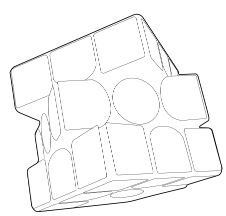
Linear Algebra I

Snochi Song



This page intentionally left blank.

Contents

This page intentionally left blank.

1.
Vector Spaces

Vector Spaces

Def'n. Vector Space over a Field

A ***vector space*** over a field $\mathbb{F}$ is a set with two binary operations, addition $V \times V \to V$ and scalar multiplication $\mathbb{F} \times V \to V$ such that the following holds. Let $x, y, z \in V$ and $a, b, c \in \mathbb{F}$.

(a) Commutativity of addition:
$$x + y = y + x.$$

(b) Associativity of addition:
$$(x + y) + z = x + (y + z).$$

(c) Existence of additive identity: There exists $0 \in V$ such that
$$0 + v = v + 0 = v$$
for all $v \in V$.

(d) Existence of additive inverse: There exists $-v \in V$ such that
$$-v + v = v + (-v) = 0$$
for all $v \in V$.

(e) Existence of identity of scalar multiplication:
$$1x = x,$$
where $1 \in \mathbb{F}$ is the unity of $\mathbb{F}$.

(f) Compatibility of scalar multiplication with field multiplication:
$$a(bx) = (ab)x.$$

(g) Distributivity of scalar multiplication with respect to addition:
$$a(x + y) = ax + ay.$$

(h) Distributivity of scalar multiplication with respect to field addition:
$$(a + b)x = ax + bx.$$

Remark 1.1. For convenience, we shall consistently use $\mathbb{F}$ to denote an arbitrary field. Moreover, unless otherwise specified, $\mathbb{F}$ is the underlying field of any vector space that we are going to discuss.

Def'n. Vector, Scalar

Let V be a vector space over $\mathbb{F}$. We call an element $v \in V$ a ***vector*** and $c \in \mathbb{F}$ a ***scalar***.

Proposition 1.1. Cancellative Property of Vector Addition

Let V be a vector space and $x, y, z \in V$. Suppose $x + z = y + z$. Then $x = y$.

Proof. By definition, there exists $v \in V$ such that $z + v = 0$, the additive inverse of z. Thus,
$$x = x + 0 = x + (z + v) = (x + z) + v = (y + z) + v = y + (z + v) = y.$$
♠

Subspaces

Remark 1.2. In study of algebraic structure, often times it is of interest to examine subsets that possess the same structure as its superset.

> **Def'n. Subspace** of a Vector Space
> Let V be a vector space over $\mathbb{F}$. We say a subset $W \subseteq V$ is a ***subspace*** of V if W is a vector space over the same field $\mathbb{F}$ with operations on V.

Proposition 1.2. Subspace Test

Let V be a vector space and W be a subset of V. Then W is a subspace of V if and only if the following hold.

(a) $0 \in W$.

(b) $x+y \in W$ *whenever* $x \in W$ *and* $y \in W$.

(c) $cx \in W$ *whenever* $c \in \mathbb{F}$ *and* $x \in W$.

Proof. For the forward direction, suppose W is a subspace of V. But this means that the operations on W are closed, so (2) and (3) hold. Furthermore, there is $0' \in W$ such that for every $x \in W$, $x+0=x$. But since V is a vector space, there is $0 \in V$, so $0'+0=0=0'$, proving (1). Moreover, for the reverse direction, suppose (1), (2), and (3) hold. Then we only have to ensure that the existance of additive inverse for each element. But clearly,

$$(-1)x = -x \in W$$

whenever $x \in W$ by (3). ♠

Proposition 1.3. Intersection of Subspaces Is a Subspace

Any intersection of subspaces of a vector space V is a subspace of V.

