MA-GY 7043: Linear Algebra II

Deane Yang

Courant Institute of Mathematical Sciences
New York University

January 23 2025

Outline I

Course Requirements: Assignments

- All homework assignments and exams will be handled using Gradescope
- Homework
  - Every one or two weeks
  - Provided as Overleaf project and Gradescope assignment
  - Solutions must be typed up using LaTeX
  - Submissions uploaded as PDF to Gradescope
- Midterm and Final
  - In person
  - 150 minutes
  - Graded exams uploaded to Gradescope

Course Requirements: Grading Policy

- Course grade
  - Homework: 20%
  - Midterm: 30%
  - Final: 50%
  - Tweaks
- Homework and Exams
  - Partial credit for correct and relevant logical reasoning
  - Full credit for correct and relevant logical reasoning and correct answer
  - No credit for correct answer but incorrect logical reasoning
  - Incorrect logic and calculations wil be severely penalized

Course Information

- Web Pages
  - My homepage: https://math.nyu.edu/∼yangd
  - Course Homepage
  - Course Calendar
- Textbook
  - Yisong Yang, **A Concise Text on Advanced Linear Algebra**, Cambridge University Press
  - PDF available in Ed Discussion Resources

Prerequisites: Mathematical Grammar

- Always write in complete English or mathematical sentences
- A sentence must have a subject and verb
- A mathematical sentence usually contains an object
- Sample sentences
  - *(subject)* **equals** *(object)*
    - *(subject)* $=$ *(object)*
  - *(subject)* **is less than** *(object)*
    - *(subject)* $<$ *(object)*
  - **If** *(sentence)*, **then** *(sentence)*
    - *(assumption)* $\implies$ *(consequence)*
  - **There exists** *(object)* **such that** *(sentence)*
    - $\exists$ *(object)*, *(mathematical sentence about object)*
  - **For any** *object*, **(sentence)**
    - $\forall$ *(object)*, *(mathematical sentence about object)*

Prerequisites: Basic Deductive Logic

- You are expected to know how to use deductive logic
- Suppose A and B are English or mathematical sentences
- You are expected to know the meaning of the following phrases:

$$A \text{ and } B$$
$$A \text{ or } B$$
$$A \text{ is false}$$
$$A \implies B$$
$$A \iff B$$

Prerequisites: Converse and Contrapositive

- The **converse** of the sentence

  $$A \implies B$$

  is

  $$B \implies A$$

  These two are **not** equivalent
- The **contrapositive** of the sentence

  $$A \implies B$$

  is

  $$(B \text{ is false}) \implies (\ \text{is false})$$

  These two sentences are equivalent

Prerequisites: Quantifiers

- Sentence holds for all objects

  For each *(object)*, *(sentence)*,

  i.e.,

  $$\forall \textit{(object), (sentence)}$$

- Sentence holds for at least one object

  There exists *(object)*, such that *(sentence)*,

  i.e.,

  $$\exists \textit{(object), (sentence)}$$

Prerequisites: Nested Quantifiers

- The sentence

  $$\forall \text{ (object1)}, \exists \text{(object2) such that (sentence)},$$

  is **not** equivalent to

  $$\exists \text{ (object2)}, \forall \text{(object1) such that (sentence)},$$

Prerequisites: Negations

- The negation of
  - *A is true* **and** *B is true*

  is
  - *A is false* **or** *B is false*
- The negation of
  - *A is true* **or** *B is true*

  is
  - *A is false* **and** *B is false*
- The negation of
  - **If** *A is true*, **then** *B is true*

  is
  - *A is true* **and** *B is false*

Prerequisites: Negations With Quantifiers

- The negation of

$$\forall \textit{(object)}, \textit{(sentence)}$$

is

$$\exists \textit{(object)}, \text{ such that } \textit{(negation of sentence)}$$

- The negation of

$$\exists \textit{(object)} \text{ such that } \textit{(sentence)}$$

is

$$\forall \textit{(object)}, \textit{(negation of sentence)}$$

Prerequisites: Modus Ponens

- All calculations and proofs **must** proceed as follows:
  - Known to be true (by definition, assumption, theorem, or proof)
    - A
    - $A \implies B$
  - True by deduction
    - B

Prerequisites: Definitions Versus Theorems

- **VERY VERY IMPORTANT:** When studying theorems or doing problems, make sure you know the definitions of every word and symbol
- Always try to solve problem (e.g., doing a proof) using **ONLY** definitions
- Use a theorem **ONLY** if absolutely necessary

Prerequisites: Functions and Maps

- We will use the following notation when defining a function or map:

$$\begin{aligned} \textit{function} : \textit{domain} &\rightarrow \textit{codomain} \\ \textit{input} &\mapsto \textit{output} \end{aligned}$$

