MATH 146 Personal Notes

Marcus Chan

Taught by Ross Willard and Giang Tran

UW Math '25

Chapter 1: Vector Spaces (S1.1)

KEY			
S :	section	P :	proposition
D :	definition	A :	assignment
R :	remark		
E :	example		
T :	theorem		
L :	lemma		
C :	corollary		

Let $\mathbb{F}$ be a field.
Then, we say V is a "vector space" over $\mathbb{F}$ if there exists
① an addition $+: (V\times V)\to V$ by $+(x,y) = x+y$; and
② a scalar multiplication $\times: (\mathbb{F}\times V)\to V$ by $\times(a,x) = ax$;
and the following conditions hold:
① V is an abelian group with respect to addition; (VS 1 = commutativity; 2 = associativity; 3 = identity; 4 = inverse)
② $1_{\mathbb{F}} x = x \quad \forall x\in V$; (VS 5)
③ multiplication is associative; ie $a(bx) = (ab)x \quad \forall a,b\in\mathbb{F},\ x\in V$; and (VS 6)
④ the left and right distributive laws hold; ie $a(x+y) = ax+ay$ and $(a+b)x = ax+bx \quad \forall a,b\in\mathbb{F},\ x\in V$. (D2)
(VS 7 = former, VS 8 = latter)

$\mathbb{F}^n$ IS A VECTOR SPACE OVER $\mathbb{F}$ (E2(1))

E_1 We can show that the Cartesian product
$$\mathbb{F}^n = \{(a_1,a_2,\dots,a_n) : a_i\in\mathbb{F}\ \forall i\in\{1,\dots,n\}\}$$
is a vector space over $\mathbb{F}$ with respect to the addition operation
$$(a_1,a_2,\dots,a_n) + (b_1,b_2,\dots,b_n) = (a_1+b_1, a_2+b_2,\dots,a_n+b_n)$$
and multiplication operation
$$c(a_1,a_2,\dots,a_n) = (ca_1, ca_2,\dots,ca_n).$$

Proof. This follows from verifying each of the conditions above. ∎

R_2 Note that we generally say "the vector space $\mathbb{F}^n$" to refer to the vector space $\mathbb{F}^n$ over $\mathbb{F}$. (R3(4))

COLUMN VECTOR NOTATION (E2(2))

Note that we can also write elements of $\mathbb{F}^n$ as "column vectors"; ie of the form
$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, \quad \text{where } a_1,a_2,\dots,a_n\in\mathbb{F}.$$

$\mathbb{Q}^n$ IS A VECTOR SPACE OVER $\mathbb{Q}$, $\mathbb{R}^n$ IS A VECTOR SPACE OVER $\mathbb{R}$, & $\mathbb{C}^n$ IS A VECTOR SPACE OVER $\mathbb{C}$ (R3(1))

We can show
① $\mathbb{Q}^n$ is a vector space over $\mathbb{Q}$;
② $\mathbb{R}^n$ is a vector space over $\mathbb{R}$; and
③ $\mathbb{C}^n$ is a vector space over $\mathbb{C}$.

Proof. This directly follows from the fact that $\mathbb{Q}$, $\mathbb{R}$ and $\mathbb{C}$ are fields (MATH 145), and substituting the respective fields into the above lemma. ∎

$\mathbb{R}^n$ IS A VECTOR SPACE OVER $\mathbb{Q}$, & $\mathbb{C}^n$ IS A VECTOR SPACE OVER $\mathbb{R}$ (R3(2))

Moreover, we can also show that
① $\mathbb{R}^n$ is a vector space over $\mathbb{Q}$; and
② $\mathbb{C}^n$ is a vector space over $\mathbb{R}$.

Proof. Essentially, this stems from the fact that we can "multiply" vectors in $\mathbb{R}^n$ by scalars in $\mathbb{Q}$, and vectors in $\mathbb{C}^n$ by scalars in $\mathbb{R}$. The formal proof is left to the reader. ∎

MATRICES (D3(1))

D_1 Let $\mathbb{F}$ be a field, and $m,n\in\mathbb{Z}^+$. Then, we say A is an "$m\times n$ matrix" with entries from $\mathbb{F}$ if it is of the form
$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \\ a_{m1} & a_{m2} & & a_{mn} \end{pmatrix},$$
where $a_{ij}\in\mathbb{F} \quad \forall i\in\{1,2,\dots,m\},\ j\in\{1,2,\dots,n\}$.

R_2 Alternatively, we can represent A via the notation
$$A = (a_{ij}), \quad \text{ie } \{1,2,\dots,m\},\ j\in\{1,2,\dots,n\}.$$

ij-ENTRY OF A MATRIX (D3(2))

Given a $m\times n$ matrix A, the "ij-entry" of A, or "a_{ij}", is defined to be the entry in A at the i^{th} row and j^{th} column.

ZERO MATRIX (D3(3))

The "$m\times n$ zero matrix", or more simply the "zero matrix", denoted as "O", is defined to be
$$O = \underbrace{\begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \\ 0 & 0 & & 0 \end{pmatrix}}_{n} \Big\} m,$$
ie the $m\times n$ matrix where which entry equals 0.

MATRIX EQUALITY (D3(4))

We say two matrices A and B are equal if and only if $a_{ij} = b_{ij} \quad \forall i\in\{1,2,\dots,m\},\ j\in\{1,2,\dots,n\}$.

MATRIX ADDITION (D3(5))

Let A and B be $m\times n$ matrices with entries from some field $\mathbb{F}$.
Then, the "addition" of A and B, denoted by "$A+B$", is defined to be the matrix where
$$(a+b)_{ij} = a_{ij} + b_{ij} \quad \forall i\in\{1,2,\dots,m\},\ j\in\{1,2,\dots,n\}.$$

MATRIX SCALAR MULTIPLICATION (D3(6))

Let A be a $m\times n$ matrix with entries from some field $\mathbb{F}$, and $c\in\mathbb{F}$ be arbitrary.
Then the "scalar multiplication" of A by c, denoted by "cA", is defined to be the matrix where
$$(ca)_{ij} = c(a_{ij}) \quad \forall i\in\{1,2,\dots,m\},\ j\in\{1,2,\dots,n\}.$$

SPACE OF $m\times n$ MATRICES (E3)

D_1 Let $\mathbb{F}$ be a field.
Then the "space of all $m\times n$ matrices" with entries from $\mathbb{F}$, denoted by "$M_{m\times n}(\mathbb{F})$", is defined to be the set of all $m\times n$ matrices with entries from $\mathbb{F}$.

R_2 Note that $M_{m\times n}(\mathbb{F})$ is a vector space over $\mathbb{F}$ with respect to the matrix addition and scalar multiplication operations.

Proof. This follows from verifying each of the conditions in D2. ∎

FUNCTION SPACES (E4)

1. Let the set $D \neq \emptyset$ be arbitrary, and let $\mathbb{F}$ be a field.
Then the "space of all functions" from D to $\mathbb{F}$, denoted by "$\mathbb{F}^D$", is defined to be the set of all functions of the form $f: D \to \mathbb{F}$.

2. Similarly, we can show that $\mathbb{F}^D$ is a vector space over $\mathbb{F}$ with respect to the operations of function addition

$$(f+g)(x) := f(x) + g(x) \quad \forall f, g \in \mathbb{F}^D, \ x \in \mathbb{F}$$

and function scalar multiplication

$$(cf)(x) := cf(x) \quad \forall f \in \mathbb{F}^D, \ x, c \in \mathbb{F}.$$

Proof. Similar strategy to E3: verify each condition in D2 holds.

POLYNOMIALS (D4)

SET OF ALL POLYNOMIALS OF DEGREE AT MOST n (D4 (1))

1. Let $\mathbb{F}$ be a field.
Then, we denote $P_n(\mathbb{F})$ to be the set of all polynomials with coefficients from $\mathbb{F}$ and of degree at most n; ie

$$P_n(\mathbb{F}) = \left\{ \sum_{i=0}^{n} a_n x^n \ : \ a_j \in \mathbb{F} \ \ \forall j \in \{0, 1, \dots, n\} \right\}.$$

POLYNOMIAL SPACES (D4 (2))

1. Let $\mathbb{F}$ be a field.
Then, we denote "$\mathbb{F}[x]$" to be the set of all polynomials with coefficients from $\mathbb{F}$; ie

$$\mathbb{F}[x] = \left\{ \sum_{i=0}^{\infty} a_i x^i \ : \ a_j \in \mathbb{F} \ \ \forall j \in \mathbb{N} \cup \{0\} \right\}.$$

2. Then, we can show that $\mathbb{F}[x]$ is a vector space over $\mathbb{F}$ with respect to the operations of polynomial addition

$$(f+g)(x) = \sum_{i=0}^{n} (a_n + b_n) x^n \quad \forall f, g \in \mathbb{F}[x]$$

and polynomial scalar multiplication

$$cf(x) = \sum_{i=0}^{n} (ca_n) x^n \quad \forall f \in \mathbb{F}[x], \ c \in \mathbb{F}.$$

Proof. Similar strategy to E4.

BASIC PROPERTIES OF VECTOR SPACES (S1.2)

CANCELLATION PROPERTY FOR VECTOR ADDITION (T1.1)

Let V be a vector space.
Suppose there exists some $x, y, z \in V$ such that $x + z = y + z$.
Then necessarily $x = y$.

Proof. Note that

$$\begin{aligned} x &= x + 0 \\ &= x + (z + (-z)) \\ &= (x+z) + (-z) \\ &= (y+z) + (-z) \\ &= y + (z + (-z)) \end{aligned}$$

and so $x = y$, as required. ☒

UNIQUENESS OF THE ZERO VECTOR IN VECTOR SPACES (C1.1.1 (1))

Let V be a vector space.
Suppose $0_1, 0_2 \in V$ are both zero vectors.
Then necessarily $0_1 = 0_2$.

Proof. This follows from the fact that V is an abelian group under addition. ☒

UNIQUENESS OF ADDITIVE INVERSES IN VECTOR SPACES (C1.1.1 (2))

Let V be a vector space.
Then for any $x \in V$, there exists one and only one vector $y \in V$ that satisfies $x + y = 0$.

Proof. This also follows from the fact that V is an abelian group under addition. ☒

$0x = 0 \quad \forall x \in V$ (T1.2 (1))

Let V be a vector space over some field $\mathbb{F}$, and let 0 be the additive identity of $\mathbb{F}$.
Then, for any $x \in V$, necessarily $0 \cdot x = 0$.

Proof. This, again, follows from the fact that V is an abelian group under addition. ☒

$a0 = 0 \quad \forall a \in \mathbb{F}$ (T1.2 (2))

Let V be a vector space over some field $\mathbb{F}$, and let 0 be the zero vector of V.
Then, for any $a \in \mathbb{F}$, necessarily $a \ 0 = 0$.

Proof. This, again, follows from the fact that V is an abelian group under addition. ☒

$(-a)x = -(ax) = a(-x) \quad \forall a \in \mathbb{F}, \ x \in V$ (T1.2 (3))

Let V be a vector space over some field $\mathbb{F}$, and let $a \in \mathbb{F}$, $x \in V$ be arbitrary.
Then necessarily $(-a)x = -(ax) = a(-x)$.

Proof. Proof is similar to the analog of this statement for rings (MATH 145). ☒

SUBSPACES (S1.3)

Let V be a vector space over some field $\mathbb{F}$.
Then we say the subset $W \subseteq V$ is a "subspace" of V if

① $W \neq \emptyset$;
*we usually check whether $0 \in W$ to verify this claim. (R4)

② If $x \in W$ and $y \in W$, then $(x+y) \in W$; and

③ If $c \in \mathbb{F}$ and $x \in W$, then $cx \in W$. (D6)

SUBSPACES ARE VECTOR SPACES OVER $\mathbb{F}$ WITH RESPECT TO THE OPERATIONS OF V (T1.3)

Let W be a subspace of a vector space V over some field $\mathbb{F}$.
Then W is also a vector space over $\mathbb{F}$ under the operations of V restricted to W.

Proof. This follows from verifying the conditions in D2, taking into account the properties of subspaces. ▨

$\{0\}$ AND V ARE SUBSPACES OF V (E8 (1))

Let V be a vector space.
Then $\{0\}$ and V itself are always subspaces of V.

Proof. $\{0\}$ is vacuously a subspace, and V is trivially a subspace. ▨

$P_2(\mathbb{R})$ IS A SUBSPACE OF $\mathbb{R}[x]$ (E8 (2))

We can show that $P_2(\mathbb{R})$ is a subspace of $\mathbb{R}[x]$.

Proof. This stems from the fact that:
- $P_2(\mathbb{R}) \subset \mathbb{R}[x]$ by definition;
- $0 \in P_2(\mathbb{R})$; and
- $P_2(\mathbb{R})$ is closed under the addition & scalar multiplication defined on $\mathbb{R}[x]$. ▨

$\{(a_{ij}) \in M_{n\times n}(\mathbb{F}) \mid \sum_{k=1}^{n} a_{kk} = 0\}$ IS A SUBSPACE OF $M_{n\times n}(\mathbb{F})$ (E8 (3))

We can show that the set $\{(a_{ij}) \in M_{n\times n}(\mathbb{F}) \mid \sum_{k=1}^{n} a_{kk} = 0\}$ is a subspace of $M_{n\times n}(\mathbb{F})$, where $n \in \mathbb{N}$ is arbitrary.

Proof. Similar proof to the above.

$\{(a_{ij}) \in M_{n\times n}(\mathbb{F}) \mid \sum_{k=1}^{n} a_{kk} = 1\}$ IS NOT A SUBSPACE OF $M_{n\times n}(\mathbb{F})$ (E8 (4))

We can show the set $\{(a_{ij}) \in M_{n\times n}(\mathbb{F}) \mid \sum_{k=1}^{n} a_{kk} = 1\}$ is not a subspace of $M_{n\times n}(\mathbb{F})$.

Proof. Let $a, b \in \{(a_{ij}) \in M_{n\times n}(\mathbb{F}) \mid \sum_{k=1}^{n} a_{kk} = 1\}$ be arbitrary.
Then, notice that

$$\begin{aligned}\sum_{k=1}^{n} (a+b)_{kk} &= \sum_{k=1}^{n} (a_{kk} + b_{kk}) \\ &= 1+1 \\ &= 2,\end{aligned}$$

so that $a+b \notin \{(a_{ij}) \in M_{n\times n}(\mathbb{F}) \mid \sum_{k=1}^{n} a_{kk} = 1\}$. ▨

SUBSPACES OF $\mathbb{R}^2$ (E9 (1))

Note that the subspaces of $\mathbb{R}^2$ are
① $\mathbb{R}^2$ itself;
② $\{0_{\mathbb{R}^2}\} = \{(0,0)\}$; and
③ all lines in $\mathbb{R}^2$ that pass through $(0,0)$.

SUBSPACES OF $\mathbb{F}^2$ (E9 (4a))

In general, for any field $\mathbb{F}$, the subspaces of $\mathbb{F}^2$ are
① $\mathbb{F}^2$ itself;
② $\{0\}$; and
③ all the "lines" in $\mathbb{F}^2$ through 0.
ie of the form $\{(x,y) \in \mathbb{F}^2 \mid \binom{x}{y} = k\binom{x_1}{y_1},\ \binom{x_1}{y_1} \in \mathbb{F}^2\}$

SUBSPACES OF $\mathbb{R}^3$ (E9 (2))

Similarly, the subspaces of $\mathbb{R}^3$ are
① $\mathbb{R}^3$ itself;
② $\{0_{\mathbb{R}^3}\} = \{(0,0,0)\}$;
③ all lines in $\mathbb{R}^3$ that pass through $(0,0,0)$; and
④ all planes in $\mathbb{R}^3$ that pass through $(0,0,0)$.

SUBSPACES OF $\mathbb{F}^3$ (E9 (4b))

Similarly, for any field $\mathbb{F}$, the subspaces of $\mathbb{F}^3$ are
① $\mathbb{F}^3$ itself;
② $\{0\}$;
③ all the "lines" in $\mathbb{F}^3$ through 0; and
ie of the form $\{(x,y,z) \in \mathbb{F}^3 \mid \begin{pmatrix}x\\y\\z\end{pmatrix} = k\begin{pmatrix}x_1\\y_1\\z_1\end{pmatrix},\ \begin{pmatrix}x_1\\y_1\\z_1\end{pmatrix} \in \mathbb{F}^3\}$ (E9(3))
④ all the "planes" in $\mathbb{F}^3$ through 0.
ie of the form $\{(x,y,z) \in \mathbb{F}^3 \mid \exists a,b,c \in \mathbb{F}$ such that $ax+by+cz=0.\}$

LINEAR COMBINATIONS & SYSTEM OF LINEAR EQUATIONS (S1.4)

* knowledge of elimination method is assumed.

LINEAR COMBINATION (D7(1))

1. Let V be a vector space over a field $\mathbb{F}$, and let the subset $S \subseteq V$ be such that $S \neq \emptyset$.
Then, we say a vector $x \in V$ is a "linear combination" of vectors from S if there exists a finite number of vectors $u_1, u_2, \dots, u_n \in S$ and scalars $a_1, a_2, \dots, a_n \in \mathbb{F}$ such that

$$x = a_1u_1 + a_2u_2 + \dots + a_nu_n,$$

where $n \geq 1$. (D7(1))

2. In this case, we also say that x is a linear combination of the vectors $u_1, u_2, \dots, u_n$.

COEFFICIENTS OF A LINEAR COMBINATION (D7 (2))

1. Let V be a vector space over some field $\mathbb{F}$, and let the vector $x \in V$ be a linear combination of the vectors $u_1, u_2, \dots, u_n \in S$, where $S \subseteq V$ and $S \neq \emptyset$.
Assume that $x = a_1u_1 + a_2u_2 + \dots + a_nu_n$, where $a_1, a_2, \dots, a_n \in \mathbb{F}$.
Then we denote the scalars $a_1, a_2, \dots, a_n \in \mathbb{F}$ as the "coefficients" of the linear combination.

