Then β spans F^n, thus, by thm 2.2

$R(L_A) = \text{span}(L_A(\beta)) = \text{span}(\{L_A(e_1), \dots, L_A(e_n)\})$. But! by thm 2.13b for any j, $L_A(e_j) = Ae_j = a_j$, where a_j is the jth column of A. Hence, $R(L_A) = \text{span}(\{a_1, \dots, a_n\})$ & thus: $\text{rank}(A) = \dim(R(L_A))$

$= \dim(\text{span}(\{a_1, a_2, \dots, a_n\}))$

Theorem 3.6 & its Long Ass Proof

Let A be an $m \times n$ matrix of rank r. Then $r \leq m$, $r \leq n$ & by means of a finite number of EROs & ECOs A can be transformed into the matrix $D = \begin{bmatrix} I_r & O_1 \\ O_2 & O_3 \end{bmatrix}$ Where $O_1, O_2, \& \ O_3$ are zero matricies

Thus, $D_{ij} = 1$ for $i, j \leq r$ & 0 otherwise. *Proof:*

If A is the zero matrix, then it follows that $r = 0$. Thus D is also the zero.

Next, suppose $A \neq O$ & $r = \text{rank}(A)$; then $r > 0$ by induction suppose $m = 1$. By at most one type 1 column operation, & at most one type 2 column operation, A can be transformed into a matrix with a 1 in the 1,1 position. By means of at most $n-1$ type 3 column Operations, this matrix can in turn be transformed into the matrix: $[1\ 0 \cdots 0]$ Note there is only one linearly independent column in D so $\text{rank}(D) = \text{rank}(A) = 1$ & thus $m = 1$

Now, Suppose $n > 1$. Since $A \neq O$, $A_{ij} \neq 0$ for some i, j By means of at most one elementary row operation & at most one elementary column operation (both type I) we can move the nonzero entry to 1,1. By means of at most $m-1$ type III row operations & at most $n-1$ type III column operations we can eliminate all nonzero entries in the first row & first column with the exception of the 1 in position (1,1), Thus with a finite number of elenmentary Operations A can be transformed into the matrix

$$B = \left[\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & B' \\ 0 & \end{array}\right]$$

where B' has a rank one less than B & since $\text{Rank}(A) = \text{rank}(B) = r$, $\text{Rank}(B') = r-1$ Therefore $r-1 \leq m-1$ & $r-1 \leq n-1$ Hence $r \leq m$ & $r \leq n$

Repeating this process for B' results in D thus we have obtained D from A!

Elementary Operations & Rank Example

Let

$$A = \begin{bmatrix} 0 & 2 & 4 & 2 & 2 \\ 4 & 4 & 4 & 8 & 0 \\ 8 & 2 & 0 & 10 & 2 \\ 6 & 3 & 2 & 9 & 1 \end{bmatrix}$$

Theorem 3.6 states that successively apply elementary operations to A can transform A into $D = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix}$

$$\begin{bmatrix} 0 & 2 & 4 & 2 & 2 \\ 4 & 4 & 4 & 8 & 0 \\ 8 & 2 & 0 & 10 & 2 \\ 6 & 3 & 2 & 9 & 1 \end{bmatrix} \xrightarrow{\text{I}: R_2 \to R_1} \begin{bmatrix} 4 & 4 & 4 & 8 & 0 \\ 0 & 2 & 4 & 2 & 2 \\ 8 & 2 & 0 & 10 & 2 \\ 6 & 3 & 2 & 9 & 1 \end{bmatrix} \xrightarrow{\text{II}: \frac{1}{2} \cdot R_1} \begin{bmatrix} 1 & 1 & 1 & 2 & 0 \\ 0 & 2 & 4 & 2 & 2 \\ 8 & 2 & 0 & 10 & 2 \\ 6 & 3 & 2 & 9 & 1 \end{bmatrix} \xrightarrow[\text{III}: R_4 + -6 \cdot R_1]{\text{III}: R_3 + -8 \cdot R_1}$$

$$\begin{bmatrix} 1 & 1 & 1 & 2 & 0 \\ 0 & 2 & 4 & 2 & 2 \\ 0 & -6 & -8 & -6 & 2 \\ 0 & -3 & -4 & -9 & 1 \end{bmatrix} \xrightarrow[\text{I}: R_3 \to R_1]{\text{III}: R_3 + 4R_4 \quad \text{III}: R_3 + R_1} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 2 & 4 & 2 & 2 \\ 0 & -6 & -8 & -6 & 2 \\ 0 & -3 & -4 & -3 & 1 \end{bmatrix} \xrightarrow{\text{II}: \frac{1}{2} R_2} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 2 & 1 & 1 \\ 0 & -6 & -8 & -6 & 2 \\ 0 & -3 & -4 & -3 & 1 \end{bmatrix}$$

Rank, Composition, & Invertibility

Theorem 3.7: Let $T: V \to W$, & $U: W \to Z$ be linear transformations on finite-dimensional vector spaces, & Let A & B be matricies such that AB is defined. Then:

a. $\text{rank}(UT) \leq \text{rank}(U)$
b. $\text{rank}(UT) \leq \text{rank}(T)$
c. $\text{rank}(AB) \leq \text{rank}(A)$
d. $\text{rank}(AB) \leq \text{rank}(B)$

Proof:

We will prove them in the order a, c, d, then b

a. Clearly, $R(T) \subseteq W$. Hence,

$R(UT) = UT(V) = U(T(V)) = U(R(T)) \subseteq U(W) = R(U)$

Thus, $\text{rank}(UT) = \dim(R(UT)) \leq \dim(R(U)) = \text{rank}(U)$.

C. By a., $\text{rank}(AB) = \text{rank}(L_{AB}) = \text{rank}(L_A L_B) \leq \text{rank}(L_A) = \text{rank}(A)$

d. $\text{rank}(AB) = \text{rank}((AB)^t) = \text{rank}(B^t A^t) \leq \text{rank}(B^t) = \text{rank}(B)$

b.

Rank & Nullity Revisited

A **Linear Transformation** is any $T: V \to W$ such that for $x, y \in V$ $T(x) + T(y) = T(x+y)$ & $T(cx) = cT(x)$ → it preserves <u>vector addition</u> & <u>scalar multiplication</u>

Note: "preimage" & "image" are preferential to "domain" & "codomain"

$\text{rank}(T) + \text{nullity}(T) = \dim(T)$ <u>or</u>
$\dim(R(T)) + \dim(N(T)) = \dim(T)$

Matrix Representation:

Proof of the Rank-Nullity Theorem (Review)

Let V & W be vector spaces & let $T: V \to W$ be linear. If V is finitely dimensional then, $\text{nullity}(T) + \text{rank}(T) = \dim(V)$

Proof:

Suppose $\dim(V) = n$, $\dim(N(T)) = k$, & $\{v_1, v_2, \cdots, v_k\}$ is a basis for $N(T)$. We can extend $\{v_1, v_2, \cdots, v_n\}$ into a basis for V, $\beta = \{v_1, v_2, \cdots, v_n\}$. Now, let $S = \{T(v_{k+1}), T(v_{k+2}), \cdots, T(v_n)\}$ be a basis for $R(T)$ → We first have to prove that S generates $R(T)$, using the fact that $T(v_i) = 0$ for $1 \le i \le k$ we have

$$R(T) = \text{span}(\{T(v_1), T(v_2), \dots, T(v_n)\})$$
$$= \text{span}(\{T(v_{k+1}), T(v_{k+2}), \cdots, T(v_n)\}) = \text{span}(S)$$

Next we need to prove S is linearly independent. Suppose that

$$\sum_{i=k+1}^{n} b_i T(v_i) = 0 \text{ for } b_{k+1}, b_{k+2}, \cdots, b_n \in F$$

since T is linear

$$T\left(\sum_{i=k+1}^{n} b_i v_i\right) = 0 \quad \text{So,} \quad \sum_{i=k+1}^{n} b_i v_i \in N(T)$$

Hence, there exists $c_1, c_2, \cdots, c_k \in F$ such that $\sum_{i=k+1}^{n} b_i v_i + \sum_{i=1}^{k} (-c_i) v_i = 0$

Since, β is a basis for V, we have $b_i = 0$ for all i. Hence, S is linearly Independent. Notice that this argument also shows that $T(v_{k+1}), T(v_{k+2}), \cdots, T(v_n)$ are distinct; Therefore $\text{rank}(T) = n - k$

Systems of Linear Equations

This portion is dedicated to solving systems of linear equations by using elementary row operations. Lets review the basics. The following system:

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

(where each a_{ij}, b_i are scalars in a field F, & $x_1, x_2, \ldots, x_n$ are n variables with values in F.)

The **Coefficient Matrix** of the system is $\longrightarrow$ $A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$

Then, if we let

$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ & $b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$

The system can be written as a single matrix equation: $Ax = b$

The **Solution Set** of the system is the set of all vectors s, such that $As = b$ denoted S. Any system with a non-empty S is said to be **consistent**, otherwise in the case that the system does not have any valid solutions it is said to be **inconsistent**

Additionally, in the case that $b = \mathbf{0}$ the system is said to be **homogeneous**, otherwise the system is **nonhomogeneous**

We turn our attention to the homogeneous case as it is the simpler of the two

Systems of Linear Equations

Theorem 3.8: Let $Ax = \mathbb{0}$ be a homogeneous system of m linear equations in n unknowns over a field F. Let K denote the set of all solutions to $Ax = \mathbb{0}$. Then, $K = N(L_A)$. Hence K is a subspace of F^n of dimension $n - \text{rank}(L_A) = n - \text{rank}(A)$. Additionally, if $m < n$ the system has a nonzero solution.

Example: Consider the following system

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

Then the coefficient matrix of this system is given by $A = \begin{bmatrix} 1 & 2 & 1 \\ 1 & -1 & -1 \end{bmatrix}$

The rank of A is 2

Thus, if K is a solution set to this system, then $\dim(K) = 3 - 2 = 1$. Thus, any nonzero solution is a basis for K. For example, since the following vector is a solution to the system it is also a basis of K $\begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$

as such, any vector in K is of the form

$$t \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix} = \begin{bmatrix} t \\ -2t \\ 3t \end{bmatrix}$$

Theorem 3.9: Let K be the solution space of a consistent system of linear equations $Ax = b$, & let K_H be the solution set of the corresponding homogenous system $Ax = \mathbb{0}$. Then for any solution s to $Ax = b$ $\quad K = \{s\} + K_H = \{s + k : k \in K_H\}$

Proof:

Let s be any solution to $Ax = b$. If $w \in K$, then $Aw = b$. Hence, $A(w - s) = Aw - As = b - b = \mathbb{0}$ so $w - s \in K_H$. Thus their exists $k \in K_H$ such that $w - s = k$. it follows that $w = s + k \in \{s\} + K_H$ & therefore $K \subseteq \{s\} + K_H$

The theorem above demonstrates that the solution set to any nonhomogeneous system can be described in terms of a **Corresponding Homogeneous System**

Systems of Linear Equations

Theorem 3.10: Let $Ax = b$ be a system of n linear equations in n unknowns. If A is invertible, then the system has exactly one solution, namely, $A^{-1}b$. Conversely, if the system has exactly one solution then A is invertible.

Example:

Consider the following system:

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

Then $A = \begin{bmatrix} 0 & 2 & 4 \\ 2 & 4 & 2 \\ 3 & 3 & 1 \end{bmatrix}$ & $A^{-1} = \begin{bmatrix} \frac{1}{8} & \frac{5}{8} & 3/4 \\ -\frac{1}{4} & 3/4 & -1/2 \\ \frac{3}{8} & -3/8 & 1/4 \end{bmatrix}$

Thus,

$$A^{-1}b = \begin{bmatrix} 1/8 & 5/8 & 3/4 \\ -1/4 & 3/4 & -1/2 \\ 3/8 & -3/8 & 1/4 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} -7/8 \\ 5/4 \\ -1/8 \end{bmatrix}$$

Theorem 3.11: Let $Ax = b$ be a system of linear equations. Then, the system is consistent if & only if $\text{rank}(A) = \text{rank}(A|b)$

Proof: To say that $Ax = b$ has a solution is equivalent to saying $b \in R(L_A)$ & by theorem 3.5: $R(L_A) = \text{span}(\{a_1, a_2, \cdots, a_n\})$ the span of the columns of A. Thus, $Ax = b$ has a solution if & only if $b \in \text{span}(\{a_1, a_2, \cdots, a_n\})$. But, $b \in \text{span}(\{a_1, a_2, \cdots, a_n\})$ if & only if $\text{span}(\{a_1, a_2, \cdots, a_n\}) = \text{span}(\{a_1, a_2, \cdots, a_n, b\})$ meaning $\dim(\text{span}(\{a_1, a_2, \cdots, a_n\})) = \dim(\text{span}(\{a_1, a_2, \cdots, a_n, b\}))$ Thus, we have $\text{rank}(A) = \text{rank}(A|b)$

Two systems of linear equations are said to be **equivalent** if they have the same solution set

Systems of Linear Equations

Theorem 3.13: Let $Ax=b$ be a system of m linear equations in n unknowns, & let C be an invertible $m\times m$ matrix. The system $(CA)x=Cb$ is equivalent to $Ax=b$

Proof: Let K be the solution set for $Ax=b$ & K' be the solution set for $(CA)x=Cb$. If $w\in K$, then $Aw=b$. So $(CA)w=Cb$, & hence, $w\in K'$. Thus $K\subseteq K'$. Conversely, if $w\in K'$, then, $(CA)w=Cb$. Hence, $Aw=C^{-1}(CAw)=C^{-1}(Cb)=b$

Additionally, if $(A'|b')$ is obtained from $(A|b)$ by elementary row operations, then the system $A'x=b'$ is equivalent to the original system $Ax=b$.

