MATH429 Linear Algebra

Albert Peng

January 18, 2023

Note: This set of notes does NOT contain everything being taught in class. In particular, I might not include the full details of materials previously covered in Matrix Algebra and might only include examples that help solidify more abstract concepts.

Contents

Proof Techniques

Induction

Usually, induction proof concerns proving properties about numbers.

Example. Prove that

$$\forall n \in \mathbb{N}, \sum_{k=1}^{n} k = \frac{n(n+1)}{2}, \sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$$

For $\sum_{k=1}^{n} k = n(n+1)/2$, denote P_n as the proposition n to prove.

First, We verify P_1, where $\sum_{k=1}^{1} k = 1$, implying that P_1 is true.

Then we verify $P_n \implies P_{n+1}$ (If P_n is true, then P_{n+1} is true). Suppose P_n holds. Then for P_{n+1},

$$\sum_{k=1}^{n+1} k = \frac{(n+1)(n+2)}{2} = \sum_{k=1}^{n} k \; + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2}$$

$\therefore P_{n+1}$ is true. Induction complete.

Proof by Contradiction/Contrapositive

To prove $P \implies Q$, we can either prove non$P \implies$ non$Q$ **or** suppose Q is false (is non$Q$) and find contradiction with P.

Example. Suppose that $n \in \mathbb{N}$ s.t. n^2 is even. Prove that n is even.

Proof: Suppose n is not even. we can prove that n^2 is not even.

Example. Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. Then, prove that if $f^2 = 1$, then either $f = 1$ or $f = -1$.

Proof: For the contradiction: suppose $\exists\, x_0 \in \mathbb{R}$ such that $f(x) \neq 1$ **and** $\exists\, x_1 \in \mathbb{R}$ such that $f(x) \neq -1$. By the Intermediate value theorem, (IVT) all values between 1 and -1 are taken by f due to its continuity. Thus, we can find $a \in [x_0, x_1]$ such that $f(a) = 0 \implies f(a^2) = 0$. There is a contradiction, so $f(x)^2 = 1$.

Double Inclusion and Distinction of Cases

Example. $A, B, C \subset E$. Prove $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

Proof: Let $x \in A \cap (B \cup C)$. This means that $x \in A$ and $x \in (B \cup C)$ (which is equivalent to $x \in B$ or $x \in C$). To prove $\subseteq$, We have two cases:

$$\begin{cases} x \in A \text{ and } x \in B \iff x \in A \cap B \subseteq (A \cap B) \cup (A \cap C), \quad x \in (A \cap B) \cup (A \cap C) \\ x \in A \text{ and } x \in C \iff x \in A \cap C \subseteq (A \cap B) \cup (A \cap C), \quad x \in (A \cap B) \cup (A \cap C) \end{cases}$$

Therefore, $x \in (A \cap B) \cup (A \cap C)$. Then $A \cap (B \cap C) \subseteq (A \cap B) \cup (A \cap C)$

To prove the other way ($\supseteq$) let $x \in (A \cap B) \cup (A \cap C)$, we have

$$\begin{cases} x \in A \cap B \iff x \in A \text{ and } x \in B \implies x \in A \text{ and } x \in B \cup C \quad x \in A \cap (B \cup C) \\ \text{Similarly, if } x \in A \cap C, \text{ then } x \in A \cap (B \cup C) \end{cases}$$

Exercise: Prove $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

0 Linear Systems and Matrices

Example. A linear system could be:

$$\begin{cases} x + 2y + 3z = 0 \\ 4x + 5y + 6z = 0 \\ 7x + 8y + 9z = 0 \end{cases}$$

We want to arrange the corresponding system into *upper triangular* form by eliminating coefficients:

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 6 \\ 0 & 6 & 12 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 3 \\ 0 & 3 & 6 \\ 0 & 0 & 0 \end{bmatrix}$$

We have 1 free variable, so we can set $z = t \in \mathbb{R}$. $y = -2t\,; x = 4t - 3t = t$. This can thus be written as

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix}, \; t \in \mathbb{R}$$

In this example, 1 free variable means that solutions can be "generated" by exactly 1 vector.

Definition. Rank is defined as the number of rows in **echelon row** after it is reduced.

This simple example also demonstrates the rank-nullity theorem, where (not fully stated yet in class)

$$\text{rank}A + dim(\text{null}A) = n$$

The rows in the echelon form could be interpreted geometrically where each row is the equation of a plane through $\vec{0}$. For a 3D space (3 by 3 matrix) of rank 2, there are hence 2 planes that intersect into a line.

If we consider the columns of the matrix which writes:

$$\begin{bmatrix} 1 \\ 4 \\ 7 \end{bmatrix} x + \begin{bmatrix} 2 \\ 5 \\ 8 \end{bmatrix} y + \begin{bmatrix} 3 \\ 6 \\ 9 \end{bmatrix} z = \vec{c}_1 x + \vec{c}_2 y + \vec{c}_3 z = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

we can conclude that

$$\text{The solution } \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \iff \vec{c}_1 - 2\vec{c}_2 + \vec{c}_3 = \vec{0}$$

The space generated by $\vec{c}_1, \vec{c}_2, \vec{c}_3$ is the **span** of the vectors.

0.1 Matrices

Definition. An $m \times n$ matrix is a $m \times n$ grid of numbers ($\mathbb{R}$ or $\mathbb{C}$)

$$M_{m,n}(\mathbb{R}) \text{ is the space of } m \times n \text{ matrices}$$

If $A \in M_{m,n}(\mathbb{R})$, denote

$$A = [a_{i,j}]_{1 \le i \le m, 1 \le j \le n} = \begin{bmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & \ddots & & a_{2,n} \\ \vdots & & \ddots & \vdots \\ a_{m,1} & \dots & \dots & a_{m,n} \end{bmatrix}$$

Definition. The **pivot** of a matrix is the first nonzero coefficient occurring in the row. In a **row echelon matrix (REM)**: $\forall i \in [1, m-1]$, the **pivot** in row $i+1$ occurs strictly after the pivot in row i. Everything below the pivots should have coefficient 0.

Definition. A **reduced row echelon matrix (RREM)** matrix keeps the properties of REM. In addition, its pivots are 1 and coefficients above pivots are 0.

We can associate linear systems in general with an augmented matrix:

$$\begin{cases} a_{1,1}x_1 + a_{1,2}x_2 + ... + a_{1,n}x_n = b_1 \\ \vdots \\ a_{m,1}x_1 + a_{m,2}x_2 + ... + 1_{m,n}x_n = b_n \end{cases} \rightarrow \left[A = (a_{i,j}) \mid \vec{b}\right], \text{ where } \vec{b} \in M_{m,1}(\mathbb{R}) \simeq \mathbb{R}^m$$

Definition. When $\vec{b} = 0$, we say that the system is **homogeneous**. The **kernel** of A, $ker(A)$, is the solution set of the linear system only when $\vec{b} = 0$

Definition. Two matrices $A, B \in M_{m,n}(\mathbb{R})$ are **row-equivalent** if we can get B from A via a sequence of elementary row operations.

Proposition. 2 linear systems with row equivalent matrices have the same solution set. In order words, elementary row operations preserves kernel.

Therefore, to solve a linear system, we try to boil down to an echelon form with row operations to solve.

Proposition. Let $A \in M_{m,n}(\mathbb{R})$.

1. A is row equivalent to a REM.
2. If A, B is row equivalent, then they have the same reduced row echelon form.

Proof: see other notes... I give up

Definition. The rank of A, $rank(A)$, where $A \in M_{m,n}(\mathbb{K})$ ($\mathbb{K}$ is either $\mathbb{R}$ or $\mathbb{C}$) is the number of nonzero rows in an echelon form of A.

Example.

$$rank \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = rank \begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} = 2$$

Remark in regards to column ranks. Everything we did so far could be done with columns, such as column operations, column echelon form (lower triangular), and column rank. However, the problem is that column operations *change the solutions of a linear system.*

Theorem. $rank =$ column $rank$
Proof: : later (Think about it as an exercise)

We can also define **column equivalence** if we can get from matrix B to A with a series of elementary column operations.

Definition. $A, B \in \mathbb{M}_{m,n}(\mathbb{K})$ are **equivalent** if we can go from one to the other via row and column operations.

Theorem. Suppose matrix $A \in M_{m,n}(\mathbb{K})$. Then A has rank $r \iff A$ is equivalent to a matrix $B \in M_{m,n}(\mathbb{K})$ of all zeroes, except the top left is an identity matrix of size I_r.

Proof: For $\implies$: First, we row reduce a matrix to rref(A). Then, column reduce rref (A) to reduced column echelon form and we get the result.
For $\impliedby$, (to be proved later)

0.2 Geometric Interpretations of Ranks

In rows, **rank** is the minimum number of equations to describe L. The **kernel** of A is the intersection of the planes associated with each equation.

In columns, suppose the $\vec{c}_1, \vec{c}_2, \vec{c}_3$ are the columns of A. $span(\vec{c}_1, \vec{c}_2, \vec{c}_3)$ is space generated by the vectors. In other words, this is the smallest space through $\vec{c}_1, \vec{c}_2, \vec{c}_3$. The column rank is thus the dimensions of the span of the vectors. The **rank nullity theorem** is where $n = rank(A) +$ free variables of L

Proof: Obvious from echelon form of A

Comment: $\dim ker(A)$ is equal to the number of free variables.

Example. For the linear system $x + y + z = 0$ (plane in $\mathbb{R}^3$), $A = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}$. If we set y, z as free variables such that $y = t, z = s$, then

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = t \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} + s \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \forall t, s \in \mathbb{R}$$

Thus in this case, number of free variables = 2 = dim plane = minimum number of vectors to generate a plane.

1 Matrices (Again?...)

1.1 Basic Operations

Definition. Matrix **addition** for $A, B \in \mathbb{M}_{m,n}(\mathbb{K})$ is defined as

$$A + B = [a_{i,j} + b_{i,j}]_{1 \le i \le m, 1 \le j \le n}$$

Definition. Matrix **dilation**, or **multiplication** by scalar, is defined for $\lambda \in \mathbb{K}$ and $A \in \mathbb{M}_{m,n}(\mathbb{K})$ where

$$\lambda A = \left[\lambda a_{i,j}\right]$$

Definition. Matrix **transpose** for $A \in \mathbb{M}_{m,n}(\mathbb{K})$ is the operation A^T or A^t, where

$$({}^t A)_{i,j} = A_{j,i}. \text{ So, } A^t \in \mathbb{M}_{m,n}(\mathbb{K})$$

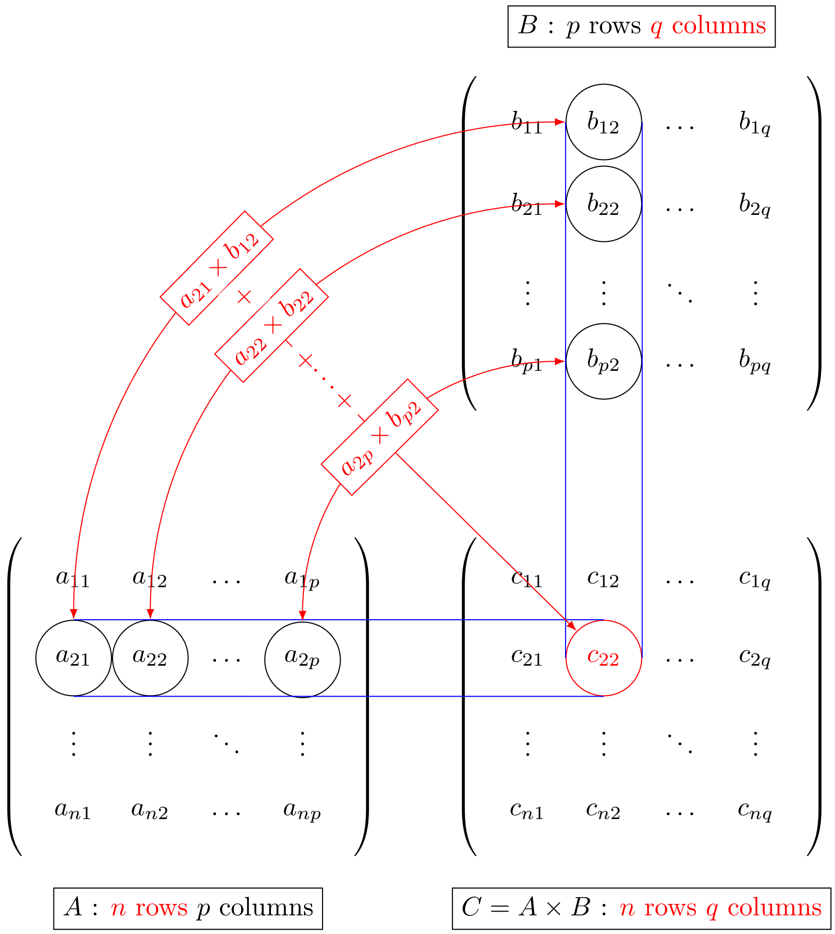
Definition. If $A \in \mathbb{M}_{m,n}(\mathbb{K})$, we call A **symmetric** if $A^T = A$, and A **skew symmetric** if $A^T = -A$.

Exercise: Prove that A skew symmetric $\implies$ Diagonal of A is made of 0.

1.2 Multiplication

Definition. Let $A \in \mathbb{M}_{m,n}(\mathbb{K}), B \in \mathbb{M}_{n,p}(\mathbb{K})$. AB is the matrix with

$$(AB)_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}; \qquad i \in [1, m], j \in [1, p]$$



Credit to https://texample.net/tikz/examples/matrix-multiplication/.