SPAN (D7(3))

1. Let V be a vector space over some field $\mathbb{F}$, and let the subset $S \subseteq V$ be such that $S \neq \emptyset$.
Then, we define the "span" of S, denoted as "$\operatorname{span}(S)$", to be the set of all linear combinations of vectors in S.

2. Note that, for convenience, we define

$$\operatorname{span}(\emptyset) = \{0\}.$$

EXAMPLE 1: SPAN OF (1,0,0) & (0,1,0) IN $\mathbb{R}^3$ (E10(1))

1. Observe that in $\mathbb{R}^3$, the span of $(1,0,0)$ & $(0,1,0)$ in $\mathbb{R}^3$ is

$$\{a(1,0,0) + b(0,1,0) : a, b \in \mathbb{R}\}$$

or more simply

$$\{(a, b, 0) : a, b \in \mathbb{R}\}.$$

EXAMPLE 2: SPAN($\{x^n : n \geq 1\}$) IN $\mathbb{Q}[x]$ (E10(2))

1. We can show that for the vector space $\mathbb{Q}[x]$, the span of $S = \{x, x^2, \dots, x^n, \dots\}$ is the set of all polynomials in $\mathbb{Q}[x]$ whose constant coefficient equals 0.

SPAN OF A FINITE AMOUNT OF VECTORS (E10(3.1))

1. Suppose V is a vector space over some field $\mathbb{F}$, and let $S \subseteq V$. Further assume that

$$S = \{v_1, v_2, \dots, v_n\},$$

ie the size of S is finite.
Then, it follows that

$$\operatorname{span}(S) = \{a_1v_1 + a_2v_2 + \dots a_nv_n : a_i \in \mathbb{F} \ \forall i \in \{1, 2, \dots, n\}\}.$$

SPAN OF A COUNTABLE AMOUNT OF VECTORS (E10 (3.2))

1. Suppose V is a vector space over some field $\mathbb{F}$, and let $S \subseteq V$. Further assume that

$$S = \{v_1, v_2, \dots, v_n, \dots\},$$

ie $|S| = |\mathbb{N}|$.
Then, we can show that

$$\operatorname{span}(S) = \operatorname{span}(\{v_1\}) \cup \operatorname{span}(\{v_1, v_2\}) \cup \dots \cup \operatorname{span}(\{v_1, v_2, \dots, v_n\}) \cup \dots$$

or in other words, that

$$\operatorname{span}(S) = \bigcup_{n=1}^{\infty} \operatorname{span}\{v_1, v_2, \dots, v_n\}.$$

SPAN OF AN UNCOUNTABLE AMOUNT OF VECTORS (E10 (3.3))

1. Suppose V is a vector space over some field $\mathbb{F}$, and let $S \subseteq V$. Further assume that $|S| > |\mathbb{N}|$; ie the size of S is uncountable.
Then note that there are no "obvious" simplifications to the formula for $\operatorname{span}(S)$.

SPAN OF A SET IS ALWAYS A SUBSPACE OF THE PARENT VECTOR SPACE (T1.4)

1. Let V be a vector space over some field $\mathbb{F}$, and let $S \subseteq V$.
Then necessarily $\operatorname{span}(S)$ is a subspace of V.
Proof. This follows from verifying each subspace condition for $\operatorname{span}(S)$. ☒

2. Moreover, $\operatorname{span}(S)$ is the "smallest possible" subspace of V that contains S, in the sense that
   1. $S \subseteq \operatorname{span}(S)$; and
   2. If W is any other subspace of V containing S, then $\operatorname{span}(S) \subseteq W$.

"GENERATES/SPANS" (D8)

1. Let V be a vector space, and let $S \subseteq V$.
Then, we say S "generates" V, or S "spans" V, if $\operatorname{span}(S) = V$.

2. Note to prove $\operatorname{span}(S) = V$, we just need to prove every vector in V can be written as a linear combination of vectors in S, since $\operatorname{span}(S) \subseteq V$ by definition. (This follows from extensionality.) (R6)

LINEAR INDEPENDENCE & DEPENDENCE (S1.5)

LINEARLY DEPENDENT (D9 (1))

1. Let V be a vector space over some field $\mathbb{F}$, and let $S \subseteq V$.
Then, we say S is "linearly dependent" if there exists a finite number of distinct vectors $u_1, u_2, \dots, u_n \in S$ and scalars $c_1, c_2, \dots, c_n \in \mathbb{F}$, where $c_1, c_2, \dots, c_n$ are <u>all not zero</u>, such that
$$c_1 u_1 + c_2 u_2 + \cdots + c_n u_n = 0.$$

2. In this case, we also say the vectors of S are linearly dependent.

3. Note that if S is finite, say $S = \{u_1, u_2, \dots, u_n\}$, then S is linearly dependent <u>if and only if</u> there exists a $(c_1, c_2, \dots, c_n) \in \mathbb{F}^n$, where $(c_1, c_2, \dots, c_n) \neq (0, 0, \dots, 0)$, such that
$$c_1 u_1 + c_2 u_2 + \cdots + c_n u_n = 0. \quad \text{(R7 (4a))}$$

LINEARLY INDEPENDENT (D9 (2))

1. Let V be a vector space over some field $\mathbb{F}$, and let $S \subseteq V$.
Then, we say S is "linearly independent" if it is <u>not</u> "linearly dependent"; ie for every choice of distinct $u_1, u_2, \dots, u_n \in S$, if $c_1, c_2, \dots, c_n \in \mathbb{F}$ are scalars such that
$$c_1 u_1 + c_2 u_2 + \cdots + c_n u_n = 0,$$
then necessarily $c_1 = c_2 = \cdots = c_n = 0$.

2. Similarly, if S is finite, say $S = \{u_1, u_2, \dots, u_n\}$, then S is linearly independent <u>if and only if</u> whenever $(c_1, c_2, \dots, c_n) \in \mathbb{F}^n$ are such that
$$c_1 u_1 + c_2 u_2 + \cdots + c_n u_n = 0,$$
then necessarily $c_1 = c_2 = \cdots = c_n = 0$.

TRIVIAL REPRESENTATION OF 0 (R7 (1))

Note that for any vector space V and the vectors $u_1, u_2, \dots, u_n \in V$, we denote "trivial representation of $0 \in V$" as a linear combination of $u_1, u_2, \dots, u_n$ by
$$0u_1 + 0u_2 + \cdots + 0u_n = 0.$$

EMPTY SET IS LINEARLY INDEPENDENT (R7(2))

Note that the empty set, $\emptyset$, is <u>vacuously</u> linearly independent.

* since linearly dependent sets must be non-empty by definition.

$\{0\}$ IS LINEARLY DEPENDENT (R7 (3))

Note that the set $\{0\}$ is linearly dependent, since $1(0) = 0$ is a <u>non-trivial</u> representation of 0 as a linear combination of finitely many distinct vectors in S.

$0 \in S \Rightarrow S$ IS LINEARLY DEPENDENT (R7 (5))

Note that any subset of a vector space that contains the zero vector is linearly dependent.

EXAMPLE 1: $S = \{(0,1,1), (1,0,1), (1,2,3)\}$ IS LINEARLY DEPENDENT IN $\mathbb{R}^3$ (E(4))

We can show that the set $S = \{(0,1,1), (1,0,1), (1,2,3)\}$ is linearly dependent in $\mathbb{R}^3$.

Proof. We search for scalars $a, b, c \in \mathbb{R}$, not all 0, such that
$$a\begin{pmatrix}0\\1\\1\end{pmatrix} + b\begin{pmatrix}1\\0\\1\end{pmatrix} + c\begin{pmatrix}1\\2\\3\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}.$$
This reduces to solving the system
$$\begin{cases} b + c = 0 \\ a + 2c = 0 \\ a + b + 3c = 0. \end{cases}$$
Simplifying, we get that $a = -2t$, $b = -t$ and $c = t$, where $t \in \mathbb{R}$.
For instance, $(a, b, c) = (-2, -1, 1)$ is a solution in which not all of a, b, c are 0.
It follows that S is linearly dependent. ☒

EXAMPLE 2: $S = \{1, x, x^2, x^3\}$ IS LINEARLY INDEPENDENT IN $\mathbb{Z}_5[x]$ (E(5))

We can show that the set $S = \{1, x, x^2, x^3\}$ is linearly independent in $\mathbb{Z}_5[x]$.

Proof. Note that if there exist $a_0, a_1, a_2, a_3 \in \mathbb{Z}_5$ such that
$$a_0(1) + a_1 x + a_2 x^2 + a_3 x^3 = 0,$$
then by definition necessarily $a_0 = a_1 = a_2 = a_3 = 0$, and this is sufficient to prove the claim. ☒

S IS LINEARLY DEPENDENT <=> S={0} OR SOME VECTOR IN S IS A LINEAR COMBINATION OF OTHER VECTORS IN S (T1.5)

Let V be a vector space, and let $S \subseteq V$. Then S is linearly dependent if and only if $S=\{0\}$ or some vector in S is a linear combination of other vectors in S.

Proof. We first prove the backward argument.
First, note we know why $\{0\}$ is linearly dependent from a previous section.
So, suppose there exists a vector $v \in S$ such that

$$v = c_1u_1 + c_2u_2 + \cdots + c_nu_n,$$

where $c_i \in \mathbb{F}$ and $u_i \in V$ $\forall i \in \{1,2,\dots,n\}$.
Without loss in generality, assume $u_1,u_2,\dots,u_n$ are distinct.
By assumption, since $v \notin \{u_1,u_2,\dots,u_n\}$, necessarily $u_1,u_2,\dots,u_n,v$ are distinct.
Finally, since

$$0 = (-1)v + c_1u_1 + c_2u_2 + \cdots + c_nu_n,$$

and $-1 \neq 0$, it follows S is linearly dependent. *

Next, we prove the forward argument.
Assume S is linearly dependent, so that there exist distinct $u_1,u_2,\dots,u_n \in S$ and $a_1,a_2,\dots,a_n \in \mathbb{F}$ (not all 0) such that

$$a_1u_1 + a_2u_2 + \cdots + a_nu_n = 0.$$

Without loss in generality, assume $a_i \neq 0$ $\forall i \in \{1,2,\dots,n\}$.

Case 1: $n=1$.
Then $a_1u_1=0$, and since $a_1\neq 0$ it follows that $u_1=0$ (since fields are integral domains, so the cancellation property applies.)
Hence $0 \in S$. If $S=\{0\}$ we are done; otherwise, we can pick a $v \in S\setminus\{0\}$, and we can write $0=0v$, proving some vector in S, 0, can be written as a linear combination of another vector, v, in S.

Case 2: $n>1$.
Then since $a_n \neq 0$, we can solve for u_n:

$$u_n = -a_n^{-1}[a_1u_1 + a_2u_2 + \cdots + a_{n-1}u_{n-1}],$$

$$\therefore u_n = (-a_n^{-1}a_1)u_1 + (-a_n^{-1}a_2)u_2 + \cdots + (-a_n^{-1}a_{n-1})u_{n-1},$$

showing u_n can be expressed as a linear combination of other elements in S.

BASES & DIMENSION (S1.6)

BASIS (D10)

Let V be a vector space.
Then, we say a subset $S \subseteq V$ is a "basis" for V if
1. S is linearly independent; and
2. S spans V.

In this case, we also say that the vectors of S form a basis for V.

STANDARD BASIS (E17)

In $\mathbb{F}^n$, define the "standard basis" for $\mathbb{F}^n$ the subset

$$S = \{e_1, e_2, \dots, e_n\},$$

where $e_j \in \mathbb{F}^n$ is the vector with j^{th} coordinate 1 and other coordinates 0.

(It is easy to prove S is indeed a basis for $\mathbb{F}^n$.)

In $P_n(\mathbb{F})$, define the "standard basis" for $P_n(\mathbb{F})$ as the set

$$S = \{1, x, x^2, \dots, x^n\}.$$

(It is also easy to prove S is indeed a basis for $P_n(\mathbb{F})$).

UNIQUE REPRESENTATION OF ELEMENTS IN VECTOR SPACES UNDER A BASIS (T1.6)

Let $\{v_1, v_2, \dots, v_n\}$ be a basis for a vector space V.
Then for every $x \in V$, x can be uniquely represented as a linear combination of $v_1, v_2, \dots, v_n$;
ie there exists a unique n-tuple $(a_1, a_2, \dots, a_n) \in \mathbb{F}^n$ such that

$$x = a_1 v_1 + a_2 v_2 + \cdots + a_n v_n.$$

Proof. Existence: this follows from the fact that $\{v_1, v_2, \dots, v_n\}$ spans V by definition.
Uniqueness: suppose there exists some $b_1, b_2, \dots, b_n \in \mathbb{F}$ such that

$$x = a_1 v_1 + \cdots + a_n v_n = b_1 v_1 + \cdots + b_n v_n.$$

It follows that

$$0 = (a_1 - b_1) v_1 + \cdots + (a_n - b_n) v_n,$$

and since $\{v_1, v_2, \dots, v_n\}$ is linearly independent, necessarily $a_k = b_k$ $\forall k \in \{1, 2, \dots, n\}$. ∎

V IS GENERATED BY S, $|S| = |\mathbb{N}|$ ⇒ T ⊂ S IS ALSO A BASIS FOR V (T1.7)

Let V be a vector space, and assume that V is generated by a countable set S.
Then there exists a subset of S that is a basis for V.

Proof. If $S = \emptyset$ or $S = \{0\}$, then $\emptyset$ is a basis for V trivially.
Otherwise, S contains at least a non-zero vector.
Hence, we can write S as

$$S = \{v_1, v_2, \dots, v_n\} \quad \text{or} \quad S = \{v_1, v_2, \dots\}.$$

By the WOP, we can pick the smallest index $i_1 \geq 1$ such that $v_{i_1} \neq 0$.
Then $\{v_{i_1}\}$ is linearly independent.
Let i_2 be the smallest index such that $v_{i_2} \in S$ and $v_{i_2} \notin \operatorname{span}(\{v_{i_1}\})$.
Continue this "process" until we obtain the set

$$T = \{v_{i_k} \in S \mid v_{i_k} \notin \operatorname{span}(\{v_{i_1}, \dots, v_{i_{k-1}}\}),\ k \geq 1\}.$$

Finally, we can prove T is a basis for V.

1. Assume T is linearly dependent.
   Then there exists $a_1, a_2, \dots, a_k$, all not 0, such that
   $$a_1 v_{i_1} + \cdots + a_k v_{i_k} = 0.$$
   It follows that
   $$v_{i_k} = -a_k^{-1} a_1 v_{i_1} - \cdots - a_k^{-1} a_{k-1} v_{i_{k-1}},$$
   contradicting the construction of T.
2. We can prove by induction that $\operatorname{span}(S_k) = \operatorname{span}(T_k)$ $\forall k \geq 1$, where
   $$S_k = \{v_1, v_2, \dots, v_k\} \quad \text{and} \quad T_k = T \cap S_k = \{v_{i_q} \mid i_q \leq k\}.$$
   Then, let $x \in V = \operatorname{span}(S)$. Then $x \in \operatorname{span}(S_m)$ for some large m, so that $x \in \operatorname{span}(T_m) \subset \operatorname{span}(T)$.
   Hence $V \subseteq \operatorname{span}(T)$, and it follows that $V = \operatorname{span}(T)$. ∎

EVERY VECTOR SPACE HAS A BASIS (T1.8)

We can prove that every vector space has a basis.
(The proof uses Zorn's lemma & maximal linearly independent subsets.)

REPLACEMENT THEOREM (T1.9)

Suppose V is a vector space with a finite spanning set S.

Let T be a linearly independent subset in V. Then

① $|T| \leq |S|$; and

② There exists a set $H \subseteq S$ containing exactly $(|S| - |T|)$ vectors such that $T \cup H$ generates V.

Proof. Let $n = |S|$, and let $m = |T|$.

Then, when $m = 0$, clearly $m = 0 \leq |S|$.

Next, assume the statement is true for some $m \geq 0$.

This implies that if $T_m \subset V$ is any linearly independent subset in V of size m, then $m \leq n$ and there exists a set $H_m \subseteq S$ containing exactly $n-m$ vectors such that $T_m \cup H_m$ generates V.

Let $T_m = \{v_1, v_2, \dots, v_m\}$ and $T = T_m \cup \{v_{m+1}\}$, such that T is linearly independent and a subset of V.

Note that this implies T_m is also linearly independent.

Now, apply the induction hypothesis on T_m to get that $n \geq m$, and there exist $(n-m)$ vectors $w_{m+1}, \dots, w_n \in S$ such that $\{v_1, \dots, v_m, w_{m+1}, \dots, w_n\}$ generates V.

Then, since $n \geq m$, either $n = m$ or $n > m$.

If $n = m$, $\{v_1, \dots, v_m, w_{m+1}, \dots, w_n\} = \{v_1, \dots, v_m\}$.

Thus, $v_{m+1} \in \operatorname{span}\{v_1, \dots, v_m\}$, so by Theorem 1.5, the set $\{v_1, \dots, v_m, v_{m+1}\}$ is linearly dependent.

But this is a contradiction; hence, it follows that $n > m$, so that $n \geq m+1$, proving ①.

Subsequently, write

$$v_{m+1} = a_1 v_1 + \dots + a_m v_m + a_{m+1} w_{m+1} + \dots + a_n w_n$$

for some scalars $a_1, \dots, a_n \in \mathbb{F}$.

Then, if $a_{m+1} = \dots = a_n = 0$, then we would get that $v_{m+1} = a_1 v_1 + \dots + a_m v_m$, which is a contradiction; hence, at least one of the scalars $a_{m+1}, \dots, a_n$ must be non-zero.

Then, without loss in generality, assume $a_{m+1} \neq 0$.