Proof. Let C be the set of some arbitrary subspaces of V and W be the intersection of the subspaces in C. Since every subspace has 0, $0 \in W$. Moreover, let $x, y \in W$ and $a \in \mathbb{F}$. Since every subspace is closed under addition and scalar multiplication, every subspace has $x+y$ and ax, so $(x+y), ax \in W$. ♠

Linear Combinations and Span

> **Def'n. Linear Combination** of Vectors
> Let V be a vector space and let $v_1, v_2, \ldots, v_n \in V$. We say $v \in V$ is a ***linear combination*** of $v_1, v_2, \ldots, v_n$ if there exists some $c_1, c_2, \ldots, c_n \in \mathbb{F}$ such that
> $$v = \sum_{i=1}^{n} c_i v_i.$$

Remark 1.3. From the definition, it is implicitly stated that any linear combination involves only finite number of vectors.

> **Def'n. Span** of a Set of Vectors
> Let V be a vector space and let $S \subseteq V$ be a nonempty subset. We define the ***span*** of S, denoted as $\operatorname{span}(S)$, to be the set of linear combinations of vectors in S.

Proposition 1.4.
span(S) **Is a Subspace**

Let V be a vector space and let $S \subseteq V$ be a subset. Then span$(S) \subseteq V$ *is a subspace and any subspace $W \subseteq V$ with $S \subseteq W$ also satisfies* span$(S) \subseteq W$.

Proof. Notice that the result is trivial for $S = \emptyset$, since $\text{span}(\emptyset) = \{0\}$, which is a subspace for any vector space. So suppose that $S \neq \emptyset$. First notice that

$$0 \in \text{span}(S),$$

since $0v$ for any $v \in S$ is a linear combination of v. Moreover, let $x, y \in \text{span}(S)$ and c be any scalar. Since x and y are linear combinations of vectors in S, clearly $x + y$ and cx are both linear combinations of vectors in S as well. So span(S) is a subspace of V. For the second part of the proposition, suppose W is a subspace of V which contains S. For the sake of contradiction, further assume that W does not contain span(S). Then, for some vectors $s_1, s_2, \ldots, s_n \in S$ and scalars $a_1, a_2, \ldots, a_n \in \mathbb{F}$,

$$a_1 s_1 + a_2 s_2 + \cdots + a_n s_n \notin W.$$

But clearly each $a_i s_i \in W$, since W contains S and W is closed under scalar multiplication. So $a_1 s_1 + a_2 s_2 + \cdots + a_n s_n \in W$, which is a contradiction. Thus span$(S) \subseteq W$, as desired. ♠

Linear Independence

Remark 1.4. Let V be a vector space and suppose that $S \subseteq V$ is a spanning set of V. Then, we have some number of expression to describe vectors in V. For instance, if S has n elements $s_1, s_2, \ldots, s_n \in S$, then for any $v \in V$,

$$v = \sum_{i=1}^{n} c_i s_i$$

for some $c_1, c_2, \ldots, c_n \in \mathbb{F}$. But the question is, is S the minimal spanning subset of V? That is, we are interested to find out if there is a subset of V which has less than n elements and spans V. To answer this question, we introduce the following definition.

Def'n. Linearly Independent, Linearly Dependent Vectors

Let V be a vector space. We say $v_1, v_2, \ldots, v_n \in V$ are ***linearly dependent*** if there exist nonzero $(c_1, c_2, \ldots, c_n) \in \mathbb{F}^n$ ($\mathbb{F}^n$ is the set of n-tuples where each entry is an element of $\mathbb{F}$; we say $x \in \mathbb{F}^n$ is zero if every entry of x is zero) such that

$$\sum_{i=1}^{n} c_i v_i = 0.$$

We say $v_1, v_2, \ldots, v_n$ are ***linearly independent*** otherwise.

Remark 1.5. We also say that a subset $S \subseteq V$ is linearly dependent if there exist finite number of elements $s_1, s_2, \ldots, s_n \in S$ such that

$$\sum_{i=1}^{n} c_i s_i = 0$$

for some nonzero $(c_1, c_2, \ldots, c_n) \in \mathbb{F}$. Of course, S is linearly independent if no such elements exist.