Proof: Suppose that $(A'|b')$ is obtained from $(A|b)$ using EROs. These may be represented by multiplying $(A|b)$ by elementary $m\times m$ matricies. Since all of the elementary matricies are invertible by the previous theorem, the two systems are equivalent.

Our next concern is describing a method for solving <u>any</u> system of linear equations. We should first more rigorous definition of the familiar method of **Gaussian Elimination** which consists of two seperate parts

1. **Forward Pass:** (step 1-5), the augmented matrix is transformed into an upper triangular matrix in which the first nonzero entry in each row is one & it occurs in a column to the right of the first nonzero entry in the previous column
2. **Backward Pass:** (step 6-7): the upper triangular matrix is transformed into its RREF by making the first nonzero entry in each row is the only nonzero entry in its column

Solving Systems of Linear Equations

Consider the following system. To solve the system we first need to express it as an augmented matrix of scalars

$$3x_1 + 2x_2 + 3x_3 - 2x_4 = 1$$
$$x_1 + x_2 + x_3 = 3$$
$$x_1 + 2x_2 + x_3 - x_4 = 2$$

$$A = \left[\begin{array}{cccc|c} 3 & 2 & 3 & -2 & 1 \\ 1 & 1 & 1 & 0 & 3 \\ 1 & 2 & 1 & -1 & 2 \end{array}\right]$$

Now we can use EROs to solve

1. In the leftmost nonzero column, create a 1 in the first row

$$\xrightarrow{R_1 \leftrightarrow R_3} \left[\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 1 & 1 & 1 & 0 & 3 \\ 3 & 2 & 3 & -2 & 1 \end{array}\right]$$

2. Add multiples of the first row to the other rows to obtain zeros in the remaining entries of the first column

$$\xrightarrow{R_2 - R_1 \,\&\, R_3 - 3R_1} \left[\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 0 & -1 & 0 & 1 & 1 \\ 0 & -4 & 0 & 1 & 5 \end{array}\right]$$

3. Create a 1 in the next <u>row</u> in the leftmost possible column, without using the previous rows

$$\xrightarrow{-1 \cdot R_2} \left[\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & -4 & 0 & -3 & -9 \end{array}\right]$$

4. Obtain zeros below the 1 created in the previous step

$$\xrightarrow{R_3 + 4R_2} \left[\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & -3 & -9 \end{array}\right]$$

5. Repeat steps 3 & 4 on each subsequent row until no nonzero rows remain

$$\xrightarrow{-\frac{1}{3} \cdot R_3} \left[\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 2 \\ 0 & 1 & 0 & -1 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right]$$

Solving Systems of Linear Equations

6. Beginning with the last row working upwards add multiples of each row to the rows above to create zeros above the first nonzero entry in each row

$$\xrightarrow{\substack{R_1+R_3 \\ R_2+R_3}} \left[\begin{array}{cccc|c} 1 & 2 & 1 & 0 & 5 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right]$$

7. Repeat step 6 for the remaining rows

$$\xrightarrow{R_1-2R_2} \left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right] \rightarrow \begin{bmatrix} 1-t \\ 2 \\ t \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 0 \\ 3 \end{bmatrix} + t \begin{bmatrix} -1 \\ 0 \\ 1 \\ 0 \end{bmatrix}$$

A matrix is said to be in **Reduced Row Echelon Form (RREF)** if the following conditions are satisfied

a. If there are any rows of all zero are last

b. The first nonzero entry in each row is the only nonzero entry in its column

c. The first nonzero entry in each row is one & it occurs in a column to the right of the first nonzero entrys of the previous row

Theorem 3.14: Gaussian elimination transforms any matrix into its reduced row echelon form

To solve a system a system in RREF.

1. Divide the variables into two sets: the left most variables of each equation, & the second set consists of the remaining variables
2. To each variable in the second set assign a paremetric variable & solve for the variables of the first set in terms of the new paremetric variables

Determinants of Order 2

When we say "order 2" we are reffering to the determinant of a 2×2 matrix

Recall, from math 165...

Let $A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ be a matrix with entries from a field F, then, we define the **determinant** of A, denoted $\det(A)$ or $|A|$, to be the scalar value defined by $ad-bc$

Theorem 4.1: The function $\det: M_{2\times2}(F)\to F$ is a linear function of each row of a 2×2 matrix when the other row is held fixed. That is, if $u, v,$ & w are in F^2 & k is a scalar then,

$$\det\begin{bmatrix} u+kv \\ w \end{bmatrix} = \det\begin{bmatrix} u \\ w \end{bmatrix} + k\det\begin{bmatrix} v \\ w \end{bmatrix} \quad \&$$

$$\det\begin{bmatrix} w \\ u+kv \end{bmatrix} = \det\begin{bmatrix} w \\ u \end{bmatrix} + k\det\begin{bmatrix} w \\ v \end{bmatrix}$$

Proof: Let $u=(a_1, a_2)$, $v=(b_1, b_2)$, & $w=(c_1, c_2)$ be in F^2 & k be a scalar. Then,

$$\det\begin{pmatrix} u \\ w \end{pmatrix} + k\det\begin{pmatrix} v \\ w \end{pmatrix} = \det\begin{bmatrix} a_1 & a_2 \\ c_1 & c_2 \end{bmatrix} + k\det\begin{bmatrix} b_1 & b_2 \\ c_1 & c_2 \end{bmatrix}$$

$$= (a_1c_2 - a_2c_1) + k(b_1c_2 - b_2c_1)$$

$$= (a_1 + kb_1)c_2 - (a_2 + kb_2)c_1$$

$$= \det\begin{bmatrix} a_1+kb_1 & a_2+kb_2 \\ c_1 & c_2 \end{bmatrix} = \det\begin{bmatrix} u+kv \\ w \end{bmatrix}$$

A similar calculation shows that

$$\det\begin{bmatrix} w \\ u \end{bmatrix} + k\det\begin{bmatrix} w \\ v \end{bmatrix} = \det\begin{bmatrix} w \\ u+kv \end{bmatrix}$$

Determinants & Invertibility

Theorem 4.2: Let $A \in M_{2\times 2}(F)$. Then, the determinant of A is nonzero if & only if A is invertible. Moreover, if A is invertible then, $A^{-1} = \frac{1}{\det(A)}\begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}$ *Proof:*

If $\det(A) \neq 0$, then we can define a matrix $M = \frac{1}{\det(A)}\begin{bmatrix} A_{22} & -A_{12} \\ -A_{21} & A_{11} \end{bmatrix}$ A straight forward calculation shows that $AM = MA = I_2$ & thus A is invertible & the inverse of A is given by M.

Conversely, suppose that A is invertible meaning its rank is two. Hence, $A_{11} \neq 0$ or $A_{21} = 0$. If $A_{11} \neq 0$, add $-A_{21}/A_{11}$ times row one to row two, which yields the following:

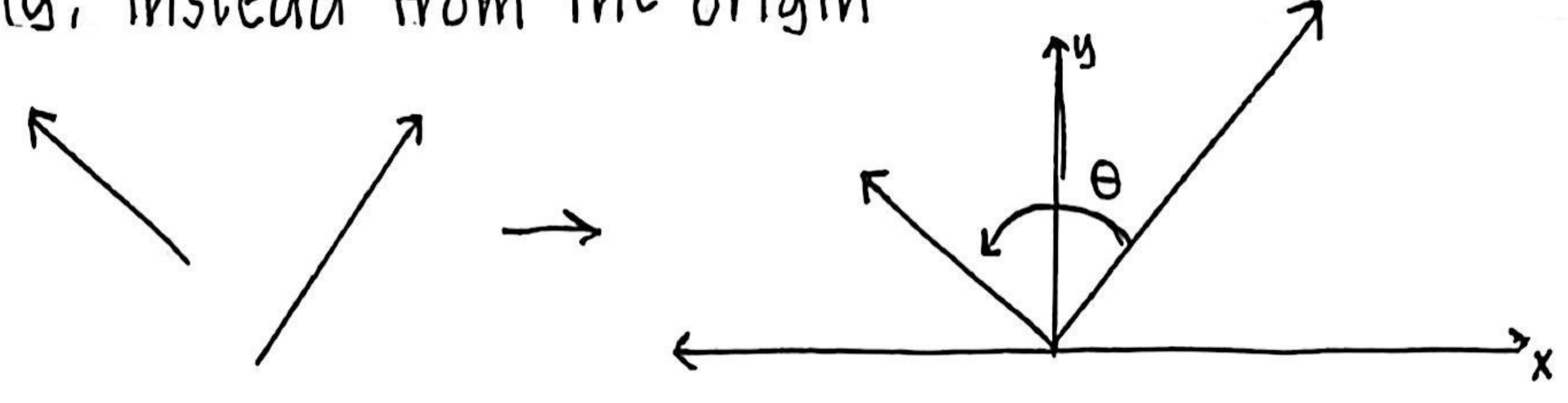
$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} - \frac{A_{12}A_{21}}{A_{11}} \end{bmatrix}$ Because EROs are rank preserving (Theorem 3.4) it follows that $A_{22} - \frac{A_{12}A_{21}}{A_{11}} \neq 0$

Therefore, $\det(A) = A_{11}A_{21} - A_{21}A_{12} \neq 0$. On the other hand, if $A_{21} \neq 0$, we see that $\det(A) \neq 0$ by adding $-A_{11}/A_{21}$ times row two to row one & applying a similar process. Thus, in either case, $\det(A) \neq 0$

Fret not, we will be expanding this proof to account for larger matricies ~~b~~briefly. But, before we do that lets briefly return to the familiar cartesian plane to discuss the geometric implications of the determinant using matrix representation & 2D coordinate points

Area of a Parallelogram Using Determinants

The angle between two vectors in $\mathbb{R}^2$ is measured by θ $(0 \leq \theta \leq \pi)$ that is formed from vectors that have the same directom & magnitude as the original vectors but emmenating, instead from the origin

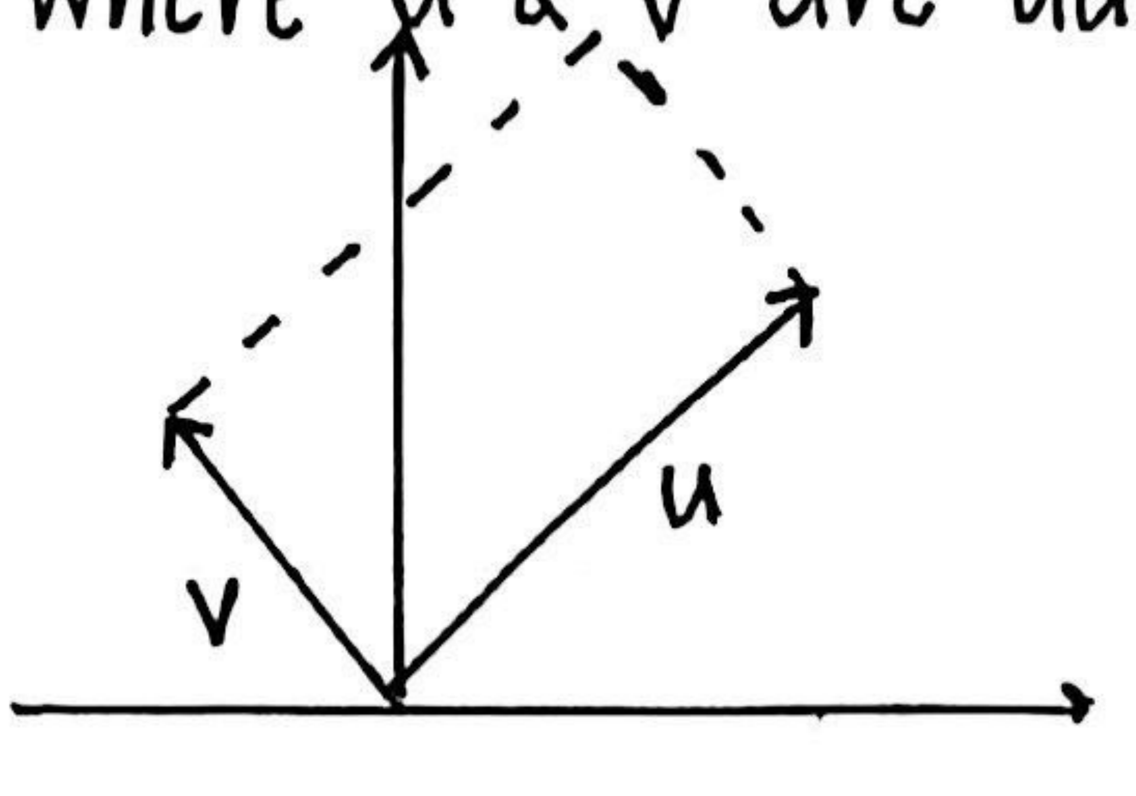


If $\beta = \{u, v\}$ is an ordered basis for $\mathbb{R}^2$ we define the **Orientation** of β to be the real number given by

$$O\begin{bmatrix} u \\ v \end{bmatrix} = \frac{\det\begin{bmatrix} u \\ v \end{bmatrix}}{\left|\det\begin{bmatrix} u \\ v \end{bmatrix}\right|}$$

meaning $-1 \leq O\begin{bmatrix} u \\ v \end{bmatrix} \leq 1$ where $O\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = 1$ & $O\begin{bmatrix} e_1 \\ -e_2 \end{bmatrix} = -1$