It follows that

$$w_{m+1} = -a_{m+1}^{-1} a_1 v_1 - \dots - a_{m+1}^{-1} a_m v_m - a_{m+1}^{-1} a_{m+1} v_{m+1}$$
$$- a_{m+1}^{-1} a_{m+2} w_{m+2} - \dots - a_{m+1}^{-1} a_n w_n.$$

Let $H = \{w_{m+2}, \dots, w_n\} \subset S$. The above shows that $w_{m+1} \in \operatorname{span}(T \cup H)$.

Moreover, since $v_1, \dots, v_m \in T \subset T \cup H$ and $w_{m+2}, \dots, w_n \in H \subset T \cup H$, it follows that

$$V = \operatorname{span}\{w_{m+1}, v_1, \dots, v_m, w_{m+2}, \dots, w_n\} \subseteq \operatorname{span}(T \cup H).$$

But since $\operatorname{span}(T \cup H) \subseteq V$, it follows that $V = \operatorname{span}(T \cup H)$, completing the proof. ▨

V IS FINITELY SPANNED ⇒ ALL BASES OF V & HAVE EQUAL CARDINALITIES (C1.9.1)

Suppose V is a finitely spanned vector space.

Then all bases of V are finite and have the same amount of elements.

Proof. Let S be a finite spanning set for V, and let B be an arbitrary basis for V.

Then by definition, B is linearly independent.

By the Replacement Theorem, $|B| \leq |S| < \infty$.

Next, let B_1 and B_2 be two bases of V.

Then, since B_1 is linearly independent and B_2 is a finite spanning set for V, by the Replacement Theorem necessarily $|B_1| \leq |B_2|$.

Similarly, since B_2 is linearly independent and B_1 is a finite spanning set for V, by the Replacement Theorem necessarily $|B_2| \leq |B_1|$.

It follows that $|B_1| = |B_2|$, and we are done.

DIMENSION

FINITE/INFINITE-DIMENSIONAL (D12)

1: We say a vector space V is "finite-dimensional" if it has a basis consisting of a finite number of vectors.

2: Otherwise, we say V is "infinite-dimensional".

DIMENSION (D12)

1: Let V be a finite-dimensional vector space.

Then, the "dimension" of V, denoted as "$\dim V$", is defined to be the unique number of vectors in each basis for V.

2: By convention, we let $\dim\{0\} = 0$.

3: Examples:

① $\dim \mathbb{F}^n = n$;

② $\dim \mathbb{C}^n = 2n$;

③ $\dim M_{m \times n}(\mathbb{F}) = mn$; and

④ $\dim P_n(\mathbb{F}) = n+1$. (E18)

ANY FINITE SPANNING SET FOR V CONTAINS AT LEAST n VECTORS (C1.9.2 (1))

Let V be a vector space with $\dim V = n$. Then if S is a finite spanning set for V, necessarily $|S| \geq n$.

Proof. By the Existence Theorem (T1.7), there exists a subset T of S that is a basis for V.
Therefore $|T| = \dim V = n$, which implies that $|S| \geq |T| = n$. ▨

S GENERATES V, |V| = n ⇒ S IS A BASIS FOR V (C1.9.2 (2))

Let V be a vector space with $\dim V = n$, and suppose S generates V, with $|S| = n$.
Then S is a basis for V.

Proof. Again, by the Existence Theorem (T1.7), there exists some subset $T \subseteq S$ such that T is a basis for V.
By the above corollary, $|T| = n$, so that if $|S| = n$, necessarily $S = T$.
It follows that S is a basis for V. ▨

S IS LINEARLY INDEPENDENT ⇒ S CONTAINS AT MOST n VECTORS (C1.9.2 (3))

Let V be a vector space, with $\dim V = n$.
Suppose the subset $S \subseteq V$ is linearly independent.
Then S contains at most n vectors.

Proof. Applying the Replacement Theorem for the spanning set β, it follows that $|S| \leq |\beta|$, and since $|\beta| = n$, this tells us that $|S| \leq n$, as needed. ▨

S IS LINEARLY INDEPENDENT, |S| = n ⇒ S IS A BASIS FOR V (C1.9.2 (4))

Let V be a vector space, with $\dim V = n$.
Suppose the subset $S \subseteq V$ is linearly independent and $|V| = n$.
Then S is a basis for V.

Proof. Applying the Replacement Theorem for the spanning set β and the linearly independent set S, there must exist a subset $H \subseteq \beta$ containing $|\beta| - |S| = n - n = 0$ vectors such that $S \cup H$ generates V.
But since $|H| = 0$, hence $H = \emptyset$, so that S generates V (and hence is a basis for V.) ▨

EVERY LINEARLY INDEPENDENT SUBSET OF V CAN BE "EXTENDED" TO A BASIS OF V (C1.9.2 (5))

Let V be a vector space, with $\dim V = n$.
Suppose $L = \{v_1, \dots, v_k\}$ is a linearly independent subset of V, where $1 \leq k \leq n$.
Then there exists a $H \subset V$ such that $L \cup H$ is a basis of V.

Proof. If $k = n$, by C1.9.2(4) L is trivially a basis for V.
If $k < n$, then by the Replacement Theorem for the spanning set β and L, there necessarily exists a subset $H \subset \beta$ containing $|\beta| - |L| = n-k$ vectors such that $L \cup H$ generates V.
By C1.9.2 (1), $|L \cup H| \geq n$. But
$$|L \cup H| \leq |L| + |H| = k + (n-k) = n,$$
so that $|L \cup H| = n$.
It follows by C1.9.2 (2) that $L \cup H$ is a basis for V. ▨

W IS A SUBSPACE OF V ⇒ dim W ≤ dim V ; dim W = dim V ⟺ W = V (C1.9.2 (6))

Let W be a subspace of the vector space V.
Then $\dim W \leq \dim V$, with equality occurring if and only if $V = W$.

Proof. If $W = \{0\}$, then $\dim W = 0 \leq \dim V$.
Otherwise, W contains a non-zero vector w_1.
Then $\{w_1\}$ is linearly independent.
Continue to choose the vectors $w_1, \dots, w_n \in W$ such that $\{w_1, \dots, w_k\}$ is linearly independent.
Note that this process cannot go on indefinitely, since $\{w_1, \dots, w_k\}$ is also linearly independent in V.
This implies that $k \leq n$.
Next, by T1.5, $W \subset \operatorname{span}(\{w_1, \dots, w_k\}) = \operatorname{span}(T)$.
Then, since $T \subseteq W$, necessarily $\operatorname{span}(T) \subset \operatorname{span}(W) = W$.
It follows that $W = \operatorname{span}(T)$, so that T is a basis (since it is also linearly independent), and
$$\dim W = |T| = k \leq n = \dim V.$$
Note that if $\dim V = n = \dim W$, then a basis for W is also a linearly independent set containing n elements.
Hence, by C1.9.2 (4), that set is also a basis for V. ▨

W IS A SUBSPACE FOR V ⇒ ANY BASIS OF W CAN BE "EXTENDED" TO A BASIS IN V (C1.9.2 (7))

Let W be a subspace of the vector space V, and let S be a basis of W.
Then we can "extend" S to a basis in V.

Proof. By C1.9.2 (6), $\dim W \leq \dim V$.
Let $T = \{w_1, \dots, w_k\}$ be a basis for W, so that T is linearly independent in W, which in turn implies T is linearly independent in V.
So, by C1.9.2 (5), we can "extend" T to a basis in V. ▨

QUOTIENT SPACES (S1.7)

COSET & REPRESENTATIVE (D13)

Let V be a vector space, and W be a subspace of V.
Then, for a given $x \in V$, its corresponding "coset" of W in V, denoted as "$x+W$", is defined to be the set

$x+W = \{x+w : w \in W\}$.

* note that $x+W \subseteq V$.

In this case, we call "x" a "representative" of the coset $x+W$.

$x \equiv y \pmod{W}$ (D13)

Let V be a vector space, and let W be a subspace of V.
Then, we write "$x \equiv y \pmod{W}$" if and only if $x-y \in W$.

V/W (D13)

Let V be a vector space, and W a subspace of V.
Then, we denote "V/W" (ie "V mod W") as the set

$V/W = \{x+W : x \in V\}$;

ie let V/W be the collection of cosets of W in V.

$V/\{0\} = V$ (E19 (2))

For any vector space V, necessarily $V/\{0\} = V$.

Proof. $V/\{0\} = \{0+x : x \in V\}$
$= \{x : x \in V\}$
$\therefore V/\{0\} = V$. ☒

COSET TEST (P1)

Let W be a subspace of a vector space V, and let $x, y \in V$ be arbitrary.
Then $x+W = y+W$ if and only if $x-y \in W$.

Proof. Similar to test for cosets in MATH 145.

$\equiv \pmod{W}$ IS AN EQUIVALENCE RELATION ON V (R8)

Note that the relation "$\equiv \pmod{W}$" is an equivalence relation on V.

ADDITION & MULTIPLICATION IN V/W (D14)

Let V be a vector space over a field $\mathbb{F}$, and let W be a subspace of V.
Then, we can define an addition on V/W by

$(x+W) + (y+W) := ((x+y) + W)$;

and a scalar multiplication on V/W by

$a(x+W) := (ax) + W$;

for any $a \in \mathbb{F}$ and $x, y \in W$.

Note that these addition and multiplication operations are well-defined. (L1)

Proof. Similar to proof for quotient groups/rings.

V/W IS A VECTOR SPACE (THE QUOTIENT SPACE OF V BY W) (T1.10)

Let V be a vector space, and W a subspace of V.
Then the set V/W is a vector space over $\mathbb{F}$ with the operations of coset addition and scalar multiplication, denoted as "the quotient space of V by W".

Proof. Verify all 8 conditions. (VS 1-8). ☒

BASIS FOR QUOTIENT SPACES (T1.11)

Let V be a vector space with $\dim V = n$, and let W be a subspace of V such that $\dim W = k$.
Let $\{v_1, \dots, v_n\}$ be a basis for V, such that $\{v_1, \dots, v_k\}$ is a basis for W.
Then,
① The set $\{v_{k+1}+W, \dots, v_n+W\}$ is a basis for V/W; and
② $\dim(V/W) = \dim V - \dim W$.

Proof. To prove ①, we show $\{v_{k+1}+W, \dots, v_n+W\}$ is both linearly independent and generates V/W, giving us our basis.
It follows that
$\dim(V/W) = |\{v_{k+1}+W, \dots, v_n+W\}|$
$= n-(k+1)+1$
$= n-k$
$\therefore \dim(V/W) = \dim V - \dim W$. ☒

$\dim V \geq \infty$, $\dim W \geq \infty$ $\not\Rightarrow$ $\dim V/W \geq \infty$ (R9)

Let V be an infinite-dimensional vector space, and let W be an infinite-dimensional subspace of V.
Then, note that it is not necessarily the case that $\dim(V/W) \geq \infty$.

Example: let $V = \mathbb{F}^\infty$ & $W = \{(0, x_2, \dots) : x_k \in \mathbb{F}\}$.
Note that each element of V/W is simply "determined" by the value of the first coordinate x_1, so that $\dim(V/W) = 1$.

SUMS & INTERNAL DIRECT SUMS OF SUBSPACES (SL8)

SUM OF SUBSPACES (DIS)

Let V be a vector space over $\mathbb{F}$, and let W_1, W_2 be subspaces of V.
Then, we define the "sum" of W_1 and W_2, denoted as "W_1+W_2", to be the set

$$W_1+W_2 := \{v_1+v_2 : v_1 \in W_1,\ v_2 \in W_2\}.$$

INDEPENDENT / DISJOINT (DIS)

Let V be a vector space, and let W_1, W_2 be subspaces of V.
Then, we say W_1 and W_2 are "independent", or "disjoint", if and only if $W_1 \cap W_2 = \{0\}$.

(INTERNAL) DIRECT SUM (DIS)

Let V be a vector space, and let W_1, W_2 be independent subspaces of V.
Then, we define the "(internal) direct sum" of W_1 and W_2", denoted as $W_1 \oplus W_2$, to be the set

$$W_1 \oplus W_2 = W_1 + W_2.$$

*ie "$\oplus$" is the notation for "+" used when W_1 & W_2 are independent.

Note that $W_1 \oplus W_2$ is well-defined as long as $W_1 \cap W_2 = \{0\}$. (R10)

W_1+W_2 IS THE "SMALLEST" SUBSPACE CONTAINING W_1 & W_2 (L2 (2))

Let V be a vector space, and let W_1, W_2 be subspaces of V.
Then W_1+W_2 is necessarily the smallest subspace of V containing W_1 and W_2.

Proof. First, we prove W_1+W_2 is a subspace of V.
Let $(v_1+v_2), (u_1+u_2) \in W_1+W_2$ and $a \in F$, where $v_1, u_1 \in W_1$ and $v_2, u_2 \in W_2$.
Then, since W_1 and W_2 are subspaces of W_1+W_2, necessarily $v_1+u_1 \in W_1$ and $v_2+u_2 \in W_2$, so that

$$(v_1+v_2)+(u_1+u_2) = (v_1+u_1)+(v_2+u_2) \in W_1+W_2.$$

Moreover, since $av_1 \in W_1$ and $av_2 \in W_2$, necessarily

$$a(v_1+v_2) = av_1 + av_2 \in W_1+W_2,$$

proving W_1+W_2 is closed under addition and scalar multiplication.
Then, since $v_1 = v_1 + 0 \in W_1+W_2\ \forall v_1 \in W_1$ & $v_2 = 0+v_2 \in W_1+W_2\ \forall v_2 \in W_2$, it follows that $W_1 \subseteq W_1+W_2$ and $W_2 \subseteq W_1+W_2$.
Finally, let Y be a subspace of V that contains both W_1 & W_2.
Since Y is closed under addition, $v_1+v_2 \in Y$ for every $v_1 \in W_1$ and $v_2 \in W_2$ necessarily.
It follows that $W_1+W_2 \subseteq Y$, completing the proof.

$V = W_1 \oplus W_2 \iff \forall v \in V: \exists$ UNIQUE $w_1 \in W_1$, $w_2 \in W_2 \ni v = w_1+w_2$ (L2 (3))

Let V be a vector space, and let W_1 and W_2 be subspaces of V.
Then $W_1 \oplus W_2 = V$ if and only if for every vector $v \in V$, there exist unique elements $w_1 \in W_1$ and $w_2 \in W_2$ such that $v = w_1 + w_2$.

Proof. ($\Rightarrow$) Since $V = W_1 \oplus W_2$, necessarily $V = W_1+W_2$ and $W_1 \cap W_2 = \{0\}$.
Let $v \in V$, and note that since $V = W_1+W_2$, it implies that $v \in W_1+W_2$.
So, by definition, there exist some $w_1 \in W_1$, $w_2 \in W_2$ such that $v = w_1+w_2$.
Next, suppose we have $v = w_1'+w_2'$ for some $w_1' \in W_1$ and $w_2' \in W_2$.
Then

$$0 = (w_1+w_2) - (w_1'+w_2') = (w_1-w_1') + (w_2-w_2').$$

Since $w_1, w_1' \in W_1$ & $w_2, w_2' \in W_2$, necessarily $w_1 - w_1' \in W_1$ & $w_2-w_2' \in W_2$ also, so that

$$(w_1-w_1') = w_2'-w_2 \in W_1 \cap W_2 = \{0\},$$

Hence $w_1-w_1' = w_2'-w_2 = 0$, implying that $w_1 = w_1'$ & $w_2 = w_2'$, proving uniqueness. #

($\Leftarrow$) By assumption, every vector $v \in V$ can be written as $v = w_1+w_2$ for some $w_1 \in W_1$ & $w_2 \in W_2$.
Hence $V \subseteq W_1+W_2$, and by L2(2) necessarily $W_1+W_2 \subseteq V$; so $V = W_1+W_2$.
Next, let $x \in W_1 \cap W_2$. Then $-x \in W_1 \cap W_2$.
Then, note that

$$0 = 0+0 = x+(-x) \in W_1+W_2,$$

and due to the uniqueness assumption, necessarily $x=0$.
Thus $W_1 \cap W_2 = \{0\}$, so that $V = W_1 \oplus W_2$. ∎

$\dim(W_1), \dim(W_2) < \infty \Rightarrow \dim(W_1+W_2) < \infty$ & $\dim(W_1) + \dim(W_2) = \dim(W_1+W_2) + \dim(W_1 \cap W_2)$ (T1.12 (1))

Let V be a vector space over some field $\mathbb{F}$, and let W_1, W_2 be finite dimensional subspaces of V.
Then necessarily W_1+W_2 is finite dimensional, and
$\dim(W_1) + \dim(W_2) = \dim(W_1+W_2) + \dim(W_1 \cap W_2)$.

Proof. First, note $W_1 \cap W_2$ is a subspace of W_1 (A2), so that
$\dim(W_1 \cap W_2) \leq \dim(W_1) < \infty$ (C1.9.2 (6)).
Next, let $\{u_1, u_2, \dots, u_k\}$ be a basis for $W_1 \cap W_2$.
Extend this basis to get the bases
$S_1 = \{u_1, \dots, u_k, v_1, \dots, v_m\}$ of W_1 and
$S_2 = \{u_1, \dots, u_k, z_1, \dots, z_p\}$ of W_2, which we can always do by C1.9.2 (5)).
Let $S = \{u_1, \dots, u_k, v_1, \dots, v_m, z_1, \dots, z_p\}$.
We claim S is a basis for W_1+W_2.
Indeed, consider

$$a_1u_1 + \dots + a_ku_k + b_1v_1 + \dots + b_mv_m + c_1z_1 + \dots + c_pz_p = 0 \quad \text{— ②}$$

for some scalars $a_1, \dots, a_k, b_1, \dots, b_m, c_1, \dots, c_p$.
Then

$$b_1v_1 + \dots + b_mv_m = -a_1u_1 - \dots - a_ku_k - c_1z_1 - \dots - c_pz_p.$$

Since the RHS is a linear combination of vectors in W_2, the RHS $\in W_2$; and since the LHS is a linear combination of vectors in W_1, the LHS $\in W_1$.