Remark 1.6. Another way to think linear independence is the following. Let $v_1, v_2, \ldots, v_n \in V$ for some vector space V. Then $v_1, v_2, \ldots, v_n$ are linearly independent if and only if the only linear combination of $v_1, v_2, \ldots, v_n$ equal to 0 is the trivial representation. That is,

$$\sum_{i=1}^{n} 0 v_i = 0.$$

Proposition 1.5. *Let V be a vector space. If $S_1 \subseteq S_2 \subseteq V$ and S_1 is linearly dependent, then S_2 is linearly dependent.*

Proof. Suppose $S_1 = \{v_1, \dots, v_n\} \in V$ be linearly dependent. Then there is a nonzero $(a_1, \dots, a_n) \in \mathbb{F}^n$ such that
$$\sum\nolimits_{i=1}^{n} a_i v_i = 0.$$
Therefore, if we define $a_{n+1} = a_{n+2} = \cdots = a_m = 0$, where $m = |S_2| \in \mathbb{N}$, then
$$\sum\nolimits_{i=1}^{m} a_i v_i = \sum\nolimits_{i=1}^{n} a_i v_i = 0,$$
which is not the trivial representation. ♠

Corollary 1.5.1. *Let V be a vector space. If $S_1 \subseteq S_2 \subseteq V$ and S_2 is linearly independent, then S_1 is linearly independent.*

Remark 1.7. Now suppose $S_n \subseteq V$ is a subset of V containing n elements and spans V. If S_n is linearly dependent, then there must be a vector s_n which can be written as a linear combination of other vectors in S_n. So it turns out that
$$\operatorname{span}(S_{n-1}) = \operatorname{span}(S_n) = V,$$
where $S_{n-1} = S_n \setminus \{s_n\}$. We may continue this process until S_k is independent. But once we hit here, there is no way $\operatorname{span}(S_{k-1}) = \operatorname{span}(S_k)$. Thus it turns out that the smallest spanning set of V must be independent. This idea can be written alternatively as the following proposition.

Proposition 1.6. *Let S be linearly independent subset of a vector space V, and let $v \in V$ with $v \notin S$. Then $S \cup \{v\}$ is linearly dependent if and only if $v \in \operatorname{span} S$.*

Proof. First, write $S = \{v_1, \dots, v_n\} \subseteq S$ for convenience. For the forward direction, suppose that $S \cup \{v\}$ is linearly dependent. Then there must exist nonzero $(a_1, \dots, a_{n+1}) \in \mathbb{F}^{n+1}$ such that
$$\sum\nolimits_{i=1}^{n} a_i v_i + a_{n+1} v = 0.$$
But this means
$$-a_{n+1} v = \sum\nolimits_{i=1}^{n} a_i v_i \iff v = \sum\nolimits_{i=1}^{n} -\frac{a_i}{a_{n+1}} v_i$$
so $v \in \operatorname{span}\{v_1, \dots, v_n\}$. For the reverse direction, suppose that $v \in \operatorname{span}(S)$. Then there exists $(a_1, \dots, a_n) \in \mathbb{F}^n$ such that
$$v = \sum\nolimits_{i=1}^{n} a_i v_i,$$
so we have nontrivial representation of zero
$$\sum\nolimits_{i=1}^{n} a_i v_i + 1v = 0.$$
♠

Bases and Dimension

Remark 1.8. From the last section, we have seen that the smallest spanning set of any vector space must be linearly independent. Indeed, there are many pleasurable behaviors of linearly independent spanning sets that would be discussed in this section.

Def'n. Basis for a Vector Space

Let V be a vector space and $\beta \subseteq V$. We say β is a ***basis*** for V if β is linearly dependent and spans V. We also say vectors of β form a basis for V.