Any ordered set $\{u, v\}$ in $\mathbb{R}^2$ determines a parallelogram where u & v are adjacent sides



Note that if u & v are linearly dependent we are left with a line segment rather than a parallelogram

Determinants of Order n

Let $A \in M_{n\times n}(F)$. For $n \geq 2$ the **Determinant** of A is defined recursively as the following:

$$\det(A) = \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \cdot \det(\tilde{A}_{1j})$$

it is also common to see $\det(A)$ written as $|A|$

each term in this summation is called the **Cofactor** of the jth entry of the ith row of A

If we let $c_{ij} = (-1)^{i+j}\det(\tilde{A}_{ij})$ denote the cofactor of the jth entry of the ith row of A we can rewrite $\det(A)$ as the following

$$\det(A) = A_{11}c_{11} + A_{12}c_{12} + \cdots + A_{1n}c_{1n}$$

This representation of the determinant is called the **Cofactor Expansion** (along the first row)

Example:

Let $A = \begin{bmatrix} 1 & 3 & -3 \\ -3 & -5 & 2 \\ -4 & 4 & -6 \end{bmatrix}$ Using cofactor expansion along the first row of A produces the following

$$\begin{aligned}
\det(A) &= (-1)^{1+1}A_{11}\cdot\det(\tilde{A}_{11}) + (-1)^{1+2}A_{12}\cdot\det(\tilde{A}_{12}) + (-1)^{1+3}A_{13}\cdot\det(\tilde{A}_{13}) \\
&= (-1)^2(1)\cdot\det\begin{bmatrix} -5 & 2 \\ 4 & -6 \end{bmatrix} + (-1)^3(3)\cdot\det\begin{bmatrix} -3 & 2 \\ -4 & -6 \end{bmatrix} + (-1)^4(-3)\cdot\det\begin{bmatrix} -3 & -5 \\ -4 & 4 \end{bmatrix} \\
&= (-5(-6)-2(4)) - 3(-3(-6)-2(-4)) - 3(-3(4)-(-5)(-4)) \\
&= 1(22) - 3(26) - 3(32) \\
&= 40
\end{aligned}$$

The determinant of the $n\times n$ Identity Matrix is, intuitively, equal to one

$$\begin{aligned}
\det(I) &= (-1)^2(1)\cdot\det(\tilde{I}_{11}) + (-1)^3(0)\cdot\det(\tilde{A}_{12}) + \cdots + (-1)^{1+n}(0)\cdot\det(\tilde{A}_{1n}) \\
&= 1(1) + 0 + \cdots + 0, \quad \text{then by induction,} \\
&= 1
\end{aligned}$$

Theorem 4.3 & Its Long ass Proof

Theorem 4.3: The determinant of an $n \times n$ matrix is a linear function of each row when the remaining rows are held fixed. That is, for $1 \le r \le n$, we have:

$$\det\begin{bmatrix} a_1 \\ \vdots \\ a_{r-1} \\ u+kv \\ a_{r+1} \\ \vdots \\ a_n \end{bmatrix} = \det\begin{bmatrix} a_1 \\ \vdots \\ a_{r-1} \\ u \\ a_{r+1} \\ \vdots \\ a_n \end{bmatrix} + k\det\begin{bmatrix} a_1 \\ \vdots \\ a_{r-1} \\ v \\ a_{r+1} \\ \vdots \\ a_n \end{bmatrix}$$

wherever k is a scalar & u, v, & each a_i are row vectors $\in F^n$

Proof:

The proof is through mathematical Induction on n.
The result is immediatly clear if $n=1$.
Assume that for some interger $n \ge 2$ the determinant of any $(n-1) \times (n-1)$ matrix is the linear function of each row when the remaining rows are held fixed. Let A be an $n \times n$ matrix with rows $a_1, a_2, \ldots, a_n$, & suppose that for some r $(1 \le r \le n)$, we have $a_r = u + kv$ for some $u, v \in F^n$ & some scalar K
Let $u = (b_1, b_2, \ldots, b_n)$ & $v = (c_1, c_2, \cdots, c_n)$ & let B & C be the matricies obtained from A by replacing row r of A by u & v respectively
For $r > 1$ & $1 \le j \le n$, the rows of $\tilde{A}_{1j}$, $\tilde{B}_{1j}$, & $\tilde{C}_{1j}$ are the same except for row $r-1$. Moreover, row $r-1$ of $\tilde{A}_{1j}$ is given by $(b_1+kc_1, \cdots, b_{j-1}+kc_{j-1}, b_{j+1}+kc_{j+1}, \cdots, b_n+kc_n)$ which is the sum of row $r-1$ of $\tilde{B}_{1j}$ & k times row $r-1$ $\tilde{C}_{1j}$. Since $\tilde{B}_{1j}$ & $\tilde{C}_{1j}$ are $(n-1)\times(n-1)$ matricies we have $\det(\tilde{A}_{1j}) = \det(\tilde{B}_{1j}) + k\det(\tilde{C}_{1j})$ by the induction hypothesis. Thus since $A_{1j} = B_{1j} = C_{1j}$

$$\det(A) = \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \cdot \det(\tilde{A}_{1j}) = \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \cdot \left(\det(\tilde{B}_{1j}) + k\det(\tilde{C}_{1j})\right)$$

$$= \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \cdot \det(\tilde{B}_{1j}) + k \sum_{j=1}^{n} (-1)^{1+j} A_{1j} \cdot \det(\tilde{C}_{1j})$$

$$= \det(B) + k\det(C)$$

Thus, the theorem holds for any $n \times n$ matrix by mathematical induction!

Properties of Determinants

Corollary: If A has a row where every entry is zero, then $\det(A) = 0$

Corollary: If A has two identical rows, then $\det(A) = 0$

Theorem 4.5: If $A \in M_{n\times n}(F)$ & B is a matrix obtained from A by interchanging any two rows of A, then $\det(B) = -\det(A)$

Proof: Let the rows of $A \in M_{n\times n}(F)$ be $a_1, a_2, \cdots, a_n$ & let B be the matrix obtained by interchanging rows r & s where $r<s$

Thus,

$$A = \begin{bmatrix} a_1 \\ \vdots \\ a_r \\ \vdots \\ a_s \\ \vdots \\ a_n \end{bmatrix} \quad \& \quad B = \begin{bmatrix} a_1 \\ \vdots \\ a_s \\ \vdots \\ a_r \\ \vdots \\ a_n \end{bmatrix}$$

Consider the matrix obtained by replacing rows r & s in A with $a_r + a_s$. Then by theorem 3.3 we have the following

$$0 = \det\begin{bmatrix} a_1 \\ \vdots \\ a_r + a_s \\ \vdots \\ a_r + a_s \\ \vdots \\ a_n \end{bmatrix} = \det\begin{bmatrix} a_1 \\ \vdots \\ a_r \\ \vdots \\ a_r + a_s \\ \vdots \\ a_n \end{bmatrix} + \det\begin{bmatrix} a_1 \\ \vdots \\ a_s \\ \vdots \\ a_r + a_s \\ \vdots \\ a_n \end{bmatrix}$$

$$= 0 + \det(A) + \det(B) + 0$$

Therefore, $\det(B) = -\det(A)$

Theorem 4.6: Let $A \in M_{n\times n}(F)$, & let B be the matrix obtained by adding a multiple of a row of A to another row of A, then $\det(A) = \det(B)$

Properties of Determinants

Theorem 4.7: If $A, B \in M_{n\times n}(F)$, $\det(AB) = \det(A)\cdot\det(B)$

Proof: We begin by establishing the result when A is an elementary matrix. If A is an elementary matrix obtained by interchanging two rows of I_n, then $\det(A) = -1$. By theorem 3.1, AB is a matrix obtained by interchanging two rows of B. Then, by theorem 4.5 $\det(AB) = -\det(B) = \det(A)\cdot\det(B)$. Similar arguments show the same result with the other types of elementary matricies. Furthermore, if A is an $n\times n$ matrix with rank less than n, then $\det(A) = 0$ by the corollary to theorem 4.6. Since $\operatorname{rank}(AB) \leq \operatorname{rank}(A) \leq n$ by theorem 3.7 we have $\det(AB) = 0$. Thus, $\det(AB) = \det(A)\cdot\det(B)$ in this case.

On the otherhand, if A has rank n, then A is invertible & hence is the product of elementary matricies (corollary to theorem 3.6), say, $A = E_m \cdots E_2E_1$. Thus, as shown above,

$$
\begin{aligned}
\det(AB) &= \det(E_m \cdots E_2E_1B) \\
&= \det(E_m)\cdot\det(E_{m-1} \cdots E_2E_1B) \\
&\;\;\vdots \\
&= \det(E_m \cdots E_2E_1)\cdot\det(B) \\
&= \det(A)\cdot\det(B)
\end{aligned}
$$

Corollary: A matrix $A \in M_{n\times n}(F)$ is invertible if & only if $\det(A) \neq 0$. Furthermore, if A is invertible, then $\det(A^{-1}) = \dfrac{1}{\det(A)}$

Proof: if $A \in M_{n\times n}(F)$ is not invertible, then the $\operatorname{rank}(A) < n$, so $\det(A) = 0$ by the corollary to theorem 4.6

On the other hand, if $A \in M_{n\times n}(F)$ is invertible, then $\det(A)\cdot\det(A^{-1}) = \det(AA^{-1}) = \det(I_n) = 1$

Thus, by the theorem above $\det(A) \neq 0$ & $\det(A^{-1}) = \dfrac{1}{\det(A)}$

Therefore, $x_k = [\det(A)]^{-1}\cdot\det(M_k)$

Properties of Determinants

To summarize the effects of the EROs on the determinant

a. Interchanging any two rows of a matrix flips the sign of the determinant

b. Multiplying a single row by a constant multiplies the determinant by the same constant

c. Adding a multiple of one row to another row preserves the determinant.

~~Theorem 4.7: For any $A, B \in M_{n\times n}(F)$, $\det(AB) = \det(A)\cdot$ de-~~
~~Proof: If A is not invertible, then $\text{rank}(A) < n$. But ra~~
~~$= \text{rank}(A)$ & thi~~

Theorem 4.1 (Rewritten): The function $\det: M_{2\times2}(F) \to F$ is a linear function of each row of a 2×2 matrix when the other row is held fixed. That is, if $u, v, w \in F^2$ & k is a scalar then

$$\det\begin{bmatrix} u+kv \\ w \end{bmatrix} = \det\begin{bmatrix} u \\ w \end{bmatrix} + k\det\begin{bmatrix} v \\ w \end{bmatrix}$$

Proof

Let $u = (u_1, u_2)$, $v = (v_1, v_2)$ & $w = (w_1, w_2)$ then

$$\det\begin{bmatrix} u \\ w \end{bmatrix} + \det\begin{bmatrix} v \\ w \end{bmatrix} = \det\begin{bmatrix} u_1 & u_2 \\ w_1 & w_2 \end{bmatrix} + k\det\begin{bmatrix} v_1 & v_2 \\ w_1 & w_2 \end{bmatrix}$$

$$= (u_1w_2 - w_1u_2) + k(v_1w_2 - w_1v_2)$$

$$= ($$

Properties of Determinants

Theorem 4.8: For any $A \in M_{n\times n}(F)$, $\det(A) = \det(A^t)$

Proof: If A is not invertible, then $\text{rank}(A) < n$. But $\text{rank}(A) = \text{rank}(A^t)$ by the corollary to theorem 3.6. & so A^t is not invertible, thus, in this case $\det(A) = \det(A^t) = 0$
On the other hand, if A is invertible, then A is the product of elementary matricies, say $A = E_m \cdots E_2 E_1$. Since $\det(E_i) = \det(E_i^t)$ for every i, then by theorem 4.7 we have

$$\begin{aligned}
\det(A^t) &= \det(E_1^t E_2^t \cdots E_m^t) \\
&= \det(E_1^t)\cdot\det(E_2^t)\cdot \cdots \cdot\det(E_m^t) \\
&= \det(E_1)\cdot\det(E_2)\cdots\det(E_m) \\
&= \det(E_m \cdots E_2 E_1) \\
&= \det(A)
\end{aligned}$$

Thus, in either case $\det(A^t) = \det(A)$

Theorem 4.9 (Cramer's Rule): Let $Ax = b$ be the matrix form of a system of linear equations in n unknowns, where $x = (x_1, x_2, \cdots, x_n)^t$ If $\det(A) \neq 0$, then this system has a unique solution, and for each k $(k = 1, 2, \dots, n)$, $x_k = \dfrac{\det(M_k)}{\det(A)}$ where M_k is the matrix obtained by replacing the kth column of A with b.