- When doing calculations and proofs, It is important to keep track of the domain and codomain of a function
- If you make sure that each input to a function really is an element of the domain and each output really is treated as an element of the codomain, youu will catch 90% of your errors

Abstract Vector Space

- Let $\mathbb{F}$ be either the reals (denoted $\mathbb{R}$) or the complex numbers (denoted $\mathbb{C}$)
- A vector space over $\mathbb{F}$ is a set V with the following:
  - A special element called the **zero vector**, which we will write as $\vec{0}$, 0_V, or simply 0
  - An operation called **vector addition**:

$$\begin{aligned} V \times V &\to V \\ (v_1, v_2) &\mapsto v_1 + v_2 \end{aligned}$$

  - An operation called **scalar multiplication**:

$$\begin{aligned} V \times \mathbb{F} &\to V \\ (v, r) &\mapsto rv = vr \end{aligned}$$

- The zero vector, vector addition, and scalar multiplication must satisfy fundamental properties that are listed below

Properties of Vector Addition

- Associativity

$$(v_1 + v_2) + v_3 = v_1 + (v_2 + v_2)$$

- Commutativity

$$v_1 + v_2 = v_2 + v_1$$

- Identity element:

$$v + \vec{0} = v$$

- Inverse element: For each $v \in V$, there exists an element, denoted $-v$, such that

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

Scalar Multiplication

- Properties
  - Associativity

$$(f_1 f_2)v = f_1(f_2 v)$$

  - Distributivity

$$(f_1 + f_2)v = f_1 v + f_2 v$$
$$f(v_1 + v_2) = fv_1 + fv_2$$

  - Identity element

$$1v = v$$

Consequences

$$\vec{0}v = v$$
$$(-1)v = -v$$

Valid and Invalid Expressions

- Valid expressions

$$
\begin{aligned}
(\text{vector}) + (\text{vector})\\
(\text{scalar}) + (\text{scalar})\\
(\text{scalar})(\text{vector})\\
(\text{vector})(\text{scalar})\\
(\text{scalar})(\text{scalar})
\end{aligned}
$$

- Invalid expressions

$$
\begin{aligned}
(\text{vector}) + (\text{scalar})\\
(\text{scalar}) + (\text{vector})\\
(\text{vector})(\text{vector})
\end{aligned}
$$

Linear Combination of Vectors

- Given a finite set of vectors $v_1, \ldots, v_m \in V$ and scalars $f^1, \ldots, f^m$, the vector

$$f^1 v_1 + \cdots + f^m v_m$$

is called a **linear combination** of $v_1, \ldots, v_m$
- Given a subset $S \subset V$, not necessarily finite, the **span** of S is the set of all possible linear combinations of vectors in S

$$[S] = \{f^1 v_1 + \cdots + f^m v_m \ : \ \forall\ f^1, \ldots, f^m \in \mathbb{F} \text{ and } v_1, \ldots, v_m \in S\}$$

- A vector space V is called **finite dimensional** if there is a finite set S of vectors such that

$$[S] = V$$

Such a set S is called by some a spanning system, generating system, or complete system

Basis of a Vector Space

- A set $\{v_1, \ldots, v_k\} \subset V$ is **linearly independent** if

$$f^1 v_1 + \cdots f^m v_m = \vec{0} \implies f^1 = \cdots = f^m = 0, \qquad (1)$$

- A finite set $S = (v_1, \ldots, v_m) \subset V$ is called a **basis** of V if it is linearly independent and

$$[S] = V$$

- For such a basis, if $v \in V$, then there exist a unique set of scalar coefficients $(a^1, \ldots, a^m)$ such that

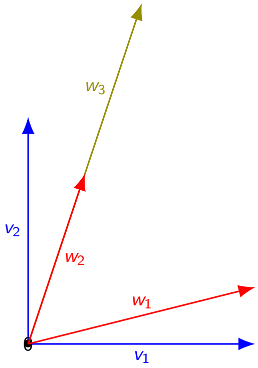
$$v = a^k v_k$$

- In other words, the map

$$\mathbb{F}^m \to V$$
$$\langle f^1, \ldots, f^m \rangle \mapsto f^1 v_1 + \cdots + f^m v_m$$

is bijective

Examples of Bases



- $\{v_1, v_2\}$ is a basis
- $\{w_1, w_2\}$ is a basis
- $\{w_1, w_3\}$ is a basis
- $\{w_2, w_3\}$ is NOT a basis