Remark 1.9. One important property of basis β for a vector space V is that any $v \in V$ can be uniquely written as a linear combination of vectors in β.

Proposition 1.7. Unique Representation of a Vector

Let V be a vector space and $\beta = \{v_1, v_2, \dots, v_n\} \subseteq V$. Then β is a basis for V if and only if there exists unique scalars $c_1, c_2, \dots, c_n \in \mathbb{F}$ such that

$$v = c_1 v_1 + c_2 v_2 + \cdots + c_n v_n$$

for all $v \in V$.

Proof. For the forward direction, suppose that β is a basis for V, and for the sake of contradiction, suppose that there exist $d_1, d_2, \dots, d_n \in \mathbb{F}$ such that

$$v = \sum_{i=1}^{n} d_i v_i$$

and $d_j \neq c_j$ for some $i \in \{1, 2, \dots, n\}$. But this means

$$\sum_{i=1}^{n} (d_i - c_i) v_i = v - v = 0$$

where $d_j - c_j \neq 0$, so we have a contradiction. For the reverse direction, suppose that we have unique representation of each vector in V by β. Then $\operatorname{span}(\beta) = V$, and β is linearly independent, since

$$\sum_{i=1}^{n} 0 v_i = 0$$

is the unique representation of $0 \in V$. ♠

Proposition 1.8. Maximal Linearly Independent Subset

Let V be a vector space. If $\beta \subseteq V$ is a maximal linearly independent subset, then β is a basis for V.

Proof. For the sake of contradiction, suppose that $\operatorname{span}(\beta) \subsetneq V$. Then there exists $v \in V \setminus \operatorname{span}(\beta)$. But this means $\beta \cup \{v\}$ is linearly independent, which violates the maximality of β, so we have a contradiction. ♠

Theorem 1.9. Existence of Basis

Let V be a vector space. Then there exists a basis β for V.

Proof. Let $S \subseteq \mathcal{P}(V)$ be the set of every linearly independent subsets of V. Then S is nonempty, since $\emptyset \in S$. Moreover, $(S, \preceq)$ is a partially ordered set. Let $C \subseteq S$ be a chain and let

$$u = \bigcup_{c \in C} c.$$

We claim that u is an upper bound for C. To verify this, we have to show that $c \in u$ for any $c \in C$ (which is clear from the definition) and that $u \in S$. So for the sake of contradiction, suppose that $u \notin S$, which means u is linearly dependent. Then there exist $v_1, v_2, \dots, v_n \in u$ such that

$$\sum_{i=1}^{n} a_i v_i = 0$$

for some nonzero $(a_1, a_2, \ldots, a_n) \in \mathbb{F}^n$. But for each $i \in \{1, 2, \ldots, n\}$, there exist $c_i \in C$ such that

$$v_i \in c_i.$$

Since C is a chain, there must exist $c \in C$ which contains $c_1, c_2, \ldots, c_n$. But $c \in C \subseteq S$, so c is linearly independent, and we have a contradiction. Thus by Zorn's lemma, there exists a maximal $\beta \in S$, which is a maximal linearly independent subset of V. By Proposition 1.8, β is a basis for V. ♠

Theorem 1.10. Replacement Theorem *Let V be a vector space. Suppose $G \subseteq V$ with $|G| = n \in \mathbb{N}$ is a spanning set and let $L \subseteq V$ be a linearly independent subset with $|L| = m \in \mathbb{N}$. Then $m \leq n$ and there exists $H \subseteq G$ with $|H| = n - m$ such that* $\operatorname{span}(L \cup H) = V$.