1.2.1 Properties

Matrix multiplication is **associative**, where $(AB)C = A(BC)$. Matrix multiplication also **distributes** over addition, so that $A(B+C) = AB + AC$ and $(B+C)A = BA + CA$.

Example. Let $A \in \mathbb{M}_{m,n}(\mathbb{R}), B \in \mathbb{M}_{n,p}(\mathbb{R}), C \in \mathbb{M}_{p,q}(\mathbb{R})$. Prove the associative property.

$$\begin{aligned} ((AB)C)_{i,j} &= \sum_{k=1}^{p} (AB)_{i,k} C_{k,j} \\ &= \sum_{k=1}^{p} \sum_{l=1}^{n} (A_{i,l} B_{l,j}) C_{k,j} = \sum_{k=1}^{p} \sum_{l=1}^{n} A_{i,j} (B_{l,k} C_{k,j}) \\ &= \sum_{l=1}^{n} A_{l,k} \left(\sum_{k=1}^{p} B_{l,k} C_{k,j} \right) = A(BC)_{i,j} \end{aligned}$$

Exercise: $A, B \in \mathbb{M}_n(\mathbb{R})$. Prove that $tr(AB) = tr(BA)$.

Proof:

$$tr(AB) = \sum_{i=1}^{n} (AB)_{i,i} = \sum_{i=1}^{n} \sum_{k=1}^{n} A_{i,k} B_{k,i} = \sum_{k=1}^{n} \sum_{i=1}^{n} A_{k,i} B_{i,k} = \sum_{k=1}^{n} (BA)_{k,k} = tr(BA)$$

Let I_n be an identity matrix of size $n \times n$. If $A \in M_{m,n}(\mathbb{K})$, then $I_m A = A I_n = A$.

Let $\lambda \in \mathbb{K}$, and A, B be multipliable matrices. Then

$$\lambda(AB) = (\lambda A)B = A(\lambda B) = (AB)\lambda \qquad \lambda A = (\lambda I_m)A$$

The puropse of this is that we can suppose L is a linear system with matrix $A = (a_{i,j}) \in M_{m,n}(\mathbb{K})$. Then

$$L \iff A\vec{x} = \vec{b}, \text{ where } \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{K}^n, \vec{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{K}^m$$

1.2.2 Elementary Matrices

Definition. Denote $(E_{i,j})_{1 \le i \le n, 1 \le j \le n} \in M_n(\mathbb{R})$ as the **canonical (standard) basis** of $M_{m,n}(\mathbb{K})$, such that the matrix has 1 at the (i,j) spot and 0 everywhere else.

Let $A \in_{m,n} (\mathbb{K})$. Then $E_{i,j}A$ is a matrix with $R_i = R_j(A)$with 0 everywhere else. The operation $AE_{i,j}$ is a $m \times n$ matrix with $C_j = C_i(A)$ and 0 everywhere else.

Proof: (First part) $(E_{i,j}A)_{p,q}$ where p$\in [1,m], q \in [1,n]$.

$$(E_{i,j}A)_{p,q} = \sum_{k=1}^{m} (E_{i,j})_{p,k} A_{k,q}, \text{ where for } (E_{i,j})_{p,k}, \begin{cases} 1, & p = i \text{ and } k = j \\ 0, & \text{otherwise} \end{cases}$$

$$\therefore A_{j,q} = \begin{cases} A_{j,q}, & \text{if } p = i \\ 0, & \text{otherwise} \end{cases}$$

Exercise: Prove $AE_{i,j}$ as matrix with $C_j = C_i(A)$

In other words, for matrix $A \in M_{m,n}(\mathbb{R})$,

$$E_{i,j}A = \begin{bmatrix} & \cdots & \\ & \vdots & \\ \cdots & R_j(A) & \cdots \\ 0 & \cdots & 0 \end{bmatrix}$$

Addendum to properties. In general, $AB \neq BA$.

Theorem. Elementary Matrices

Left Multiplication by Elementary Matrix	Row Operation
$I_n + \lambda E_{i,j}$	$R_i \leftarrow R_i + \lambda R_j$

$$\begin{bmatrix} 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & & \vdots \\ \vdots & & \ddots & \lambda & \ddots & & \vdots \\ \vdots & & & \ddots & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 1 \end{bmatrix}$$

Moreover, right multiplication by elementary matrix $\iff$ column operation

Proof: For $I_n + \lambda E_{i,j} \iff$ row operation $R_i \leftarrow R_i + \lambda R_j$. Let $A \in M_{n,m}(\mathbb{R})$:

$$(I_n + \lambda E_{i,j})A = A + \lambda E_{i,j}(A) = ...$$

1.2.3 Block Matrix Multiplication

Setting $M \in \mathbb{M}_{m,n}(\mathbb{R}), M' \in \mathbb{M}_{n,p}(\mathbb{R})$, we can decompose it so that

$$M = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \qquad M' = \left[\begin{array}{c|c} A' & B' \\ \hline C' & D' \end{array}\right]$$

$$A = \left[\begin{array}{c|c} A & B \\ \hline C & C \end{array}\right]$$

test

$$\left[\begin{array}{c|c} \begin{matrix} & \\ & A & \\ & \end{matrix} & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline 0 \cdots\cdots\cdots 0 & 0 \end{array}\right]$$

Note: Summations can commonly be transformed with

$$\sum_{k=r+1}^{n} a_k = \sum_{k=1}^{n-r} a_{k+r}$$

We can establish a relationship between $rref(A)$ and the original matrix A, with

$$rref(A) = E_k \dots E_2 E_1 A$$

In particular, A and B are row equivalent iff $\exists E_1, \dots, E_k$ elementary matrices $\ni A = E_k...E_1 B$. Similarly, A and B are column equivalent iff $\exists E'_1, ..., E'_p$ elementary matrices $\ni A = BE'_1...E'_p$.

1.3 Inverse of Matrices (Square Only)

Definition. $A \in M_n(\mathbb{K})$. A is **left invertible** if $\exists\, B \in M_n(\mathbb{R})$ such that $BA = I_n$. A is **right-invertible** if $\exists\, C \in M_n(\mathbb{K})$ such that $AC = I_n$. A is **invertible** if it is right and left invertible and $B = C$. In this case, denote $B = C = A^{-1}$.

We need this idea of left and right invertibility to solve for inverses of linear systems.

Proposition. If $A \in M_n(\mathbb{R})$ is both left and right invertible, with B being right inverse and C being left inverse, then $B = C$.

Proof: Given $BA = I_n$ and $AC = I_n$, then $B = BI_n = B(AC) = (BA)C = I_nC = C$

Theorem. If we do not assume left and right invertibility, then A left invertible $\iff$ A right invertible.

Proof: Done at end of section

Denote $GL_n(\mathbb{K})$ as the set of $n \times n$ invertible matrices. Then, $A, B \in GL_n(\mathbb{K}) \implies AB \in GL_n(\mathbb{K})$. Morevoer, $(AB)^{-1} = B^{-1}A^{-1}$.

<u>Uniqueness of Inverse:</u> If $A \in GL_n(\mathbb{K})$ and has 2 inverses B, C, then $B = C$

Looking at the inverse of elementary matrices, we have

1. $$(I_n + \lambda E_{i,j})^{-1} = I_n - \lambda E_{i,j}, \quad \exists i \neq j, \lambda \in \mathbb{K}$$

2. $$\begin{bmatrix} 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & & \vdots \\ \vdots & & \ddots & \lambda & \ddots & & \vdots \\ \vdots & & & \ddots & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \ddots & \ddots & & & & \vdots \\ \vdots & \ddots & 1 & \ddots & & & \vdots \\ \vdots & & \ddots & \frac{1}{\lambda} & \ddots & & \vdots \\ \vdots & & & \ddots & 1 & \ddots & \vdots \\ \vdots & & & & \ddots & \ddots & 0 \\ 0 & \dots & \dots & \dots & \dots & 0 & 1 \end{bmatrix}$$

3. $$E_{i\Leftrightarrow j}^{-1} = E_{i\Leftrightarrow j} \qquad \text{involution; inverse = itself}$$

To prove (1), we check that operation

$$(I_n - \lambda E_{i,j})(I_n + \lambda E_{i,j}) = ...$$

Lemma. (Avatar of the dimennon theorem). Denote $J_r = J_{m,n}^r =$

$$J_r = \left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right]$$

Then J_r is equivalent to $J_{r'}$ iff $r = r'$.

Proof: $\Longleftarrow$: $r = r' \implies J_r = J_{r'}$

$\implies$ Suppose J_r equivalent to $J_{r'}$.

Proof. Without loss of generality, we can assume $r \leq r'$.

$$J_r \sim J_{r'} \iff PJ_rQ = J_{r'}$$

□

Theorem. For $A \in M_{m,n}(\mathbb{K})$

1. $rank(A) = r \iff A$ is equivalent to J_r
2. column $rank(A) = rank(A)$
3. $rank(A) = rank(B) \iff A$ equivalent to B (matrix operations preserve rank)

Proof: For (1) $\Longleftarrow$, If $A \sim J_r$, then $A = PJ_rQ$. Moreover, if $r' = rank(A)$, we can row reduce + column reduce where $A = P'J_{r'}Q' \implies J_r \sim J_{r'} \implies r = r' = rank(A)$

For (2), let $r' = col\, rk(A), r = rk(A)$. $A = P'J_{r'}Q', A = PJ_rQ \implies J_{r'} \sim J_r \implies r = r'$

(3) is left as exercise

1.4 Invertibility vs Rank

Theorem. Let $A \in M_n(\mathbb{K})$. Then

$$A \text{ invertible } \iff A\vec{x} = \vec{0} \text{ has only solution } \vec{x} = \vec{0} \iff rk(A) = n$$

Proof: $1 \to 2$: If A invertible, then $A\vec{x} = \vec{0} \implies \vec{x} = A^{-1}\vec{0} = \vec{0}$.

$2 \to 3$: If only solution to $A\vec{x} = 0$ is $\vec{0}$, then it means that rref(A) has all diagonal entries as pivots $\implies rank(A) = n$

$3 \to 1$: $rank(A) = n \implies$ rref(A) . We have rref(A)$= I_n = PA = AQ \implies P = Q = A^{-1}$

1.4.1 Finding inverse of matrix

How to find A^{-1} if $A \in GL_n(\mathbb{K})$?

$$\text{Convert } \left[\begin{array}{c|c} A & I_n \end{array}\right] \to \left[\begin{array}{c|c} rref(A) = I_n & A^{-1} \end{array}\right]$$

This can be explained as

$$E_k...E_2E_1A = rref(A) = I_n \implies BA = I_n \implies .....$$

Corollary. Invertible matrices are products of elementary matrices.

Final notes on equivalence of matrices: We saw that all invertible matrices are products of elementary matrices. Moreover, elementary matrices are invertible. Thus, equivalent matrices can be defined for $A, B \in M_{m,n}$as

$$A \sim B \text{ iff } \exists\, P \in GL_m(\mathbb{K}),\ \exists Q \in GL_n(\mathbb{K}) \text{ s.t. } B = PAQ$$

2 Vector Spaces (Finally)...

2.1 General Vector Space Theory

2.1.1 Motivation

In $\mathbb{R}^n$, **linear combinations** of $\vec{v}_1, \vec{v}_2, ..., \vec{v}_k$ is a vector of the form

$$\vec{v} = \lambda_1\vec{v}_1 + \lambda_2\vec{v}_2 + ... + \lambda_k\vec{v}, \qquad \lambda_i \in \mathbb{R}$$

Why are they special?

1. Calculating rank involves linear combinations of rows or columns.
2. Solutions of linear systems of the form $A\vec{x} = \vec{0} \,\exists\, \vec{x} \in \mathbb{R}^n$ is given by linear combinations of a given set of vectors, where the number of vectors is equal to $n - rk(A)$

With vector spaces, we try to reconstruct everything using the idea of linear combinations; vector space is essentially the space where linear combinations make sense. It begs the question of *what is needed to form linear combinations?*

- Addition
- Dilations

"But wait... what is an operation?"

Definition. An **internal operation**/composition law on a set E is a function $*$; $E \times E \to E$, where $E \times E = \{(x, y) \,|\, x \in E, y \in E\}$ and $(x, y) \longmapsto x * y \in E$ where $*$ is the operation

Definition. An **external operation** between 2 sets $\mathbb{K}, E$ where $\cdot\,;\ \mathbb{K} \times E \to E$, $(\lambda, x) \longmapsto \lambda \cdot x$

Example. Suppose we have $E = \mathbb{R}^2, \mathbb{K} = \mathbb{R}$. The dilation operation can be expressed as

$$\cdot\,;\ \mathbb{K} \times E \to E \qquad (\lambda, \vec{x}) \longmapsto \lambda \cdot \vec{x} = (\lambda x_1, \lambda x_2)$$

Definition. $(E, +, \cdot)$ is a **vector space** over $\mathbb{K}$ ($\mathbb{K}$-v. s.) if $+$ is an *internal operation* and $\cdot$ is an external operation $\mathbb{K} \times E \to E$ that satisfies the following axioms:

1. There is a "0 element" 0_E where $u + 0_E = 0_e + u = u \,\forall\, u \in E$
2. $+$ is associative, $\forall\, u, v, w, \in E$, $(u + v) + w = u + (v + w)$
3. Every element $u \in E$ has an inverse/opposite $(-u) \in E \ni u + (-u) = (-u) + u = 0$
4. $+$ is commutative, $\forall u, v, \in E$, $u + v = v + u \iff (E, +)$ is an *abelian* group
5. $\cdot$ distributes over $+$, where $\lambda \cdot (u + v) = \lambda \cdot u + \lambda \cdot v$
6. $+$ distributes over $\cdot$, where $(\lambda + \mu) \cdot u = \lambda \cdot u + \mu \cdot u$
7. $\lambda \cdot (\mu \cdot \vec{u}) = (\lambda \cdot \mu)\vec{u}$
8. $I_{\mathbb{K}} \cdot u = u$, $\forall\, \lambda, \mu \in \mathbb{K}$; $u, v \in K$

Example. Suppose function $\mathcal{F} : \mathbb{R} \to \mathbb{R}$. Then addition is if $f, g \in \mathcal{F}(\mathbb{R}, \mathbb{R})$:

$$f + g \text{ is the function defined as } (f + g)(x) = f(x) + g(x), \forall\, x \in \mathbb{R}$$

Dilation is where if $\lambda \in \mathbb{R}, f \in \mathcal{F}(\mathbb{R}, \mathbb{R})$,

$$(\lambda \cdot f)(x) = \lambda f(x), \forall\, x \in \mathbb{R}$$

Therefore $(\mathcal{F}(\mathbb{R}, \mathbb{R}), +, \cdot)$ is a vector space.