Proof: If $\det(A) \neq 0$, then the system $Ax = b$ has a unique solution by the corollary to theorem 4.7 & theorem 3.10 For each interger k $(1 \leq k \leq n)$. Let a_k denote the kth column, of A, & X_k denote the matrix obtained by replacing the kth column of I_n with x. Then by theorem 2.13, AX_k is the $n\times n$ matrix whose ith column is

$Ae_i = a_i$ if $i \neq k$ & $Ax = b$ if $i = k$

Thus $AX_k = M_k$. Evaluating X_k by cofactor expansion along row k produces: $\det(X_k) = x_k \cdot \det(I_{n-1}) = x_k$ Hence, by theorem 4.7: $\det(M_k) = \det(AX_k) = \det(A)\cdot\det(X_k) = \det(A)\cdot x_k$
Therefore, $x_k = [\det(A)]^{-1}\cdot\det(M_k)$

Cramer's Rule Example

Consider the following system of equations

$$x_1 + 2x_2 + 3x_3 = 2$$
$$x_1 \quad\quad + x_3 = 3$$
$$x_1 + x_2 - x_3 = 1$$

The matrix representation $Ax=b$ where

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 0 & 1 \\ 1 & 1 & -1 \end{bmatrix} \quad \& \quad b = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$$

because $\det(A) = 6 \neq 0$ we can solve using Cramer's rule

$$x_1 = \frac{\det(M_1)}{\det(A)} = \frac{\det\begin{bmatrix} 2 & 2 & 3 \\ 3 & 0 & 1 \\ 1 & 1 & -1 \end{bmatrix}}{\det(A)} = \frac{15}{6} = \frac{5}{2}$$

$$x_2 = \frac{\det(M_2)}{\det(A)} = \frac{\det\begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 1 \\ 1 & 1 & -1 \end{bmatrix}}{\det(A)} = \frac{-6}{6} = -1$$

$$x_3 = \frac{\det(M_3)}{\det(A)} = \frac{\det\begin{bmatrix} 1 & 2 & 2 \\ 1 & 0 & 3 \\ 1 & 1 & 1 \end{bmatrix}}{\det(A)} = \frac{3}{6} = \frac{1}{2}$$

Thus, the unique solution to this sustem is
$(x_1, x_2, x_3) = (\frac{5}{2}, -1, \frac{1}{2})$

Summary of Determinants

1. Interchanging any two rows of a matrix makes the determinant negative
2. If B is a matrix obtained by multiplying each entry of of a single row by a scalar k then $det(B) = k\,det(A)$
3. If B is a matrix obtained by adding a multiple of one row to another then the determinant remains unchanged

4. The determinant of an upper triangular matrix is the product of its diagonal enteries, particularly, $det(I_n) = 1$
5. If two rows (or columns) are identical within a matrix then its determinant is zero
6. For any $n \times n$ matricies A & B, $det(AB) = det(A) \cdot det(B)$
7. An $n \times n$ matrix is invertible if & only if $det(A) \neq 0$. Furthermore, if A is invertible then $det(A^{-1}) = 1/det(A)$
8. For any $n \times n$ matrix A, $det(A) = det(A^t)$
9. If A & B are similar matricies then $det(A) = det(B)$

Theorem 4.4: The determinant of a square matrix can be evaluated by cofactor expansion along any row, that is if $A \in M_{n \times n}(F)$, then for any interger i $(1 \leq i \leq n)$

$$det(A) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \cdot det(\tilde{A}_{ij})$$

Proof: Cofactor expansion along the first row of A gives the determinant by definition. Beyond that, if $i > 1$ row i of A can be written as $\sum_{j=1}^{n} A_{ij} e_j$. For $1 \leq j \leq n$, let B_j denote the matrix obtained from A by replacing row i of A by e_j. Then by theorem 4.3 we have

$$det(A) = \sum_{j=1}^{n} A_{ij} \cdot det(B_j) = \sum_{j=1}^{n} (-1)^{i+j} A_{ij} \cdot det(\tilde{A}_{ij})$$

The Area of a Parallelogram

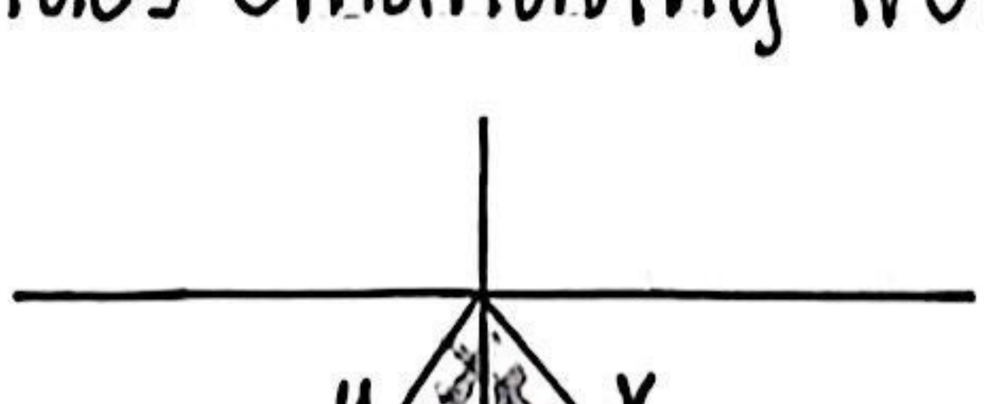
Let $\beta = \{u, v\}$ be an ordered basis of $\mathbb{R}^2$. We define the **Orientation** of β to be the following: which only has two possible values; ± 1

$$\sigma\begin{bmatrix} u \\ v \end{bmatrix} = \frac{\det\begin{bmatrix} u \\ v \end{bmatrix}}{\left|\det\begin{bmatrix} u \\ v \end{bmatrix}\right|}$$

> Recall: A coordinate system is called **Right-Handed** if u can be rotated counterclockwise through an angle to coincide with v, otherwise its considered **Left-Handed**

In general, if $\{u, v\}$ forms a right-handed coordinate system then, $\sigma\begin{bmatrix} u \\ v \end{bmatrix} = 1$ which, for convience, we will say is also true if $\{u, v\}$ is linearly dependent

Additionally, any ordered set $\{u, v\}$ where $u, v \in \mathbb{R}^2$ defines a parallelogram with the vectors u & v acting as adjacent sides emanating from the origin. Observe, that if $\{u, v\}$ is linearly dependent (i.e. they are parallel. Then the parallelogram collapses to a line segment



Right-Handed

Left-Handed

The area of the parallelogram is defined by the following

$$\mathbb{A}\begin{bmatrix} u \\ v \end{bmatrix} = \sigma\begin{bmatrix} u \\ v \end{bmatrix} \cdot \det\begin{bmatrix} u \\ v \end{bmatrix}$$

→ Since $\mathbb{A}$ must be positive but the determinant can be negative, we know that $\det\begin{bmatrix} u \\ v \end{bmatrix} \neq \mathbb{A}\begin{bmatrix} u \\ v \end{bmatrix}$ but we will show that

$$\mathbb{A}\begin{bmatrix} u \\ v \end{bmatrix} = \left|\det\begin{bmatrix} u \\ v \end{bmatrix}\right| = \sigma\begin{bmatrix} u \\ v \end{bmatrix} \cdot \det\begin{bmatrix} u \\ v \end{bmatrix}$$

Area of a Parallelogram 2

The technique we will use to show that $\mathbb{A}\begin{bmatrix}u\\v\end{bmatrix} = \sigma\begin{bmatrix}u\\v\end{bmatrix}\cdot\det\begin{bmatrix}u\\v\end{bmatrix}$ is indirect but it can be generalized to higher dimensional $\mathbb{R}^n$

Recall: a function $\delta : M_{2\times2}(\mathbb{R}) \to \mathbb{R}$ if & only if it satisfies

I. δ is a linear function of each row of the matrix when the other is held fixed

II. If the two rows are identical then $\delta(A) = 0$

III. If the two rows form the identity matrix then $\delta(A) = 1$

We need to show that the function

Q E D

$$\delta\begin{bmatrix}u\\v\end{bmatrix} = \sigma\begin{bmatrix}u\\v\end{bmatrix}\cdot\mathbb{A}\begin{bmatrix}u\\v\end{bmatrix}$$

satisfies the conditions specified here →

a. We begin by showing that for any real number c

$$\delta\begin{bmatrix}u\\cv\end{bmatrix} = c\cdot\delta\begin{bmatrix}u\\v\end{bmatrix}$$

If $c = 0$ then $\delta\begin{bmatrix}u\\cv\end{bmatrix} = \sigma\begin{bmatrix}u\\0\end{bmatrix}\cdot\mathbb{A}\begin{bmatrix}u\\0\end{bmatrix}$ $1\cdot 0 = 0$

alternatively, when $c \neq 0$ we can think of cu as scaling one of the points of of our parallelogram

$$\delta\begin{bmatrix}u\\cv\end{bmatrix} = \sigma\begin{bmatrix}u\\cv\end{bmatrix}\cdot\mathbb{A}\begin{bmatrix}u\\cv\end{bmatrix} = \left(\frac{c}{|c|}\cdot\sigma\begin{bmatrix}u\\v\end{bmatrix}\right)\cdot\left(|c|\cdot\mathbb{A}\begin{bmatrix}u\\v\end{bmatrix}\right)$$

$$= c\cdot\sigma\begin{bmatrix}u\\v\end{bmatrix}\cdot\mathbb{A}\begin{bmatrix}u\\v\end{bmatrix}$$

↑ by our definition of orientation

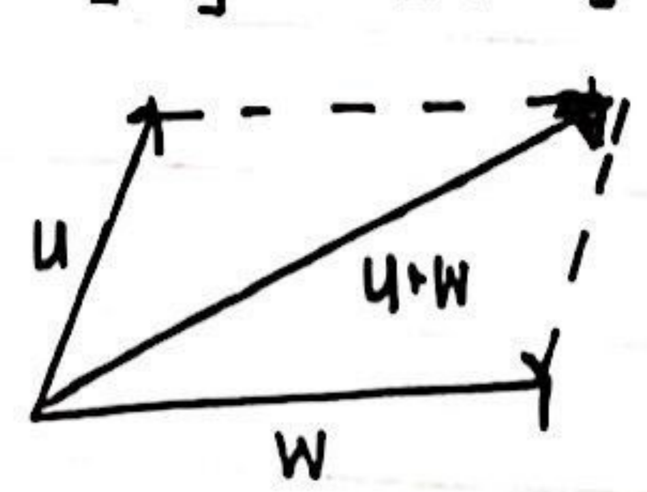
$$= c\cdot\delta\begin{bmatrix}u\\v\end{bmatrix}$$

We need to show that for any real numbers a & b & $u, w \in \mathbb{R}^2$

$$\delta\begin{bmatrix}u\\au+bw\end{bmatrix} = b\cdot\delta\begin{bmatrix}u\\w\end{bmatrix}$$

because the parallelograms defined by $\{u, w\}$ & $\{u, u+w\}$ have u for the first element we know they have the same height so it follows geometrically that

$\mathbb{A}\begin{bmatrix}u\\w\end{bmatrix} = \mathbb{A}\begin{bmatrix}u\\u+w\end{bmatrix}$ since $\mathbb{A}\begin{bmatrix}u\\w\end{bmatrix}$ = length of w · height where w is the base of the parallelogram



Area of a Parallelogram 3

If $a=0$, then $\delta\begin{bmatrix} u \\ au+bw \end{bmatrix} = \delta\begin{bmatrix} u \\ bw \end{bmatrix} = b\cdot\delta\begin{bmatrix} u \\ w \end{bmatrix}$.

Otherwise, if $a\neq 0$, then $\delta\begin{bmatrix} u \\ au+bw \end{bmatrix} = a\cdot\delta\begin{bmatrix} u \\ u+\frac{b}{a}w \end{bmatrix} = a\cdot\delta\begin{bmatrix} u \\ \frac{b}{a}w \end{bmatrix} = b\cdot\delta\begin{bmatrix} u \\ w \end{bmatrix}$

Thus, in either case we have what we need in order to prove that δ is closed under addition which is demonstrated by verifying that

$$\delta\begin{bmatrix} u \\ v_1+v_2 \end{bmatrix} = \delta\begin{bmatrix} u \\ v_1 \end{bmatrix} + \delta\begin{bmatrix} u \\ v_2 \end{bmatrix}$$ for all $u, v_1, v_2 \in \mathbb{R}^2$.

If $u=0$ the result is trivial

Thus for $u\neq 0$ let $w\in\mathbb{R}^2$ such that $\{u,w\}$ is linearly independent & any $v_1, v_2 \in \mathbb{R}^2$ there exists scalars a_i & b_i such that $v_i = a_i u + b_i w$ $(i=1,2)$. Thus,

$$\delta\begin{bmatrix} u \\ v_1+v_2 \end{bmatrix} = \delta\begin{bmatrix} u \\ (a_1+a_2)u+(b_1+b_2)w \end{bmatrix} = (b_1+b_2)\,\delta\begin{bmatrix} u \\ w \end{bmatrix}$$

$$= \delta\begin{bmatrix} u \\ a_1u+b_1w \end{bmatrix} + \delta\begin{bmatrix} u \\ a_2u+b_2w \end{bmatrix} = \delta\begin{bmatrix} u \\ v_1 \end{bmatrix} + \delta\begin{bmatrix} u \\ v_2 \end{bmatrix}$$

A similar argument shows that for all $u_1, u_2 \in \mathbb{R}^2$

$$\delta\begin{bmatrix} u_1+u_2 \\ v \end{bmatrix} = \delta\begin{bmatrix} (a_1+a_2)v+(b_1+b_2)w \\ v \end{bmatrix} = (b_1+b_2)\,\delta\begin{bmatrix} w \\ v \end{bmatrix}$$

$$= \delta\begin{bmatrix} a_1v+b_1w \\ v \end{bmatrix} + \delta\begin{bmatrix} a_2v+b_2w \\ v \end{bmatrix} = \delta\begin{bmatrix} u_1 \\ v \end{bmatrix} + \delta\begin{bmatrix} u_2 \\ v \end{bmatrix}$$

b. Since $A\begin{bmatrix} u \\ u \end{bmatrix} = 0$ for any $u\in\mathbb{R}^2$ it follows that $\delta\begin{bmatrix} u \\ u \end{bmatrix} = \sigma\begin{bmatrix} u \\ u \end{bmatrix}\cdot A\begin{bmatrix} u \\ u \end{bmatrix} = 0$

& finally...