Proof. Write $G = \{v_1, v_2, \ldots, v_n\}$ and let $L = \{u_1, u_2, \ldots, u_m\}$. For the sake of contradiction, suppose $n < m$. Since G is a spanning set, there exist $a_1, a_2, \ldots, a_n \in \mathbb{F}$ such that

$$u_1 = \sum\nolimits_{i=1}^{n} a_i v_i.$$

Since $u_1 \neq 0$, some $a_i \neq 0$. Without loss of generality, suppose $a_1 \neq 0$. Then

$$v_1 = \frac{\sum_{i=2}^{n} a_i v_i - u_1}{-a_1},$$

which means $\{u_1, v_2, v_3, \ldots, v_n\}$ spans V. Then, for u_2, there exist $b_1, b_2, \ldots, b_n \in \mathbb{F}$ such that

$$u_2 = \sum\nolimits_{i=2}^{n} b_i v_i + b_1 u_1$$

where $b_i \neq 0$ for some $i \in \{2, 3, \ldots, n\}$, since u_1 and u_2 are linearly independent and $u_2 \neq 0$. So suppose $b_2 \neq 0$ without loss of generality. Then again, $\{u_1, u_2, v_3, v_4, \ldots, v_n\}$ spans V. By continuing this process, we obtain $\{u_1, u_2, \ldots, u_n\}$ as a spanning set for V. But this means there exist $c_1, c_2, \ldots, c_n \in \mathbb{F}$ such that

$$\sum\nolimits_{i=1}^{n} c_i u_i = u_{n+1},$$

which is a contradiction, since L is linearly independent. So $m \leq n$. Observe that H can be found in an analogous way. For, we may obtain a spanning set

$$\{u_1, u_2, \ldots, u_m, v_{m+1}, \ldots, v_n\}$$

for any $m \leq n$. ♠

Def'n. Finite-Dimensional, Infinite-Dimensional Vector Space
Let V be a vector space. We say V is ***finite-dimensional*** if there exists a finite spanning set for V. We say V is ***infinite-dimensional*** otherwise.

Corollary 1.10.1. *Let V be finite-dimensional. Then every basis for V has the same number of vectors.*

Remark 1.10. Corollary 1.10.1 enables the following definition.

Def'n. Dimension of a Finite-Dimensional Vector Space
Let V be a finite-dimensional vector space. Then the unique number of elements $n \in \mathbb{N}$ of any basis for V, denoted as $\dim(V)$, is called the ***dimension*** of V.

Proposition 1.11. Properties of Finite-Dimensional Vector Space

Let V be a vector space with dimension n. Then the following holds.

(a) Any finite generating set for V contains at least n vectors, and a generating set for V containing n vectors is a basis of V.

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

(c) Every linearly independent subset of V can be extended to a basis for V.

Proof. The result follows immediately from replacement theorem (Theorem 1.10), Corollary 1.10.1, and the definition of basis. ♠

Proposition 1.12.

Let W be a subspace of a finite-dimensional vector space V. Then W is finite-dimensional and $\dim(W) \leq \dim(V)$. *Moreover, if* $\dim(W) = \dim(V)$, *then* $V = W$.

Proof. Let $\beta = \{v_1, v_2, \ldots, v_n\}$ be a basis for V. Since the result is clear when $W = \{0\}$, assume that $W \neq \{0\}$. Then there exist some v_i such that $v_i \in W$. Without loss of generality, suppose $v_1, v_2, \ldots, v_m \in W$ for some $m \leq n$. We claim that $\alpha = \{v_1, v_2, \ldots, v_m\}$ is a basis for W. To verify this, observe that α is linearly independent. Moreover, for the sake of contradiction, suppose there exists $w \in W$ such that w cannot be expressed as a linear combination of $v_1, v_2, \ldots, v_m$. Since $w \in V$, there exist $c_1, c_2, \ldots, c_n \in \mathbb{F}$ such that

$$\sum_{i=1}^{n} c_i v_i = w,$$

and there exists $i \in \{m+1, \ldots, n\}$ such that $c_i \neq 0$ by assumption. But this means $v_i \in W$, which is a contradiction. Thus $\operatorname{span}(\alpha) = W$, so $\dim(W) \leq \dim(V)$. If $\dim(W) = \dim(V)$, then given a basis $\alpha = \{w_1, w_2, \ldots, w_n\}$ for W and $\beta = \{v_1, v_2, \ldots, v_n\}$ for V, we may replace each $w_i \in \alpha$ with $v_i \in \beta$. That is, $\operatorname{span}(\beta) = V = W$. ♠

Theorem 1.13. Basis Extension Theorem

If W is a subspace of finite-dimensional vector space V, then any basis for W can be extended to a basis for V.