Example. $(\mathbb{R}, *, \cdot)$ where $x * y = x + y + 1$ is not a vector space.

1. Doesn't distribute over $*$

2.2 Subspaces

Fix $(E, +, \cdot)$ = vector space over $\mathbb{K}$. Then,

Definition. A subspace $\mathcal{F} \subseteq E$ is a subset satisfying 2 conditions.

Vector space properties. For $\lambda \in \mathbb{K}, u \in E$:

1. $\lambda \cdot u = 0_E \iff \lambda = 0$ or $u = 0$
2. $-(\lambda \cdot u) = (-\lambda) \cdot u$

Proof. For $\implies$ (1):

$$\forall v \in E, v + 0_E = v.u + v = \frac{1}{\lambda}(\lambda v + \lambda u) = \frac{1}{\lambda}(\lambda v) = v + 0_E$$

For $\impliedby$, suppose $\lambda = 0$ or $u = 0_E$. If $\lambda = 0$, we want $\lambda \cdot u = 0$.

$$\lambda \cdot u = 0 \cdot u = (0 + 0) \cdot u = 0 \cdot u + 0 \cdot u \implies 0 \cdot u = 0 \cdot u + 0 \cdot u \implies 0 = 0 \cdot u$$

For (2): (exercise) prove that

$$\lambda \cdot u + (-\lambda \cdot u) = 0 \text{ and } \lambda \cdot u + (-\lambda) \cdot u = 0$$

$$\lambda \cdot u + (-\lambda) \cdot u = (\lambda + (-\lambda)) \cdot u = 0$$

□

Definition. A subspace $F \in E$ is a subset satisfying 2 conditions:

1. $\forall\, u, v \in F$, $u + v \in F$.
2. $\forall \lambda \in \mathbb{K}, \forall\, u \in F, \lambda \cdot u \in F$

Proposition. $\forall\, A \in M_{m,n}(\mathbb{R})$, $ker(A)$ is a subspace of $\mathbb{R}^n$

Proof. Let $\vec{x}, \vec{y} \in ker(A) \subseteq \mathbb{R}^n$, let $\lambda \in \mathbb{R} \iff A\vec{x} = A\vec{y} = 0$ Then, $\lambda\vec{x} + \vec{y} \in ker(A) = \lambda(A\vec{x}) + A\vec{y} = \vec{0}$ □

Example. $F = \{M \in \mathbb{M}_n(\mathbb{R}), tr(M) = 0\}$ is a subspace.

Proof. $\text{tr}(\lambda A + B) = \sum_{i=1}^{n}(\lambda a_{ii} + b_{ii}) = 0$ □

Example. $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$ be function $\mathbb{R} \to \mathbb{R}$. $F = \{f \in E, f(0) = 0\}$ is a subspace but $G = \{g \in E, f(0) = 0\}$ is not. Important subspaces of E:

- $F = C([a, b], \mathbb{R})$ are continuous functions that are real valued.
- $F = \mathbb{R}_n[x]$ are polynomials with degrees $\leq n$

Example. E is vector space over $\mathbb{K}$. Fix $e_1, ..., e_k \in E$. For $F = span(e_1, ..., e_k)$, it is the smallest subspace containing $E_1, ..., e_k$ Why is it the intersection of all subspaces of E containing $e_1, ..., e_k$?

2.3 Linear Independence

Let E be a $\mathbb{K}$ vector space. If we fix $e_1, ..., e_k \in E$. Then the family of vectors in E is said to be **linearly dependent** if $\exists\, \lambda_1, ..., \lambda_k \in \mathbb{K}$ not all zero such that $\lambda_1 e_1 + ... + \lambda_k e_k = 0_E$.

Thus, they are **linearly independent** if

$$\lambda_e e_1 + ... + \lambda_e e_k = 0 \implies \lambda_1 = \lambda_2 = ... = 0$$

Example. Pick $A \in M_{n,k}(\mathbb{R})$, set $\vec{\Lambda} = \{\lambda_1, ..., \lambda_k\}$

$$A\vec{\Lambda} = 0 \iff \lambda_1 C_1(A) + \lambda_2 C_2(A) + ... + \lambda_k C_k(A) = 0$$

Therefore, $A\Lambda = 0$ tells if columns are linearly independent or not.

Example. $\mathcal{S} = (p_0, ..._n)$ within E with $p_k(x) = x^k$. We can pick $\lambda_0, ..., \lambda_n \in \mathbb{R}$ such that $\lambda_0 p_0 + ... + \lambda_n p_n = 0$. If $\lambda_0 = 0$, I get $\lambda_1 x + ... + \lambda_n x^n = 0$. Then we differentiate in x so that $\forall\, x, \lambda_1 + 2\lambda_2 x + 3\lambda_3 x^2 + ... n\lambda_n x^{n-1} = 0$, so $\lambda_1 = 0$. Iterate this reasoning to show that $\forall\, i, \lambda_i = 0$. So $\mathcal{S}$ is linearly independent.

Exercise: Check that any subset of a linearly independent set is linearly independent.

Proof. The original independent set $\mathcal{S} = (e_1, ..., e_k)$. Let the subset $\mathcal{S}_0 = (e_1, ..., e_p)$ with $p \leq k$.

Let $\lambda_1, ..., \lambda_p \in \mathbb{K}$. such that $\lambda_1 e_1 + ... + \lambda_p e_p = 0 \iff$ set $\lambda_{p+1} = ... = \lambda_k = 0$, $\lambda_1 e_1 + ... + \lambda_p e_p + \lambda_{p+1} e_{p+1} + ... + \lambda_k e_k = 0$.

Then, $(e_1, ..., e_k)$ linearly independent $\implies \lambda_1 = ... = \lambda_p = \lambda_{p+1} = ... = \lambda_k = 0$, so $\mathcal{S}_0$ is linearly independent. □

Corollary. By contrapositive, a set of vectors containing a linearly dependent set is also linearly dependent.

2.4 Spanning Sets

Definition. We have $E = \mathbb{K}$ vector space. A set of vectors $\mathcal{S} = (e_1, ..., e_n)$ is spanning for E if $\forall\, x \in E, \exists\, x_i \in \mathbb{K}$ such that $x = \sum_{i=1}^n x_i e_i$. In other words, every vector $x \in E$ can be written as linear combination of $e_1, ..., e_n$, or

$$span(e_1, ..., e_n) = E$$

Example. $E = \{(x, y, z) \in \mathbb{R}^3, x + y + z = 0\}$ The spanning set can be found by solving the equation, where

$$x + y + z = 0 \iff \begin{bmatrix} x \\ y \\ z \end{bmatrix} = s \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} + t \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} \implies \text{ the two vectors span } E$$

Example. Find a spanning set of $\mathcal{E} = \{A \in M_2(\mathbb{R}), tr(A) = 0\}$.

$$A = \begin{bmatrix} a & b \\ c & -a \end{bmatrix} \in \mathcal{E} \implies a \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} + b \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} + c \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \implies A \in span(E_{1,1} - E_{2,2}, E_{1,2}, E_{2,1})$$

Example. $E = \mathbb{R}[x]$, $\mathcal{S} = (1, x, x^2, ..., x^n, ...)$ is a spanning set for E of finite linear combinations.

2.5 Basis

Definition. Let E be a $\mathbb{K}$ vector space. A basis $\mathcal{B}$ of $E \iff \mathcal{B}$ is linearly independent and spanning for E.

Example. $E = \mathbb{R}^n$. $(e_i)_{i\in[1,n]}$ are the canonical or standard basis.

Example. Suppose

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}$$

Find the basis for $ker(A)$ and $col(A) = span(C_1(A), C_2(A), C_3(A), C_4(A))$.

Basis of

$$ker(A), x = \begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} \in ker(A) \iff Ax = 0$$

Row reducing,

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} \implies \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Thus, after simplifying,

$$\begin{bmatrix} x \\ y \\ z \\ t \end{bmatrix} = a \begin{bmatrix} 1 \\ -2 \\ 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 2 \\ -3 \\ 0 \\ 1 \end{bmatrix}$$

Set e_1, e_2 as the two previous vectors, then $ker(A) = span(e_1, e_2)$. Since (e_1, e_2) are linearly independent, $(e_1, e_2) =$ basis of $ker(A)$.

For the basis of $span(C_1(A), ..., C_4(A))$, we know that $ker(A)$ tells which vectors can be thrown away without altering the space. Thus,

$$e_1 \in ker(A) \implies C_1(A) - 2C_2(A) + C_3(A) = 0, \qquad e_2 \in ker(A) \implies 2C_1(A) - 3C_2(A) + C_4(A) = 0$$

I can conclude that C_3, C_4 are linear combinations of C_1, C_2 so $span(C_1, ...C_4) = span(C_1, C_2)$. Then, we can also check by hand that C_1, C_2 are linearly independent. Thus, $(C_1, C_2) =$ basis of column space of A.

Remark: This basically explains the rank nullity theorem.

Theorem. Let $E = \mathbb{K}$ vector space, $\mathcal{B} = (e_1, ..., e_n)$ a finite basis of E. Then

$$\forall\, x \in E, \exists\, !x_1 \in \mathbb{K} \text{ such that } x = \sum_{i=1}^{n} x_i e_i$$

Proof. Let $x \in E$. There is existence because $\mathcal{B}$ spans $E \implies \forall\, i \in \{1, ..., n\}$, $\exists x_i \in \mathbb{K}$ such that $x = \sum_i x_i e_i$.

For uniqueness, suppose we can also write $x = \sum_i y_i e_i$, then

$$x = \begin{cases} \sum_i x_i e_i \\ \sum_i y_i e_i \end{cases} \implies \sum_i x_i e_i - \sum_i y_i e_i = 0 \implies \sum_i (x_i - y_i) e_i = 0.$$

$$\mathcal{B} \text{ linearly independent} \implies (x_i - y_i = 0\, \forall\, i \in [1, n]) \implies x_i = y_i$$

□

2.6 Dimensions

Definition. Vector space E is finite dimensional if it has a finite generating/spanning set.

Theorem. Existence of a Basis. Let $g = (e_1, ..., e_n)$ be a generating set. Let $I = (e_1, ..., e_k)$ be a linearly independent set within g. Then, one can turn I into a basis, by adding up elements of g (not in I).

Proof. Let $N = \{\#J, J \text{ linearly independent set containing } I, \text{ within } g\} \subseteq \mathbb{N}$.

- $N \neq \emptyset, I$ satisfies the conditions and $\#J = k \implies k \in N$.
- N bounded below by k, above by n. This is proven by least upper bound property in $\mathbb{N}$, where every bounded, non $\emptyset$ subset of $\mathbb{N}$ has a maximum. We can denote the maximum of N as $p = \max(N)$.

We can conclude that $(e_1, ..., e_p)$ is a linearly independent set within g.

We claim that $\mathcal{B}$ is a basis of E, which is proved by showing $\forall\, i \in [1, n], e_i \in span(e_1, ..., e_p)$. This is obvious if $i \in [1, p]$. If not, then $(e_1, ..., e_p, e_{p+1})$ is linearly dependent; otherwise, $(e_1, ..., e_p, e_{p+1})$ linearly independent, $\subseteq g$ with $p + 1 > p = \max(N)$ elements, which is impossible.

So, $(e_1, ..., e_p, e_{p+1})$ is linearly dependent. Thus $\exists\, \lambda_1, ..., \lambda_p, \lambda_{p+1} \in \mathbb{K}$ not all 0, such that $\lambda_1 e_1 + ... + \lambda_p e_p + \lambda_{p+1} e_{p+1} = 0$. $\lambda_{p+1} \neq 0$ because $(e_1, ..., e_p, e_{p+1})$ linearly dependent.

This implies that $e_{p+1} \in span(e_1, ..., e_p) = span(\mathcal{B})$, this holds for $e_j \,\exists\, j \geq p + 1$. $\implies g \subseteq span(\mathcal{B}) \implies span(g) \subseteq span(\mathcal{B})$ since g generates $E \implies E = span(\mathcal{B})$ and $\mathcal{B}$ generates E.

□

Corollary. From any finite generative set, one can extract a basis of E.

Proof. Apply theorem with $I = \emptyset$. □

Remark: Vector space $\{0\}$ has $\emptyset$ as basis.