C. let $\{e_1, e_2\}$ be the standard ordered basis of $\mathbb{R}^2$, then

$$\delta\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \sigma\begin{bmatrix} e_1 \\ e_2 \end{bmatrix}\cdot A\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = 1\cdot 1 = 1$$

Thus, the area of a parallelogram given by $\{u,v\}$ is given by $\delta\begin{bmatrix} u \\ v \end{bmatrix} = \det\begin{bmatrix} u \\ v \end{bmatrix}$ $\blacksquare$

Eigenvalues & Eigenvectors

A linear operator T on a finite-dimensional vector space V is said to be **diagonalizable** if there is an ordered basis β of V such that $[T]_\beta$ is a diagonal matrix. Additionally a square matrix A is called diagonalizable if L_A is diagonalizable

If $D=[T]_\beta$ is a diagonal matrix, then for each vector $v_j \in \beta$ we have:

$$T(v_j) = \sum_{i=1}^{n} D_{ij} v_i = D_{jj} v_j = \lambda_j v_j \quad \text{where } \lambda_j = D_{jj}$$

Conversely, if $\beta = \{v_1, v_2, \dots, v_n\}$ is an ordered basis of V such that $T(v_j) = \lambda_j v_j$ for some scalars $\lambda_1, \lambda_2, \dots, \lambda_n$ then,

$$[T]_\beta = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

Each vector $v \in \beta$ satisfies the condition that $T(v) = \lambda(v)$ for some scalar λ. Moreover, since β is a basis $v \neq 0$

Let T be a linear operator on a vectorspace V. a non-zero vector $v \in V$ is called an **eigenvector** of T if there exists a scalar λ such that $T(v) = \lambda v$. This scalar λ is called the **eigenvalue** corresponding to v

Alternatively...

Let $A \in M_{n\times n}(F)$. A vector $v \in F^n$ is called an **eigenvector** of A if $Av = \lambda v$ (meaning v is an eigenvector of L_A) for some scalar λ. This scalar λ is the corresponding **eigenvalue** of v

Note: You may see characteristic vector or characteristic value used in place of eigenvector & eigenvalue respectively

Theorem 5.1: A linear operator T on a finite-dimensional vector space V is diagonalizable if & only if there exists an ordered basis β for V containing eigenvectors of T. Furthermore, if T is diagonalizable, $\beta = \{v_1, v_2, \cdots, v_n\}$ is an ordered basis of eigenvectors of T, & $D = [T]_\beta$, then D is a diagonal matrix & D_{jj} is the eigenvalue corresponding to v_j for $1 \le j \le n$

Corollary: A matrix $A \in M_{n \times n}(F)$ is diagonalizable if & only if there exists an ordered basis for F^n consisting of eigenvectors of A. Furthermore, if $\{v_1, v_2, \ldots, v_n\}$ is an ordered basis for F^n consisting of eigenvectors of A, & Q is the $n \times n$ matrix whose jth column is v_j for $j = 1, 2, \ldots, n$, then $D = Q^{-1}AQ$ is a diagonal matrix such that D_{jj} is the eigenvalue of A corresponding to v_j. Hence, A is diagonalizable if & only if it is similar to a diagonal matrix.

Example: Let $A = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}$, $v_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ & $v_2 = \begin{bmatrix} 3 \\ 4 \end{bmatrix}$ Then

$L_A(v_1) = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} 1 \\ -1 \end{bmatrix} = -2\begin{bmatrix} 1 \\ -1 \end{bmatrix} = -2v_1$ Thus, v_1 is an eigenvector of A with eigenvalue -2

$L_A(v_2) = \begin{bmatrix} 1 & 3 \\ 4 & 2 \end{bmatrix}\begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 15 \\ 20 \end{bmatrix} = 5\begin{bmatrix} 3 \\ 4 \end{bmatrix} = 5v_2$ Thus, v_2 is an eigenvector with eigen value 5

& since $\beta = \{v_1, v_2\}$ is an ordered basis of $\mathbb{R}^2$ consisting of eigenvectors, both A & L_A are diagonalizable

If $Q = \begin{bmatrix} 1 & 3 \\ -1 & 4 \end{bmatrix}$ then $Q^{-1}AQ = [L_A]_\beta = \begin{bmatrix} -2 & 0 \\ 0 & 5 \end{bmatrix}$

Theorem 5.2: Let $A \in M_{n\times n}(F)$. Then a scalar λ is an eigenvalue of A if & only if $\det(A-\lambda I_n)=0$

Proof: A scalar λ is an eigenvalue of A if & only if there exists a nonzero vector $v \in F^n$ such that $Av=\lambda v$, that is, $(A-\lambda I_n)(v)=0$. By theorem 2.5, this is true if & only if $A-\lambda I_n$ is not invertible which is equivalent to the statement $\det(A-\lambda I_n)=0$

Let $A \in M_{n\times n}(F)$. The polynomial given by $f(t)=\det(A-tI_n)$ is called the **characteristic polynomial**

Example: Let $A=\begin{bmatrix}1 & 1\\ 4 & 1\end{bmatrix} \in M_{2\times 2}(\mathbb{R})$ to find the eigenvalues of A we first compute the characteristic polynomial of A

$\det(A-tI_2)=\det\begin{bmatrix}1-t & 1\\ 4 & 1-t\end{bmatrix}=(1-t)^2-4=t^2-2t-3=(t-3)(t+1)$ Thus, it follows from theorem 5.2 that the only eigenvalues of A are 3 & -1

Example: Let T be the linear operator on $P_2(\mathbb{R})$ defined by $T(f(x))=f(x)+(x+1)f'(x)$ & let β be the standard ordered basis for $P_2(\mathbb{R})$, & let $A=[T]_\beta$ then,

$A=\begin{bmatrix}1 & 1 & 0\\ 0 & 2 & 2\\ 0 & 0 & 3\end{bmatrix}$ & the charecteristic polynomial of A is:

$\det(A-tI_3)=\det\begin{bmatrix}1-t & 1 & 0\\ 0 & 2-t & 2\\ 0 & 0 & 3-t\end{bmatrix}=(1-t)(2-t)(3-t)$
$=-(t-1)(t-2)(t-3)$ hence, the only eigenvalues of A are $\lambda=1$, $\lambda=2$, & $\lambda=3$

Theorem 5.3: Let $A \in M_{n\times n}(F)$

a. The characteristic polynomial of A is a polynomial of degree n with leading coefficient $(-1)^n$

b. A has at most n distinct eigenvalues.

Theorem 5.4: Let $A \in M_{n\times n}(F)$ & let λ be an eigenvalue of A. $v \in F^n$ is an eigenvector of A corresponding to λ if & only if $v \neq \emptyset$ & $(A-\lambda I_n)v = \emptyset$

To find the eigenvectors of a linear operator T on an n-dimensional vector space V, select an ordered basis β of V & let $A = [T]_\beta$. Recall that for $v \in V$, $\phi_\beta(v) = [v]_\beta$, the coordinate vector of v relative to β. Suppose v is an eigenvector of T corresponding to λ. Then, $T(v) = \lambda v$. Hence,

$$A\phi_\beta(V) = L_A\phi_\beta(V) = \phi_\beta T(v) = \phi_\beta(\lambda V) = \lambda\phi_\beta(V)$$

Thus, $\phi_\beta(V) \neq \emptyset$ since ϕ_β is an isomorphism; hence, $\phi_\beta(V)$ is an eigenvector of A

Example: On the previous page we demonstrated that if $T(f(x)) = f(x) + (x+1)f'(x)$ & β is the standard basis of $P_2(\mathbb{R})$

$$A = [T]_\beta = \begin{bmatrix} 1 & 1 & 2 \\ \emptyset & 2 & \emptyset \\ \emptyset & \emptyset & 3 \end{bmatrix}$$ then $\lambda_1 = 1$, $\lambda_2 = 2$, & $\lambda_3 = 3$

Consider $\lambda_1 = 1$

$$B_1 = A - \lambda_1 I_3 = \begin{bmatrix} \emptyset & 1 & \emptyset \\ \emptyset & 1 & 2 \\ \emptyset & \emptyset & 2 \end{bmatrix}$$ → The solution to the homogeneous system expressed as a column vector is an eigenvector corresponding to $\lambda_1 = 1$ (assuming that the solution is not trivial)

↓

$$\begin{aligned} x_2 \quad &= \emptyset \\ x_2 + 2x_3 &= \emptyset \\ 2x_3 &= \emptyset \end{aligned}$$

⟶ x_1 is a free variable & thus $\begin{bmatrix} 1 \\ \emptyset \\ \emptyset \end{bmatrix} t$ where $t \neq \emptyset$ represents all the eigenvectors corresponding to $\lambda_1 = 1$

Diagonalizability

Theorem 5.5: Let T be a linear operator on a vectorspace V, & let $\lambda_1, \lambda_2, \cdots, \lambda_n$ be distinct eigenvalues of T. For each $i = 1, 2, \cdots, k$, let S_i be a finite set of eigenvectors of T corresponding to λ_i. If each S_i $(i = 1, 2, \cdots, k)$, is linearly independent, then $S_1 \cup S_2 \cup \cdots \cup S_k$ is linearly independent

Proof: The proof is by mathematical induction on k. If $k = 1$, there is nothing to prove. So assume the theorem holds for $k-1$ distinct eigenvalues, where $k-1 \geq 1$, & that we have k distinct eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_k$ of T. For each $i = 1, 2, \cdots, k$ let $S_i = \{v_{i1}, v_{i2}, \cdots, v_{in_i}\}$ be a linearly independent set of eigenvectors of T corresponding to λ_i. We want to show that $S = S_1 \cup S_2 \cup \cdots \cup S_k$ is linearly independent.

Consider any scalars $\{a_{ij}\}$, where $i = 1, 2, \cdots, k$ & $j = 1, 2, \cdots, n_i$ such that $\sum_{i=1}^{k} \sum_{j=1}^{n_i} a_{ij} v_{ij} = \emptyset$. Because v_{ij} is an eigenvector of T corresponding to λ_i, applying $T - \lambda_k I_n$ to both sides of this equation yields the following $\sum_{i=1}^{k-1} \sum_{j=1}^{n_i} a_{ij} (\lambda_i - \lambda_k) v_{ij} = \emptyset$

But, since $S_1 \cup S_2 \cup \cdots \cup S_{k-1}$ is linearly independent, by the induction hypothesis which implies that $a_{ij}(\lambda_i - \lambda_k) = \emptyset$ for $i = 1, 2, \cdots, k-1$ & $j = 1, 2, \cdots, n_i$. Since $\lambda_1, \lambda_2, \cdots, \lambda_k$ are distinct it follows that $\lambda_i - \lambda_k \neq \emptyset$ for $1 \leq i \leq k-1$. Hence $a_{ij} = \emptyset$ for $i = 1, 2, \cdots, k-1$ & $j = 1, 2, \cdots, n_i$ & therefore, the summation reduces to

$\sum_{j=1}^{n_k} a_{kj} v_{kj} = \emptyset$ But, S_k is also linearly independent & so

$a_{kj} = \emptyset$ for $j = 1, 2, \cdots, n_k$. Consequently, $a_{ij} = \emptyset$ for $i = 1, 2, \cdots, k$ & $j = 1, 2, \cdots, n_i$ proving that S is linearly Independent

Corollary: Let T be a linear operator on an n-dimensional vector space V. If T has n distinct eigenvalues, then T is diagonalizable.