Every Finite Dimensional Vector Space Has a Basis

- Assume that T is a finite dimensional vector space
- There exists a finite set $S = \{s_1, \dots, s_p\}$ that spans T
- If S is linearly independent, then S is a basis
- If not, then there exists $f^1, \dots, f^p \in \mathbb{F}$, not all zero, such that

$$f^1 s_1 + \cdots f^p s_p = \vec{0}$$

- If $f^p \neq 0$, then

$$s_p = \frac{f^1}{f^p} s_1 + \cdots + \frac{f^{p-1}}{f^p} s_{p-1}$$

- It follows that $S' = \{s_1, \dots, s_{p-1}\}$ spans T
- If S' is not a basis, then repeat previous steps
- After a finite number of steps, you get either a basis or $S = \{\vec{0}\}$

Triangular Change of Basis

- Let $E = (e_1, \ldots, e_m)$ be a basis of V
- A subset $F = (f_1, \ldots, f_m)$ is **triangular** with respect to E if

$$
\begin{aligned}
f_1 &= e_1 + e_2 M_1^2 + \cdots + e_m M_1^m \\
f_2 &= e_2 + e_3 M_2^3 + \cdots + e_m M_2^m \\
\vdots &\quad \vdots \\
f_k &= e_k + e_{k+1} M_k^{k+1} + \cdots e_m M_k^m \\
\vdots &\quad \vdots \\
f_m &= e_m
\end{aligned}
$$

- Observe that for each $1 \leq k \leq m$, $\{f_1, \ldots, f_k)\}$ is linearly independent and

$$[f_1, \ldots, f_k] = [e_1, \ldots, e_k]$$

- It follows that E is a basis of V if and only if F is a basis of V

Existence of Triangular Change of Basis (Part 1)

- Let $E(e_1, \ldots, e_m)$ be a basis of V
- Let $F = (f_1, \ldots, f_n)$ be a basis of V, where for each $1 \leq k \leq n$,

$$f_k = e_1 M_k^1 + \cdots + e_m M_k^m$$

- Rearranging and rescaling the basis vectors $e_1, \ldots, e_m$, we can assume that $M_1^1 = 1$, i.e.,

$$f_1 = e_1 + M_1^2 e_2 + \cdots + M_1^m e_m$$

- Suppose for each $1 \leq j \leq k$,

$$f_j = e_j + e_{j+1} M_j^{j+1} + \cdots + e_m M_j^m$$

and

$$f_{k+1} = e_1 M_{k+1}^1 + \cdots + e_m M_{k+1}^m$$

Existence of Triangular Change of Basis (Part 2)

- If $f_{k+1} \notin [e_1, \ldots, e_k]$, then

$$\begin{aligned}\hat{f}_{k+1} &= f_{k+1} - (e_1 M^1_{k+1} + \cdots e_k M^k_{k+1}) \\ &= e_{k+1} M^{k+1}_{k+1} + \cdots + e_m M^m_{k+1} \notin [e_1, \ldots, e_k]\end{aligned}$$

- Rearranging and rescaling $e_{k+1}, \ldots, e_m$, we can assume

$$f_{k+1} = e_{k+1} + e_{j+2} M^{j+2}_{k+1} + \cdots + e_m M^m_{k+1}$$

- Observe that for each $1 \leq k \leq m$, $\{f_1, \ldots, f_k\}$ is linearly independent and

$$[f_1, \ldots, f_k] = [e_1, \ldots, e_k]$$

- It follows that $m = n$

Dimension of a Vector Space

- Every basis of a finite dimensional vector space V has the same number of elements
- The **dimension** of a finite dimensional vector space V to be the number of elements in a basis
- The dimension of V is denoted $\dim V$

Product of Row Matrix and Column Matrix

- A row matrix looks like this:

$$R = (r_1, \ldots, r_m) = \begin{bmatrix} r_1 & \cdots & r_m \end{bmatrix}$$

- A column matrix looks like this:

$$C = \langle c^1, \ldots, c^m \rangle = \begin{bmatrix} c^1 \\ \vdots \\ c^m \end{bmatrix}$$

- The matrix product of R and C is the 1-by-1 matrix

$$RC = \begin{bmatrix} r_1 & \cdots & r_m \end{bmatrix} \begin{bmatrix} c^1 \\ \vdots \\ c^m \end{bmatrix} = r_1 c^1 + \cdots + r_m c^m$$

Generalized Matrix Products

- This notation id valid if
  -
    - Each r_i is a scalar
    - Each c^j is a scalar
    - And therefore RC is a scalar
  -
    - Each r_i is a scalar
    - Each c^j is a vector
    - And therefore RC is a vector
  -
    - Each r_i is a vector
    - Each c^j is a scalar
    - And therefore RC is a vector
- The notation is invalid if
  - Each r_i is a vector
  - Each c^j is a vector
- Order matters: $CR \neq RC$!
- We will use only items 1 and 3 above