2.6.1 Number of elements in a Basis

Theorem. Steinitz's Exchange Lemma Let $I = (e_1, ..., e_p)$ be any linearly independent set and $g = (f_1, ..., f_q)$ be any generating set, then $p \leq q$ and up to renumbering, $(e_1, ..., e_p, f_{p+1}, ..., f_q)$ generates E. The notation of 'up to renumbering' does not necessarily corresponds to the number in g.

Proof. By induction, we can start with $p = 0$ with $I = \emptyset$, so there are nothing to do as f is already generating.

With the induction step, suppose this is true with linearly independent sets with cardinal (number of elements) $p-1$, or $I = (e_1, ..., e_p)$ independent $\implies (e_1, ..., e_{p-1})$ independent. We can apply induction hypothesis to have $p - 1 \leq q$, and can construct generating set of the form $(e_1, ..., e_{p-1}, f_p, f_{p+1}, ..., f_q)$ up to renumbering.

If $p-1 = q \implies (e_1, ..., e_{p-1})$ is generating $\implies e_p$ is linearly combination of $(e_1, ..., e_{p-1})$, impossible because I linear independent. So $p - 1 < q \implies p \leq q$.

We want to find f_{i_o} to exchange with e_p, so $(e_1, ..., e_{p-1}, e_p, f_{p+1}, ..., \hat{f_{i_o}}, ..., f_q)$ is generating where $\hat{f_{i+o}}$ symbols the element to omit.

It is given that g_{p-1} generating implies $\exists\, \lambda_1, ..., \lambda_{p-1}, \mu_p, ..., \mu_q \in \mathbb{K}$ such that

$$e_p = \sum_{i=1}^{p-1} \lambda_i e_i + \sum_{j=p}^{q} \mu_j f_j$$

The idea is the swap with f_{j_o} such that $\mu_{j_o} \neq 0$. This is possible since if $\mu_j = 0\, \forall\, j \in [p, q]$, $\implies e_p = \sum_{i=1}^{p-1} \lambda_i e_i$, which is impossible since I is linearly independent. So, $\exists\, j_o \in [p, q]$ such that $\mu_{j_o} \neq 0$. Therefore,

$$f_{j_o} = \frac{1}{\mu_{j_o}} \left(e_p - \sum_{i=1}^{p-1} \lambda_i e_i - \sum_{j=p, j \neq j_o}^{q} \mu_j f_j \right)$$

Then, $g_p(e_1, ..., e_{p-1}, e_p, f_p, ..., f_{j_o}, \hat{...}, f_q)$ is generating. f_{j_o} is linear combination of them. g_{p-1} is generating, and $f_{j_o} \in g_{p-1}$. Therefore, all members of g_{p-1} are linear combinations of elements of $g_p + g_{p-1}$ generating $\implies g_p$ generating. □

Corollary.

1. All basis of E finite dimensional has same *cardinal* (number of elements).
2. Define $\dim_{\mathbb{K}} E$ = dimensions of E over $\mathbb{K}$ = number of elements of any basis.
3. $\dim_{\mathbb{K}} E$ is the maximum number of linearly independent elements in E and the minimum number of elements in a generating set.
4. If $n = \dim E$, then any set with $n + 1$ vectors is linearly dependent.
5. $\mathcal{B}$ basis $\iff \mathcal{B}$ linearly independent with number of $\mathcal{B} = n \iff .\mathcal{B}$ generating set with number of $\mathcal{B} = n$

Example. $\mathbb{C}$ = complex numbers, vector space of $\mathbb{C}$ and over $\mathbb{R}$.

$$\dim_{\mathbb{C}} \mathbb{C} = 1, \text{ but } \dim_{\mathbb{R}} \mathbb{C} = 2$$

Example. $\dim_{\mathbb{K}} \mathbb{K}^n = n$; $\dim M_{m,n}(\mathbb{R}) = nm$; $\dim \mathbb{R}_n[X] = n + 1$

Example. $\dim \mathbb{R}[x] = \infty$. More generally, E is infinite dimensional if we can find sequence of vectors $(x_i)_{i \in \mathbb{N}}$, such that $\forall n \in \mathbb{N}$, $(e_0, ..., e_n)$ linearly independent.

Proposition. Let F be subspace of E, $\dim E < \infty$. Then F is finite dimensional and $F = E$ iff $\dim(F) = \dim(E)$.

Proof. Consider I_n be the set of linearly independent sets within F. Pick a maximum linearly independent set in F. By least upper bound proposition, the number of elements of this is $\leq n = \dim E$. Moreover, it's a basis of F.

The " $\implies$ " direction is obvious. For "$\impliedby$", let $\dim(F) = p, \dim(E) = n$. Pick a basis of F, $(e_1, ..., e_p)$, complete into a basis of E. Since $p = n, (e_1, ..., e_n)$, it is already basis of E. □

2.6.2 Construction of Spaces

Cartesian product of 2 spaces E and F: $E \times F$:

$$E \times F = \{(x, y);\, x \in E, y \in F\}$$

Addition: $(x, y) + (x' + y') = (x + x', y + y')$ Dilation: $\lambda \cdot (x, y) = (\lambda \cdot x, \lambda \cdot y)$

Proposition. If $\mathcal{B}_E(e_1, ..., e_p)$ and $\mathcal{B}_F(f_1, ..., f_q)$ are basis of E and F respectively. Then, $((e_1, 0), ..., (e_p, 0), (0, f_1), ..., (0, f_q))$ is a basis of $E \times F$. This also implies that $\dim E \times F = \dim E + \dim F$

V is vector space. Let E, F be 2 subspaces f V. The *sum* of E and F is

$$E + F = \{x + y; x \in E, y \in F\}$$

For *direct sums*, we say that V is a direct sum of E and F, denoted as $E \oplus F$ if every vector $v \in V$ admits a unique decomposition. where $v = x + y; x \in E, y \in F$.

Proposition.

$$V = E \oplus F. \iff \begin{cases} \text{Every } v \in V \text{ writes } v = x + y, x \in E, y \in F \\ E \cap F = \{0\} \end{cases}$$

Proof. If $E \cap F = \{0\}$, and $v = x + y = x' + y'$. When their intersection is 0, $x - x' = y - y'$ and each element is in E and F respectively. Then, since

$$\begin{cases} x - x' \in E \cap F = \{0\} \\ y - y' \in E \cap F = \{0\} \end{cases} \implies x - x' = 0, y - y' = 0 \implies x = x', y = y'$$

□

Thus if $v = E \oplus F; (e_i)_{i \in [1,p]}$ basis of E, $(f_j)_{j \in [1,q]}$ basis of F, then $(e_1, ..., e_p, ..., f_1, ..., f_q)$ is basis of V. Such a basis is said adopted to the **decomposition** $V = E \oplus F$ and call F a **complimentary subspace** of E.

Remark: If E is subspace of $V(\dim V < \infty)$, then, E admits a complimentary subspace.

Grassmann Theorem. $\dim(E + F) = \dim E + \dim F - \dim(F \cap F)$

Proof. Let E' be a complimentary subspace to $E \cap F$ within $E \iff E' \oplus (E \cap F) = E$. Thus, we can rewrite $\dim E = \dim(E \cap F) + \dim E'$.

Observe that E' is also a complimentary subspace of F inside $E + F$, so

$$\begin{cases} E + F = E' \oplus F \\ E' \cap F = \{0\} \end{cases} \implies \dim(E + f) = \dim E' + \dim F = \dim E + \dim F - \dim(E \cap F)$$

To prove the observation $E' \cap F = \{0\}$, if $x \in E' \cap F$, then $x \in F$ and $x \in E' \subseteq E \implies x \in E \cap F$ and $x \in E' \implies x \in (E \cap F) \cap E' = \{0\}$.

To prove $E + F = E' + F$, use double inclusion. It is obvious that $\supseteq$. For $\subseteq$, pick $x + y \in E + F$. Check $x + y = x' + y'$ with $x' \in E', y' \in F$. $x \in E = E' \oplus E \cap F \implies x = x_{E'} + x_{E \cap F} \implies x + y = x_{E'} + (x_{E \cap F} + y)$. With $x_{E \cap F} + y$, we conclude that $x_{E \cap F} \subseteq F$, $y \in F$. We can write any element in $E + F$ as element in $E' + F$, so $E + F \subseteq E' + F$. □

3 Linear Maps

3.1 Generalities

Definition. Let E, F be 2 vector spaces over $\mathbb{K}$, a **linear map** (or linear operator) $f : E \to F$ is a map satisfying the following conditions:

- $f(\lambda u) = \lambda f(u); \forall \lambda \in \mathbb{K}, u \in E$
- $f(u + v) = f(u) + f(v); \forall u, v \in E$

This could be compressed into a single axiom: $\forall \lambda \in K$, $\forall u, v \in E$, $f(\lambda_u + v) = \lambda f(u) + f(v)$.

Algebraically, this could be interpreted as

$$f\left(\sum_i \lambda_i u_i\right) = \sum_i \lambda_i f(u_i)$$

Geometrically, this also tells us that $span(u, v)$ is mapped to $span(f(u), f(v))$ This implies that $f(span(u_1, ..., v_n)) = span(f(u_1), ..., f(u_n))$.

On a higher level, we can think that $f : E \to F$ preserves the vector space structure of E, F since it sends linear combination in E to F.

Proposition. Let $f : E \to F$. Then $ker(f) = \{x \in E;\ f(x) = 0\}$ is a subspace of E and $im(f) = \{f(x);\ x \in E\} \subseteq F$ is a subspace of F.

Exercise: Check the above propositions

Notation: $\mathcal{L}(E, F)$ or $\text{Hom}(E, F)$ (homomorphism) is the space of linear maps $E \to F$.
Exercise: Prove that this is also a vector space (subspace of functions $E \to F$).

Example. Let $A \in M_{m,n}(\mathbb{R})$, where $x \in R^n = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}^T$. Then $f(x) = Ax$, we have $f : \mathbb{R}^n \to \mathbb{R}^m is$ <u>linear</u>.

Example. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a linear map. Pick $(e_i)_{i \in [1,n]}$ to be the basis of $\mathbb{R}^n$. To know f entirely, it suffices to know it only for a finite number of vectors, namely $f(e_i), i \in [1, n]$. This is because if we let $\mathbf{x} \in \mathbb{R}^n$, we can write

$$\mathbf{x} = \sum_i x_i e_i \implies f(x) = f\left(\sum_i x_i e_i\right) = \sum_i x_i f(e_i)$$

If e_i are standard basis of $\mathbb{R}^n$ and f_i are standard basis of $\mathbb{R}^m$, we can set
matrix $A = \begin{bmatrix} f(e_1) & \dots & f(e_n) \end{bmatrix}$ where $f_{e_i} = \begin{bmatrix} y_{1,i} & \dots & y_{m,i} \end{bmatrix}^T$

Sub Example. We have

$$f(x_1, x_2) = \begin{bmatrix} x_1 + 2x_2 \\ 3x_1 + 4x_2 \end{bmatrix} = A \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \text{ with } A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}$$

$$\implies f(e_1) = f(1, 0) = \begin{bmatrix} 1 \\ 3 \end{bmatrix} = C_1(A) \qquad f(e_2) = f(0, 1) = \begin{bmatrix} 2 \\ 4 \end{bmatrix} = C_2(A)$$

<u>Consequence:</u>

$$im(f) = \text{ column space } (A) = span(C_1(A), \dots, C_n(A))$$

because any $f(x)$ is linear combination of $f(e_i)$:

$$f\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n x_i f(e_i) = \sum_{i=1}^n x_i C_i(A)$$

Definition. When a linear function lands in scalers, we call it **linear form / functional**

Example. Show trace function : $M_n(\mathbb{K}) \to \mathbb{K}$ is a linear form. Formally, we can write this as tr$\in \mathcal{L}(M_n(\mathbb{K}), \mathbb{K})$.

$$\forall A, B \in M_n(\mathbb{K}), \lambda \in \mathbb{K}.\, tr(\lambda A + B) = \sum_{i=1}^{n} (\lambda A + B)_{ii} = \lambda \sum_i A_{ii} + \sum_i B_{ii} = \lambda tr(A) + tr(B)$$

Example. $E = \mathbb{R}[x]$. $f = \frac{d}{dx} \in \mathcal{L}(E, E)$. If $P(x) = a_0 + a_1 x + \ldots + a_n x^n$, then $f(P) = a_1 + 2a_2 x + \ldots + n a_n x^{n-1}$. This is *linear*.

Example. Let $E = C([a, b])$ be continuous functions on $[a, b]$. $f = \int_a^b \cdot\, dt$, $E \to \mathbb{R}$. If $\varphi \in E$ is continuous function, then $f(\varphi) = \int_a^b \varphi(t)\, dt$.

Definition. Let $f : E \to F$. It is

- **surjective** if $\forall y \in F$, $f(x) = y$ has a solution $x \in E \iff \exists\, x \in E, f(x) = y$ (implies existence of solution)
- **injective** if $f(x) = f(x')\, \exists\, x, x' \in E \implies x = x'$ (implies uniqueness of solution when they exists)
- **bijective** if it is injective and surjective $\iff \forall\, y \in E, \exists\, !x \in E, f(x) = y$

Proposition. Let $f : E \to F$.