↳ Note: if T is diagonalizable that does not necessarily mean that T has n distinct eigenvalues

Diagonalizability

Example: Let $A=\begin{bmatrix}1&1\\1&1\end{bmatrix}\in M_{n\times n}(\mathbb{R})$. The charecteristic polynomial of A is given by $\det(A-tI_2)=\det\begin{bmatrix}1-t&1\\1&1-t\end{bmatrix}=t(t-2)$ & thus the eigenvalues of A are 0 & 2

since $n=2$ & A has 2 eigenvalues, A is diagonalizable

A polynomial $f(t)$ in $P(F)$ **splits over** F if there are scalars $c, a_1, \cdots, a_n$ (not necessarily distinct) in F such that $f(t)=c(t-a_1)(t-a_2)\cdots(t-a_n)$

Theorem 5.6: The characteristic polynomial of any diagonalizable linear operator on a vector space V over a field F splits over F

Proof: Let T be a diagonalizable linear operator on the n-dimensional vectorspace V, & let β be an ordered basis for V such that $[T]_\beta=D$ is a diagonal matrix. Suppose that

$$D=\begin{bmatrix}\lambda_1&0&\cdots&0\\0&\lambda_2&\cdots&0\\\vdots&\vdots&&\vdots\\0&0&\cdots&\lambda_n\end{bmatrix}$$

& let $f(t)$ be the characteristic polynomial of T. Then,

$$f(t)=\det(D-tI_n)=\det\begin{bmatrix}\lambda_1-t&0&\cdots&0\\0&\lambda_2-t&\cdots&0\\\vdots&\vdots&&\vdots\\0&0&\cdots&\lambda_n-t\end{bmatrix}=$$

$$(\lambda_1-t)(\lambda_2-t)\cdots(\lambda_n-t)=(-1)^n(t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_n)=$$

If T is a linear operator on an n-dimensional vector space & T is diagonalizable but has less than n distinct eigenvalues then the characteristic polynomial must have at least one repeated solution

Let T be a linear operator on a vectorspace V & let λ be an eigenvalue of T. Define $E_\lambda = \{x \in V : T(x) = \lambda x\} = N(T - \lambda I_V)$
The set E_λ is called an **eigenspace** of T corresponding to the eigenvalue λ.
Analogously, the **eigenspace** of a square matrix A corresponding to the eigenvalue λ to be the eigenspace of L_A corresponding to λ.

Theorem 5.7: Let T be a linear operator on a finite-dimensional vector space V, & let λ be an eigenvalue having multiplicity m. Then, $1 \leq \dim(E_\lambda) \leq m$

Proof: Choose an ordered basis $\{v_1, v_2, \cdots, v_p\}$ for E_λ, extend it into an ordered basis $\beta = \{v_1, v_2, \cdots, v_p, v_{p+1}, \cdots, v_n\}$ for V & let $A = [T]_\beta$. Observe that $v_i \ (1 \leq i < p)$ is an eigenvector of T corresponding to λ, & therefore $A = \begin{bmatrix} \lambda I_p & B \\ O & C \end{bmatrix}$
Thus, the characteristic polynomial of T is found to be $f(t) = \det(A - tI_n) = \det \begin{bmatrix} (\lambda - t) I_p & B \\ O & C - tI_{n-p} \end{bmatrix}$
$= \det((\lambda - t) I_p) \cdot \det(C - tI_{n-p})$
$= (\lambda - t)^p g(t)$ Where $g(t)$ is a polynomial. Thus, $(\lambda - t)^p$ is a factor of $f(t)$, & hence, the multiplicity of λ is at least p
But, $\dim(E_\lambda) = p$, so $\dim(E_\lambda) \leq m$

Eigenspace Example

Let T be the linear operator on $\mathbb{R}^3$ defined by

$$T\begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 4a_1 + a_3 \\ 2a_1 + 3a_2 + 2a_3 \\ a_1 + 4a_3 \end{bmatrix} \quad [T]_\beta = \begin{bmatrix} 4 & 0 & 1 \\ 2 & 3 & 2 \\ 1 & 0 & 4 \end{bmatrix}$$

Hence the chacteristic polynomial of T is

$$\det([T]_\beta - tI_3) = \det\begin{bmatrix} 4-t & 0 & 1 \\ 2 & 3-t & 2 \\ 1 & 0 & 4-t \end{bmatrix} = -(t-5)(t-3)^2$$

So the eigenvalues of T are $\lambda_1 = 5$, & $\lambda_2 = 3$ with multiplicity 2

$$E_{\lambda_1} = N(T-\lambda_1 I_3) = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 : \begin{bmatrix} -1 & 0 & 1 \\ 2 & -2 & 2 \\ 1 & 0 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$$

Thus, E_{λ_1} is the solution space to the system

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

Thus, $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$ is a basis for E_{λ_1}. Hence, $\dim(E_{\lambda_1}) = 1$

Similarly,

$$E_{\lambda_2} = N(T-\lambda_2 I_3) = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 : \begin{bmatrix} 1 & 0 & 1 \\ 2 & 0 & 2 \\ 1 & 0 & 1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}$$

Thus,

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = s\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + t\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

, for all $s, t \in \mathbb{R}$ it follows that $\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ is a basis of E_{λ_2} & $\dim(E_{\lambda_2}) = 2$

Observe that by theorem 5.5 $\begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ & $\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$ are linearly independent & hence the three vectors form a basis of $\mathbb{R}^3$ consisting of eigenvectors, hence T is diagonalizable!

Theorem 5.8: Let T be a linear operator on a finite-dimensional vectorspace V such that the characteristic polynomial of T splits. Let $\lambda_1, \lambda_2, \dots, \lambda_k$ be distinct eigenvalues of T. Then,

a. T is diagonalizable if & only if the multiplicity of λ_i is equal to $\dim(E_{\lambda_i})$ for all i

b. If T is diagonalizable & β_i is an ordered basis for E_{λ_i} for each i, then $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ is an ordered basis for V consisting of eigenvectors of T

Proof: For each i, let m_i denote the multiplicity of λ_i, $d_i = \dim(E_{\lambda_i})$ & $n = \dim(V)$

First, suppose that T is diagonalizable. Let β be a basis for V consisting of eigenvectors of T. For each i, let $\beta_i = \beta \cap E_{\lambda_i}$, the set of vectors in β that are eigenvectors corresponding to λ_i, & let n_i denote the number of vectors in β_i. Then, $n_i \leq d_i$ for each i because β_i is a linearly independent subset of a subspace of dimension d_i, & $d_i \leq m_i$ by theorem 5.7. The n_is sum to n because β contains n vectors. The m_is also sum to n because the degree of the charecteristic polynomial of T is equal to the sum of the multiplicities of the eigenvalues. Thus,

$$n = \sum_{i=1}^{k} n_i \leq \sum_{i=1}^{k} d_i \leq \sum_{i=1}^{k} m_i = n$$

Then it follows that

$$\sum_{i=1}^{k} (m_i - d_i) = 0$$

Since $(m_i - d_i) \geq 0$ for all i, we conclude that $m_i = d_i$ for all i

Conversely, suppose that $m_i = d_i$ for all i. We simultaneously show that T is diagonalizable & prove b. For each i, let β_i be an ordered basis for E_{λ_i}, & let $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$. By theorem 5.5, β is linearly independent Furthermore, since $d_i = m_i$ for all i, β contains

$$\sum_{i=1}^{k} d_i = \sum_{i=1}^{k} m_i = n$$

vectors. Therefore, β is an ordered basis of V consisting of eigenvectors of V, thus we can conclude that T is diagonalizable

Tests for Diagonalizability

Let T be a linear operator on an n-dimensional vector space V. Then T is diagonalizable if & only if both of the following conditions hold.

1. The characteristic polynomial splits
2. For each eigenvalue λ of T, the multiplicity of λ equals nullity$(T-\lambda I_n)$, that is, the multiplicity of λ equals n-rank$(T-\lambda I_n)$

When testing T for diagonalizability, it is usually easiest to choose a convienient basis for V & work with $B=[T]_\beta$

Example: Let

$$A=\begin{bmatrix}3 & 1 & 0\\ 0 & 3 & 0\\ 0 & 0 & 4\end{bmatrix} \in M_{3\times 3}(\mathbb{R})$$

To test for diagonalizability we first find the charecteristic polynomial $\det(A-It)$

$$\det(A-tI)=\det\begin{bmatrix}3-t & 1 & 0\\ 0 & 3-t & 0\\ 0 & 0 & 4-t\end{bmatrix} = -(t-4)(t-3)^2$$

$3-\text{rank}(A-\lambda_2 I)=1$ Which is not the multiplicity of $t_2=3$

Thus, A is not diagonalizable!

Diagonalizability Example

Let T be the linear operator on $P_2(\mathbb{R})$ defined by
$T(f(x)) = f(1) + f'(0)x + (f'(0) + f''(0))x^2$

First, lets verify that T is diagonalizable. Let α denote the standard ordered basis of $P_2(\mathbb{R})$ ($\alpha = \{1, x, x^2\}$) & $B = [T]_\alpha$ then,

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

The characteristic polynomial of T is given by $p(t) = -(t-1)^2(t-2)$ since T has eigenvalues of $\lambda_1 = 1$ & $\lambda_2 = 2$ with multiplicities 2 & 1 respectively & since $\lambda_2 = 2$ has multiplicity of 1 we know $\dim(E_{\lambda_2}) = 1$. But, we still need to verify that $\dim(E_{\lambda_1}) = 2$

$$3 - \operatorname{rank}(B - \lambda_1 I) = 3 - \operatorname{rank}\begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} = 3 - 1 = 2$$

Thus, T is diagonalizable!

Next, we need to find an ordered basis of $\mathbb{R}^3$ γ consisting of eigenvectors of B. We consider each eigenvalue seperately. The eigenspace corresponding to $\lambda_1 = 1$ is given by

$$E_{\lambda_1} = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \;\middle|\; \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \right\}$$

which is the solution space of $x_2 + x_3 = 0$ which has a basis

$$\gamma_1 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix} \right\}$$

The eigenspace corresponding to $\lambda_2 = 2$ is given by

$$E_{\lambda_2} = \left\{ \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \in \mathbb{R}^3 \;\middle|\; \begin{bmatrix} -1 & 1 & 1 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0 \right\}$$

which is the solution space to the system

$$\begin{aligned} -x_1 + x_2 + x_3 &= 0 \\ x_2 &= 0 \end{aligned}$$

Thus, $\gamma_2 = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$

Then, $\gamma = \left\{ \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix} \right\}$ is an ordered basis of $\mathbb{R}^3$ consisting of eigen vectors of B

Finally, note that the vectors in γ are the coordinate vectors relative to α of the vectors in that set $\beta = \{1, -x + x^2, 1 + x^2\}$ which is an ordered basis of $P_2(\mathbb{R})$ consisting of eigenvectors of T. Thus $\longrightarrow$

$$[T]_\beta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Extending a Solution Space Basis

Let $V = \{(x_1, x_2, x_3, x_4, x_5) \in \mathbb{R}^5 \mid x_1 + 7x_2 + 5x_3 - 4x_4 + 2x_5 = 0\}$

$S = \{(-2, 0, 0, -1, -1), (1, 1, -2, -1, -1), (-5, 1, 0, 1, 1)\}$ is a linearly independent subset of V

To extend S into a basis for V lets first find a basis β of V in general. We can find this by solving the "system" of equation which can be written as $x_1 = -7x_2 - 5x_3 + 4x_4 - 2x_5$. We can assign paremetric variables to x_2, x_3, x_4 & x_5. $\rightarrow (x_1, x_2, x_3, x_4, x_5) = (-7t_1 - 5t_2 + 4t_3 - 2t_4, t_1, t_2, t_3, t_4) =$

5.3*: Matrix Limits & Markov Chains

The *limit* of a sequence of powers $A, A^2, \ldots, A^n, \ldots$ where A is a square matrix with complex entries can be applied in statistics/probability

The limit of a sequence of complex numbers $\{z_m : m = 1, 2, \ldots\}$ can be defined in terms of the limits of the sequences of the real & imaginary parts.: If $z_m = r_m + i s_m$, where r_m & s_m are real numbers & $i = \sqrt{-1}$, then

$\lim_{m\to\infty} z_m = \lim_{m\to\infty} r_m + i \lim_{m\to\infty} s_m$ provided that the limits exist.

Let $L, A_1, A_2, \ldots$ be $n \times p$ matricies with complex enteries. The sequence $A_1, A_2, \ldots$ is said to **converge** to the $n \times p$ matrix L, called the **limit** of the sequence if

$\lim_{m\to\infty} (A_m)_{ij} = L_{ij}$ for all $1 \le i \le n$ & $1 \le j \le p$ to denote that L is the limit of the sequence, we write $\lim_{m\to\infty} A_m = L$

Example #1: if

$$A_m = \begin{bmatrix} 1 - \frac{1}{m} & \left(\frac{-3}{4}\right)^m & \frac{3m^2}{m^2+1} + i\left(\frac{2m+1}{m-1}\right) \\ \left(\frac{i}{2}\right)^m & 2 & \left(1 + \frac{1}{m}\right)^m \end{bmatrix}$$ Then,

$$\lim_{m\to\infty} A_m = \begin{bmatrix} 1 & 0 & 3+2i \\ 0 & 2 & e \end{bmatrix}$$

If $\lim_{m\to\infty} a_m$ exists, then,

$$\lim_{m\to\infty} c a_m = c\left(\lim_{m\to\infty} a_m\right)$$

Diagonalizability & Systems of Equations

Consider the system of differential equations

$x_1' = 3x_1 + x_2 + x_3$
$x_2' = 2x_1 + 4x_2 + 2x_3$
$x_3' = -x_1 - x_2 + x_3$

where, for each i, $x_i = x_i(t)$ is a differentiable real-valued function of the real variable t. Clearly, this system has a valid solution, namely, the solution in which each $x_i(t)$ is the zero function. Let's find all the solutions.