Product of Column Matrix and Row Matrix

- Consider a column matrix

$$C = \begin{bmatrix} c^1 \\ \vdots \\ c^n \end{bmatrix}$$

and ra ow matrix

$$R = \begin{bmatrix} r_1 & \cdots & r_m \end{bmatrix}$$

- The matrix product of C and R looks like this

$$CR = \begin{bmatrix} c^1 \\ \vdots \\ c^n \end{bmatrix} \begin{bmatrix} r_1 & \cdots & r_m \end{bmatrix} = \begin{bmatrix} c^1 r_1 & \cdots & c^1 r_m \\ \vdots & & \vdots \\ c^n r_1 & \cdots & c^n r_m \end{bmatrix}$$

Product of Two Matrices

- The matrix product of the matrices

$$M = \begin{bmatrix} M_1^1 & \cdots & M_k^1 \\ \vdots & & \vdots \\ M_1^m & \cdots & M_k^m \end{bmatrix} = \begin{bmatrix} R^1 \\ \vdots \\ R^m \end{bmatrix}$$

$$N = \begin{bmatrix} N_1^1 & \cdots & M_n^1 \\ \vdots & & \vdots \\ N_1^k & \cdots & N_m^n \end{bmatrix} = \begin{bmatrix} C_1 & \cdots & C_n \end{bmatrix}$$

is the m-by-n matrix

$$MN = \begin{bmatrix} R^1 C_1 & \cdots & R^1 C_n \\ \vdots & & \vdots \\ R^m C_1 & \cdots & R^m C_n \end{bmatrix}$$

- This formula can be used if
  - Components of both M and N are scalars
  - Components of M are scalars, components of N are vectors
  - Components of M are vectors, components of N are scalars

Abstract Matrix Notation for Vector With Respect to Basis

- A basis $(f_1, \ldots, f_m)$ of a vector space V will always be written as a row matrix of vectors,

$$F = \begin{bmatrix} f_1 & \cdots & f_m \end{bmatrix}$$

- Any vector is a unique linear combination of the basis vectors

$$v = f_1 b^1 + \cdots + f_m b^m \in V$$

- This can be written as the matrix product of the basis written as a row matrix and the coefficients written as a column matrix

$$v = f_1 b^1 + \cdots + f_m b^m = \begin{bmatrix} f_1 & \cdots & f_m \end{bmatrix} \begin{bmatrix} b^1 \\ \vdots \\ b^m \end{bmatrix} = Fb,$$

Standard Basis of $\mathbb{F}^3$

- Denote the standard basis vectors of $\mathbb{F}^3$ by

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \; e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \; e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

- The basis can be written as a row matrix of column vectors:

$$E = \begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix} = \left[\begin{array}{c|c|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right] = I$$

Change of Basis Example

- Consider a basis

$$F = \begin{bmatrix} f_1 & f_2 & f_3 \end{bmatrix} = \left[\begin{array}{c|c|c} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 1 & 1 \end{array}\right]$$

- Given a vector $v = (1, 2, 3)$, there are coefficients b^1, b^2, b^3 such that

$$\begin{aligned} (1, 2, 3) &= b^1(1, -1, 1) + b^2(0, 1, 1) + b^3(0, 0, 1) \\ &= (b^1, -b^1 + b^2, b^1 + b^3 + b^3) \end{aligned}$$

or, equivalently,

$$\begin{aligned} b^1 &= 1 \\ -b^1 + b^2 &= 2 \\ b^1 + b^2 + b^3 &= 3 \end{aligned}$$

- Unique solution is $(b^1, b^2, b^3) = (1, 3, -1)$

Change of Basis

- Consider two different bases of an n-dimensional vector space V,

$$E = \begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix} \text{ and } F = \begin{bmatrix} f_1 & \cdots & f_n \end{bmatrix}$$

- Since F is a basis, we can write each vector in F as a linear combination of the vectors in E

$$\begin{aligned} F &= \begin{bmatrix} f_1 & \cdots & f_n \end{bmatrix} \\ &= \begin{bmatrix} e_1 M_1^1 + \cdots + e_n M_1^n & \cdots & e_1 M_n^1 + \cdots + e_n M_n^n \end{bmatrix} \\ &= \begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix} \begin{bmatrix} M_1^1 & \cdots & M_n^1 \\ \vdots & & \vdots \\ M_1^n & \cdots & M_n^n \end{bmatrix} \\ &= EM \end{aligned}$$

Change of Coefficients

- Any vector v can be written as either a linear combination of the basis E,

$$v = e_1 a^1 + \cdots + e_n a^n = \begin{bmatrix} e_1 & \cdots & e_n \end{bmatrix} \begin{bmatrix} a^1 \\ \vdots \\ a^n \end{bmatrix} = Ea$$

or as a linear combination of the basis F,

$$v = f_1 b^1 + \cdots + f_n b^n = \begin{bmatrix} f_1 & \cdots & f_n \end{bmatrix} \begin{bmatrix} b^1 \\ \vdots \\ b^n \end{bmatrix} = Fb$$