Thus $b_1v_1 + \dots + b_mv_m \in W_1 \cap W_2$.
Next, since $\{u_1, \dots, u_k\}$ is a basis for $W_1 \cap W_2$, there exist scalars $d_1, \dots, d_k$ such that

$$b_1v_1 + \dots + b_mv_m = d_1u_1 + \dots + d_ku_k.$$

So

$$b_1v_1 + \dots + b_mv_m - d_1u_1 - \dots - d_ku_k = 0.$$

Since $\{u_1, \dots, u_k, v_1, \dots, v_m\}$ is a basis for W_1, necessarily $b_1 = \dots = b_m = d_1 = \dots = d_k = 0$.
Substitute $b_1 = \dots = b_m$ into ② to get that

$$a_1u_1 + \dots + a_ku_k + c_1z_1 + \dots + c_pz_p = 0.$$

Then, since $\{u_1, \dots, u_k, z_1, \dots, z_p\}$ is a basis for W_2, we have

$$a_1 = \dots = a_k = c_1 = \dots = c_p = 0,$$

proving S is linearly independent.

Subsequently, let $x+y \in W_1+W_2$ be arbitrary, where $x \in W_1$ and $y \in W_2$.
Then, since S_1 and S_2 are bases for W_1 and W_2 respectively, we can write x and y as linear combinations of vectors in S_1 and S_2, respectively:

$$x = a_1u_1 + \dots + a_ku_k + b_1v_1 + \dots + b_mv_m \text{ ; \&}$$
$$y = d_1u_1 + \dots + d_ku_k + c_1z_1 + \dots + c_pz_p,$$

where $a_1, \dots, a_k, b_1, \dots, b_m, d_1, \dots, d_k, c_1, \dots, c_p \in \mathbb{F}$.
Hence

$$x+y = (a_1+d_1)u_1 + \dots + (a_k+d_k)u_k + b_1v_1 + c_1z_1 + \dots + b_mv_m + c_pz_p,$$

which is sufficient to show $x+y \in \text{span}(S)$.
Thus $W_1+W_2 \subseteq \text{span}(S)$, and since $\text{span}(S) \subseteq W_1+W_2$ by definition, it follows that $W_1+W_2 = \text{span}(S)$, verifying that S is indeed a basis for W_1+W_2.
In particular,

$$\begin{aligned}\dim(W_1+W_2) + \dim(W_1 \cap W_2) &= |S| + k \\ &= m+p+k+k \\ &= (m+k) + (p+k) \\ &= \dim W_1 + \dim W_2.\end{aligned}$$

∎

$\dim(V) < \infty$, $W_1 \oplus W_2 = V \Rightarrow \dim W_1 + \dim W_2 = \dim V$ (T1.12 (2))

Let V be a vector space over $\mathbb{F}$, and let W_1, W_2 be finite-dimensional subspaces of V.
Suppose further that V itself is finite dimensional, and $W_1 \oplus W_2 = V$.
Then necessarily $\dim W_1 + \dim W_2 = \dim V$.

Proof. Since $W_1 \oplus W_2 = V$, necessarily $W_1 \cap W_2 = \{0\}$.
So, by T1.12 (1), it follows that

$$\begin{aligned}\dim W_1 + \dim W_2 &= \dim(W_1+W_2) + \dim(W_1 \cap W_2) \\ &= \dim(V) + 0\end{aligned}$$

$\therefore \dim W_1 + \dim W_2 = \dim(V)$. ∎

COMPLEMENTARY SUBSPACES (D15)

Let V be a vector space, and let W be a subspace of V.
Then a subspace W' of V is said to be a "complementary subspace" to W if $W \oplus W' = V$; ie

① $W \cap W' = \{0\}$; and
② $W + W' = V$.

$\dim W + \dim W' = \dim V$

Let V be a vector space, and let W be a subspace of V.
Let W' be a complementary subspace to W.
Then necessarily $\dim W + \dim W' = \dim V$.

Proof. Follows directly from T1.12 (2).

EXISTENCE OF COMPLEMENTARY SUBSPACES (R11 (1))

Let V be a vector space, and let W be a subspace of V.
Then there always exists a complementary subspace W' to W of V such that $W \oplus W' = V$.

Proof. First, note that every finite linearly independent set can be extended to a basis V that has a countable spanning set (A3).
Hence, every linearly independent subset of V can be extended to a basis for V.
It follows that every subspace W of V has a complementary subspace W'. ∎

NON-UNIQUENESS OF COMPLEMENTARY SUBSPACES (R11 (2))

Note that complementary subspaces of a given vector space V are not necessarily unique.

eg $V = \mathbb{R}^3$, $W = \{(1,0,0), (0,1,0)\}$, $W_1' = \{(0,0,1)\}$,
$W_2' = \{(0,0,-1)\}$;
observe that both W_1' and W_2' are complementary subspaces to W.

Chapter 2: Linear Transformations and Matrices

LINEAR TRANSFORMATIONS (§2.1)

Let V and W be vector spaces over the same field $\mathbb{F}$.
Then, we say the function $T: V \to W$ is a "linear transformation" from V to W if

(L1) → ① $T(x+y) = T(x) + T(y) \quad \forall x, y \in V$; and

(L2) → ② $T(cx) = cT(x) \quad \forall x \in V,\ c \in \mathbb{F}$. (D16)

In this case, we say the function $T: V \to W$ is "linear".

T IS LINEAR ⟺ $T(cx+y) = cT(x) + T(y)$ (P2)

Let the function $T: V \to W$, where V and W are vector spaces over the same field $\mathbb{F}$.
Then T is linear if and only if $T(cx+y) = cT(x) + T(y)$ for all $x, y \in V$ and $c \in \mathbb{F}$.

ZERO TRANSFORMATION (E23 (1a))

For any vector spaces V and W, the "zero transformation", given by "$T_0: V \to W$", is defined by $T_0(x) = 0 \quad \forall x \in V$.

IDENTITY TRANSFORMATION (E23 (1b))

For any vector space V, the "identity transformation" $I_V: V \to V$ is given by $I_V(x) = x \quad \forall x \in V$.

$T: V \to \mathbb{F}^n$ BY $T(a_1v_1 + \cdots + a_nv_n) := (a_1, \dots, a_n)$ (E23 (3))

Let V be a finite-dimensional vector space over $\mathbb{F}$, and let $\{v_1, v_2, \dots, v_n\}$ be a basis for V.
Then, the mapping

$$T: V \to \mathbb{F}^n \text{ by } T(a_1v_1 + \cdots + a_nv_n) := (a_1, \dots, a_n)$$

is linear.

$T: \mathbb{F}^n \to \mathbb{F}^k$, $T(x_1, \dots, x_n) := (x_1, \dots, x_k)$ (E23 (4))

Let $\mathbb{F}$ be a field, and suppose $1 \leq k < n$.
Then the projection mapping

$$T: \mathbb{F}^n \to \mathbb{F}^k \text{ by } T(x_1, \dots, x_n) := (x_1, \dots, x_k)$$

is linear.

$T(0) = 0$ (P3 (1))

Let $T: V \to W$ be linear.
Then necessarily $T(0) = 0$.

Proof. $T(0) = T(0+0) = T(0) + T(0)$;
thus $0 = T(0) + T(0) - T(0) = T(0)$. ∎

$T(x-y) = T(x) - T(y)$ (P3(2))

Let $T: V \to W$ be linear.
Then necessarily $T(x-y) = T(x) - T(y) \quad \forall x, y \in V$.

Proof. $T(x-y) = T(x) + T(-y)$
$= T(x) + (-1)T(y)$
$\therefore T(x-y) = T(x) - T(y)$. ∎

$T(a_1x_1 + \cdots + a_nx_n) = a_1T(x_1) + \cdots + a_nT(x_n)$ (P3(3))

Let T be linear, and $a_1, \dots, a_n \in \mathbb{F}$ and $x_1, \dots, x_n \in V$ be arbitrary.
Then necessarily

$$T(a_1x_1 + \cdots + a_nx_n) = a_1T(x_1) + \cdots + a_nT(x_n).$$

$\{v_1, \dots, v_n\}$ IS A BASIS FOR V, $\{w_1, \dots, w_n\}$ ARE ELEMENTS FOR W ⟹ ∃ A UNIQUE LINEAR MAPPING $T: V \to W$ ∋ $T(v_k) = w_k$ (T2.1)

Let $\{v_1, \dots, v_n\}$ be a basis for a vector space V, and let $\{w_1, \dots, w_n\}$ be arbitrary elements of another vector space W.
Then there exists a unique linear mapping $T: V \to W$ such that

$$T(v_1) = w_1, \quad \dots, \quad T(v_n) = w_n.$$

Proof. Let $v \in V$ be arbitrary. Since $\{v_1, v_2, \dots, v_n\}$ is a basis for V, there must exist $a_1, \dots, a_n \in \mathbb{F}$ such that

$$v = a_1v_1 + \cdots + a_nv_n.$$

Let $T(v) = T(a_1v_1 + \cdots + a_nv_n) = a_1w_1 + \cdots + a_nw_n$ (by P3(3)).
Then, by construction, for any $1 \leq k \leq n$, we have

$$\begin{aligned} T(v_k) &= T(0v_1 + \cdots + 0v_{k-1} + 1v_k + 0v_{k+1} + \cdots + 0v_n) \\ &= 0w_1 + \cdots + 0w_{k-1} + 1w_k + 0w_{k+1} + \cdots + 0w_n \\ &= w_k, \end{aligned}$$

proving uniqueness.
Next, suppose there exists another linear mapping $L: V \to W$ satisfying $L(v_1) = w_1, \dots, L(v_n) = w_n$.
Let $v = a_1v_1 + \cdots + a_nv_n$, where $v \in V$ and $a_1, \dots, a_n \in \mathbb{F}$.
Then

$$\begin{aligned} L(v) &= L(a_1v_1 + \cdots + a_nv_n) \\ &= a_1L(v_1) + \cdots + a_nL(v_n) \\ &= a_1w_1 + \cdots + a_nw_n \\ &= a_1T(v_1) + \cdots + a_nT(v_n) \\ &= T(a_1v_1 + \cdots + a_nv_n) \end{aligned}$$

$\therefore L(v) = T(v)$.

Hence $L(v) = T(v) \quad \forall v \in V$, so that $T = L$, proving uniqueness. ∎

It also follows that we have

$$T(a_1v_1 + \cdots + a_nv_n) = a_1w_1 + \cdots + a_nw_n. \qquad (C2.1.1)$$

NULL SPACE / KERNEL (D17(1))

Let V and W be vector spaces, and let $T: V \to W$ be linear.
Then the "null space" of T, or the "kernel" of T, denoted as "$N(T)$", is defined to be the set

$$N(T) := \{x \in V \mid T(x) = 0\}.$$

RANGE / IMAGE (D17(2))

Let V and W be vector spaces, and let $T: V \to W$ be linear.
Then the "range" of T, or the "image" of T, denoted as "$R(T)$", is defined to be the set

$$R(T) := \{T(x) : x \in V\}.$$

N(T) IS A SUBSPACE OF V (T2.2)

Let $T: V \to W$ be linear.
Then necessarily $N(T)$ is a subspace of V.

R(T) IS A SUBSPACE OF W (T2.2)

Let $T: V \to W$ be linear.
Then necessarily $R(T)$ is a subspace of W.

$\{v_1, \dots, v_n\}$ IS A BASIS FOR $V \Rightarrow$ $\text{span}(\{T(v_1), \dots, T(v_n)\}) = R(T)$ (T2.3)

Let V and W be vector spaces, and let $T: V \to W$ be linear.
Suppose the set $\{v_1, \dots, v_n\}$ is a basis for V.
Then necessarily $\{T(v_1), \dots, T(v_n)\}$ generates $R(T)$.

NULLITY (D18)

Let $T: V \to W$ be linear, and suppose that $\dim(N(T)) < \infty$.
Then, we define the "nullity" of T, denoted by "nullity(T)", to be equal to

$$\text{nullity}(T) = \dim(N(T)).$$

RANK (D18)

Let $T: V \to W$ be linear, and suppose that $\dim(R(T)) < \infty$.
Then, we define the "rank" of T, denoted by "rank(T)", to be equal to

$$\text{rank}(T) = \dim(R(T)).$$

RANK-NULLITY THEOREM (T2.4)

Let V and W be vector spaces, and let $T: V \to W$ be linear with $\dim(V) < \infty$.
Then necessarily

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

Proof. Since $N(T)$ is a subspace of V (T2.2) and $\dim V < \infty$, by C1.9.2 (6) necessarily $\text{nullity}(T) \leq \dim(V) < \infty$.
Then, let $\text{nullity}(T) = k$, and suppose that $\{v_1, \dots, v_k\}$ is a basis for $N(T)$.
We know that we can "extend" $\{v_1, \dots, v_k\}$ to get a basis for V, $\{v_1, \dots, v_n\}$, so let us do so.
Next, we claim $\{T(v_{k+1}), \dots, T(v_n)\}$ is a basis for $R(T)$.
First, we show $\{T(v_{k+1}), \dots, T(v_n)\}$ spans $R(T)$.
By T2.2, $R(T) = \text{span}(\{T(v_1), \dots, T(v_k), T(v_{k+1}), \dots, T(v_n)\})$.
Then, since $\{v_1, \dots, v_k\}$ is a basis for $N(T)$, necessarily

$$T(v_1) = \dots = T(v_k) = 0.$$

Hence,

$$R(T) = \text{span}(\{T(v_{k+1}), \dots, T(v_n)\}),$$

as needed.
Next, we show $\{T(v_{k+1}), \dots, T(v_n)\}$ is linearly independent.
Consider

$$c_{k+1}T(v_{k+1}) + \dots + c_nT(v_n) = 0, \quad \text{where } c_{k+1}, \dots, c_n \in \mathbb{F}$$

$$\Rightarrow T(c_{k+1}v_{k+1} + \dots + c_nv_n) = 0.$$

Hence $c_{k+1}v_{k+1} + \dots + c_nv_n \in N(T)$; then, since $\{v_1, \dots, v_k\}$ is a basis for $N(T)$, there exist $d_1, \dots, d_k \in \mathbb{F}$ such that

$$c_{k+1}v_{k+1} + \dots + c_nv_n = d_1v_1 + \dots d_kv_k.$$

$$\Rightarrow -d_1v_1 - \dots - d_kv_k + c_{k+1}v_{k+1} + \dots + c_nv_n = 0.$$

Since $\{v_1, \dots, v_n\}$ is a basis for V, consequently

$$d_1 = \dots = d_k = c_{k+1} = \dots = c_n = 0,$$

showing $\{T(v_{k+1}), \dots, T(v_n)\}$ is linearly independent.
Consequently,

$$\begin{aligned} \text{rank}(T) + \text{nullity}(T) &= \dim(R(T)) + \dim(N(T)) \\ &= k + (n - (k+1) + 1) \\ &= n \end{aligned}$$

$\therefore \text{rank}(T) + \text{nullity}(T) = \dim(V).$ ∎

ONE-TO-ONE (1-1) (D19)

Let $T: V \to W$ be linear.
Then, we say T is "one-to-one" if, for any $x, y \in V$, $T(x) = T(y)$ implies $x = y$.

ONTO (D19)

Let $T: V \to W$ be linear.
Then, we say T is "onto" if $R(T) = V$.

ISOMORPHISM (D19)

Let $T: V \to W$ be linear.
Then, we say T is an "isomorphism" if it is both one-to-one and onto.

We say V is "isomorphic" to W if an isomorphism $T: V \to W$ exists, (D20) and denote this by the notation "$V \cong W$".

T IS 1-1 $\iff N(T)=\{0\}$ (L3)

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

Proof. ($\Rightarrow$) Suppose T is one-to-one.
Let $x \in V$ be such that $T(x)=0$.
Then since $T(0) = 0 = T(x)$, by definition $x=0$, so that $N(T) = \{0\}$.
($\Leftarrow$) Suppose $N(T) = \{0\}$. Consider $x, y \in V$ such that $T(x) = T(y)$.
Then

$$T(x-y) = T(x) - T(y) = 0,$$

so that $x-y \in N(T)$; hence $x-y=0$, so that $x=y$ (and hence T is 1-1). ∎

$\{v_1, \dots, v_n\}$ IS A BASIS FOR $V \Rightarrow$ T IS ISOMORPHIC $\iff \{T(v_1), \dots, T(v_n)\}$ IS A BASIS FOR W (T2.5)

Let V and W be vector spaces over a field $\mathbb{F}$, with $\dim V < \infty$.
Let $\{v_1, \dots, v_n\}$ be a basis for V, and let $T: V \to W$ be linear.
Then T is an isomorphism if and only if $\{T(v_1), \dots, T(v_n)\}$ is a basis for W.