Let $x: \mathbb{R} \to \mathbb{R}^3$ be the function defined by $x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix}$. The derivative of x, denoted x', is defined as

$$x'(t) = \begin{bmatrix} x_1'(t) \\ x_2'(t) \\ x_3'(t) \end{bmatrix}$$

Let $A = \begin{bmatrix} 3 & 1 & 1 \\ 2 & 4 & 2 \\ -1 & -1 & 1 \end{bmatrix}$ be the coefficient matrix of the given system so that we can rewrite the system as the matrix equation $x' = Ax$. Additionally, if

$$Q = \begin{bmatrix} -1 & 0 & -1 \\ 0 & -1 & -2 \\ 1 & 1 & 1 \end{bmatrix} \quad \& \quad D = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix}$$

Then, $Q^{-1}AQ = D$.

Substituting $A = QDQ^{-1}$ into $x' = Ax$ to obtain $x' = QDQ^{-1}x$ or, equivalently, $Q^{-1}x' = DQ^{-1}x$. The function $y: \mathbb{R} \to \mathbb{R}^3$ defined by $y(t) = Q^{-1}x(t)$ is also differentiable & $y' = Q^{-1}x'$. Hence, the original system can be written as $y' = Dy$. Since D is a diagonal matrix, the system $y' = Dy$ is simple to solve. Setting $y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix}$ allows us to rewrite $y' = Dy$ as

$$\begin{bmatrix} y_1'(t) \\ y_2'(t) \\ y_3'(t) \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix} = \begin{bmatrix} 2y_1(t) \\ 2y_2(t) \\ 4y_3(t) \end{bmatrix}$$

The three equations $y_1'(t)$, $y_2'(t)$, $y_3'(t)$ are independent & thus, can be solved individually. The general solution to these equations is $y_1(t) = c_1e^{2t}$, $y_2(t) = c_2e^{2t}$ & $y_3(t) = c_3e^{4t}$ where $c_1, c_2, \& c_3$ are arbitrary constants. Finally

$$\begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \end{bmatrix} = x(t) = Qy(t) = \begin{bmatrix} -1 & 0 & -1 \\ 0 & -1 & -2 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} c_1e^{2t} \\ c_2e^{2t} \\ c_3e^{4t} \end{bmatrix} = \begin{bmatrix} -c_1e^{2t} - c_3e^{4t} \\ -c_2e^{2t} - c_3e^{4t} \\ c_1e^{2t} + c_2e^{2t} + c_3e^{4t} \end{bmatrix}$$

yields the general sol. to the original equation

Invariant Subspaces

Let T be a linear operator on a vector space V. A subspace W of V is called a T-**Invariant Subspace** of V if $T(W) \subseteq W$, that is, if $T(v) \in W$ for all $v \in W$

Suppose T is a linear operator on a vector space V. Then the following are examples of T-Invariant subspaces

$\{0\}$

V

$R(T)$

$N(T)$

E_λ, for any eigenvalue λ of T

Let T be a linear operator on a vector space V, & let x be a nonzero vector in V. The subspace $W = \text{span}(\{x, T(x), T^2(x), \cdots\})$ is called the T-**cyclic Subspace** of V, generated by x. W is a T-Invariant Subspace of V. Additionally, any T-invariant subspace that contains x must also contain W.

These cyclic subspaces have many applications. Right now our concern is applying them to establish the Cayley-Hamilton theorem.

An example of a cyclic subspace. If T is the linear operator defined by $(P_2(\mathbb{R})$ $T(f(x)) = f'(x)$. Then the T-cyclic subspace generated by x^2 is $\text{span}(\{x^2, 2x, 2\}) = P_2(\mathbb{R})$

Theorem 5.20: Let T be a linear operator on a finite-dimensional vector space V. & let W be a T-Invariant subspace of V. Then the charecteristic polynomial of T_W divides the charecteristic polynomial of T

Proof: Choose an ordered basis $\gamma = \{v_1, v_2, \dots, v_k\}$ for W, & extend it to an ordered basis $\beta = \{v_1, v_2, \dots, v_k, v_{k+1}, \dots, v_n\}$ for V. Let $A = [T]_\beta$ & $B_1 = [T_W]_\gamma$. Then, A can be written in the form $A = \begin{bmatrix} B_1 & B_2 \\ O & B_3 \end{bmatrix}$ Let $f(t)$ be the charecteristic polynomial of T & $g(t)$ be the chachteristic polynomial of T_W

Then,

$$f(t) = \det(A - tI_n) = \det\begin{bmatrix} B_1 - tI_k & B_2 \\ O & B_3 - tI_{n-k} \end{bmatrix} = g(t) \cdot \det(B_3 - tI_{n-k})$$

Thus, $g(t)$ divides $f(t)$

We may use the charecteristic polynomial of T_W to gain Information about the charecteristic polynomial of T itself. In this regard, cyclic subspaces are useful because the charecteristic polynomial of the restricted linear operator T to a cyclic-subspace is readily computable.

Theorem 5.21: Let T be a linear operator on a finite-dimensional vector space V & let W denote the T-cyclic subspace of V generated by a nonzero vector $v \in V$, & let $k = \dim(W)$. Then

a. $\{v, T(v), T^2(v), \dots, T^{k-1}(v)\}$ is a basis of W

b. If $a_0 v + a_1 T(v) + \dots + a_{k-1} T^{k-1}(v) + T^k(v) = 0$ then the charecteristic polynomial of T_W is given by $f(t) = (-1)^k (a_0 + a_1 t + \dots + a_{k-1} t^{k-1} + t^k)$

Proof:

a. Since $v \neq 0$, the set $\{v\}$ is linearly independent let j be the largest positive interger for which $\beta = \{v, T(v), \dots, T^{j-1}(v)\}$ is linearly independent. Such a j must exist because V is finite-dimensional Let $Z = \text{span}(\beta)$. Then β is a basis of Z. Furthermore, $T^j(v) \in Z$ by theorem 1.7. Since $w \in Z$ is a linear combination of the vectors in β, there exists scalars $b_0, b_1, \dots, b_{j-1}$ such that

$w = b_0 v + b_1 T(v) + \dots + b_{j-1} T^{j-1}(v)$ & hence,

$T(w) = b_0 T(v) + b_1 T^2(v) + \dots + b_{j-1} T^j(v)$ Thus, $T(w)$ is a linear combination of vectors in Z, & hence, belongs to Z. So Z is T-invariant. Furthermore, $v \in Z$. W is the smallest T-invariant subspace of V that contains v, so that $W \subseteq Z$. Clearly, $Z \subseteq W$, & so we can conclude that $Z = W$. It follows that β is a basis for W, & therefore, $\dim(W) = j$. Thus, $j = k$ which proves a.

b. Now, view β as an ordered basis of W. Let $a_0, a_1, \dots, a_{k-1}$ be scalars such that $a_0 v + a_1 T(v) + \dots + a_{k-1} T^{k-1}(v) + T^k(v) = 0$ observe that

$$[T_W]_\beta = \begin{bmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -a_{k-1} \end{bmatrix}$$

has the charecteristic polynomial given in b.

Theorem 5.22: The Cayley-Hamilton Theorem: Let T be a linear operator on a finite-dimensional vector space V, & let $f(t)$ be the charecteristic polynomial of T. Then, $f(t) = T_0$, the zero transformation. That is, T "satisfies" its charecteristic polynomial.

Proof: We show that $f(T)(v) = 0$ for all $v \in V$. This is obvious if $v = 0$ since $f(T)$ is linear. Suppose $v \neq 0$. Let W be the T-cyclic subspace generated by v, & suppose $\dim(W) = k$. By Theorem 5.21a there exist scalars, $a_0, a_1, \dots, a_{k-1}$ such that

$$a_0 v + a_1 T(v) + \cdots + a_{k-1} T^{k-1}(v) + T^k(v) = 0$$

Hence, Theorem 5.21b implies that

$$g(t) = (-1)^k (a_0 + a_1 t + \cdots + a_{k-1} t^{k-1} + t^k)$$

is the charecteristic polynomial of T_W. Combining these two equations yields the following

$$g(T)(v) = (-1)^k (a_0 I + a_1 T + \cdots + a_{k-1} T^{k-1} + T^k)(v) = 0$$

By Theorem 5.20, $g(t)$ divides $f(t)$; hence there exists a polynomial $q(t)$ such that $f(t) = q(t) g(t)$ So,

$$f(T)(v) = q(T) g(T)(v) = q(T)(g(T)(v)) = q(T)(0) = 0$$

Example: Let T be the linear operator on $\mathbb{R}^2$ defined by $T(a,b) = (a+2b, -2a+b)$ & let $\beta = \{e_1, e_2\}$. Then,

$A = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$ where $A = [T]_\beta$. The charecteristic polynomial of T is therefore given by

$$f(t) = \det(A - tI) = \det\begin{bmatrix} 1-t & 2 \\ -2 & 1-t \end{bmatrix} = t^2 - 2t + 5$$

It is easily verified that $T_0 = f(T) = T^2 - 2T + 5I$

$$f(A) = A^2 - 2A + 5I = \begin{bmatrix} -3 & 4 \\ -4 & -3 \end{bmatrix} + \begin{bmatrix} -2 & -4 \\ 4 & -2 \end{bmatrix} + \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

Direct Sums

Let T be a linear operator on a vector space V
There is a way of decomposing V into simpler subspaces that offer insight into the behavior of T.

Let $W_1, W_2, \dots, W_k$ be subspaces of a vector space V we define the **sum** of these subspaces to be the set
$\{v_1+v_2+\dots+v_k \mid v_i \in W_i \text{ for } 1 \le i \le k\}$

Example:

Let $V=\mathbb{R}^3$, let W_1 denote the xy-plane, & let W_2 denote the yz-plane. Then $\mathbb{R}^3 = W_1+W_2$ because for any vector $(a,b,c) \in \mathbb{R}^3$ we have $(a,b,c)=(a,0,0)+(0,b,c)$ where $(a,0,0) \in W_1$ & $(0,b,c) \in W_2$ $\longleftarrow$ Note that this representation is not unique; for example, $(a,b,c)=(a,b,0)+(0,0,c)$ is another representation.

Let $W, W_1, W_2, \dots, W_k$ be subspaces of a vector space V such that $W_i \in W$ for $i=1,2,\dots,k$. We call W the **direct sum** of the subspaces $W_1, W_2, \dots, W_k$ & we can write $V = W_1 \oplus W_2 \oplus \dots \oplus W_k$ if $V=\sum_{i=1}^{k} W_i$ & $W_j \cap \sum_{i \ne j} W_i = \{0\}$ for each $j (1 \le j \le k)$

Example:

In $\mathbb{R}^5$, let $W=\{(x_1,x_2,x_3,x_4,x_5) \mid x_5=0\}$, $W_1=\{(a,b,0,0,0) \mid a,b \in \mathbb{R}\}$
$W_2=\{0,0,c,0,0 \mid c \in \mathbb{R}\}$ & $W_3=\{(0,0,0,d,0) \mid d \in \mathbb{R}\}$
$(a,b,c,d,0)=(a,b,0,0,0)+(0,0,c,0,0)+(0,0,0,d,0)$
Thus, $W=\sum_{i=1}^{3} W_i$

Direct Sums

Theorem 5.9: Let $W_1, W_2, W_3, \dots, W_k$ be subspaces of a finite-dimensional vector space V. The following are equivalent

a. $V = W_1 \oplus W_2 \oplus \cdots \oplus W_k$

b. $V = \sum_{i=1}^{k} W_i$ & for any vectors $v_1, v_2, \dots, v_k$ such that $v_i \in W_i$

c. Each vector $v \in V$ can be uniquely written as $v = v_1 + v_2 + \cdots + v_k$ where $v_i \in W_i$

d. If γ_i is an ordered basis for W_i, then $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ is an ordered basis of V

e. For each $i = 1, 2, 3, \dots, k$ there exists an ordered basis γ_i for W_i such that $\gamma_1 \cup \gamma_2 \cup \cdots \cup \gamma_k$ is an ordered basis of V

Inner Products & Norms

For the majority of applications that follow we are primarily concerned with vector spaces are over the field of real numbers or the field of complex numbers

Let V be a vector space over a field F. Then, an **Inner Product** on V is a function that assigns to every ordered pair of vectors x & y in V a scalar in F, denoted $\langle x, y\rangle$, such that for all $x, y,$ & z in V & all c in F

$$\left.\begin{array}{l} a.\ \langle x+z, y\rangle = \langle x, y\rangle + \langle z, y\rangle \\ b.\ \langle cx, y\rangle = c\langle x, y\rangle \end{array}\right\} \text{These require the inner product to be linear in the first component}$$

c. $\overline{\langle x, y\rangle} = \langle y, x\rangle$, where the bar denotes complex conjugation

d. If $x \neq 0$, then $\langle x, x\rangle$ is a positive real number

Note: c. reduces to $\langle x, y\rangle = \langle y, x\rangle$ if $F = \mathbb{R}$

Additionally, if $a_1, a_2, \ldots, a_n \in F$ & $y, v_1, v_2, \ldots, v_n \in V$, then

$$\left\langle \sum_{i=1}^{n} a_i v_i, y \right\rangle = \sum_{i=1}^{n} a_i \langle v_i, y\rangle$$

Example#1

For $x = (a_1, a_2, \ldots, a_n)$ & $y = (b_1, b_2, \ldots, b_n)$ in F^n, define

$\langle x, y\rangle = \sum_{i=1}^{n} a_i \overline{b_i}$ The verification that $\langle \cdot, \cdot \rangle$ satisfies a-d is straight forward. Suppose that $z = (c_1, c_2, \ldots, c_n)$, we have for a.

$$\langle x+z, y\rangle = \sum_{i=1}^{n} (a_i + c_i)\overline{b_i} = \sum_{i=1}^{n} a_i \overline{b_i} + \sum_{i=1}^{n} c_i \overline{b_i} = \langle x, y\rangle + \langle z, y\rangle$$

Thus, for $x = (1+i, 4)$ & $y = (2-3i, 4+5i)$ in $\mathbb{C}^2$

$$\langle x, y\rangle = (1+i)(2+3i) + 4(4-5i) = 15-15i$$

↑ This inner product is called the **standard inner product** on F^n. When $F = \mathbb{R}$, the conjugations are not needed; in early courses this standard inner product is usually called the *dot Product* & is denoted $x \cdot y$ instead of $\langle x, y\rangle$

Example #2:

If $\langle x, y \rangle$ is any inner product on a vector space V & $r > 0$, we may define another inner product by the $\langle x, y \rangle' = r\langle x, y \rangle$. If $r \leq 0$, then $\langle x, x \rangle$ could be negative which violates d. on the previous page.