Proof. This is a direct result of the replacement theorem (Theorem 1.10). For, if α is a basis for W, then α is a linearly independent subset of V. Thus, for any basis β for V, we may find a set $\gamma \subseteq \beta$ with $|\gamma| = \dim(V) - \dim(W)$ vectors such that $\alpha \cup \gamma$ is a basis for V. ♠

2.

Linear Transformations and Matrices

Linear Transformations

Remark 2.1. In previous section, we developed the theory of abstract vector space. It is now natural to consider those functions defined on vector spaces that, in some sense, preserve the structure.

Def'n. Linear Transformation on Vector Spaces
Let V, W be vector spaces over the same field $\mathbb{F}$. We say a function $T : V \to W$ is a ***linear transformation*** from V to W if T preserves the operations. That is, for any $v \in V$, $w \in W$, and $c \in \mathbb{F}$,

(a) $T(v+u) = Tv + Tu$ and

(b) $Tcv = cTv$.

Def'n. Linear Operator on a Vector Space
Let V be a vector space. A linear transformation $T : V \to V$ from V to V is called a ***linear operator***.

Remark 2.2. We denote the set of linear transformations from V to W by $\mathcal{L}(V,W)$,

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

In case of $V = W$, we write $\mathcal{L}(V)$ for simplicity.

Proposition 2.1. Properties of Linear Transformations

Let V, W be vector spaces over the same field $\mathbb{F}$ and let $T : V \to W$.

(a) If T is linear, then $T(0) = 0$.

(b) T is linear if and only if $T(cv+u) = cTv + Tu$ for all $v, u \in V$ and $c \in \mathbb{F}$.

(c) If T is linear, then $T(v-u) = Tv - Tu$ for all $v, u \in V$.

(d) T is linear if and only if

$$T\left(\sum_{i=1}^{n} c_i v_i\right) = \sum_{i=1}^{n} c_i T v_i$$

for all $c_1, c_2, \ldots, c_n \in \mathbb{F}$ and $v_1, v_2, \ldots, v_n \in V$. That is, T preserves linear combinations.

Proof. For (a), let $v \in V$. Observe that $T(0) = T(0v) = 0Tv = 0$. For (b), suppose that T is linear. Then

$$T(cv+u) = Tcv + Tu = cTv + Tu.$$

On the other hand, if $T(cv+u) = cTv + Tu$ for all $v, u \in V$ and $c \in \mathbb{F}$, then for any $x, y \in V$,

$$T(x+y) = T(1x+y) = 1Tx + Ty = Tx + Ty$$

and for any $d \in \mathbb{F}$,

$$Tdx = T(dx+0) = dTx + T(0) = dTx$$

by (a). (c) and (d) are direct consequences of (b). ♠

Def'n. Null Space (Kernel), Range (Image) of a Linear Transformation
Let V, W be vector spaces and let $T : V \to W$ be a linear transformation. We define the ***null space*** (or ***kernel***) of T, denoted as $\ker(T)$, by

$$\ker(T) = \{v \in V : Tv = 0\} \subseteq V.$$

In other words, $\ker(T)$ is the set of vectors in V which are mapped to 0 by T. Moreover, we define the

range (or ***image***) of T, denoted as image(T), by

$$\text{image}(T) = \{w \in W : \exists v \in V\,[w = Tv]\} \subseteq W.$$

Remark 2.3. An important property of the null space and range of a linear transformation is that they are subspaces of their respective supersets.