Proof. ($\Rightarrow$) Consider

$$c_1 T(v_1) + \cdots + c_n T(v_n) = 0$$

$$\Rightarrow T(c_1 v_1 + \cdots + c_n v_n) = 0.$$

Since T is one-to-one by definition, hence $c_1 v_1 + \cdots + c_n v_n = 0$,
and as $\{v_1, \dots, v_n\}$ is a basis for V, necessarily $c_1 = \cdots = c_n = 0$;
hence $\{T(v_1), \dots, T(v_n)\}$ is a basis for W. #

($\Leftarrow$) If $\{T(v_1), \dots, T(v_n)\}$ is a basis for W, by definition $\{T(v_1), \dots, T(v_n)\}$ generates W.
Thus $W = \text{span}(\{T(v_1), \dots, T(v_n)\}) = R(T)$,
where the second equality comes from T2.3
Then, since $W = R(T)$, T is necessarily onto.
Let $x \in N(T)$. Since $\{v_1, \dots, v_n\}$ is a basis for V, there must exist some $a_1, \dots, a_n \in \mathbb{F}$ such that

$$x = a_1 v_1 + \cdots + a_n v_n.$$

Hence

$$0 = T(x) = a_1 T(v_1) + \cdots + a_n T(v_n).$$

Since $\{T(v_1), \dots, T(v_n)\}$ is a basis for W by assumption, thus $\{T(v_1), \dots, T(v_n)\}$ is linearly independent, so that

$$a_1 = \cdots = a_n = 0,$$

and so

$$x = 0v_1 + \cdots + 0v_n = 0.$$

Consequently $N(T) = \{0\}$, so that (by L3) T is 1-1. ∎

CONSTRUCTING AN ISOMORPHISM FROM V TO W

Let V and W be vector spaces.
Then, we can construct an isomorphism from V to W as follows:

1. Choose a basis $\{v_1, \dots, v_n\}$ for V, and a basis $\{w_1, \dots, w_n\}$ for W.
2. Let the linear transformation $T: V \to W$ be such that $T(v_k) = w_k$ $\forall k \in \{1, 2, \dots, n\}$.
   (T exists ; this follows from T2.1)
3. Then, by T2.5, T is also an isomorphism.

$V \cong W \iff \dim V = \dim W$ (T2.6)

Let V and W be two finite-dimensional vector spaces over a field $\mathbb{F}$.
Then V is isomorphic to W if and only if $\dim V = \dim W$.

$\dim V = \dim W < \infty$; T IS 1-1 $\iff$ T IS ONTO $\iff \text{rank}(T) = \dim(V)$ (T2.7)

Let V and W be two vector spaces over a field $\mathbb{F}$, and assume $\dim V = \dim W < \infty$.
Let $T: V \to W$ be linear.
Then the following are equivalent to one another:

1. T is one-to-one ;
2. T is onto ; and
3. $\text{rank}(T) = \dim(V)$.

SET OF ALL LINEAR TRANSFORMATIONS (D21)

Let V and W be vector spaces over $\mathbb{F}$.
Then, we let $\mathcal{L}(V, W) \subseteq W^V$ denote the set of all linear transformations $T: V \to W$.

$\mathcal{L}(V, W)$ IS A SUBSPACE OF W^V (T2.8)

Let V and W be vector spaces over some field $\mathbb{F}$.
Then necessarily $\mathcal{L}(V, W)$ is a subspace of W^V.

Proof. Clearly $\mathcal{L}(V, W) \subset W^V$, so we only need to show that it is non-empty and is closed under the addition & scalar multiplication operations of W^V.
Also note the zero transformation $T_0: V \to W$ is in $\mathcal{L}(V, W)$, so that $\mathcal{L}(V, W)$ is non-empty.
Next, assume $T, U \in \mathcal{L}(V, W)$. Note that for any $x, y \in V$ & $c \in \mathbb{F}$:

$$\begin{aligned}(T+U)(cx+y) &= T(cx+y) + U(cx+y) \\ &= cT(x) + T(y) + cU(x) + U(y) \\ &= c(T+U)(x) + (T+U)(y),\end{aligned}$$

showing $T+U$ is linear (by P2), so that $T+U \in \mathcal{L}(V, W)$
A similar argument shows $cT \in \mathcal{L}(V, W)$ as well $\forall c \in \mathbb{F}$.
Thus $\mathcal{L}(V, W)$ is a subspace of W^V, and we are done. ∎

MORE ON MATRICES

TRANSPOSITION OF A MATRIX

Let $A \in M_{m\times n}(\mathbb{F})$ be arbitrary, and write

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \\ a_{m1} & a_{m2} & & a_{mn} \end{pmatrix}.$$

Then the "transposition" of A, denoted as "A^T" (or "A^t"), is defined to be the matrix

$$A^T = \begin{pmatrix} a_{11} & a_{21} & \cdots & a_{m1} \\ a_{12} & a_{22} & \cdots & a_{m2} \\ \vdots & \vdots & \ddots & \\ a_{1n} & a_{2n} & & a_{mn} \end{pmatrix} \in M_{n\times m}(\mathbb{F}).$$

MATRIX VECTOR MULTIPLICATION (D22)

Let $A \in M_{m\times n}(\mathbb{F})$ and $x \in \mathbb{F}^n$ be arbitrary, where $\mathbb{F}$ is some field.

We define "Ax" to be equal to

$$Ax = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \\ a_{m1} & a_{m2} & & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^n a_{1k}x_k \\ \sum_{k=1}^n a_{2k}x_k \\ \vdots \\ \sum_{k=1}^n a_{mk}x_k \end{pmatrix};$$

ie the i^{th} entry of Ax is obtained by multiplying the entries in the i^{th} row of A by the entries of x, and then summing up the resultant products.

$L_A(x) = Ax$ (D23)

Let $\mathbb{F}$ be a field, and let $A \in M_{m\times n}(\mathbb{F})$ be arbitrary.

Then, we let the function $L_A : \mathbb{F}^n \to \mathbb{F}^m$ be defined by $L_A(x) = Ax \ \ \forall x \in \mathbb{F}^n$.

"a_j" MATRIX NOTATION

Let $A \in M_{m\times n}(\mathbb{F})$, and write

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \\ a_{m1} & a_{m2} & & a_{mn} \end{pmatrix}.$$

Then, we use the notation "a_j" to denote

$$a_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix},$$

and we can also write A as

$$A = (a_1 \ a_2 \ \cdots \ a_j).$$

$Ax = x_1a_1 + x_2a_2 + \cdots + x_na_n$ (L4(1))

Let $A \in M_{m\times n}(\mathbb{F})$ be arbitrary, and write $A = (a_1 \ a_2 \ \cdots \ a_n)$.

Then for any $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{F}^n$, we have

$$Ax = x_1a_1 + x_2a_2 + \cdots + x_na_n.$$

$a_j = Ae_j$ (L4 (2))

Let $A \in M_{m\times n}(\mathbb{F})$ be arbitrary, and write $A = (a_1 \ a_2 \ \cdots \ a_n)$.

Suppose $\{e_1, e_2, \ldots, e_n\}$ are the standard basis vectors for $\mathbb{F}^n$.

Then necessarily $Ae_j = a_j$.

MATRIX EQUALITY THEOREM (C2.8.1)

Let $A, B \in M_{m\times n}(\mathbb{F})$ be arbitrary.

Then $A = B$ if and only if $Ax = Bx \ \forall x \in \mathbb{F}^n$.

Proof. ($\Rightarrow$) is obvious.

($\Leftarrow$) Suppose $Ax = Bx \ \forall x \in \mathbb{F}^n$.
This implies $Ae_j = Be_j \ \forall j \in \{1, \ldots, n\}$, which tells us (by L4(2)) that $a_j = b_j \ \forall j \in \{1, \ldots, n\}$.
It follows that $A = B$, as needed. ∎

L_A IS A LINEAR TRANSFORMATION (T2.9)

Let $A \in M_{m\times n}(\mathbb{F})$ be arbitrary.

Then $L_A : \mathbb{F}^n \to \mathbb{F}^m$ is necessarily a linear transformation.

Proof. We prove $L_A(cx+y) = cL_A(x) + L_A(y) \ \forall x,y \in \mathbb{F}^n$ & $c \in \mathbb{F}$; the result follows from P2.
Write $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$, and $A = (a_1 \ a_2 \ \cdots \ a_n)$.

Then

$$\begin{aligned} L_A(cx+y) &= A(cx+y) \\ &= (cx_1+y_1)a_1 + (cx_2+y_2)a_2 + \cdots + (cx_n+y_n)a_n \quad \text{(by L4(1))} \\ &= c(x_1a_1 + \cdots + x_na_n) + (y_1a_1 + \cdots + y_na_n) \\ &= c(Ax) + Ay \end{aligned}$$

$\therefore L_A(cx+y) = cL_A(x) + L_A(y)$, as needed. ∎

$L: M_{m\times n}(\mathbb{F}) \to \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ BY $L(A) = L_A$ IS A 1-1 LINEAR TRANSFORMATION (P4)

Let $L : M_{m\times n}(\mathbb{F}) \to \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ by $L(A) = L_A \ \ \forall A \in M_{m\times n}(\mathbb{F})$, where $\mathbb{F}$ is a field.

Then L is necessarily a one-to-one linear transformation.

Proof. We first show L is linear.
By P2, we just need to show $L(cA+B) = cL(A) + L(B)$ $\forall A, B \in M_{m\times n}(\mathbb{F})$, $c \in \mathbb{F}$; ie $L_{cA+B} = cL_A + L_B$.
To do this, let $x \in \mathbb{F}^n$ be arbitrary.
Write $A = (a_1 \ a_2 \ \cdots \ a_n)$ and $B = (b_1 \ b_2 \ \cdots \ b_n)$,
so that $cA + B = (ca_1 + b_1 \ \ ca_2 + b_2 \ \ \cdots \ \ ca_n + b_n)$.
So

$$\begin{aligned} L_{cA+B}(x) &= (cA+B)x \\ &= x_1(ca_1+b_1) + x_2(ca_2+b_2) + \cdots + x_n(ca_n+b_n) \quad \text{(by L4(1))} \\ &= c(x_1a_1 + \cdots + x_na_n) + (x_1b_1 + \cdots + x_nb_n) \\ &= c(Ax) + Bx \\ &= cL_A(x) + L_B(x) \end{aligned}$$

$\therefore L_{cA+B}(x) = (cL_A + L_B)(x)$,

and since $x \in \mathbb{F}^n$ was arbitrary this is sufficient to prove $L_{cA+B} = cL_A + L_B$, as needed. ∗

Next, we prove L is 1-1.
Assume for some $A, B \in M_{m\times n}(\mathbb{F})$, we have $L_A = L_B$.
This means $L_A(x) = L_B(x) \ \forall x \in \mathbb{F}^n$, or $Ax = Bx \ \forall x \in \mathbb{F}^n$.
So by the Matrix Equality Theorem, $A = B$, which is sufficient to prove L is 1-1. ∎

COORDINATES (S2.2)

ORDERED BASIS (D24)

Let V be a vector space with $\dim V < \infty$. Then, an "ordered basis" for V is a basis $\{v_1, \dots, v_n\}$ with a total order.

eg $\{e_1, e_2, e_3\}$ is the standard ordered basis for $\mathbb{F}^3$, since we can define a "total order" by saying the indexes must be in "increasing order" (E30 (1))

COORDINATE VECTOR (D25)

Let $\beta = \{u_1, \dots, u_n\}$ be an "ordered basis" for a finite-dimensional vector space V. By T1.6, we can write any $x \in V$ in the form $x = \sum_{k=1}^{n} a_k u_k$, where $a_1, \dots, a_n \in \mathbb{F}$. Then, we define the "coordinate vector" of x relative to β, denoted as "$[x]_\beta$", to be

$$[x]_\beta := \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{F}^n.$$

eg for $V = P_2(\mathbb{R})$, $\beta = \{1, x, x^2\}$, $p(x) = 2 - 3x + 4x^2 \in V$, $[p(x)]_\beta = \begin{pmatrix} 2 \\ -3 \\ 4 \end{pmatrix}$.

$[\ ]_\beta : V \to \mathbb{F}^n$ IS AN ISOMORPHISM (T2.10)

Let V be a vector space over some field $\mathbb{F}$, with $\dim V = n$, and let β be an ordered basis for V. Then, the map $[\ ]_\beta : V \to \mathbb{F}^n$ is an isomorphism.

MATRIX REPRESENTATION OF LINEAR TRANSFORMATIONS (S2.3)

Let V and W be finite-dimensional vector spaces over $\mathbb{F}$, and let $T: V \to W$ be a linear transformation.

Let $\beta = \{v_1, \dots, v_n\}$ be an ordered basis for V, and let $\gamma = \{w_1, \dots, w_n\}$ be an ordered basis for W.

Then, the "matrix representation" of T in the ordered bases β and γ, denoted as $[T]_\beta^\gamma$, is defined as the matrix

$$[T]_\beta^\gamma := \left([T(v_1)]_\gamma \quad [T(v_2)]_\gamma \quad \cdots \quad [T(v_n)]_\gamma\right).$$

In particular, if $T: V \to V$ is linear and β is an ordered basis of the finite-dimensional vector space V, we denote

$$[T]_\beta := [T]_\beta^\beta.$$

Note that $[T]_\beta^\gamma \in M_{m\times n}(\mathbb{F})$, where $m = \dim W$ and $n = \dim V$. (R12 (1))

Also, we have

$$T(v_j) = \sum_{k=1}^{m} a_{kj} w_k,$$

where a_{kj} denotes the element at the k^{th} row and j^{th} column in the matrix $[T]_\beta^\gamma$. (R12 (2))

eg If $T: P_2(\mathbb{R}) \to \mathbb{R}^2$ by $T(a+bx+cx^2) = \begin{pmatrix} a \\ b+4c \end{pmatrix}$, we can verify T is linear.

Let $\beta = \{1, (x+1), (x+1)^2\}$ and $\gamma = \left\{\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}\right\}$.

Then

$$[T]_\beta^\gamma = \left([T(1)]_\gamma \quad [T(x+1)]_\gamma \quad [T(x+1)^2]_\gamma\right)$$

$$= \begin{pmatrix} 1 & 1 & 1 \\ 0+4(0) & 1+4(0) & 2+4(1) \end{pmatrix}$$

$$\therefore [T]_\beta^\gamma = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 6 \end{pmatrix}. \qquad \text{(E32)}$$

$[L_A]_\beta^\gamma = A$ (E33)

Let $A \in M_{m\times n}(\mathbb{F})$, where $\mathbb{F}$ is a field.

Let β be the standard ordered basis for $\mathbb{F}^n$, and γ the standard ordered basis for $\mathbb{F}^m$.

Then necessarily $[L_A]_\beta^\gamma = A$.

$[T(x)]_\gamma = [T]_\beta^\gamma \cdot [x]_\beta$ (T2.11)

Let $T: V \to W$ be linear, and let $\beta = \{v_1, \dots, v_n\}$ and $\gamma = \{w_1, \dots, w_n\}$ be ordered bases of V and W respectively.

Then necessarily $[T(x)]_\gamma = [T]_\beta^\gamma \cdot [x]_\beta \quad \forall x \in V$.

Proof. Let $x \in V$ be arbitrary. Take $x = \sum_{k=1}^{n} a_k v_k$, where $[x]_\beta = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$.

Then, since T is linear,

$$T(x) = T\left(\sum_{k=1}^{n} a_k v_k\right) = \sum_{k=1}^{n} a_k T(v_k).$$

Thus

$$[T(x)]_\gamma = \left[\sum_{k=1}^{n} a_k T(v_k)\right]_\gamma = \sum_{k=1}^{n} a_k [T(v_k)]_\gamma. \quad \text{(by linearity of } [\,]_\gamma \text{ (T2.10))}$$

Note that

$$\sum_{k=1}^{n} a_k [T(v_k)]_\gamma = \left([T(v_1)]_\gamma \quad [T(v_2)]_\gamma \cdots [T(v_n)]_\gamma\right)\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

$$= [T]_\gamma^\beta \cdot [x]_\beta,$$

so that

$$[T(x)]_\gamma = [T]_\gamma^\beta \cdot [x]_\beta, \quad \text{as needed. } \blacksquare$$

$[\;]_\beta^\gamma : \mathcal{L}(V,W) \to M_{m\times n}(\mathbb{F})$ IS AN ISOMORPHISM (P5)

Let V and W be finite-dimensional vector spaces over $\mathbb{F}$, and let β and γ be ordered bases of V and W respectively.

Then the map $[\;]_\beta^\gamma : \mathcal{L}(V,W) \to M_{m\times n}(\mathbb{F})$ is an isomorphism, where $m = \dim W$ and $n = \dim V$;

in other words,

① For any $T, U \in \mathcal{L}(V,W)$ and $c \in \mathbb{F}$, we have that
$$[cT + U]_\beta^\gamma = c[T]_\beta^\gamma + [U]_\beta^\gamma; \quad \text{and}$$

② For any $C \in M_{m\times n}(\mathbb{F})$, there exists a unique $T \in \mathcal{L}(V,W)$ such that $[T]_\beta^\gamma = C$.

Proof. We first prove ①.

Let $\beta = \{v_1, \dots, v_n\}$. Then

$$[T+U]_\beta^\gamma = \left([(T+U)(v_1)]_\gamma \quad [(T+U)(v_2)]_\gamma \quad \cdots \quad [(T+U)(v_n)]_\gamma\right)$$

$$= \left([T(v_1)+U(v_1)]_\gamma \quad [T(v_2)+U(v_2)]_\gamma \quad \cdots \quad [T(v_n)+U(v_n)]_\gamma\right)$$

$$= \left(([T(v_1)]_\gamma + [U(v_1)]_\gamma) \quad ([T(v_2)]_\gamma + [U(v_2)]_\gamma) \quad \cdots \quad ([T(v_n)]_\gamma + [U(v_n)]_\gamma)\right)$$

$$= \left([T(v_1)]_\gamma \quad [T(v_2)]_\gamma \quad \cdots \quad [T(v_n)]_\gamma\right) + \left([U(v_1)]_\gamma \quad [U(v_2)]_\gamma \quad \cdots \quad [U(v_n)]_\gamma\right)$$

$$\therefore [T+U]_\beta^\gamma = [T]_\beta^\gamma + [U]_\beta^\gamma,$$

and a similar proof shows $[cT]_\beta^\gamma = c[T]_\beta^\gamma$, which is sufficient to show $[cT+U]_\beta^\gamma = c[T]_\beta^\gamma + [U]_\beta^\gamma$, and hence that the map $[\;]_\beta^\gamma : \mathcal{L}(V,W) \to M_{m\times n}$ is linear. ※

We next prove ②.