Example #3

Let $V = C([0,1])$, the vector space of real-valued continious functions on $[0,1]$. For $f, g \in V$, define $\langle f, g \rangle$ as $\langle f, g \rangle = \int_0^1 f(t)g(t)\,dt$ Since this integral is linear in f if $f \neq 0$, f^2 is bounded away from zero on some $[0,1]$, & hence $\langle f, f \rangle = \int_0^1 [f(t)]^2\,dt > 0$

Let $A \in M_{m\times n}(F)$. We define the **conjugate transpose** or **adjoint** of A to be the $n \times m$ matrix A^* such that $(A^*)_{ij} = \overline{A_{ij}}$ for all i, j. In the case that A has real entries A^* is simply the transpose of A.

Example #5

Let $V = M_{n\times n}(F)$ & define $\langle A, B \rangle = tr(B^*A)$ for $A, B \in V$. Suppose $A, B, C \in V$. Then $\langle A+B, C \rangle = tr(C^*(A+B))$

$= tr(C^*A + C^*B) = tr(C^*A) + tr(C^*B)$

Also... $= \langle A, C \rangle + \langle B, C \rangle$

$$\langle A, A \rangle = tr(A^*A) = \sum_{i=1}^{n} (A^*A)_{ii} = \sum_{i=1}^{n}\sum_{k=1}^{n} (A^*)_{ik} A_{ki}$$

$$= \sum_{i=1}^{n}\sum_{k=1}^{n} \overline{A_{ki}} A_{ki} = \sum_{i=1}^{n}\sum_{k=1}^{n} |A_{ki}|^2$$

Now if $A \neq 0$, then $A_{ki} \neq 0$ for some k & i

So $\langle A, A \rangle > 0$

This inner product on $M_{n\times n}(F)$ is called the **Frobenius inner product**

Inner Product Spaces

A vector space V over a field F endowed with a specific inner product is called an **inner product space**
If $F=\mathbb{C}$, we call V a **complex inner product space** whereas if $F=\mathbb{R}$ we call V a **real inner product space**

It is clear that if V has an inner product $\langle x,y\rangle$ & W is a subspace of V, then W is also an inner product space when the same function $\langle x,y\rangle$ is restricted to the vectors $x,y\in W$

Note that two distinct inner products on a given vector space yield two disinct inner product spaces

Theorem 6.1: Let V be an inner product space. Then for $x,y,z\in V$ & $c\in F$, the following statements are true

a. $\langle x, y+z\rangle = \langle x,y\rangle + \langle x,z\rangle$
b. $\langle x, cy\rangle = \bar{c}\langle x,y\rangle$
c. $\langle x,0\rangle = \langle 0,x\rangle = 0$
d. $\langle x,x\rangle = 0$ if & only if $x=0$
e. If $\langle x,y\rangle = \langle x,z\rangle$ for all $x\in V$, then $y=z$

Let V be an inner product space over F. For $x\in V$ we define the **norm** or **length** of x as $\|x\|=\sqrt{\langle x,x\rangle}$

Example: Let $V=F^n$. If $x=(a_1,a_2,\ldots,a_n)$ then

$$\|x\| = \|(a_1,a_2,\ldots,a_n)\| = \left[\sum_{i=1}^{n}|a_i|^2\right]^{1/2}$$ is the Euclidean definition of length. Note that if $n=1$, we have $\|a\|=|a|$
As expected the properties of Euclidean length in $\mathbb{R}^3$ hold in general.

Complex Numbers: Review

a **complex number** is an expression of the form $z = a + bi$, where a & b are real numbers called the **real part** & the **imaginary part**. We denote the **conjugate** of the complex number z as $\bar{z} = a - bi$

Theorem: Let z & w be complex numbers, then the following statements are true

a. $\bar{\bar{z}} = z$
b. $\overline{(z+w)} = \bar{z} + \bar{w}$
c. $\overline{zw} = \bar{z} \cdot \bar{w}$
d. $\overline{\left(\frac{z}{w}\right)} = \frac{\bar{z}}{\bar{w}}$ if $w \neq 0$
e. z is a real number if & only if $\bar{z} = z$

Let $z = a + bi$, where $a, b \in \mathbb{R}$. The **modulus** (or absolute value of z is denoted $|z| = \sqrt{a^2 + b^2}$

Theorem: Let z & w be complex numbers, then the following statements are true

a. $|zw| = |z| \cdot |w|$
b. $\left|\frac{z}{w}\right| = \frac{|z|}{|w|}$ if $w \neq 0$
c. $|z+w| \leq |z| + |w|$
d. $|z-w| \leq |z+w|$

The Fundemental Theorem of Algebra: Suppose that $p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$ is a polynomial in $P(\mathbb{C})$ of degree $n \geq 1$. Then $p(z)$ has a zero

Theorem 6.2: Let V be an inner product space over F. Then for all $x, y \in V$ & $c \in F$, the following statements are true:

a. $\|cx\| = |c| \cdot \|x\|$

b. $\|x\| = 0$ if & only if $x = 0$. In any case $\|x\| \geq 0$

c. $|\langle x, y\rangle| \leq \|x\| \cdot \|y\|$ (Cauchy-Schwars Inequality)

d. $\|x+y\| \leq \|x\| + \|y\|$ (Triangle Inequality)

Proof: c. If $y = 0$ then the result is immediate. So assume $y \neq 0$. For any $c \in F$, we have

$$0 \leq \langle x-cy, x-cy\rangle = \langle x, x-cy\rangle - c\langle y, x-cy\rangle$$
$$= \langle x,x\rangle - \bar{c}\langle x,y\rangle + c\bar{c}\langle y,y\rangle$$

In particular if we set $c = \dfrac{\langle x,y\rangle}{\langle y,y\rangle}$ then for each

$\bar{c}\langle x,y\rangle$, $c\langle x,y\rangle$, & $c\bar{c}\langle y,y\rangle$ equals $\dfrac{\langle x,y\rangle\langle y,x\rangle}{\langle y,y\rangle} = \dfrac{|\langle x,y\rangle|^2}{\|y\|^2}$ from which the Cauchy-Schwars Inequality follows.

d. We have $\|x+y\|^2 = \langle x+y, x+y\rangle = \langle x,x\rangle + \langle y,x\rangle + \langle x,y\rangle + \langle y,y\rangle$

$$= \|x\|^2 + 2\Re\langle x,y\rangle + \|y\|^2$$
$$\leq \|x\|^2 + 2|\langle x,y\rangle| + \|y\|^2$$
$$\leq \|x\|^2 + 2\|x\| \cdot \|y\| + \|y\|^2$$
$$= (\|x\| + \|y\|)^2$$

Wher $\Re\langle x,y\rangle$ denotes the real part of the complex number $\langle x,y\rangle$. Thus, the Triangle Inequality follows from the Cauchy-Schwars Inequality

For F^n $\left|\sum_{i=1}^{n} a_i \bar{b}_i\right| \leq \left[\sum_{i=1}^{n} |a_i|^2\right]^{1/2} \left[\sum_{i=1}^{n} |b_i|^2\right]^{1/2}$ &

$$\left[\sum_{i=1}^{n} |a_i + b_i|^2\right]^{1/2} \leq \left[\sum_{i=1}^{n} |a_i|^2\right]^{1/2} + \left[\sum_{i=1}^{n} |b_i|^2\right]^{1/2}$$

Let V be an inner product space. Vectors x & y in V are **Orthogonal** or **Perpendicular** if $\langle x, y \rangle = 0$. A subset S of V is **Orthogonal** if any two distinct vectors in S are orthogonal.

Additionally, a vector x in V

Theorem 6.4 (Gram-Schmidt)

Let V be an inner product space & $S = \{w_1, w_2, \dots, w_n\}$ be a linearly independent subset of V.
Define $S' = \{v_1, v_2, \dots, v_n\}$ where $v_1 = w_1$ &

$$v_k = w_k - \sum_{j=1}^{k-1} \frac{\langle w_k, v_j \rangle}{\|v_j\|^2} v_j \quad \text{for } 2 \leq k \leq n$$

Then S' is an orthogonal set of nonzero vectors such that $\operatorname{span}(S') = \operatorname{span}(S)$

$$v_2 = w_2 - \frac{\langle w_2, v_1 \rangle}{\|v_1\|^2} v_1 \quad \& \quad v_3 = w_3 - \frac{\langle w_3, v_1 \rangle}{\|v_1\|^2} v_1 - \frac{\langle w_3, v_2 \rangle}{\|v_2\|^2} v_2$$

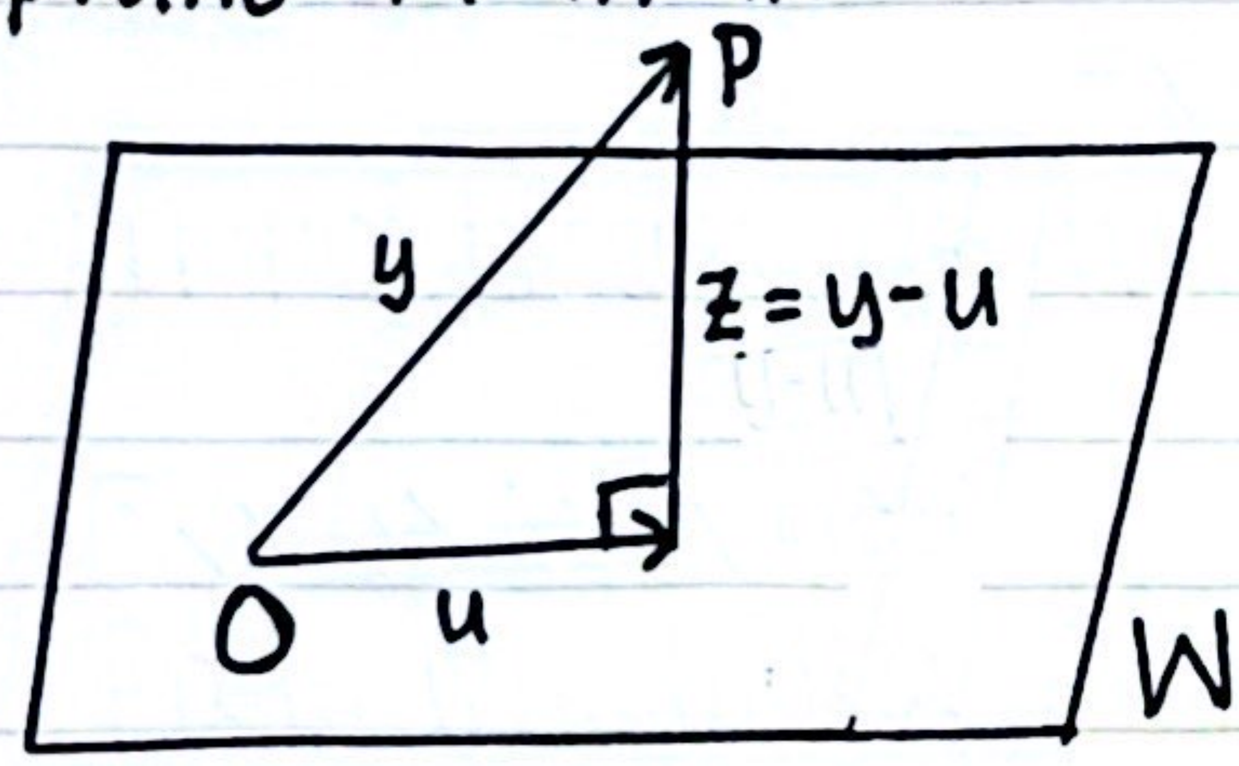
Gram-Schmidt Process:

Let S be a non-empty subset of an inner. We define $S^\perp$ to be the set of all vectors in V that are orthogonal to every vector in S; that is,

$$S^\perp = \{ x \in V \mid \langle x, y \rangle = 0 \text{ for all } y \in S \}$$

The subspace $S^\perp$ is called the **Orthogonal Complement** of S

Consider finding the distance from a point P to a plane W in $\mathbb{R}^3$



If we let y be the vector from 0 to P we can restate this problem as the following: Determine the vector $u \in W$ that is "closest" to y. Then, the desired distance is given by $\|y - u\|$

Note also, that $z = y - u$ is orthogonal to every vector in W. thus $z \in W^\perp$