- If $F = EM$, then

$$v = Fb = E(Mb) = Ea$$

- Therefore,

$$a = Mb \text{ and } b = M^{-1}a$$

Change of Basis Formula

- Let E and F be bases of V such that

$$F = EM,$$

- If $v = Ea = Fb$, then

$$a = Mb \text{ and } b = M^{-1}a$$

- The matrix that transforms old coefficients into new coefficients is the inverse of the matrix that transforms the old basis into the new basis
- **WARNING:** This works only if you write a basis as a row matrix of vectors and the coefficients as a column matrix of scalars

Linear Functions

- If V is a vector space, then a function

$$\ell : V \to \mathbb{F}$$

is **linear**, if for any $v_1, v_2 \in V$

$$\ell(v_1 + v_2) = \ell(v_1) + \ell(v_2)$$

and for any $v \in V$ and $s \in \mathbb{F}$,

$$\ell(vs) = \ell(v)s$$

- Consequences:

$$\ell(0_V) = 0$$
$$\ell(-v) = -\ell(v)$$

Linear Maps

- If V and W are vector spaces, then

$$L : V \to W$$

is a **linear map** or **linear transformation**, if for any $v, v_1, v_2 \in V$ and $s \in \mathbb{F}$,

$$L(v_1 + v_2) = L(v_1) + L(v_2)$$
$$L(sv) = sL(v)$$

- Consequences:

$$L(0_V) = 0_W$$
$$L(-v) = -L(v)$$

Properties of Linear Maps

- If $K : U \to V$ and $L : V \to W$ are linear maps, then so is

$$L \circ K : U \to W$$

- If $L : V \to W$ is bijective, it is called a **linear isomorphism**
- If $L : V \to W$ is a linear isomorphism, then so is

$$L^{-1} : W \to V$$

n-Dimensional Vector Spaces are Isomorphic

- Let $\dim V = \dim W = m$
- Let $E = (e_1, \dots, e_m)$ be a basis of V
- Let $F = (f_1, \dots, f_m)$ be a basis of W
- The map

$$\begin{aligned} L_{E,F} &: V \to W \\ e_1 a^1 + \cdots + e_m a^m &\mapsto f_1 a^1 + \cdots + f_m a^m \end{aligned}$$

  is a linear isomorphism
- Given any basis $(e_1, \dots, e_m)$ of V, there is a linear isomorphism

$$\begin{aligned} L_V &: \mathbb{F}^m \to V \\ (a^1, \dots, a^m) &\mapsto e_1 a^1 + \cdots + e_m a^m \end{aligned}$$

Vector Space of Linear Maps

- Given vector spaces V and W, let

$$\mathcal{L}(V, W) = \{L : V \to W \ : \ L \text{ is linear}\}$$

- $\mathcal{L}(V, W)$ is itself a vector space, because
  - If $A, B \in \mathcal{L}(V, W)$ and $s \in \mathbb{F}$, then

$$A + B,\ sA \in \mathcal{L}(V, W)$$

- Let $\text{gl}(n, m, \mathbb{F})$ denote the vector space of n-by-m matrices with components in $\mathbb{F}$
  - $\dim \text{gl}(n, m, \mathbb{F}) = nm$
- Let $\text{gl}(n, \mathbb{F}) = \text{gl}(n, n, \mathbb{F})$
- Let $\text{gl}(n) = \text{gl}(n, \mathbb{R})$

Matrix as Linear Map

- Let $E = (e_1, \ldots, e_m)$ be a basis of V
- Let $F = (f_1, \ldots, f_n)$ be a basis of W
- For each $M \in \mathrm{gl}(n, m, \mathbb{F})$, let $L : V \to W$ be the linear map where

$$\forall\ 1 \leq k \leq m,\ \ L(e_k) = f_1 M_k^1 + \cdots + f_n M_k^n$$

and therefore for any $v = e_1 a^1 + \cdots e_m a^m = Ea$,

$$\begin{aligned} L(v) &= L(e_1 a^1 + \cdots + e_m a^m) \\ &= L(e_1)a^1 + \cdots + L(e_m)a^m \\ &= (f_1 M_1^1 + \cdots + f_n M_1^n)a^1 + \cdots + (f_1 M_m^1 + \cdots + f_n M_m^n)a^m \\ &= f_1(M_1^1 a^1 + \cdots + M_m^1 a^m) + \cdots f_n(M_1^n a^1 + \cdots + M_m^n a^m) \\ &= f_1(Ma)^1 + \cdots + f_n(Ma)^n \end{aligned}$$

- This defines a map $I_{E,F} : \mathrm{gl}(n, m, \mathbb{F}) \to \mathcal{L}(V, W)$