1. f surjective $\iff$ f has a right inverse, where $\exists\, g : F \to E \ni f \circ g = \underbrace{id_F}_{id_F(y)=y}$
2. f injective $\iff$ f has a left inverse, where $\exists h, F \to E, h \circ f = id_E$
3. f bijective $\iff$ right and left inverse with both of them equal

Proof. f surjective $\implies$: $\forall y \in F, \exists x \in E \ni f(x) = y$. Define $g : F \to E$, where $g(y) =$ one solution of equation $f(x) = y \implies (f \circ g)(y) = f(g(y)) = y$. The converse is left as an *exercise.* ☐

Proof. f injective $\implies$: If f injective, the equation $f(x) = y$ has a solution if $\underbrace{y \in f(E)}_{\exists ! x_y \in E \ni f(x_y) = y}$ We define

$$h(y) = \begin{cases} x_y, & \text{if } y \in f(E) \\ \text{anything else}, & \text{if } y \notin f(E) \end{cases}$$

Now, we simply check that $(h \circ f)(x) = h(f(x)) = x$ ☐

Proof. f bijective: If f has left inverse g and right inverse h,

$$f \circ h \circ g = \begin{cases} h \circ \underbrace{(f \circ g)}_{id} = h \\ \underbrace{(h \circ f)}_{id} \circ g = g \end{cases} \implies h = g$$

☐

We can conclude the for finite E, F , the number of elements can be concluded as

- f surjective $\implies \#E \geq \#F$
- f injective $\implies \#E \leq \#F$
- f bijective $\implies \#E = \#F$

3.2 Back to Vector Spaces

Proposition. $f \in \mathcal{L}(E, F)$ and $\mathcal{B} = (e_1, ..., e_n)$ basis of E, so $f(\mathcal{B}) := (f(e_1), \ldots, f(e_n))$

1. f injective $\iff \ker(f) = \{0\}$
2. $f(\mathcal{B})$ generates $im(f) = f(E)$.
3. f surjective $\iff f(\mathcal{B})$ generates F
4. f injective $\iff f(\mathcal{B})$ linearly independent
5. f bijective $\iff f(\mathcal{B})$ is a basis, and $\dim(E) = \dim(F)$

Proof. 1 $\implies$: Suppose f injective. Let $x \in ker(f) \iff f(x) = 0 = f(0) \implies x = 0.$. So $\ker(f) = \{0\}$.
$\impliedby$: If $\ker(f) = 0$, let $x, x' \in E$ such that $f(x) = f(x') \iff f(x) - f(x') = 0 \implies f(x - x') = 0 \implies x - x' \in ker(f) = \{0\}$. S $x - x' = 0 \implies x = x'$. ☐

Proof. 2: Pick $y \in im(f)$, $\exists x \in E$ such that $f(x) = y$. $\mathcal{B}$ basis of E so $x = \sum_{i=1}^{n} x_i e_i$ for some scalars $x_1, \ldots, x_n \in \mathbb{K} \implies f(x) = \sum_{i=1}^{n} x_i f(e_i) \in span(f(\mathcal{B}))$ ☐

Proof. 3 $\implies$: f surjective $\iff f(E) = im(f) = F$. From 2, $f(\mathcal{B})$ generates $im(f) = F$.
$\impliedby$: *exercise* ☐

Proof. 4 $\implies$: Suppose f injective. Let $\lambda_1, \ldots, \lambda_n \in \mathbb{K}$ such that $\lambda f(e_1) + ... + \lambda_n f(e_n) = 0 \iff f\left(\sum_{i=1}^{n} \lambda_i e_i\right) = 0 \implies \sum_i \lambda_i e_i \in ker(f) = \{0\} \implies \sum_i \lambda_i e_i = 0 \implies \lambda_i = 0 \, \forall i \in [1, n]$ because $\mathcal{B}$ basis and so it is linearly independent. ☐

Definition. If $f \in \mathcal{L}(E, F)$ bijective, we call f **isomorphism** where $\dim E = \dim F$. If $f \in \mathcal{L}(E, E) = \mathcal{L}(E)$ or $End(E)$, we call f **endomorphism** (with matrices, it would be square matrices). If $f \in End(E)$ and is isomorphism, we call f an **automorphism**.

Corollary. If $f \in \mathcal{L} \in (E, F)$ with $\dim E = \dim F$, then f injective $\iff f$ surjective $\iff f$ isomorphism.

Corollary. If $f \in End(E)$ with $\dim E < \infty$, then f injective $\iff f$ surjective $\iff f$ automorphism.

Corollary. E finite dimensional $\iff E$ isomorphic to $\mathbb{K}^n$. In particular, E isomorphic to $F \iff \dim E = \dim F$.

Example. $E = \mathbb{R}_n[x], \dim(E) = n + 1$. E isomorphic to $\mathbb{R}^{n+1}$ via $\varphi : E \to \mathbb{R}^{n+1}$. $\mathcal{B} = (1, x, ..., x^n)$ basis of E where $e = (e_i)_{i \in [1, n+1]}$ standard basis of $\mathbb{R}^{n+1}$. Set φ such that $\varphi(x_i) = e_{i+1}, \forall i \in [0, n]$...

Proof. "$\impliedby$": by definition

"$\implies$": Pick a basis $\mathcal{B} = (e_1, ..., e_n)$ of E. Set $\varphi : E \to \mathbb{K}^n$ with $\varphi(e_i)$ for basis vectors. Set $\varphi(e_i) = \begin{bmatrix} 0...0 & 1 & 0...0 \end{bmatrix}^T$, so $\varphi(x) = \varphi(\sum_{i=1}^{n} x_i e_i) = \begin{bmatrix} x_1 & \ldots & x_n \end{bmatrix}^T$ ☐

Important Remark: $\simeq$ notates isomorphism. If $\varphi : E \simeq F$, then φ^{-1} *is also linear.* (*Exercise. Use injectivity*). Identifying that $f : \mathbb{R}^n \to \mathbb{R}^n$ linear $\iff f(x) = Ax$ for some $A \in M_n(\mathbb{R})$, we can look at this with square matrix. Previously, we had $A \in M_n(\mathbb{R})$ left invertible$\iff A$ right invertible $\iff A$ invertible.

Proof. $f(x) = Ax, f \in End(\mathbb{R}^n)$.

$$f \text{ left invertible } \iff f \text{ injective } \iff \underbrace{f \text{ surjective}}_{f \text{ right invertible}} \iff \underbrace{f \text{ bijective}}_{f \text{ invertible}}$$

□

$A \in GL_n(\mathbb{R}) \iff Ax = 0$ admits only $x = 0$ as solution.

Proof.

$$f(x) = Ax, f \in End(\mathbb{R}^n) \implies f \text{ isomorphism } \iff f \text{ injective } \iff ker(f) = \{0\}$$

□

3.3 Rank-nullity Theorem

Fix $f \in \mathcal{L}(E, F)$. Linear system $Ax = 0$ with $A \in M_{m,n}(\mathbb{K})$ had

$$rank(A) + \dim ker(A) = n$$

Rank-nullity Theorem. $f \in \mathcal{L}(E, F)$.

$$\dim ker(f) + \dim im(f) = \dim E$$

We call rank(f) = dim im(f).

Heuristic Proof: (see TH2A 3) $f \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^n)$, $f(x) = Ax$. We know that im(f) = column space(A). Take one solution to $Ax = 0$ so show a linear relationship between columns. Extract a basis of im(f) by erasing dim $ker(F)$ vectors.

Pick a basis $(e_1, ..., e_p)$ of ker(f) and complete into basis of E, $\mathcal{B} = (e_1, ..., e_p, e_{p+1}, ..., e_n)$. From last time, we know that $f(\mathcal{B})$ generates $im(f) \implies f(\mathcal{B}) = (\underbrace{f(e_1), ..., f(e_p)}_{=0}, \underbrace{f(e_{p+1}), ..., f(e_n)}_{n-p \text{ elements}})$. We claim that $(f(e_{p+1}), ..., f(e_n))$ is the basis of $im(f)$ since it is generating. Looking at its linear independence, we pick $\lambda_{p+1}, ..., \lambda_n \in \mathbb{K}$ such that $\lambda_{p+1} f(e_{p+1}) + ... + \lambda_n f(e_n) = 0 \iff f(\underbrace{\lambda_{p+1} e_{p+1} + ... + \lambda_n e_n}_{\text{in } ker(f) = span(e_1, ..., e_p)}) = 0$.
So, the only possibility is that $\lambda_{p+1} e_{p+1} + ... + \lambda_n e_n = 0 \implies \lambda_{p+1} = ... = \lambda_n = 0$

Second Proof (technically the same):

Proof. Let E_0 be the complimentary subspace of $ker(f)$ in $E \implies E_0 \oplus ker(f) = E$. Then, $f : E_0 \to f(E_0)$ is an isomorphism, since it is subjective by definition. To prove its injectivity, we pick $x \in ker(f_0)$, where $f(x) = 0$. So, $x \in ker(f) \cap E_0 = \{0\}$. Also, $f(E_0) = im(f)$ and f_0 isomorphism $\implies \dim E_0 = \dim(im(f))$. $\dim E = \dim E_0 + \dim ker(f) = \dim im(f) + \dim ker(f)$. □

Example. Let V be a finite dimensional vector space where E, F are subspaces of V. We have $\dim(E + F) = \dim E + \dim F - \dim E \cap F$.

Proof. Set $\varphi : E \times F \to E + F$, where $(x, y) \mapsto \varphi(x + y) = x + y$. $ker(\varphi) = \{(x, -x); x \in E \cap F\}$, so φ is surjective by definition. We see that ker(φ) is isomorphic to $E \cap F$ via $z \in E \cap F \mapsto (z, -z) \in ker(\varphi) \implies \dim ker(\varphi) = \dim E \cap F$. So, grassmann follows from rank-nullity. □

Exercise: f endomorphism of E. Then f injective $\iff$ f surjective $\iff$ f isomorphic. Prove with rank-nullity.

Example. $E \xrightarrow{\varphi} F \xrightarrow{g} G$. φ isomorphism, $g \in \mathcal{L}(F, G)$. Then $rk(g \circ \varphi) = rk(g)$.

Example. $E \xrightarrow{f} F \xrightarrow{\Psi} G$. Ψ isomorphism, $f \in \mathcal{E}, \mathcal{F}$. Then, $rk(\varphi \circ f) = rk(f)$

Exercise: Prove the statements above, which shows that multiplication by inverse matrices doesn't change rank.

3.4 Hyper Planes and Linear Forms

The purpose of this is to abstract row rank. We denote $E = \mathbb{K}$ vector space and $E^* = \mathcal{L}(E, \mathbb{K})$.

Example. If $E = \mathbb{R}^n$, any linear form is a function $\varphi(x_1, ..., x_n) = a_1 x_1 + ... + a_n x_n$. When $n = 3$, $ker(\varphi)$ is a plane. In general, by rank-nullity theorem, $\dim ker(\varphi) = n - 1$. If $\varphi_1, ..., \varphi_p \in E^*$,

$$x \in \bigcap_{i=1}^{p} ker(\varphi_i); \text{ in } E = \mathbb{R}^n \iff x \text{ solution of linear system}$$

In $\mathbb{R}^n$, $\varphi_1, ..., \varphi_p$ linear independent $\iff$ row vectors of the linear system are linearly independent.

Theorem. If $\varphi_1, ..., \varphi_p \in E^*$ linear independent, then

$$\dim \left(\bigcap_{i=1}^{p} ker(\varphi_i) \right) = \dim E - p$$

We interpret solution of the linear system as intersection of hyperplanes $ker(\varphi_i)$

3.5 Bases

$\forall x \in E$, we can write $x = \sum_j x_j e_j \implies f(x) = \sum_j x_j f(e_j)$. In turn, we can express $\forall \in [1, n]$, $f(e_j)$ with coordinates in $\mathcal{C} \implies f(e_j) = \sum_{i=1}^{m} a_{i,j} f_i \; \exists \, a_{i,j} \in \mathbb{K}$. We get a matrix $A = (a_{i,j})_{i \in [1,m], j \in [1,n]}$. In particular, we need only $a_{i,j}$ to determine f, and there are only $m \times n$ parameters $\implies \dim \mathcal{L}(E, F) = m \times n$.

Definition. We define matrix $[f]_{\mathcal{B},\mathcal{C}}$ of f relative to bases $\mathcal{B}, \mathcal{C}$ is the matrix A above.

$$[f]_{\mathcal{B},\mathcal{C}} = [f(\mathcal{B})]_{\mathcal{C}} = \begin{bmatrix} [f(e_1)]_{\mathcal{C}} & \cdots & [f(e_n)]_{\mathcal{C}} \end{bmatrix}$$

Then $f(\mathcal{B}) = (f(e_1), ..., f(e_n)) \subseteq F$ so each $f(e_j)$ can be written as coordinates in basis $\mathcal{C}$. **Example.** $f : \mathbb{R}^3 \to \mathbb{R}^2$ where

$$f(x_1, x_2, x_3) = \begin{bmatrix} x_1 - x_2 + x_3 \\ x_1 + x_2 + x_3 \end{bmatrix}$$

$\mathcal{B} = (e_1, e_2, e_3)$ are standard basis of $\mathbb{R}^3$ and $\mathcal{C} = (f_1, f_2)$ are standard basis of $\mathbb{R}^2$.

$$f(e_1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = f_1 + f_2 \qquad f(e_2) = \begin{bmatrix} -1 \\ 1 \end{bmatrix} = -f_1 + f_2 \qquad f(e_3) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = f_1 + f_2$$

Example. If $\mathcal{B}' = (e_1', e_2', e_3') = (e_1 - e_2, e_1 + e_2, e_2 + e_3)$

$$f(\mathcal{B}') : \; f(e_1') = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad f(e_2') = \begin{bmatrix} 0 \\ 2 \end{bmatrix}, \quad f(e_3') = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \implies [f]_{\mathcal{B},\mathcal{C}} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 2 \end{bmatrix}$$

Example. $f = \frac{d}{dx}$. If $P(x) = a_1 + a_1 x + a_2 x^2 \implies f(P) = a_1 + 2a_2 x$:

$$[f]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{bmatrix}$$

Theorem. If $f \in \mathcal{L}(E, F)$, $\mathcal{B} = (e_1, ..., e_n) =$ basis of E and $\mathcal{C} = (f_1, ..., f_n)$= basis of F. We denote $[x]_{\mathcal{B}} = \begin{bmatrix} x_1 & \dots & x_n \end{bmatrix}_{\mathcal{B}}^{T}$ means $x = \sum_{i=1}^{n} x_i e_i$. Then $[f]_{\mathcal{B},\mathcal{C}} [x]_{\mathcal{B}} = [f(x)]_{\mathcal{C}}$. In addition, if $g \in \mathcal{F}, \mathcal{G}$ and $\mathcal{D}$ is basis of G, then $[g \circ f]_{\mathcal{B},\mathcal{D}} = [g]_{\mathcal{C},\mathcal{D}} [f]_{\mathcal{B},\mathcal{C}}$.