Suppose $[T]_\beta^\gamma = [U]_\beta^\gamma$, so that $[T]_\beta^\gamma$ and $[U]_\beta^\gamma$ have the same j^{th} column $\forall j \in \{1, \dots, n\}$.

This means $[T(v_j)]_\gamma = [U(v_j)]_\gamma$, and since $[\;]_\gamma : W \to \mathbb{F}^n$ is a bijection (by T2.10) it follows that $T(v_j) = U(v_j) \quad \forall j \in \{1, 2, \dots, n\}$.

So, by T2.1, $T = U$, proving injectivity.

Then, let $C = (c_1 \quad c_2 \quad \cdots \quad c_n) \in M_{m\times n}(\mathbb{F})$ be arbitrary.

For each $j \in \{1, \dots, n\}$, let $w_j \in W$ be the unique vector satisfying $[w_j]_\gamma = c_j$.

By T2.10, there exists a unique linear transformation $T: V \to W$ satisfying $T(v_j) = w_j \quad \forall j \in \{1, \dots, n\}$;

it follows this T satisfies $[T]_\beta^\gamma = C$, proving surjectivity,

so we are done. $\blacksquare$

$L: M_{m\times n}(\mathbb{F}) \to \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ IS AN ISOMORPHISM (C2.11.1)

Recall that the map $L: M_{m\times n}(\mathbb{F}) \to \mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)$ is defined by $L(A) = L_A \quad \forall A \in M_{m\times n}(\mathbb{F})$.

Then, L is necessarily an isomorphism.

Proof. We know L is already 1-1 & linear by P4, so we only need to prove it is onto.

Applying P5 to $V = \mathbb{F}^n$ & $W = \mathbb{F}^m$, we get that $\mathcal{L}(\mathbb{F}^n, \mathbb{F}^m) \cong M_{m\times n}(\mathbb{F})$, so that

$\dim(\mathcal{L}(\mathbb{F}^n, \mathbb{F}^m)) = \dim(M_{m\times n}(\mathbb{F})) = mn$ (by T2.6).

So L is a 1-1 linear transformation between vector spaces of the same finite dimension;

it follows by T2.7 that L is onto. $\blacksquare$

MATRIX MULTIPLICATION & COMPOSITIONS OF LINEAR TRANSFORMATIONS (S2.4)

MATRIX PRODUCT (D27)

Let $\mathbb{F}$ be a field, and let $A \in M_{m\times n}(\mathbb{F})$ and $B \in M_{n\times p}(\mathbb{F})$ be arbitrary.

Note the number of columns in A equals the number of rows in B; this is required.

Then, the matrix product of A and B, denoted by AB, is defined to be the $m\times p$ matrix

$$AB = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{np} \end{pmatrix}$$

$$= \begin{pmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{pmatrix},$$

where $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj} = \sum_{t=1}^{n} a_{it}b_{tj}$.

In other words, c_{ij} is the sum of products formed multiplying the entries in the i^{th} row of A with the j^{th} column of B.

* an example is highlighted in blue;
$c_{21} = a_{21}b_{11} + a_{22}b_{21} + \cdots + a_{2n}b_{n1}$.

Note that c_j is the linear combination of the columns of A formed using the entries in the j^{th} column of B as coefficients. (R13(3))

ZERO MATRIX

The "zero matrix", denoted by the letter O, is defined to be the matrix with each entry being zero.

We write "$O_{m\times n}$" to denote the $m\times n$ zero matrix.

IDENTITY MATRIX

The "$n\times n$ identity matrix", denoted as I_n, is defined as the matrix (δ_{ij}) with

$$\delta_{ij} = \begin{cases} 1, & i=j \\ 0, & i\neq j. \end{cases}$$

* δ_{ij} is known as the "Kronecker delta".

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

RIGHT MATRIX DISTRIBUTIVE LAW (L5(1))

For any $A \in M_{m\times n}(\mathbb{F})$ and $B, C \in M_{n\times p}(\mathbb{F})$, we have

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

LEFT MATRIX DISTRIBUTIVE LAW (L5(2))

Similarly, for any $A \in M_{m\times n}(\mathbb{F})$ and $D, E \in M_{q\times m}(\mathbb{F})$, we have

$(D+E)A = DA + EA$.

ASSOCIATIVITY OF MATRIX SCALAR MULTIPLICATION (L5(3))

For any $\alpha \in \mathbb{F}$, $A \in M_{m\times n}(\mathbb{F})$ and $B \in M_{q\times m}(\mathbb{F})$, we have

$\alpha(AB) = (\alpha A)B = A(\alpha B)$.

$(AB)^T = B^T A^T$ (L5(4))

For any $A \in M_{m\times m}(\mathbb{F})$ and $B \in M_{n\times n}(\mathbb{F})$, we have

$(AB)^T = B^T A^T$.

$I_m A = A I_n$ (L5(5))

For any $A \in M_{m\times n}(\mathbb{F})$, we have that $I_m A = A I_n = A$.

$AO_{n\times p} = O_{m\times p}$, $O_{q\times m}A = O_{q\times n}$ (L5(6))

For any $A \in M_{m\times n}(\mathbb{F})$, we have
① $AO_{n\times p} = O_{m\times p}$; and
② $O_{q\times m}A = O_{q\times n}$.

COMPOSITION OF LINEAR TRANSFORMATIONS IS ALSO A LINEAR TRANSFORMATION (T2.12)

Let $T: V \to W$ and $U: W \to Z$ be linear transformations.
Then the composition $(U\circ T): V \to Z$ is also a linear transformation.

* we usually denote $(U\circ T)$ as UT.

MATRIX OF COMPOSITION OF LINEAR TRANSFORMATIONS (T2.13)

Let V, W and Z be finite-dimensional vector spaces having ordered bases $\alpha = \{v_1, \dots, v_p\}$, $\beta = \{w_1, \dots, w_n\}$ and $\gamma = \{z_1, \dots, z_n\}$ respectively.

Let $T: V\to W$ and $U: W\to Z$ be linear transformations.

Denote $A = [U]_\beta^\gamma \in M_{m\times n}(\mathbb{F})$, $B = [T]_\alpha^\beta \in M_{n\times p}(\mathbb{F})$ and $C = [UT]_\alpha^\gamma \in M_{m\times p}(\mathbb{F})$.

Then necessarily $C = AB$; ie $[UT]_\alpha^\gamma = [U]_\beta^\gamma \cdot [T]_\alpha^\beta$.

Proof. Note that both sides are $m\times p$ matrices.
We show that the j^{th} columns of the LHS & RHS are equal $\forall j \in \{1, \dots, p\}$.
On one hand, the j^{th} column of $[UT]_\alpha^\gamma$ is $[(UT)(v_j)]_\gamma$.
On the other hand, $[T]_\alpha^\beta = B = (b_1\ b_2 \cdots b_p)$.
Hence, the j^{th} column of $[U]_\beta^\gamma \cdot [T]_\alpha^\beta$ is $[U]_\beta^\gamma \cdot b_j$, which equals

$$\begin{aligned} [U]_\beta^\gamma \cdot b_j &= [U]_\beta^\gamma \cdot [T(v_j)]_\beta \\ &= [U(T(v_j))]_\gamma \quad (\text{by T2.11}) \\ &= [(UT)(v_j)]_\gamma. \end{aligned}$$

It follows that $[UT]_\alpha^\gamma$ and $[U]_\beta^\gamma \cdot [T]_\alpha^\beta$ have the same j^{th} columns; since j was arbitrary, it follows that $[UT]_\alpha^\gamma = [U]_\beta^\gamma \cdot [T]_\alpha^\beta$, as needed. ☒

$L_{AB} = L_A L_B$ (C2.13.1(1))

Let $A \in M_{m\times n}(\mathbb{F})$ and $B = M_{n\times p}(\mathbb{F})$ be arbitrary.
Then necessarily $L_{AB} = L_A L_B$.

Proof. Let α, β & γ denote the standard ordered bases for $\mathbb{F}^p$, $\mathbb{F}^n$ and $\mathbb{F}^m$ respectively.
By E33, $[L_A]_\beta^\gamma = A$, $[L_B]_\alpha^\beta = B$ and $[L_{AB}]_\alpha^\gamma = AB$.
On the other hand
$[L_A L_B]_\alpha^\gamma = [L_A]_\beta^\gamma \cdot [L_B]_\alpha^\beta = AB = [L_{AB}]_\alpha^\gamma$ (by T2.13).
Since the mapping $[\ ]_\alpha^\gamma$ is 1-1 (by C2.11.1), it follows that $L_A L_B = L_{AB}$, as needed. ☒

$A(BC) = (AB)C$ (C2.13.1(2))

Assume the matrix product "$A(BC)$" is defined.
Then necessarily $A(BC) = (AB)C$.

Proof. By C2.13.1(1), we get that
$L_{A(BC)} = L_A L_{BC} = L_A(L_B L_C) = (L_A L_B)L_C = L_{AB}L_C = L_{(AB)C}$,
since function composition is associative.
Then, as L is 1-1 (by P4), it follows that $A(BC) = (AB)C$, as needed. ☒

INVERTIBILITY & ISOMORPHISMS (S2.5)

INVERTIBLE MATRICES (D28)

Let $A \in M_{n \times n}(\mathbb{F})$ be arbitrary.
Then, we say A is "invertible" if there exists a matrix $B \in M_{n\times n}(\mathbb{F})$ such that $AB = BA = I_n$.

Note that if such a matrix B exists, it is uniquely determined by A.

Proof. Suppose $B, C \in M_{n\times n}(\mathbb{F})$ are such that $AB = BA = I_n$ & $AC = CA = I_n$.
Then
$$B = BI_n = B(AC) = (BA)C = I_nC = C,$$
proving uniqueness. ☒

INVERSE MATRICES (D28)

Let $A \in M_{n\times n}(\mathbb{F})$ be an invertible square matrix.
Then the "inverse" of A, denoted as "A^{-1}", is the unique $n\times n$ square matrix such that $AA^{-1} = A^{-1}A = I_n$.

INVERTIBLE MAPPING (D29)

Let $T: V \to W$ be a linear mapping between two vector spaces V and W.
Then, we say T is "invertible" if there exists a function $U: W \to V$ such that $UT = I_V$ and $TU = I_W$.

INVERSE MAPPING (D29)

Let $T: V \to W$ be an invertible linear mapping.
Then the "inverse" of T, denoted as "T^{-1}", is the mapping $T^{-1}: W \to V$ such that $TT^{-1} = I_V$ and $T^{-1}T = I_W$.

Similarly, we can show T^{-1} is unique. (T2.14(1))

Proof. Suppose there exist $U_1, U_2 : W \to V$ such that $U_1T = I_V$, $TU_1 = I_W$, $U_2T = I_V$ & $TU_2 = I_W$.
Then
$$U_1 = U_1 I_W = U_1(TU_2) = (U_1T)U_2 = I_VU_2 = U_2,$$
proving uniqueness. ☒

T IS LINEAR & INVERTIBLE ⇒ T IS AN ISOMORPHISM (T2.14(2))

Let $T: V\to W$ be linear and invertible.
Then T is necessarily an isomorphism.

Proof. Suppose $x, y \in V$ are such that $T(x) = T(y)$.
Then observe that
$$x = (T^{-1}T)(x) = T^{-1}(T(x)) = T^{-1}(T(y)) = y,$$
proving injectivity.
Then, let $z \in W$. Since $TT^{-1} = I_W$, we have that
$$z = I_W(z) = (TT^{-1})(z) = T(T^{-1}(z)).$$
Since $T^{-1}(z) \in V$, it follows that T is surjective.
Hence T is bijective, and since T is also linear, it follows that T is an isomorphism. ☒

T^{-1} IS ALSO LINEAR (T2.14(3))

Let $T: V \to W$ be linear and invertible.
Then T^{-1} is necessarily also linear.

Proof. Let $y_1, y_2 \in V$ and $c \in \mathbb{F}$ be arbitrary.
Since T is bijective (by T2.14(2)), there exist unique $x_1, x_2 \in V$ such that $T(x_1) = y_1$ and $T(x_2) = y_2$.
Then
$$T^{-1}(cy_1 + y_2) = T^{-1}(cT(x_1) + T(x_2)) = T^{-1}(T(cx_1 + x_2))$$
$$= cx_1 + x_2 = cT^{-1}(y_1) + T^{-1}(y_2),$$
and it follows from P2 that T^{-1} is linear. ☒

T IS AN ISOMORPHISM ⇔ $[T]_\alpha^\beta$ IS INVERTIBLE (T2.15(1))

Let V and W be finite-dimensional vector spaces, and let α and β be (ordered) bases of V and W respectively.
Let $T: V \to W$ be linear.
Then T is an isomorphism if and only if $[T]_\alpha^\beta$ is an invertible matrix.

Proof. (⇒) Suppose T is an isomorphism, so that $V \cong W$.
Then, by T2.6, $\dim V = \dim W = n$.
Let $A := [T]_\alpha^\beta$. By the above, A is a $n\times n$ square matrix.
By T2.14(3), $T^{-1}: W \to V$ is also linear.
Let $B := [T^{-1}]_\beta^\alpha$, which is also a $n\times n$ matrix.
Also,
$$AB = [T]_\alpha^\beta [T^{-1}]_\beta^\alpha = [TT^{-1}]_\beta^\beta \quad (T2.12)$$
$$= [I_W]_\beta^\beta$$
$$= I_n.$$
A similar proof shows $BA = [I_V]_\alpha^\alpha = I_n$. So, by D28, A is an invertible matrix, proving the forward argument.

(⇐) Suppose $A = [T]_\alpha^\beta$ is an invertible matrix with inverse A^{-1}.
In particular, A must be square, say $n\times n$, so $\dim V = \dim W = n$.
Then, let $x, y \in V$ such that $T(x) = T(y)$. By T2.11,
$$A[x]_\alpha = [T]_\alpha^\beta [x]_\alpha = [T(x)]_\beta = [T(y)]_\beta = [T]_\alpha^\beta [y]_\alpha = A[y]_\alpha.$$
Thus $A[x]_\alpha = A[y]_\alpha$. It follows that
$$A^{-1}(A[x]_\alpha) = A^{-1}(A[y]_\alpha),$$
or $[x]_\alpha = [y]_\alpha$ and so $x = y$, proving injectivity.
Then, as T is linear and $\dim V = \dim W$, by T2.7 T is also onto.
Hence T is bijective, and since T is also linear, it follows that T is an isomorphism, proving the backward argument. ☒

In particular, if T is an isomorphism, then
$$[T^{-1}]_\beta^\alpha = ([T]_\alpha^\beta)^{-1}.$$

$A \in M_{n\times n}(\mathbb{F}) \Rightarrow$ (L_A IS AN ISOMORPHISM ⇔ A IS INVERTIBLE) (T2.15(2))

Let V and W be finite-dimensional vector spaces, and let α and β be ordered bases of V and W respectively.
Then for any $A \in M_{n\times n}(\mathbb{F})$, necessarily L_A is an isomorphism if and only if A is invertible.

Proof. By T2.15(1), L_A is an isomorphism if and only if $[L_A]_{\sigma_n}^{\sigma_n}$ is invertible, where σ_n is the standard ordered basis for $\mathbb{F}^n$.
By E33, $[L_A]_{\sigma_n}^{\sigma_n} = A$, and this is sufficient to prove the claim. ☒

A IS INVERTIBLE ⇒ A^{-1} IS INVERTIBLE & $(A^{-1})^{-1} = A$ (L6 (1))

Let A be an invertible matrix.

Then A^{-1} is also invertible, and $(A^{-1})^{-1} = A$.

Proof. Since A is invertible, A^{-1} exists.

In particular, $A^{-1}A = I_n$.

By uniqueness of matrix inverses, it follows that $A = (A^{-1})^{-1}$, as needed. ☒

$(cA)^{-1} = \frac{1}{c}A^{-1}$ (L6(2))

Let A be an invertible matrix, and let $c \in \mathbb{F}$.

Then necessarily $(cA)^{-1} = \frac{1}{c}A^{-1}$.

$(A^T)^{-1} = (A^{-1})^T$ (L6(3))

Let A be an invertible matrix.

Then necessarily $(A^T)^{-1} = (A^{-1})^T$.

$(AB)^{-1} = B^{-1}A^{-1}$ (L6(4))

Let $A, B \in M_{n\times n}(\mathbb{F})$ be invertible matrices.

Then AB is also invertible, and necessarily $(AB)^{-1} = B^{-1}A^{-1}$.

Proof. $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = AI_nA^{-1} = AA^{-1} = I_n$.

By uniqueness of matrix inverses, it follows that $(AB)^{-1} = B^{-1}A^{-1}$. ☒

AB IS INVERTIBLE ⇒ A & B ARE INVERTIBLE (L6 (5))

Let $A, B \in M_{n\times n}(\mathbb{F})$ be such that AB is invertible.

Then necessarily A and B are also invertible matrices.

Proof. By T2.15, L_{AB} is invertible. By T2.14, L_AL_B is an isomorphism.

Thus, L_AL_B is 1-1 and onto. Thus L_A is surjective and L_B is injective.

Then, as L_A and L_B are both linear mappings from $\mathbb{F}^n$ to itself, by T2.7 L_A and L_B are isomorphisms.

Hence A and B are invertible by T2.15(2), and we are done. ☒

INVERSE MATRIX THEOREM, PART I (T2.16)

Let $A \in M_{n\times n}(\mathbb{F})$. Then the following statements are equivalent:

1. A is invertible;
2. There exists a matrix $C \in M_{n\times n}(\mathbb{F})$ such that $AC = I_n$; and
3. There exists a matrix $B \in M_{n\times n}(\mathbb{F})$ such that $BA = I_n$.

Proof. This follows directly from the definition of inverse matrices. ☒

THE CHANGE OF COORDINATE MATRIX (S2.6)

CHANGE OF COORDINATE MATRIX FROM α TO β (T2.17(1))

Let α and β be two ordered bases for a finite-dimensional vector space V.