Linear Map as Matrix

- Let $E = (e_1, \ldots, e_m)$ be a basis of V
- Let $F = (f_1, \ldots, f_n)$ be a basis of W
- Let $L : V \to W$ be a linear map
- For each e_k, $1 \leq k \leq m$, there exists $(M_k^1, \ldots, M_k^n) \in \mathbb{F}^n$ such that
$$L(e_k) = f_1 M_k^1 + \cdots f_n M_k^n$$
- Therefore, for any $v = e_1 a^1 + \cdots + e_m a^m \in V$,
$$\begin{aligned} L(v) &= L(e_1 a^1 + \cdots + e_m a^m) \\ &= L(e_1) a^1 + \cdots + L(e_m) e^m \\ &= (f_1 M_1^1 + \cdots f_n M_1^n) a^1 + \cdots + (f_1 M_m^1 + \cdots + f_n M_m^n) a^m \\ &= f_1 (M_1^1 a^1 + \cdots M_m^1 a^m) + \cdots + f_n (M_1^n a^1 + \cdots + M_m^n a^m) \\ &= f_1 (Ma)^1 + \cdots + f_n (Ma)^n \end{aligned}$$
- This defines a map $J_{E,F} : \mathcal{L}(V, W) \to \mathrm{gl}(n, m, \mathbb{F})$
- $J_{E,F} = I_{E,F}^{-1}$ and $I_{E,F} = J_{E,F}^{-1}$
- Therefore, $\dim \mathcal{L}(V, W) = \dim \mathrm{gl}(n, m, \mathbb{F}) = nm$

Concrete to Abstract Notation

$$L(v) = L(e_1 a^1 + \cdots + e_m a^m) = L\left(\begin{bmatrix} e_1 & \cdots & e_m \end{bmatrix} \begin{bmatrix} a^1 \\ \vdots \\ a^m \end{bmatrix}\right)$$

$$= L\left(\begin{bmatrix} e_1 & \cdots & e_m \end{bmatrix}\right) \begin{bmatrix} a^1 \\ \vdots \\ a^m \end{bmatrix} = \begin{bmatrix} L(e_1) & \cdots & L(e_m) \end{bmatrix} \begin{bmatrix} a^1 \\ \vdots \\ a^m \end{bmatrix}$$

$$= \begin{bmatrix} f_1 M_1^1 + \cdots + f_n M_1^n & \cdots & f_1 M_n^1 + \cdots + f_n M_n^n \end{bmatrix} \begin{bmatrix} a^1 \\ \vdots \\ a^m \end{bmatrix}$$

$$= \begin{bmatrix} f_1 & \cdots & f_n \end{bmatrix} \begin{bmatrix} M_1^1 & \cdots & M_m^1 \\ \vdots & & \vdots \\ M_1^n & \cdots & M_m^n \end{bmatrix} \begin{bmatrix} a^1 \\ \vdots \\ a^m \end{bmatrix} = FMa$$

Subspace and its Dimension

- A subset T of a vector space X is a **subspace** of X if for any $p, q \in \mathbb{F}$ and $a, b \in T$,

$$pa + qb \in T$$

- If a subspace has at least one nonzero vector, then it is itself a vector space
- Define the dimension of a subspace S as follows:
  - If $S = \{\vec{0}\}$ then $\dim S = 0$
  - If $S \neq \{\vec{0}\}$, then S is a vector space and $\dim S$ is its dimension as a vector space

Kernel, Image, Rank of a Linear Map

- Consider any linear map $P : Z \to Y$
- The **kernel** of P is defined to be

$$\ker P = \{z \in Z \ : \ P(z) = \vec{0}\}$$

  - $\ker(P)$ is a subspace of Z
- The **image** of P is defined to be

$$P(Z) = \{P(z) \ : \ z \in Z\} \subset Y$$

  - $P(Z)$ is a subspace of Y
- The **rank** of P is

$$\text{rank}(P) = \dim P(Z)$$

Example 0

- Define $Z : \mathbb{F}^2 \to \mathbb{F}^3$ to be

$$Z(x, y) = (x, y, 0), \text{ for all } (x, y) \in \mathbb{F}^2$$

- In other words,

$$Z\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

- $\ker Z = \{0\}$
- $Z(\mathbb{F}^2) = \{(x, y, 0) \ : x, y, \in \mathbb{F}\} \subset \mathbb{F}^n$
  - A basis of $Z(\mathbb{F}^2)$ is $\{Z(e_1), Z(e_2)\} = \{(1, 0, 0), (0, 1, 0)\}$
- Therefore,