Example. Let $\theta \in \mathbb{R}$. Let r_θ be the rotation angle $\theta \in End(\mathbb{R}^2)$ and $\mathcal{B} = (e_1, e_2)$. Then, $r_\theta(e_1) = \cos(\theta)e_1 + \sin(\theta)e_2$ and $r_\theta(e_2) = -\sin(\theta)e_1 + \cos(\theta)e_2$. If ψ is another angle and $r_{\theta+\psi} = r_\theta \circ r_\psi$. Then, s

Theorem.

$$[g \circ f]_{\mathcal{B},\mathcal{D}} = [g]_{\mathcal{C},\mathcal{D}}[f]_{\mathcal{B},\mathcal{C}}$$

Proof. $[g \circ f]_{\mathcal{B},\mathcal{D}}$ has column vectors $[(g \circ f)(\underbrace{e_j}_{\in \mathcal{B}})]_{\mathcal{D}}$. We have

$$(g \circ f)(e_j) = g(f(e_j)) = g\left(\sum_k b_{k,j} f_k\right) = \sum_k b_{k,j} g(f_k) = \sum_k b_{k,j} \sum_i a_{i,k} g_i = \sum_i \left(\sum_k a_{i,k} b_{k,j}\right) g_i$$

$$= (AB)_{i,j} \text{ and } AB = [g]_{\mathcal{C},\mathcal{D}}[f]_{\mathcal{B},\mathcal{C}}$$

□

Corollary. $[f]_{\mathcal{B},\mathcal{C}}[x]_{\mathcal{B}} = [f(x)]_{\mathcal{C}}$

3.6 Transition Matrix

When we have $\mathcal{B}, \mathcal{B}'$ as 2 bases of E, we want to *translate* $[x]_{\mathcal{B}}$ into $[x]_{\mathcal{B}'}$.

Definition. The **transition matrix** $\mathcal{B}$ to $\mathcal{B}'$ is denoted as

$$P_{\mathcal{B} \to \mathcal{B}'}$$

Corollary. When $f = id, \mathcal{C} = B', \mathcal{D} = B''$, then

$$P_{\mathcal{B} \to \mathcal{B}'}[x]_{\mathcal{B}'} = [x]_{\mathcal{B}} \qquad P_{\mathcal{B}' \to \mathcal{B}''} P_{\mathcal{B} \to \mathcal{B}'} = P_{\mathcal{B} \to \mathcal{B}''}$$

Example. For matrix $f \in \mathcal{L}(E, F), \mathcal{B}$ basis of E and $\mathcal{B}'$ basis of F, with $E = F$ and $f = id_E$. $\mathcal{B} = (e_j)$ and $\mathcal{B}' = (e'_i)$.

$$P_{\mathcal{B} \to \mathcal{B}'} = [id]_{\mathcal{B},\mathcal{B}'} \implies P_{\mathcal{B} \to \mathcal{B}} = I_n$$

The consequence is that we can conclude $P_{\mathcal{B} \to \mathcal{B}'} \in GL_n(\mathbb{K})$ and $P_{\mathcal{B}' \to \mathcal{B}} = P^{-1}_{\mathcal{B} \to \mathcal{B}'}$

Theorem. Change of Basis of Linear Maps. With basis $\mathcal{B}, \mathcal{B}'$ in E and $\mathcal{C}, \mathcal{C}'$ in F with some matrix f, we have

$$[f]_{\mathcal{B},\mathcal{C}} = P_{\mathcal{C}' \to \mathcal{C}}[f]_{\mathcal{B}',\mathcal{C}'} P_{\mathcal{B} \to \mathcal{B}'}$$

3.6.1 Applications to Equivalent Matrices

We can denote **equivalent matrices** $A \sim B \in M_{m,n}(\mathbb{K})$ when $\exists P \in GL_m(\mathbb{K}), Q \in GL_n(\mathbb{K})$ such that $A = PBQ$.

Thus, $A \sim B \iff A$ and B represent the same linear map in different bases.

Proof. "$\implies$": $A = PBQ$ as above. Set $f : \mathbb{R}^n \to \mathbb{R}^m, f(x) = Ax$. If $\mathcal{B}, \mathcal{C}$ are canonical bases of $\mathbb{R}^n, \mathbb{R}^m$, then $[f]_{\mathcal{B},\mathcal{C}} = A$. Also, $\mathcal{B}'$ are the columns of Q^{-1} according to $\mathcal{B} \implies Q^{-1} = P_{\mathcal{B}' \to \mathcal{B}}$ and $\mathcal{C}$ are columns of P according to $\mathcal{C}$, $\implies P_{\mathcal{C}' \to \mathcal{C}} = P$, and $[f]_{\mathcal{B}',\mathcal{C}'} = \mathcal{B}$.

"$\Longleftarrow$": Change of base formula. □

Exercise: Check on last example.

Revisit the following theorem: A has rank $r \iff A \sim J_r$

Proof. Set $f(x) = Ax$: " $\implies$ " From rank-nullity, basis of $\ker f = (e_1, ..., e_p)$. Here, $p = \dim\ker f, r = rank(f), p + r = n$. Complete this into basis of $\mathbb{R}^n$, which is $(\varepsilon_1, ..., \varepsilon_r, e_1, ..., e_p)$. Consider $f(\mathcal{B}) = (f(\varepsilon_1), ..., f(\varepsilon_r), \underbrace{f(e_1), ..., f(e_p)}_{=0; e_i \in ker}) = f(\mathcal{B}) = (f(\varepsilon_1), ..., f(\varepsilon_r))$. Completing this into basis of $\mathbb{R}^m$, $\mathcal{C} = (f(\varepsilon_1), ..., f(\varepsilon_r), f_{r+1}, ..., f_m)$. This becomes obvious when we look at change of basis matrix.

"$\impliedby$:" Multiplying by invertible matrix $P, \psi(x) = Px$ invertible map ($\psi^{-1}(y) = P^{-1}y$). We saw that comparing f by isomorphisms (left or right) doesn't change rank of f □

4 Abstract Theory of Determinant

Initially, determinant is motivated by calculation of volume.

Motivation Denote $\underbrace{\det(u,v)}_{\mathbb{R}^2}, \underbrace{\det(u,v,w)}_{\mathbb{R}^3}$ be area or volume (with u, v, w) being vectors. Also denote $\det = \varphi$ and consider the properties of φ.

- $\varphi(\lambda u, v) = \lambda\varphi(u, v)$ for $\lambda \neq 0$
- $\varphi(u + w, v) = \varphi(u, v) + \varphi(w, v)$. (Same for v).
- $\varphi(u, v) = -\varphi(v, u)$

4.1 General Definition

Denote $\Lambda^n E*$ as the space of alternating n-linear form (volume form).

Definition. Alternating n-linear form is function $\varphi : \underbrace{E \times ... \times E}_{n \text{ times}} \to \mathbb{K}$ with following properties:

- $\varphi(u_1, ..., u_n)$ linear in each variable: $\varphi(u_1, ..., \lambda u_i + v_i, ..., v_n) = \lambda\varphi(u_1, ..., u_i, ..., u_n) + \varphi(v_1, ..., v_n)$
- φ is **alternating**: $\varphi(u_1, ..., u_i, ..., u_j, ..., u_n) = -\varphi(u_1, ..., u_j, ..., u_i, ..., u_n)\ \forall i \neq j$

Proposition. If $u_i = u_j = u\, \exists\, i \neq j$, then $\varphi(u_1, ..., u, ..., u, ..., u_n) = -\varphi(u_1, ..., u, ..., u, ..., u_n) = 0$. Thus, two same input will result in 0.

Proposition. $(u_1, ..., u_n)$ linearly dependent $\implies \varphi(u_1, ..., u_n) = 0$.

Proposition. $\varphi(u_1, ..., u_i + span(u_1, ..., u_n), ..., u_n) = \varphi(u_1, ..., u_i, ..., u_n)$. This is an abstract form of the fact that determinants are "invariant" under this type of column operation.

Remark: These are usual properties of classical determinant of $n \times n$ matrices.

Theorem. $\boxed{\dim \Lambda^n E^* = 1}$

This means that $\forall\, \varphi, \psi \in \Lambda^n E^*$, then $\exists\, \lambda \neq 0$ such that $\varphi = \lambda\psi$. It also implies that up to a choice of unit of volume, there is only 1 choice of alternating n-linear form. Formally:

Theorem. Fix $\mathcal{B}$ = basis of E, $\mathcal{B} = (e_1, ..., e_n)$.. $\exists!\, \varphi_0 \in \Lambda^n E^*$ such that $\varphi_0(\mathcal{B}) = \varphi_0(e_1, ..., e_n) = 1$. Denote $\varphi_0 = det_{\mathcal{B}} \implies \det_{\mathcal{B}}(\mathcal{B}) = 1$. Then we have the determinant of $(u_1, ..., u_n)$ in $\mathcal{B}$.

Proof. For $n = 2$, $\mathcal{B} = (e_1, e_2)$. $u = \begin{bmatrix} a \\ b \end{bmatrix}_{\mathcal{B}} = ae_1 + be_2$; $\qquad v = \begin{bmatrix} c \\ d \end{bmatrix}_{\mathcal{B}} = ce_1 + de_2$.

Let $\varphi \in \Lambda^2 E^*$. Then

$$\begin{aligned}
\varphi(u, v) =& \varphi(ae_1 + be_2, ce_1 + de_2) \\
&= \varphi(ae_1, ce_1) + \varphi(ae_1, de_2) + \varphi(be_2, ce_1) + \varphi(be_2, de_2) \\
&= ad\varphi(e_1, e_2) - bc\varphi(e_1, e_2) \\
&= \underbrace{(ad - bc)}_{=\varphi_0(u,v)}\varphi(e_1, e_2), \text{ where } \varphi(e_1, e_2) \text{ is choice of unit of volume} \implies \Lambda^2 E^* = span(\varphi_0)
\end{aligned}$$

For $n = 3$, $\varphi(u, v, w) = \varphi_0(u, v, w)\varphi(e_1, e_2, e_3)$ □

In general, $u_j = \sum_{i=1}^{n} a_{i,j} e_i \implies A = \begin{bmatrix} u_1 & \cdots & u_n \end{bmatrix} \in M_n(\mathbb{K})$.
Then $\varphi(u_1, ..., u_n) = \varphi_0(u_1, ..., u_n)\, \varphi(e_1, ..., e_n)$ where

$$\varphi_0(u_1, ..., u_n) = \sum_{\sigma \in S_n} (-1)^{\tau(\sigma)} a_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n}$$

where S_n is the symmetric group = all permutations of the set $\{1, ..., n\}$ and $\tau(\sigma)$ is the number of transpositions (permutation that only flips 2 elements) involved in σ.

Definition. If $A \in M_n(\mathbb{K})$, then its determinant $det(A) = det_{\mathcal{B}}(C_1(A), ..., C_n(A))$

Example.

$$\det \begin{bmatrix} a & c \\ b & d \end{bmatrix} = \det_{\mathcal{B}}(ae_1 + be_2, ce_1 + de_2) = \det \left(\begin{bmatrix} a \\ b \end{bmatrix}_{\mathcal{B}}, \begin{bmatrix} c \\ d \end{bmatrix}_{\mathcal{B}} \right) = ad - bc$$

Properties

1. A invertible $\iff \det(A) \neq 0$
2. If A upper triangular, then $\det A = a_1 a_2 ... a_n$. where $a_1, ..., a_n$ are the diagonals
3. $\det A = \det(A^t)$

Proof. Number 1 $\implies (u_1, ..., u_n)$ linearly dependent $\implies \varphi(u_1, ..., u_n) = 0$ when $\varpi \in \Lambda^n E^*$.

"$\Longleftarrow$": A invertible $\iff rk(A) = n \iff$ columns of A form a bsis $\mathcal{B}'$ of $\mathbb{R}^n$.

Then, $\det_{\mathcal{B}}(u_1, ..., u_n) = \lambda \det_{\mathcal{B}}(u_1, ..., u_n)$ for some $\lambda \in \mathbb{K} \implies \det_{\mathcal{B}}(\mathcal{B}) = \lambda \det_{\mathcal{B}'}(\mathcal{B}) \implies \lambda = \det_{\mathcal{B}}(\mathcal{B}')$. Then, $\det_{\mathcal{B}}(\mathcal{B}') \det_{\mathcal{B}} \mathcal{B}' = 1$. $\det_{\mathcal{B}}(\mathcal{B}') = \det(A) \neq 0$. ☐

Proof. Number 2: Use the formula with permutations. If $\sigma \neq \text{id}$, then there's at least one i such that $(\sigma(i), i)$ spot below diagonal. ☐

Proof. Number 3: $\det(A)$ is also a volume form in the rows. n-linearity with respect to rows is becauses terms $a_\sigma = a_{\sigma(a),1}, ..., a_{\sigma(n),n}$ involve each row only once.