Then the "change of coordinate matrix from α to β" is the matrix

$$Q = [I_V]_\alpha^\beta.$$

The matrix $Q = [I_V]_\alpha^\beta$ is necessarily invertible.

Proof. Since I_V is an isomorphism, by T2.15, Q is necessarily invertible. ☒

Also, note that if we let $\alpha = \{v_1, \dots, v_n\}$ and $\beta = \{w_1, \dots, w_n\}$ and fix an $x \in V$, then

$$[I_V]_\alpha^\beta = ([v_1]_\beta \quad \cdots \quad [v_n]_\beta).$$

Then, by comparing the j^{th} column on both sides, we have that

$$v_j = \sum_{i=1}^{n} Q_{ij} w_i. \quad (R15)$$

$[x]_\beta = Q[x]_\alpha$ (T2.17 (2))

Let α and β be two ordered bases of the finite-dimensional vector space V.

Let $Q = [I_V]_\alpha^\beta$ be the change of coordinate matrix from α to β.

Then necessarily for any $x \in V$, we have $[x]_\beta = Q[x]_\alpha$.

Proof. By T2.11, we have

$$[x]_\beta = [I_V(x)]_\beta = [I_V]_\alpha^\beta [x]_\alpha = Q[x]_\alpha,$$

as needed. ☒

$T: V \to V$; $[T]_\alpha = Q^{-1}[T]_\beta Q$ (T2.18)

Let $T: V \to V$ be linear, where V is a finite-dimensional vector space.

Let α and β be two ordered bases of V, and let $Q = [I_V]_\alpha^\beta$.

Then necessarily $[T]_\alpha = Q^{-1}[T]_\beta Q$.

Proof. By T2.13, we have

$$Q[T]_\alpha = [I_V]_\alpha^\beta [T]_\alpha^\alpha = [I_V T]_\alpha^\beta = [T]_\alpha^\beta,$$

and

$$[T]_\beta Q = [T]_\beta^\beta [I_V]_\alpha^\beta = [TI_V]_\alpha^\beta = [T]_\alpha^\beta,$$

showing $[T]_\beta Q = Q[T]_\alpha$.

Then, since Q is invertible, Q^{-1} exists; hence

$$Q^{-1}[T]_\beta Q = Q^{-1}Q[T]_\alpha = [T]_\alpha,$$

as needed. ☒

SIMILAR MATRICES (D30)

Let $A, B \in M_{n\times n}(\mathbb{F})$ be arbitrary.

Then, we say B is "similar" to A if there exists an invertible matrix Q such that

$$B = Q^{-1}AQ.$$

Chapter 3: Elementary Matrix Operations and Systems of Linear Equations

ELEMENTARY MATRIX OPERATIONS & ELEMENTARY MATRICES (S3.1)

ELEMENTARY ROW/COLUMN OPERATIONS (D31)

Let $A \in M_{m \times n}(\mathbb{F})$. Then, we denote the following as "elementary row/column operations" on A:

1. Interchanging any two rows/columns of A, denoted as "$R_i \leftrightarrow R_j$";

eg $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_3} \begin{pmatrix} 7 & 8 & 9 \\ 4 & 5 & 6 \\ 1 & 2 & 3 \end{pmatrix}$

2. Multiplying any row/column by a non-zero scalar, denoted as "$R_i \leftarrow cR_i$"; and

eg $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \xrightarrow{R_2 \leftarrow 2R_2} \begin{pmatrix} 1 & 2 & 3 \\ 8 & 10 & 12 \\ 7 & 8 & 9 \end{pmatrix}$

3. Adding any scalar multiple of a row/column of A to another row/column, denoted as "$R_i \leftarrow R_i + cR_j$".

eg $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \xrightarrow{R_2 \leftarrow R_2 + 2R_1} \begin{pmatrix} 1 & 2 & 3 \\ 6 & 9 & 12 \\ 7 & 8 & 9 \end{pmatrix}$

nxn ELEMENTARY MATRIX (D32)

1. An "nxn elementary matrix" is a matrix obtained by performing an elementary operation on I_n.

eg performing "$R_3 \leftarrow R_3 + 4R_1$" on I_3 results in

$$I' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix}. \qquad \text{(E39)}$$

2. Let $A \in M_{m \times n}(\mathbb{F})$, and suppose B is obtained from A by performing an elementary row operation. Then necessarily $B = \underline{E}A$, where E is the nxn elementary matrix obtained from I_n by performing the said elementary row/column operation. (T3.1)

3. Conversely, if E is an mxm elementary matrix, then $\underline{E}A$ is the matrix obtained from A by performing the same elementary row operation as that which produces E from I_m. (T3.1)

* a similar result holds for elementary matrices formed by performing an elementary <u>column</u> operation, but in this case $B = A\underline{\underline{E}}$. (T3.2)

Proof. This can be proven by verifying each of the three elementary row/column operations "holds" under this transformation.

ELEMENTARY MATRICES ARE INVERTIBLE, & THE INVERSE OF AN ELEMENTARY MATRIX IS OF THE SAME "TYPE" (T3.3)

Note that any elementary matrix $A \in M_{m \times m}(\mathbb{F})$ is invertible, and A^{-1} is also an elementary matrix with the same "type" as A.

Proof. Suppose A is an elementary matrix obtained from I_m. Then, we verify this theorem for each of the three operations;

1. $R_i \leftrightarrow R_j$;
2. $R_i \leftarrow c \cdot R_i$; and
3. $R_i \leftarrow$

THE RANK OF A MATRIX & MATRIX INVERSES (S3.2)

RANK OF A MATRIX (D33)

Let $A \in M_{m\times n}(\mathbb{F})$ be arbitrary.
Then, we define the "rank" of A, denoted as "$\text{rank}(A)$", to be the rank of the linear transformation $L_A : \mathbb{F}^n \to \mathbb{F}^m$ by $L_A(x) = Ax$ $\forall x \in \mathbb{F}^n$.
In other words,
$\text{rank}(A) = \dim(R(L_A)) = \dim(L_A(\mathbb{F}^n))$.

Note that
① $\text{rank}(I_n) = \dim(R(I_n)) = \dim(\mathbb{F}^n) = n$; and
② $\text{rank}(O) = \dim(R(O)) = \dim(\{0\}) = 0$. (E40)
(where O denotes the zero matrix)

$\text{rank}(A) = \dim(\text{span}(\{a_1, \dots, a_n\}))$ (R16(1))

For any matrix $A \in M_{m\times n}(\mathbb{F})$, we have
$\text{rank}(A) = \dim(\text{span}(\{a_1, \dots, a_n\}))$,
where a_j denotes the j^{th} column of A.

Proof. Let $\{e_1, \dots, e_n\}$ be the standard (ordered) basis for $\mathbb{F}^n$. Then
$R(L_A) = \text{span}(\{L_A(e_1), \dots, L_A(e_n)\})$ (by T2.3)
$= \text{span}(\{Ae_1, \dots, Ae_n\})$
$\therefore R(L_A) = \text{span}(\{a_1, \dots, a_n\})$ (by L4),
so that $\dim(R(L_A)) = \text{rank}(A) = \dim(\text{span}(\{a_1, \dots, a_n\}))$, as needed. ∎

$\text{rank}(A) \le \min(m,n)$ (R16(2))

Moreover, for any matrix $A \in M_{m\times n}(\mathbb{F})$, we have that $\text{rank}(A) \le \min(m,n)$.

Proof. Since $\{a_1, \dots, a_n\}$ generates $R(L_A)$ by the above, and since any finite spanning set for $R(L_A)$ contains at least $\dim(R(L_A)) = \text{rank}(A)$ vectors, by C1.9.2 we must have that $n \ge \text{rank}(A)$.
Then, since $R(L_A)$ is a subspace of $\mathbb{F}^m$, by C1.9.2 $\text{rank}(A) = \dim(R(L_A)) \le \dim(\mathbb{F}^m) = m$.
Hence $\text{rank}(A) \le \min(m,n)$, as required. ∎

$T: V \to W$ IS 1-1 & LINEAR, V_0 IS A SUBSPACE OF V $\Rightarrow$ $T(V_0)$ IS A SUBSPACE OF W (L9(1))

Let $T: V \to W$ be a linear injective mapping between vector spaces V and W.
Let V_0 be a subspace of V.
Then necessarily $T(V_0) = \{T(v) : v \in V_0\}$ is a subspace of W.

$T: V \to W$ IS 1-1 & LINEAR, V_0 IS A SUBSPACE OF V, $\dim(V_0) < \infty$ $\Rightarrow$ $\dim(V_0) = \dim(T(V_0))$ (L9(2))

Let $T: V \to W$ be a linear injective mapping between vector spaces V and W.
Let V_0 be a finite-dimensional subspace of V.
Then necessarily $\dim(V_0) = \dim(T(V_0))$.

$\text{rank}(AQ) = \text{rank}(A)$ (T3.4(1))

Let $A \in M_{m\times n}(\mathbb{F})$, and let $Q \in M_{n\times n}(\mathbb{F})$ be an invertible matrix.
Then necessarily $\text{rank}(AQ) = \text{rank}(A)$.

Proof. Since Q is invertible, L_Q is necessarily an isomorphism.
Thus $L_Q(\mathbb{F}^n) = \mathbb{F}^n$, and so
$L_{AQ}(\mathbb{F}^n) = L_A L_Q(\mathbb{F}^n) = L_A(\mathbb{F}^n)$.
It follows that
$\text{rank}(AQ) = \dim(L_{AQ}(\mathbb{F}^n)) = \dim(L_A(\mathbb{F}^n)) = \text{rank}(A)$,
as required. ∎

$\text{rank}(PA) = \text{rank}(A)$ (T3.4(2))

Let $A \in M_{m\times n}$, and let $P \in M_{m\times m}(\mathbb{F})$ be an invertible matrix.
Then necessarily $\text{rank}(PA) = \text{rank}(A)$.

Proof. Since P is invertible, L_P is an isomorphism.
So, by L9, using $T = L_P$, $V = W = \mathbb{F}^m$ and $V_0 = L_A(\mathbb{F}^n)$, we have
$\dim(L_A(\mathbb{F}^n)) = \dim(L_P(L_A(\mathbb{F}^n)))$
$\Rightarrow \text{rank}(A) = \dim(L_{PA}(\mathbb{F}^n))$
$\Rightarrow \text{rank}(A) = \text{rank}(PA)$, as needed. ∎

$\text{rank}(PAQ) = \text{rank}(A)$ (T3.4(3))

Let $A \in M_{m\times n}(\mathbb{F})$, and let $P \in M_{m\times m}(\mathbb{F})$ and $Q \in M_{n\times n}(\mathbb{F})$ be invertible matrices.
Then necessarily $\text{rank}(PAQ) = \text{rank}(A)$.

Proof. This follows from T3.4(1) and T3.4(2). ∎

INVERTIBLE MATRIX THEOREM, PART 2 (C3.4.1)

Let $A \in M_{n\times n}(\mathbb{F})$ be arbitrary.
Then A is invertible if and only if $\text{rank}(A) = n$.

Proof. ($\Rightarrow$) If A is invertible, necessarily $I_n = AA^{-1}$.
Since A^{-1} is also invertible, by T3.4, it follows that
$\text{rank}(A) = \text{rank}(AA^{-1}) = \text{rank}(I_n) = n$,
proving the forward argument. #
($\Leftarrow$) If $n = \text{rank}(A)$, necessarily $n = \dim(L_A(\mathbb{F}^n))$.
Then, since $L_A(\mathbb{F}^n)$ is a subspace of $\mathbb{F}^n$, it follows that $L_A(\mathbb{F}^n) = \mathbb{F}^n$ (by C1.9.2(b)).
Hence L_A is onto; thus (by T2.7) it is also 1-1, and so (since L_A is linear by T2.9) L_A is an isomorphism.
It follows that A is invertible (by T2.15(2)), proving the backward argument. ∎

ELEMENTARY ROW & COLUMN OPERATIONS ON A MATRIX ARE RANK-PRESERVING (C3.4.2)

For any matrix $A \in M_{m\times n}(\mathbb{F})$, performing elementary row and column operations on A does not change the rank of the resultant matrix.

Proof. Suppose B is obtained from A by performing an elementary row operation; so, there exists an elementary matrix $E \in M_{m\times m}(\mathbb{F})$ such that $B = EA$.
Since E is invertible, by T3.4 necessarily $\text{rank}(B) = \text{rank}(A)$.
A similar result holds for elementary column operations. ∎

This result can be used to transform complicated matrices into simpler ones to determine their rank.

ANY MATRIX CAN BE TRANSFORMED INTO THE FORM $D=\begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix}$ (T3.5)

Let $A \in M_{m\times n}(\mathbb{F})$ be arbitrary.
Then there exists a finite sequence of elementary row and column operations such that when applied to A, the resultant matrix D is of the form

$$D=\begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix},$$

where O_1, O_2, O_3 are zero matrices, and $r=\operatorname{rank}(A)$.

Proof. If $A=O$, we are done.
Then, suppose $A \neq O$. Then A has a non-zero entry.
By means of at most one elementary row and at most one elementary column operation (each of the form $R_k \leftrightarrow R_\ell$ or $C_k \leftrightarrow C_\ell$), we can move the non-zero entry to the $(1,1)$ position.
By means of at most one operation of the form $(R_k \leftarrow c\cdot R_k)$ or $(C_k \leftarrow c\cdot C_k)$, we can change that entry to 1.
Then, by at most $(m-1)$ row operations of type "$R_k \leftarrow R_k + c\cdot R_\ell$" and by at most $(n-1)$ column operations of type "$C_k \leftarrow C_k + c\cdot C_\ell$", we can change all the remaining entries in the first row and in the first column to be 0.
It follows that after a finite number of elementary matrix operations, we have transformed A to a matrix A' of the form

$$A' = \left(\begin{array}{c|c} 1 & 0 \cdots 0 \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right).$$

By continuing this recursive process on B, we can continue this process to obtain a matrix of the form D after a finite number of elementary row/column operations.
Since these preserve rank, it follows that $\operatorname{rank}(A) = \operatorname{rank}(D)$.
Then, by R16,

$$\operatorname{rank}(D) = \dim(\operatorname{span}\{e_1, \dots, e_r, \underset{\sim}{0}, \dots, \underset{\sim}{0}\}) = \dim(\operatorname{span}\{e_1,\dots,e_r\}) = r,$$

where $\{e_1, \dots, e_n\}$ is the standard basis for $\mathbb{F}^n$.
It follows that $\operatorname{rank}(A) = \operatorname{rank}(D) = r$, as desired. ☒

$A\in M_{m\times n}(\mathbb{F}) \Rightarrow \exists$ INVERTIBLE $B \in M_{m\times m}(\mathbb{F})$, $C\in M_{n\times n}(\mathbb{F})$ $\ni D = BAC = \begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix} \in M_{m\times n}(\mathbb{F})$, $r=\operatorname{rank}(A)$ (C3.5.1)

Let $A\in M_{m\times n}(\mathbb{F})$ be such that $r=\operatorname{rank}(A)$.
Then there necessarily exist invertible matrices $B\in M_{m\times m}(\mathbb{F})$, $C\in M_{n\times n}(\mathbb{F})$ such that the matrix $D=BAC$ is of the form $D=\begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix} \in M_{m\times n}(\mathbb{F})$, where O_1, O_2, O_3 are zero matrices.

Proof. By T3.5, we can convert A into D via a finite number of elementary row & column operations.
It follows that

$$D = E_p \cdots E_1 A G_1 \cdots G_q,$$

where $E_1, \dots, E_p \in M_{m\times m}(\mathbb{F})$ and $G_1, \dots, G_q \in M_{n\times n}(\mathbb{F})$ are elementary matrices.
Thus they are invertible, so it follows that $B = E_p \cdots E_1$ and $C = G_1 \cdots G_q$ are invertible and $D = BAC$, completing the proof. ☒

ANY MATRIX CAN BE TRANSFORMED INTO "D_{upper}" (T3.6)

Let $A\in M_{m\times n}(\mathbb{F})$ be such that $r=\operatorname{rank}(A)$.
Then there exist a finite sequence of elementary row and column operations such that when applied to A, it transforms into the matrix

$$D_{upper} = \begin{pmatrix} 1 & d_{12} & d_{13} & \cdots & d_{1,r} & d_{1,r+1} & \cdots & d_{1n} \\ 0 & 1 & d_{23} & \cdots & d_{2,r} & d_{2,r+1} & \cdots & d_{2n} \\ 0 & 0 & 1 & \cdots & d_{3,r} & d_{3,r+1} & \cdots & d_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 1 & d_{r,r+1} & \cdots & d_{rn} \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{pmatrix}.$$

Proof. If $A=O$, we are done.
Suppose $A\neq O$, so that there exists a non-zero entry of A.
By doing at most one row and at most one column (each of type 1; ie "swapping") operation, we can move this non-zero entry to the $(1,1)$ position.
By doing at most one "type 2" operation (ie $R_k \leftarrow c\cdot R_k$ or $C_k \leftarrow c\cdot C_k$), we can change it to 1.
By at most $(m-1)$ type-3 row operations (ie $R_k \leftarrow R_k + c\cdot R_\ell$), we can change all the remaining entries in the first row to be 0.
Hence, we have transformed A to a matrix A' of the form

$$A' = \left(\begin{array}{c|c} 1 & d_{12} \cdots d_{1n} \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right).$$

By repeating this recursive process on B, we can transform A into the form of D_{upper}.
Then,

$$R(D_{upper}) = \dim(\operatorname{span}(\{e_1, d_{12}e_1 + e_2, \dots, \sum_{i=1}^{r-1} d_{i,r}e_i + e_r, d_{r+1}, \dots, d_n\}))$$

$$\therefore R(D_{upper}) = \dim(\operatorname{span}(\{e_1, \dots, e_r, d_{r+1}, \dots, d_n\})),$$

where $\{e_1, \dots, e_n\}$ is the standard (ordered) basis for $\mathbb{F}^n$, and d_k is the k^{th} column of D_{upper} for $1\le k\le n$.
Then, since $d_k = \sum_{i=1}^{r} d_{i,k}e_i$ for $r+1\le k\le n$, we have

$$\operatorname{span}(\{e_1, \dots, e_r, d_{r+1}, \dots, d_n\}) = \operatorname{span}(\{e_1, \dots, e_r\}).$$

It follows that $L_{D_{upper}}(\mathbb{F}^n) = \operatorname{span}(\{e_1, \dots, e_r\})$, so that

$$\operatorname{rank}(D_{upper}) = \dim(L_{D_{upper}}(\mathbb{F}^n)) = \dim(\operatorname{span}(\{e_1,\dots,e_r\})) = r = \operatorname{rank}(A),$$

since elementary matrix operations preserve the rank of the matrix, and we are done. ☒

METHOD TO CONVERT MATRICES TO D_{upper} (R17)

By T3.6, we can formulate a method to convert a complicated matrix A into D_{upper} to find its rank:

1. Find a non-zero entry of A;
2. Apply at most one type-1 row operation and at most one type-1 column operation to move the entry to the position $(1,1)$;
3. Apply at most one type-2 row (or column) operation to make the entry at the position $(1,1)$ to be $1_{\mathbb{F}}$;
4. Apply at most $(m-1)$ type-3 row operations so that all of the remaining entries in the first row is 0, so the new matrix is of the form

$$A' = \left(\begin{array}{c|c} 1 & d_{12} \cdots d_{1n} \\ \hline 0 & \\ \vdots & B \\ 0 & \end{array}\right).$$

5. Repeat steps ①–④ on B recursively until a matrix of the form of D_{upper} is obtained.
6. It follows that $\operatorname{rank}(A) = \operatorname{rank}(D_{upper}) = r$.