$$\dim \ker Z = 0$$
$$\operatorname{rank} Z = 2$$

Example 1

- Define $W : \mathbb{F}^2 \to \mathbb{F}^3$ to be

$$W(x, y) = (y, 0, 0), \text{ for all } (x, y) \in \mathbb{F}^2$$

- In other words,

$$W\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

- $\ker W = \{(x, 0) \; : \; x \in \mathbb{F}\}$
  - A basis of $\ker W$ is $\{(1, 0)\}$
- $W(\mathbb{F}^2) = \{(y, 0, 0) \; : \; y \in \mathbb{F}\}$
  - A basis of $W(\mathbb{F}^2)$ is $\{(1, 0, 0)\}$
- Therefore,

$$\dim \ker W = 1$$
$$\operatorname{rank} W = 1$$

Example 2

- Define $U : \mathbb{F}^2 \to \mathbb{F}^3$ to be

$$U(x, y) = (0, 0, 0), \text{ for all } (x, y) \in \mathbb{F}^2$$

- In other words,

$$U\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

- $\ker U = \mathbb{F}^2$
- $U(\mathbb{F}^2) = \{(0, 0, 0)\}$
- Therefore,

$$\dim \ker U = 2$$
$$\operatorname{rank} U = 0$$

Example 3

- Define $U : \mathbb{F}^3 \to \mathbb{F}^2$ to be

$$U(x, y, z) = (y, z), \text{ for all } (x, y, z) \in \mathbb{F}^3$$

- In other words,

$$U\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

- $\ker U = \{(x, 0, 0) \ : \ z \in \mathbb{F}\}$
  - A basis is $\{(1, 0, 0)\}$
- $U(\mathbb{F}^3) = \mathbb{F}^2$
- Therefore,

$$\dim \ker U = 1$$
$$\operatorname{rank} U = 2$$

Example 4

- Define $U : \mathbb{F}^3 \to \mathbb{F}^2$ to be

$$U(x, y, z) = (z, 0), \text{ for all } (x, y, z) \in \mathbb{F}^3$$

- In other words,

$$U\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

- $\ker U = \{(x, y, 0) \ : \ x, y \in \mathbb{F}\}$
  - A basis is $\{(1, 0, 0), (0, 1, 0)\}$
- $U(\mathbb{F}^2) = \{(z, 0) \ : \ z \in \mathbb{F}\}$
  - A basis is $\{(1, 0)\}$
- Therefore,

$$\dim \ker U = 2$$
$$\operatorname{rank} U = 1$$

Example 5

- Define $U : \mathbb{F}^3 \to \mathbb{F}^2$ to be

$$T(x, y, z) = (0, 0, 0), \text{ for all } (x, y, z) \in \mathbb{F}^3$$

- In other words,

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

- $\ker U = \mathbb{F}^3$
- $U(\mathbb{F}^3) = \{(0, 0, 0)\}$
- Therefore,

$$\dim \ker U = 3$$
$$\operatorname{rank} U = 0$$

Bases of V and W Induce Basis of $\mathcal{L}(V, W)$

- If $(e_1, \dots, e_m)$ is a basis of V and $(f_1, \dots, f_n)$ is a basis of W, then for each $1 \leq k \leq m$ and $1 \leq p \leq n$, let

$$L_k^p : V \to W$$

be the linear map where

$$L_p^k(e_j) = \begin{cases} f_p & \text{if } j = k \\ 0 & \text{otherwise} \end{cases}$$

and let $E_k^p \in \mathrm{gl}(n, m)$ be the matrix that has a 1 in the p-th row and k-th column and 0 everywhere else

- The set $\{L_p^k \ : \ 1 \leq k \leq m \text{ and } 1 \leq p \leq n\}$ is a basis of $\mathcal{L}(V, W)$ such that

$$I_{V,W}(E_k^p) = M_k^p$$

Normal Form of a Linear Map

- Let $L : V \to W$ be a linear map
- Lemma: There exists a basis $(e_1, \dots, e_m)$ of V and a basis $(f_1, \dots, f_n)$ of W such that for each $1 \le k \le m$,

$$L(e_k) = \begin{cases} f_k & \text{if } 1 \le k \le r \\ 0_W & \text{if } r+1 \le k \le m \end{cases},$$

  where $r = \operatorname{rank}(L)$
- In particular,

$$\ker(L) = \text{ span of } \{e_{r+1}, \dots, e_m\} \text{ and } L(V) = \text{ span of } \{f_1, \dots, f_r\}$$

- The matrix of L with respect to this basis is

$$M = \left[\begin{array}{c|c} I_{r\times r} & 0_{r\times m-r} \\ \hline 0_{n-r,r} & 0_{n-r,m-r} \end{array}\right]$$

Corollary: Rank-Nullity Theorem

- Theorem: $\dim \ker(L) + \operatorname{rank}(L) = \dim V$
- Proof: The normal form shows that if $\dim V = m$ and $\operatorname{rank}(L) = r$, then $\dim \ker(L) = m - r$