Alternating: $(n = 2, \varphi(u, v) = -\varphi(v, u))$ alternating. ☐

Effect of elementary row or column operations. Let A' be A after some operations.

- $R_i \leftarrow R_i + \lambda R_j$ has $\det(A') = \det(A)$
- $R_i \leftarrow \lambda R_i$ has $\det(A') = \lambda \det(A)$
- $R_i \leftrightarrow R_j$ has $\det(A') = -\det(A)$

4.2 Recursive Formula for Determinants

Definition. Let $A \in M_n(\mathbb{R})$. $\triangle_{i,j}(A) = (i, j)$ - minor of A = matrix btained from A alter removing R_i, C_j. Thus, $\triangle_{i,j} \in M_{m,n}(\mathbb{R})$. We also define (i, j) - cofactor = $\det(\triangle_{i,j})$.

Theorem. Expansion of $\det(A)$ along a row or column.

- Along R_i: $\det(A) = \sum_{j=1}^{n} (-1)^{i+j} a_{i,j} \cdot \det(\triangle_{i,j})$
- Along C_j: $\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{i,j} \cdot \det(\triangle_{i,j})$

Proof. Suppose $n = 3$, then

$$\begin{vmatrix} \cdot & \cdot & *_1 \\ \cdot & \cdot & *_3 \\ \cdot & \cdot & *_3 \end{vmatrix} = \det_\beta(C_1, C_2, *_1e_1 + *_2e_2 + *_3e_3)$$

$$= \det_\beta(C_1, C_2 + *_1e_1) + \det_\beta(C_1, C_2 + *_2e_2) + \det_\beta(C_1, C_2 + *_3e_3)$$

□

Example.

$$D_2 = \begin{vmatrix} \lambda & 1 & 1 \\ 1 & \lambda & 1 \\ 1 & 1 & \lambda \end{vmatrix} = \begin{vmatrix} \lambda & 1 & 1 \\ 1 & \lambda & 1 \\ 0 & 1-\lambda & \lambda-1 \end{vmatrix} = \begin{vmatrix} \lambda & 2 & 1 \\ 1 & \lambda+1 & 1 \\ 0 & 0 & \lambda-1 \end{vmatrix}$$

$$= (\lambda-1)\begin{vmatrix} \lambda & 2 \\ 1 & \lambda+1 \end{vmatrix} = (\lambda-1)\begin{vmatrix} \lambda+2 & 2 \\ 0 & \lambda-1 \end{vmatrix} = (\lambda-1)^2(\lambda+2)$$

4.3 Vandermonde Determinant

Let $x_1, ..., x_n \in \mathbb{K}$, then

$$V_n(x_1, ..., x_n) = \begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix}$$

To solve this,

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} = \begin{vmatrix} 1 & 1 & \dots & 1 \\ 0 & (x_2 - x_1) & \dots & x_n - x_1 \\ \vdots & \vdots & \ddots & \vdots \\ x^{n-2}() & & & \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} ... = \prod_{1 \le i < j \le n} (x_j - x_i)$$

Another more classical and concise method is the following:

Proof. Set $\varphi(x) = V_n(x_1, ..., x_{n-1}, x_n)$, where $\varphi(x)$ is a polynomial in x with degree $n-1$. If we set $x = x_i, i \in [1, n-1]$, then 2 columns equal $\implies \varphi(x_i) = 0 \implies x_1, ..., x_n$ are roots of $\varphi(x) \in \mathbb{R}_n[x] \implies \varphi(x) = \alpha(x - x_1)...(x - x_{n-1})$. Expansion along C_n shows that $\alpha = (n, n)$ minor $= V_{n-1}(x_1, ..., x_{n-1}) \implies \varphi(x) = V_{n-1}(x_1, ..., x_{n-1})$. □

4.3.1 Determinant of an Endomorphism

Definition. Let $E = \mathbb{K}$ vector space. $\mathcal{B}$ is the basis of E, $f \in End(E)$. Then

$$\det(f) = \det_\mathcal{B}(f(\mathcal{B})) = \det([f]_\mathcal{B})$$

Rewriting, let $u_1, ..., u_n \in E$. Then we have

$$\det_\mathcal{B}(f(u_1), ..., f(u_n)) = \lambda \det_\mathcal{B}(u_1, ..., u_n)$$

where λ is the same for all vectors u_i. Therefore,

$$\det_\mathcal{B}(f(\mathcal{B})) = \lambda \underbrace{\det_\mathcal{B}(\mathcal{B})}_{=1} \implies \lambda = \det(f)$$

<u>Properties:</u>

1. $\det(f)$ doesn't depend on choice of basis.
2. If $g \in End(E)$, $\det(g \circ f) = \det(g) \cdot \det(f)$
3. $A, B \in M_n(\mathbb{K})$. Then $\det(AB) = \det(A)\det(B)$
4. If $A \in GL_n(\mathbb{K})$, then $\det(A^{-1}) = \frac{1}{\det(A)}$

Proof. Property 2: Let $u_1, ..., u_n \in E$.

$$\det_{\mathcal{B}}(g(\underbrace{f(u_1)}_{=v_1}), ..., g(f(u_n))) = \begin{cases} \det(g \circ f)\det_{\mathcal{B}}(u_1, ..., u_n) \\ \det(g)\det_{\mathcal{B}}(v_1, ..., v_n) \\ \det(g)\det_{\mathcal{B}}(f(u_1), ..., f(u_n)) \end{cases}$$

Then, $\det_{\mathcal{B}}(f(u_1), ..., f(u_n)) = \det(f)\det_{\mathcal{B}}(u_1, ..., u_n) \implies \det(g \circ f) = \det(g)\det(f)$. □

Proof. Property 1: $\det(f) = \det([f]_{\mathcal{B}})$. Let $\mathcal{B}'$ be another basis of E.

$$[f]_{\mathcal{B}'} = P_{\mathcal{B}\to\mathcal{B}'}[f]_{\mathcal{B}}P_{\mathcal{B}'\to\mathcal{B}} \implies \det[f]_{\mathcal{B}'} = \det(P_{\mathcal{B}\to\mathcal{B}'})\det([f]_{\mathcal{B}})\det(P_{\mathcal{B}'\to\mathcal{B}})$$

□

In a matrix where $A = [f]_{\mathcal{B}}, P \in GL_n(\mathbb{K})$, $\mathcal{B} = [g]_{\mathcal{B}}$

- $\det(P^{-1}AP) = \det(A)$
- $\det(BA) = \det(B)\det(A)$

5 Eigenvalues

5.1 Eigenstuff

Definition. Let $f \in End(E)$, $\dim E < \infty$.

1. $\lambda \in \mathbb{K}$ is an **eigenvalue** of f if $\exists\, x \neq 0 \in E$ such that $f(x) = \lambda x \iff x \in ker(f - \lambda id_E)$
2. The x above is called a **eigenvector** with respect to the eigenvalue λ
3. $\ker(f - \lambda id_E)$ is the λ **eigenspace**, or the subspace of all λ eigenvectors.
4. Can replace the above f, id with matrices A, I_n to get matrix version, where:

$$Ax = \lambda x \iff x \in \ker(A - \lambda I_n)$$

Often, with finding eigenvalues,

$$x \neq 0 \in \ker(f - \lambda id_E) \iff (f - \lambda id_E) \text{ not injective} \iff \det(f - \lambda id_E) = 0$$

Definition. $\chi_f(\lambda) = \det(f - \lambda id_E) =$ **characteristic polynomial** of $f \in End(E)$, where its roots are eigenvalues of f. Supppose $\chi_f(\lambda)$ is **split**. Then

$$\chi_f(\lambda) = (\lambda_1 - \lambda)^{m_1} \cdots (\lambda_p - \lambda)^{m_p}$$

where $\lambda_1, ..., \lambda_p$ are eigenvalues are m_i are **algebraic multiplicity** of λ_i, or $m_{alg}(\lambda_i)$. Note that if $\mathbb{K} = \mathbb{C}$, all polynomials are split. The **geometric multiplicity** of λ_i is

$$m_{geo}(\lambda_i) = \dim \ker(f - \lambda_i id)$$

Once the eigenvalues are found, let λ be the eigenvalue and solve $f(x) - \lambda x = 0$, where its solutions are eigenvectors.

The main interest is to find basis $\mathcal{B}'$ where $[f]_{\mathcal{B}'}$ has diagonal of eigenvalues counted with multiplicity.

Example. $f \in End(\mathbb{R}^n)$,

$$f(x) = \begin{bmatrix} x_2 + x_3 \\ x_1 + x_3 \\ x_1 + x_2 \end{bmatrix} \implies [f]_{\mathcal{B}} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}. \implies \chi_f(\lambda) = \begin{vmatrix} -1 & 1 & 1 \\ 1 & -\lambda & 1 \\ 1 & 1 & -\lambda \end{vmatrix} = (\lambda + 1)(\lambda - 2)$$

5.1.1 Purpose of Notations

Example. Let $f \in End(\mathbb{R}^n)$, $f(x) = Ax$ where $A \in M_n(\mathbb{R})$ and $\mathcal{B}$ be the canonical basis of $\mathbb{R}^n$ where $[f]_{\mathcal{B}} = A$. Then we can find a basis $\mathcal{B}'$ made of eigenvectors $\mathcal{B} = (e_1', ..., e_n') \implies \forall\, i \in [1, n], \exists\, \lambda_i \in \mathbb{K} \ni f(e_i') = \lambda_i \implies [f]_{\mathcal{B}}$ is a matrix with diagonal of eigenvalues.

Definition. $f \in End(E)$ is **diagonalizable** if $\exists\, \mathcal{B}'$ basis of E made of eigenvectors $\iff [f]_{\mathcal{B}}$ is **diagonal**.

Theorem. Let $f \in End(E)$ with eigenvalues $\lambda_1, ..., \lambda_p$ all distinct. Then

$$f \text{ diagonalizable } \iff \forall\, i, m_{alg}(\lambda_i) = m_{geo}(\lambda_i)$$

Fact: Geometric multiplicity is always ≥ 1

Example. $f \in End(\mathbb{R}^2)$ such that

$$f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} -x_1 + 2x_2 \\ -3x_1 + 4x_2 \end{bmatrix}. \quad \text{Is this diagonalizable?}$$

$$\begin{bmatrix} -x_1 + 2x_2 \\ -3x_1 + 4x_2 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ -3 & 4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \implies \chi_f(\lambda) = \begin{vmatrix} -1-\lambda & 2 \\ -3 & 4-\lambda \end{vmatrix} = \begin{vmatrix} 1-\lambda & 2 \\ 1-\lambda & 4-\lambda \end{vmatrix} = \begin{vmatrix} 1-\lambda & 2 \\ 0 & 2-\lambda \end{vmatrix}$$
$$= (1-\lambda)(2-\lambda) \implies \text{eigenvalues: 1 with A.M. 1, 2 with A.M. 1}$$
$$\implies m_{alg}(1) = m_{geo}(1), m_{alg}(2) = m_{geo}(2) \implies f \text{diagonalizable}$$

To check,

$$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \in \ker(f - id) = E_1(f) \iff (A - I_2)x = 0 = \begin{bmatrix} -2 & 2 \\ -3 & 3 \end{bmatrix} x$$

By rank-nullity,

$$\underbrace{\dim \ker(A - I_2)}_{m_{geo}(1)} = 1 = m_{alg}(1) \implies \text{ basis } E_1(f) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} = v_1$$

Similar for $E_2(f)$, where its basis is $\begin{bmatrix} 2 \\ 3 \end{bmatrix} = v_2$. We can <u>conclude</u> that with eigenvector basis v_1, v_2,

$$A = PDP^{-1},\ D = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix},\ P = P_{\mathcal{B}' \to \mathcal{B}} = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \mathcal{B}' = (v_1, v_2)$$

5.2 Polynomials in Endomorphism

5.2.1 Annihilator Polynomials

Let $f \in End(E)$ (or $A \in M_n(\mathbb{K})$).

Definition. Polynomial in f if $P \in \mathbb{K}[x]$, $P(x) = a_0 + a_1 x + ... + a_p x^p$, and denote $P(f) = a_0 id + a_1 f + a_2 f^2 + ... + a_p f^p \in End(E)$. With matrices, $P(A) = a_0 I_n + a_1 A + ... + a_p A^p$.

Definition. Annihilator Polynomial of f: $P \in \mathbb{K}[x]$ such that $P(f) = 0$, or $P(A) = 0$. Then, $I_f = \{P \in \mathbb{K}[x];\ P(f) = 0\}$ = set of annihilator polynomials. This is also called annihilator <u>ideal</u> of f.

Example.

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix},\ A^2 = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix} \iff A^2 - A - 2I_2 = 0 \implies P(x) = x^2 - x - 2 \in I_A$$

$\therefore$ Here, P annihilates A. Also note in this example that $(x^2 - x - 2) = (x+1)(x-2)$, and this has the same factors as $\chi_A(A)$.

Cayley-Hamilton Theorem. Let $f \in End(E)$ such that $\dim(E) < \infty$. Then, $\chi_f(f) = 0$

Definition. Minimal Polynomial: We say that f admits/has a minimal polynomial if $\exists P \in I_f, i.e., P(f) = 0, \exists\, P \neq 0$.