$\mathrm{rank}(A^T) = \mathrm{rank}(A)$ (C3.6.1 (1))

Let $A \in M_{m\times n}(\mathbb{F})$ be arbitrary.

Then necessarily $\mathrm{rank}(A^T) = \mathrm{rank}(A)$

Proof. From C3.5.1, there exists invertible matrices B, C such that $D = BAC$.

Then $D^T = (BAC)^T = C^T A^T B^T$.

Since B and C are invertible, by L8 B^T and C^T are also invertible.

Thus $\mathrm{rank}(A^T) = \mathrm{rank}(D^T)$.

Then, as $D^T \in M_{n\times m}(\mathbb{F})$ has the form of the matrix D in C3.5.1, necessarily $\mathrm{rank}(D^T) = \mathrm{rank}(A)$.

It follows that $\mathrm{rank}(A^T) = \mathrm{rank}(A)$, as required. ☒

$\mathrm{rank}(A) = \dim(\mathrm{span}(\{R_1, \dots, R_m\})) = \dim(\mathrm{span}(\{C_1, \dots, C_n\}))$ (C3.6.1 (2))

Let $A \in M_{m\times n}(\mathbb{F})$ be arbitrary.

Then necessarily $\mathrm{rank}(A) = \dim(\mathrm{span}(\{R_1, \dots, R_m\})) = \dim(\mathrm{span}(\{C_1, \dots, C_n\}))$, where R_i and C_j denote the i^{th} row and j^{th} column of A respectively.

Proof. By R16, $\mathrm{rank}(A) = \dim(\mathrm{span}(\{C_1, \dots, C_n\}))$.

So, the rank of A^T is the dimension of the subspace generated by the columns of A^T.

But since the columns of A^T are the rows of A, and $\mathrm{rank}(A) = \mathrm{rank}(A^T)$ by C3.6.1(1), it follows that $\mathrm{rank}(A)$ is also the dimension of the subspace generated by the rows of A, as needed. ☒

$\mathrm{rank}(AB) \leq \min(\{\mathrm{rank}(A), \mathrm{rank}(B)\})$ (T3.7)

Let A and B be matrices such that AB is defined.

Then necessarily $\mathrm{rank}(AB) \leq \min(\{\mathrm{rank}(A), \mathrm{rank}(B)\})$.

Proof. Let $A \in M_{m\times n}(\mathbb{F})$ and $B \in M_{n\times p}(\mathbb{F})$. Then, since

$R(L_{AB}) = \{ABx : x \in \mathbb{F}^p\} \subset \{Ay : y \in \mathbb{F}^n\} = R(L_A)$,

we have

$\mathrm{rank}(AB) = \dim(R(L_{AB})) \leq \dim(R(L_A)) = \mathrm{rank}(A)$.

On the other hand,

$\mathrm{rank}(AB) = \mathrm{rank}((AB)^T) = \mathrm{rank}(B^T A^T) \leq \mathrm{rank}(B^T) = \mathrm{rank}(B)$.

Thus $\mathrm{rank}(AB) \leq \min(\{\mathrm{rank}(A), \mathrm{rank}(B)\})$, as needed.

FOUR FUNDAMENTAL SUBSPACES OF A MATRIX (S3.3)

COLUMN SPACE OF A MATRIX, Col(A) (D34)

Let $A \in M_{m\times n}(\mathbb{F})$. Then, we define the "column space" of A, denoted as "Col(A)", to be the vector space

$$\mathrm{Col}(A) := \{Ax : x \in \mathbb{F}^n\}$$
$$= \{\text{all linear combinations of columns in } A\}$$
$$= \mathrm{span}(\{\text{columns of } A\}).$$

We can show Col(A) is a subspace of $\mathbb{F}^m$. (T3.8 (1))

Proof. This follows from the fact that $\mathrm{Col}(A) = \mathrm{span}(\{\text{columns of } A\})$. ☒

ROW SPACE OF A MATRIX, Row(A) (D34)

Let $A \in M_{m\times n}(\mathbb{F})$. Then, we define the "row space" of A, denoted as "Row(A)", to be the vector space

$$\mathrm{Row}(A) := \mathrm{Col}(A^T)$$
$$= \{A^T y : y \in \mathbb{F}^m\}$$
$$= \{\text{all linear combinations of rows in } A\}$$
$$= \mathrm{span}(\{\text{rows of } A\}).$$

We can similarly show Row(A) is a subspace of $\mathbb{F}^n$. (T3.8 (1))

Proof. Again, this follows from the fact that $\mathrm{Row}(A) = \mathrm{span}(\{\text{rows of } A\})$.

NULL SPACE OF A MATRIX, Null(A) (D34)

Let $A \in M_{m\times n}(\mathbb{F})$. Then, we define the "null space" of A, denoted as "Null(A)", to be the vector space

$$\mathrm{Null}(A) := \{x \in \mathbb{F}^n \mid Ax = 0\}.$$

We can show that Null(A) is a subspace of $\mathbb{F}^n$. (T3.8(1))

LEFT NULL SPACE OF A MATRIX, $\mathrm{Null}(A^T)$ (D34)

Let $A \in M_{m\times n}(\mathbb{F})$. Then, we define the "left null space" of A, denoted as "$\mathrm{Null}(A^T)$", to be the vector space

$$\mathrm{Null}(A^T) := \{y \in \mathbb{F}^m \mid A^T y = 0\}.$$

We can similarly show that $\mathrm{Null}(A^T)$ is a subspace of $\mathbb{F}^m$. (T3.8(1))

NULLITY OF A MATRIX, Nullity(A) (D34)

For any matrix $A \in M_{m\times n}(\mathbb{F})$, we define the "nullity" of A, denoted as "Nullity(A)", to be

$$\mathrm{Nullity}(A) := \dim(\mathrm{Null}(A)).$$

$\mathrm{rank}(A) = \dim(\mathrm{Col}(A)) = \dim(\mathrm{Row}(A))$ (T3.8(2))

Let $A \in M_{m\times n}(\mathbb{F})$ be arbitrary.

Then necessarily $\mathrm{rank}(A) = \dim(\mathrm{Col}(A)) = \dim(\mathrm{Row}(A))$.

Proof. This follows from R16 (1), and the fact that $\mathrm{rank}(A) = \mathrm{rank}(A^T)$ by C3.6.1(1). ☒

$\mathrm{nullity}(A^T) = m - \mathrm{rank}(A)$, $\mathrm{nullity}(A) = n - \mathrm{rank}(A)$ (T3.8(3))

Let $A \in M_{m\times n}(\mathbb{F})$ be arbitrary.

Then necessarily $\mathrm{nullity}(A^T) = m - \mathrm{rank}(A)$ and $\mathrm{nullity}(A) = n - \mathrm{rank}(A)$.

Proof. By the Rank-Nullity theorem (T2.4), necessarily

$$\dim(\mathbb{F}^n) = n = \dim(R(L_A)) + \dim(N(L_A)) = \mathrm{rank}(A) + \mathrm{nullity}(A),$$

and

$$\dim(\mathbb{F}^m) = m = \dim(R(L_{A^T})) + \dim(N(L_{A^T})) = \mathrm{rank}(A) + \mathrm{nullity}(A^T).$$

The proof directly follows from this observation. ☒

$\mathbb{F}^m = \mathrm{Col}(A) \oplus \mathrm{Null}(A^T)$, $\mathbb{F}^n = \mathrm{Row}(A) \oplus \mathrm{Null}(A)$ (T3.8 (4))

Let $A \in M_{m\times n}(\mathbb{F})$ be arbitrary.

Then necessarily $\mathbb{F}^m = \mathrm{Col}(A) \oplus \mathrm{Null}(A^T)$ and $\mathbb{F}^n = \mathrm{Row}(A) \oplus \mathrm{Null}(A)$.

Proof. We first prove $\mathrm{Row}(A) \cap \mathrm{Null}(A) = \{0\}$.

Let $v \in \mathrm{Row}(A) \cap \mathrm{Null}(A)$ be arbitrary. By definition, this implies $v \in \mathrm{Col}(A^T) = \{A^T y : y \in \mathbb{F}^m\}$, and $Av = 0$.

Hence there exists $y \in \mathbb{F}^m$ such that $v = A^T y$.

Since $Av = 0$, thus $AA^T y = 0$.

This implies

$$0 = y^T A A^T y = (y^T A)(A^T y)$$
$$= (A^T y)^T (A^T y)$$
$$\therefore 0 = v^T v,$$

hence implying $v = 0$ necessarily, so that $\mathrm{Row}(A) \cap \mathrm{Null}(A) = \{0\}$.

So, by T1.12, we have

$$\dim(\mathrm{Row}(A) + \mathrm{Null}(A)) = \dim(\mathrm{Row}(A)) + \dim(\mathrm{Null}(A))$$
$$= \dim(\mathrm{Row}(A)) + \mathrm{nullity}(A).$$

Then, by C3.6.1, $\mathrm{rank}(A) = \dim(\mathrm{Row}(A))$. Thus

$$\dim(\mathrm{Row}(A)) + \mathrm{nullity}(A) = \mathrm{rank}(A) + \mathrm{nullity}(A) = n = \dim(\mathbb{F}^n).$$

Since $\mathrm{row}(A) + \mathrm{Null}(A)$ is a subspace of $\mathbb{F}^n$ and $\dim(\mathrm{Row}(A) + \mathrm{Null}(A)) = \dim(\mathbb{F}^n)$, we have $\mathrm{Row}(A) + \mathrm{Null}(A) = \mathbb{F}^n$.

Together with $\mathrm{Row}(A) \cap \mathrm{Null}(A) = \{0\}$, this tells us that

$$\mathrm{Row}(A) \oplus \mathrm{Null}(A) = \mathbb{F}^n,$$

as needed. ☒

THE INVERSE OF A MATRIX (S3.4)

INVERTIBLE MATRIX THEOREM, PART 3 (T3.9)

Let $A \in M_{n\times n}(\mathbb{F})$. Then the following statements are equivalent:

① A is invertible;
② The columns of A form a basis for $\mathbb{F}^n$;
③ The rows of A form a basis for $\mathbb{F}^n$; and
④ A is a product of elementary matrices.

Proof. (② ⟺ ①) Note that $\mathrm{rank}(A) = n \iff \dim(\mathrm{Col}(A)) = n$
$\iff$ the columns of A form a basis for $\mathbb{F}^n$, since A has n columns. (This follows from C1.9.2).
We can similarly prove (③ ⟺ ①) in this manner. #

(④ ⟹ ①) Suppose $A = E_1 \cdots E_p$, where $E_1, \dots, E_p$ are elementary matrices.
Then, since elementary matrices are invertible, and the matrix product of invertible matrices is invertible, it follows that A is invertible and $A^{-1} = E_p^{-1} \cdots E_1^{-1}$. #

(① ⟹ ④) By C3.5.1, we have $D = BAC$, where $D = \begin{pmatrix} I_r & O_1 \\ O_2 & O_3 \end{pmatrix}$, $r = \mathrm{rank}(A)$, and B and C are products of elementary matrices.
Then, since A is invertible, necessarily (by C3.4.1) $r = n$, implying $D = I_n$.
Hence $A = B^{-1} I_n C^{-1} = B^{-1}C^{-1}$.
Finally, since B and C are both products of elementary matrices, and the inverse of an elementary matrix is also an elementary matrix, it follows that A is itself the product of elementary matrices.
This is sufficient to prove the 4 statements are equivalent to one another. ▨

$A \in M_{n\times n}(\mathbb{F})$ IS INVERTIBLE ⟹ CAN TRANSFORM $(A \mid I_n)$ INTO $(I_n \mid A^{-1})$ BY ROW OPERATIONS (T3.10(1))

Let $A \in M_{n\times n}(\mathbb{F})$ be invertible. Then there exists a finite sequence of elementary row operations which can transform the matrix $(A \mid I_n)$ into the matrix $(I_n \mid A^{-1})$.

Proof. Since $AM = (Av_1 \cdots Av_p)$ for any $M = (v_1 \cdots v_p) \in M_{n\times p}(\mathbb{F})$, we have
$$A^{-1}(A \mid I_n) = (A^{-1}A \mid A^{-1}I_n) = (I_n \mid A^{-1}).$$
Then, by the invertible matrix theorem Part 3, we have
$$A^{-1} = E_p \cdots E_1,$$
where $E_1, \dots, E_p$ are elementary matrices.
It follows that
$$(E_p \cdots E_1)(A \mid I_n) = (I_n \mid A^{-1}),$$
which show (since each E_i is the result of an elementary row operation) that we can transform $A \mid I_n$ into $I_n \mid A^{-1}$ via a finite sequence of elementary row operations. ▨

$A \in M_{n\times n}(\mathbb{F})$, $\exists B \in M_{n\times n} \in \mathbb{F} \ni (A \mid I_n) \rightsquigarrow (I_n \mid B)$ BY FINITELY MANY ROW OPERATIONS ⟹ A IS INVERTIBLE & $B = A^{-1}$ (T3.10 (2))

Let $A \in M_{n\times n}(\mathbb{F})$, and suppose there exists a $B \in M_{n\times n}(\mathbb{F})$ such that we can transform the matrix $(A \mid I_n)$ into $(I_n \mid B)$ by finitely many elementary row operations.
Then necessarily A is invertible, and $B = A^{-1}$.

Proof. Let $G_1, \dots, G_q$ be the elementary matrices associated with the elementary row operations that transform $(A \mid I_n)$ into $(I_n \mid B)$, so that
$$(G_q \cdots G_1)(A \mid I_n) = (I_n \mid B).$$
Let $G = G_q \cdots G_1$, so that $G(A \mid I_n) = (GA \mid G) = (I_n \mid B)$.
It follows that $I_n = GA$ and $B = G$, so that $AB = I_n$, and hence that A is invertible and $B = A^{-1}$. ▨

GAUSS-JORDAN METHOD TO FINDING INVERSES TO SQUARE MATRICES (R18)

Using T3.10, we can formulate a method to find the inverse of a square matrix A (if it exists):

① If the first column of A is a zero vector, A is not invertible; otherwise, the first column of A has a non-zero entry.
Why? · follows from the Invertible Matrix Theorem part 3.

② In a manner similar to the process for T3.6, we can convert $(A \mid I_n)$ into a matrix of the form
$$B = \left(\begin{array}{c|ccc} 1 & d_{12} & \cdots & d_{1,2n} \\ \hline 0 & & & \\ \vdots & & Q & \\ 0 & & & \end{array}\right)$$
using only elementary row operations.
· in particular, at most one type-1, at most one type-2, and at most $(n-1)$ type-3 operations.

③ Then, we repeat steps ① and ② recursively on Q, until either
① The first column of Q is a zero vector; or
· then, by the Invertible Matrix Theorem part 3, A is not invertible and so we stop the procedure
② We get a matrix of the form
$$C' = \begin{pmatrix} 1 & d_{12} & \cdots & d_{1,n} & d_{1,n+1} & \cdots & d_{1,2n} \\ 0 & 1 & \cdots & d_{2,n} & d_{2,n+1} & \cdots & d_{2,2n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & d_{n,n+1} & \cdots & d_{n,2n} \end{pmatrix}.$$

④ Then, by at most $(n-1)$ type-3 row operations, we can convert C' to the matrix C_n, where
$$C_n = \begin{pmatrix} 1 & d_{12} & \cdots & 0 & d_{1,n+1} & \cdots & d_{1,2n} \\ 0 & 1 & \cdots & 0 & d_{2,n+1} & \cdots & d_{2,2n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & d_{n,n+1} & \cdots & d_{n,2n} \end{pmatrix};$$
ie C_n is C' but with the n^{th} column having all zero entries except the last one.

⑤ By at most $(n-2)$ type-3 row operations, we can convert C_n into the matrix C_{n-1}, which is C_n but with the $(n-1)^{th}$ column of C_n being zeros except at the $(n-1, n-1)$ position.

⑥ Continue step ⑤ until we get a matrix of the form $(I_n \mid B)$.

Then, by T3.10, necessarily $B = A^{-1}$.