Proof of Existence of Normal Form

- Let $s = \dim ker(L)$ and $r = \dim V - \dim \ker(L) = m - s$
- If $s > 0$, there exists a basis of $\ker(L)$, which will be denoted

$$(e_{m-s+1}, \ldots, e_m)$$

- This can be extended to a basis $(e_1, \ldots, e_r, e_{r+1}, \ldots, e_m)$ of V
- For each $1 \leq k \leq r$, let $f_k = L(e_k)$
- $(f_1, \ldots, f_r)$ is linearly independent
- It can be extended to a basis $(f_1, \ldots, f_n)$ of W
- It follows that

$$\begin{aligned} \dim \ker L + \operatorname{rank} L &= \dim \ker L + \dim L(V) \\ &= s + r = m \\ &= \dim V \end{aligned}$$

Injective and Surjective Maps

- Consider a linear map $L : V \to W$
- $\dim \ker L = 0 \iff L$ is injective:

$$\begin{aligned} L(v_1) = L(v_2) &\iff L(v_2) - L(v_1) = 0_W \\ &\iff L(v_2 - v_1) = 0_W \\ &\iff v_2 - v_2 \in \ker L = \{0_V\} \\ &\iff v_2 = v_1 \end{aligned}$$

- $\operatorname{rank} L = \dim W \iff L$ is surjective:

$$\begin{aligned} &\operatorname{rank} L = \dim W \\ \iff &\dim L(V) = \dim W \\ \iff &L(V) = W \end{aligned}$$

Bijective Maps

- A map $L : V \to W$ an **isomorphism** if it is **bijective**, i.e., both injective and surjective
- Therefore,

$$L : V \to W \text{ is bijective } \iff \dim\ker(L) = 0 \text{ and } \operatorname{rank}(L) = \dim W$$

- By the rank-nullity theorem, this holds if and only if

$$\operatorname{rank}(L) = \dim W$$

- Equivalently, L is an isomorphism if and only if

$$\dim V = \dim W \text{ and } \dim\ker L = 0$$

if and only if

$$\dim V = \dim W = \operatorname{rank} L$$

Example (Part 1)

- Consider the map $L : \mathbb{F}^3 \to \mathbb{F}^2$ given by

$$L\left(\begin{bmatrix} v^1 \\ v^2 \\ v^3 \end{bmatrix}\right) = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} v^1 \\ v^2 \\ v^3 \end{bmatrix} = \begin{bmatrix} v^1 + 2v^2 + 3v^3 \\ 4v^3 \end{bmatrix}$$

- $\ker L = \{(v^1, v^2, v^3) \ : \ v^1 + 2v^2 = 0\}$
- A basis of $\ker L$ is $\{(-2, 1, 0)\}$
- A basis of $\mathbb{F}^3$ is $\{(0, 1, 0), (0, 0, 1), (-2, 1, 0)\}$
- A basis of $L(\mathbb{F}^3)$ is

$$\{L(0, 1, 0), L(0, 0, 1)\} = \{(2, 0), (3, 4)\}$$

Example (Part 2)

- If

$$\begin{bmatrix} e_1 & e_2 & e_3 \end{bmatrix} = \left[\begin{array}{c|c|c} 0 & 0 & -2 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{array}\right] \text{ and } \begin{bmatrix} f_1 & f_2 \end{bmatrix} = \left[\begin{array}{c|c} 2 & 3 \\ 0 & 4 \end{array}\right]$$

- Then

$$\begin{bmatrix} L(e_1) & L(e_2) & L(e_3) \end{bmatrix} = \begin{bmatrix} f_1 & f_2 & 0 \end{bmatrix} = \begin{bmatrix} f_1 & f_2 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

- And given any vector $v = e_1 a^1 + e_2 a^2 + e_3 a^3$,

$$L(v) = L(e_1)a^1 + L(e_2)a^2 + L(e_3)a^3 = f_1 a^2 + f_2 a^3 = FMa,$$

where

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

Composition is Matrix Multiplication

- Consider vector spaces U, V, W and linear maps

$$K : U \to V, \; L : V \to W$$

- Let $(e_1, \ldots, e_k)$ be a basis of U
- Let $(f_1, \ldots, f_m)$ be a basis of V
- Let $(g_1, \ldots, g_n)$ be a basis of W
- There is an m-by-k matrix M such that

$$K(e_j) = f_p M^p_j, \; 1 \le j \le k$$

- There is an n-by-m matrix N such that

$$L(f_p) = g_a N^a_p, \; 1 \le p \le m$$

- There is an n-by-k matrix P such that

$$(L \circ K)(e_j) = g_a P^a_j, \; 1 \le j \le k$$

- On the other hand,

$$(L \circ K)(e_j) = L(K(e_j)) = L(f_p M^p_j) = L(f_p) M^p_j = g_a N^a_p M^p_j$$

- Therefore, $P^a_j = N^a_p M^p_j$.