<u>Discussion:</u> If $I_f \neq \{0\}$, so f admits a minimal polynomial. $\implies \exists\, \pi_f \in \mathbb{K}[x]$ of polynomial degree ≥ 1. We pick π_f so its leading coefficient equals 1 and call π_f "the" *minimal polynomial.*

Proposition. If $P \in I_f$, or P annihilates f, then π_f divides P. Formally, this means $P(x) = \pi_f(x)Q(x) \exists\, Q \in \mathbb{K}[x]$.

Proof. Use Long division. We get

$$P(x) = Q(X)\pi_f(x) + R(x) \implies \underbrace{P(f)}_{0} = \underbrace{Q(f) \circ \pi_f(f)}_{=0 \text{ since } Q(F) \text{ linear}} + R(f) \implies R(f) = 0$$

Hence $R_f \in I_f$ with $\deg(R_f) < \deg(\pi_f) \implies R = 0 \implies P(x) = Q(x)\pi_f(x)$ and π_f divides P. □

Consequences:

- Uniqueness of minimal polynomial by long division.
- Admitting χ_f is an annihilator polynomial $\implies$ π_f divides χ_f so some eigenvalues of f are roots of π_f.

Example. (continued) $P(x) = x^2 - x - 2 = (x-2)(x+1)$ *is* the minimal polynomial π_A since P annihilates A and π_A divides $(x-2)(x-1)$. So, $\pi_A(x) = P$ or $(x+1)$ or $(x-2)$.

$\pi_A(x) = x+1$ not possible since $\pi_A = A + I_3 \implies A = -I_3$ which is not possible. Similar argument for $(x-2)$.

Proposition. $\mathbb{K}[f]$ finite dimensional iff f admits a minimal polynomial.

Proof. $\Longleftarrow$: If f has a minimal polynomial $\pi_f(x)$ of degree $r > 0$, or $\pi_f = a_0 id + a_1 f + ... + a_{r-1} f^{r-1} + f^r \implies \pi_f(f) = 0 = a_{-} id + a_1 f + ... + f^r \implies f^r$ can be written as a non-trivial linear combination of $(id, f, ..., f^{r-1})$. Similarly, $f^k \,\exists\, k \geq r$ can also bee written as a linear combination of previous terms $\implies (id, f, ..., f^{r-1})$ generates $\mathbb{K}[f]$. This is a basis. Indeed, if it is not lienarly independent, $\exists\, \lambda_0, ..., \lambda_{r-1} \in \mathbb{K}$ not all 0 such that $\lambda_0 id + ... + \lambda_{r-1} f^{r-1} = 0 \in I_f$ with degree $< r$, so this is *impossible*.

$\implies$: By contrapositive, if f doesn't have a minimal polynomial $\iff I_f = \{0\}$. Evaluation linear map has $\mathbb{K}[x] \underset{\varphi_f}{\longrightarrow} \mathbb{K}[f], P(x) \mapsto P(f) \implies I_f = \ker \varphi_f = \{0\} \implies \varphi_f$ injective $\implies \mathbb{K}(f)$ infinite dimensional. ☐

Corollary. E finite dimensional $\implies$ f has a minimum polynomial.

Indeed, $\dim E = n \implies \dim \underbrace{End(E)}_{\supseteq \mathbb{K}[f]} = n^2 \implies \mathbb{K}[f]$ finite dimensional, then use theorem.

Proposition. Let $f \in End(E)$

1. If λ is an eigenvalue of f, then with $P \in I_f$, λ is a root of P, where $P(\lambda) = 0$.
2. For minimum polynomial, all roots of π_f in $\mathbb{K}$ are also eigenvalues of f.

Proof. 1: $P(x) = \sum_{k=0}^{p} a_k x^k \implies P(f) = \sum_{k=0}^{p} a_k f^k$. Then λ = eigenvalue $\implies \exists x \neq 0$ such that $f(x) = \lambda x \implies f^k(x) = \lambda^k x$. Apple x to $P(f)$, so

$$P(f)(x) \begin{cases} \sum_{k=0}^{p} a_k \underbrace{f^k(x)}_{\lambda^k x} = P(\lambda) \underbrace{x}_{\neq 0} \\ 0, \text{ because} P \in I_f \end{cases} \implies P(\lambda) = 0$$

2: Suppose λ root of π_f, not an eigenvalue $\iff ker(f - \lambda id) = \{0\} \implies \pi_f(x) = (x-\lambda)Q(x) \implies \pi_f(f) = 0 = (f - \lambda id) \circ Q(f) \implies Q(f) = 0 \implies Q \in I_f$ annihilates f, but $deg(Q) < r$, and this is impossible. So, λ is an eigenvalue of f. ☐

Theorem. **[Kernel Lemma]** Suppose $P \in \mathbb{K}[x]$ splitting as $P = P_1 P_2 ... P_k, P_i \in \mathbb{K}[x], i \neq j, P_i, P_j$ relatively prime. Then,

$$\ker P(f) = \bigoplus_{i=1}^{k} \ker(P_i(f))$$

Proof. For $n = 2$. Finish by induction. Let $P(x) = P_1(x)P_2(x)$ with $P_1(x), P_2(x)$ relatively prime. From number theory Bezout identity, we have P_1, P_2 relatively prime $\implies \exists U_1, U_2 \in \mathbb{K}[x], U_1P_1 + U_2P_2 = 1$. We know that we want $\ker P_1(f) \cap \ker P_2(f) = \{0\}$. Then we can pick $x \in \cap \implies P_1(f)(x) = P_2(f)(x) = 0$. Then, by Bezout $\implies U_1(f) \circ \underbrace{P_1(f)(x)}_{=0} + U_2(f) \circ \underbrace{P_2(f)(x)}_{=0} \implies x = 0$. To prove $\ker P(f) = \ker P_1(f) + \ker P_2(f)$, the $\subseteq$ direction is straightforward. For $\supseteq$, pick $x \in \ker P(f)$. By Bezout, $x = \underbrace{U_1(f) \circ P_1(f)(x)}_{\in \ker P_2(f)} + \underbrace{U_2(f) \circ P_2(f)(x)}_{\in \ker P_1(f)}$. Here,

$$P_2(f)(U_1(f) \circ P_1(f)(x)) = (P_2(f) \circ U_1(f) \circ P_1(f))(x) = (U_1(f) \circ P(f))(x) = 0$$

Similar argument for $\ker P_1(f)$ □

Remark: If $P \in I_f$, then $\ker P(f) = E$.

Consequence Suppose Cayley-Hamilton true, $\chi_f(f) = 0$. Suppose $\chi_f(\lambda) = (\lambda_1 - \lambda)^{m_1} ... (\lambda_p - \lambda)^{m_p}$. By kernel lemma,

$$E = \bigoplus_{i=1}^{p} \ker(f - \lambda id)^{m_i}$$

Corollary. Let $f \in End(E)$ with non-repeated eigenvalues $\lambda_1, ..., \lambda_p$, then

$$\sum_{i=1}^{p} E_{\lambda_i}(f) = \bigoplus_{i=1}^{p} E_{\lambda_i}(f)$$

with $E_{\lambda_i}(f) = \ker(f - \lambda_i id) = \lambda_i$-eigenspace

Proof. Apply kernel lemma to $P(\lambda) = (\lambda_1 - \lambda)...(\lambda_p - \lambda)$ □

Definition. f **invariant subspace**. We say a subspace F is f invariant if $f(F) \subseteq F \implies f|_F \in End(F)$.

Exercise: $f, g \in End(E)$ communting (i.e. $f \circ g = g \circ f$). Then $\ker(g)$ and $im(g)$ are f-invariant.

In particular, eigenspace $E_\lambda(f) = \ker(f - \lambda id)^k$ are f-invariant. Generalized eigenspaces has $\ker(f - \lambda id)^k = E_\lambda^k(f)$

Proposition. $f \in End(E)$. Suppose F be f-invariant subspace of E. Pick F' be component of F in E ($F \oplus F' = E$), $\mathcal{B} = \mathcal{B}_F \cup \mathcal{B}_{F'}$ = basis adapted to direct sum. Recall that $f_{|F} \in End(F)$. Then $[f]_\mathcal{B} =$

Proposition. If F is $f-$invariant, then

1. In some basis $\mathcal{B}$ of E where $\mathcal{B} = \mathcal{B}_F \cap \mathcal{B}'$.

$$[f]_\mathcal{B} = [...]$$

2. $\chi_{f|F}$ divides χ_f

3. $\pi_{f|F}$ divides π_f. We see that π_f annihilates $\pi_{f|F}$ and use $f^k|_F = f|_F^k$ by f-invariance.

Back to Diagonalization $f \in End(E)$ diagonalizable $\iff \exists \mathcal{B}$ basis of E made of eigenvectors of f. In this basis, $[f]_\mathcal{B}$ is diagonal, so with $\lambda_1, ..., \lambda_p$ eigenvalues without repetition, $\iff E = \oplus_{i=1}^p E_{\lambda_i}(f)$ where $\ker(f - \lambda_i id) = \lambda_i$ subspace.

Example. Let

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}, \chi_A(\lambda) = (2-\lambda)(1+\lambda)^2, m_A(\lambda) = (2-\lambda)(1+\lambda)$$

We have $\dim E_2(A) = 1, \dim E_{-1}(A) = 2$. With $f(x) = A(x), f|_{E_2(A)} = 2id|_{E_2(A)}$ and $f_{E_{-1}(A)} = -id|_{E_{-1}(A)}$. Together, $\mathbb{R}^3 = E_{-2}(A) \oplus E_1(A)$.

Theorem. $f \in End(E)$, $\dim E < \infty$, with $\lambda_1, ..., \lambda_p$ are eigenvalues without repetition.

1. f diagonalizable $\iff \underbrace{m_{geo}(\lambda_i)}_{\dim E_{\lambda_i}(f)} = m_{alg}(\lambda_i)$
2. f diagonalizable $\iff m_f$ is split with simple roots.

Proof. Part 1: $\impliedby$: Suppose $m_{geo} = m_{alg}$ $\forall$ eigenvalues. By the kernel lemma,

$$\sum_i E_{\lambda_i}(f) = \bigoplus_{i=1}^{p} E_{\lambda_i(f)} \subseteq E$$

Since $m_{geo} = m_{alg}$, $\dim \sum_i m_{geo}(\lambda_i) = n$ and $\dim E = n$. So, $\bigoplus_i E_{\lambda_i}(f)$ subspace of E with same dimension, so they are equal.

"" $\implies$ ": Lemma: If λ_i eigenvalue of f, then $1 \leq m_{geo}(\lambda_i) \leq m_{alg}(\lambda_i)$.

Proof of lemma: $f - \lambda_i id|_{E_{\lambda_i}(f)} = (f - \lambda_i id)|_{\ker((f-\lambda_i id))} = 0 \implies f|_{E_{\lambda_i}(f)} = \lambda_i id|_{E_{\lambda_i}(f)}$. So in basis $\mathcal{B}_i$ of $E_{\lambda_i}(f)$,

$$\left[f|_{E_{\lambda_i}(f)}\right]_{\mathcal{B}_i} \text{ with } \lambda_i \text{ on the diagonals} \implies \chi_{f|_{E_{\lambda_i}(f)}}(\lambda) = (\lambda_i - \lambda)^{m_{geo}(\lambda_i)}$$

But, $\chi_{f|_{E_{\lambda_i}(f)}}$ divides $\chi_f(\lambda) = \prod_{k=1}^{p}(\lambda_k - \lambda)^{m_{alg}(\lambda_i)} \implies m_{geo}(\lambda_i \leq m_{alg}(\lambda_i))$ ■

Suppose f diagonal $\iff E = \bigoplus_{i=1}^{p} E_{\lambda_i}(f)$ has dimensional $\dim n = \sum_{i=1}^{p} m_{alg}(\lambda_i)$ and $m_{geo}(\lambda_i) \leq m_{alg}(\lambda_i)$. So, this $\leq$ must be an $=$. □

Proof. "$\impliedby$": If $m_f(\lambda) = (\lambda - \alpha_1)...(\lambda - \alpha_k)$ with $\alpha_i \neq \alpha_j \forall i, j$. Then by the kernel lemma, $m_f(f) = 0 \implies \ker(m_f(f)) = E$ and $E = \bigoplus_{i=1}^{k} \ker(f - \alpha_i id)$, so E splits into eigenspaces $\implies E$ diagonalizable.

" $\implies$ " Suppose f diagonalizable $\iff E = \bigoplus_{i=1}^{p} \ker(f - \lambda_i id)$. Set $P(\lambda) = \prod_{i=1}^{p}(\lambda - \lambda_i)$. Check P annihilates f, $\forall i \in [1, p]$, pick $x \in \ker(f - \lambda_k id)$. $P(f)(x) = (\prod_{i=1}^{p}(f - \lambda_i id))(x)$. Here, $(f - \lambda_k id)(x) = 0$, so we have the previous expression equal $\prod_{i=1}^{p}(f - \lambda_i id) \circ (f - \lambda_k id)(x) = 0 \implies P((x) = 0$. □

Introduction to Jordan-type Reduction Let $A \in M_n(\mathbb{K})$. What do we do when A is not diagonalizable?

Theorem. Suppose $\chi_A(\lambda) = (\lambda_1 - \lambda)^{m_1}...(\lambda_p - \lambda)^{m_p}$. $f(x) = Ax$ with $f \in End(\mathbb{R}^n)$.