Proposition 2.2. Null Space and Range Are Subspaces

Let V, W be vector spaces and let $T : V \to W$ be linear. Then $\ker(T) \subseteq V$ and $\text{image}(T) \subseteq W$ are subspaces.

Proof. To verify that $\ker(T) \subseteq V$ is a subspace, first observe that $0 \in \ker(T)$, since $T(0) = 0$. Moreover, if $v, u \in \ker(T)$ and $c \in \mathbb{F}$, then

$$T(cv+u) = cTv + Tu = 0$$

so $cv+u \in \ker(T)$ as well. To verify that $\text{image}(T) \subseteq W$ is a subspace, observe that $T(0) = 0 \in \text{image}(T)$. Moreover, if $w, z \in \text{image}(T)$ and $c \in \mathbb{F}$, then there exist $v, u \in V$ such that $w = Tv$ and $z = Tu$. That is,

$$cw + z = cTv + Tu = T(cv+u) \in \text{image}(T).$$

♠

Remark 2.4. The following proposition provides a method for finding a spanning set for the range of a linear transformation.

Proposition 2.3.

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

$$\ker(T) = \text{span}\,(T(\beta)) = \text{span}\,\{Tv_1, Tv_2, \ldots, Tv_n\}\,.$$

Proof. Suppose $\text{image}(T) \neq \text{span}(T\beta)$. Then there exists $Tv \in \text{image}(T)$ such that

$$Tv \neq a_1 Tv_1 + a_2 Tv_2 + \cdots + a_n Tv_n$$

for any $(a_1, a_2, \ldots, a_n) \in \mathbb{F}^n$. So

$$Tv \neq T \sum_{i=1}^{n} a_i v_i,$$

which means

$$v \neq \sum_{i=1}^{n} a_i v_i$$

for any $(a_1, a_2, \ldots, a_n) \in \mathbb{F}^n$. This is a contradiction, since β is a basis for V. ♠

Def'n. Nullity, Rank of a Linear Transformation

Let V, W be vector spaces and let $T : V \to V$ be linear. We define the ***nullity*** of T, denoted as nullity(T), by

$$\text{nullity}(T) = \dim\,(\ker(T))\,,$$

if $\ker(T)$ is finite-dimensional. Moreover, we define the ***rank*** of T, denoted as rank(T), by

$$\text{rank}(T) = \dim\,(\text{image}(T))\,,$$

if image(T) is finite-dimensional. That is, the nullity and rank of a linear transformation are the dimension of the associated null space and the range.

Remark 2.5. By thinking the definition of null space and kernel, one may find it intuitive to think that, given a linear transformation, larger the nullity, smaller the rank. In other words, more vectors are mapped to 0, the smaller the range. The next theorem address the balance between the rank and nullity of a linear transformation.

Theorem 2.4.
Rank-Nullity Theorem (Dimension Theorem)

Let V and W be vector spaces, and let $T : V \to W$ be linear. If V is finite-dimensional, then

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

Proof. Let $n = \dim(V), k = \text{nullity}(T) \in \mathbb{N}$ for convenience. Let $\beta_N = \{v_1, v_2, \ldots, v_k\}$ be a basis for $\ker(T)$. By basis extension theorem, there exists vectors $v_{k+1}, v_{k+2}, \ldots, v_n \in V$ such that

$$\beta = \beta_N \cup \{v_{k+1}, v_{k+2}, \ldots, v_n\}$$

where β is a basis for V. Now the claim is that $\beta_R = \{v_{k+1}, v_{k+2}, \ldots, v_n\}$ is a basis for $\text{image}(T)$. To verify this claim, first observe that $\text{span}(T\beta) = \text{image}(T)$ implies that

$$\text{image}(T) = \text{span}(T\beta) = \text{span}(T(\beta \setminus \beta_N)) = \text{span}(T(\beta_R)),$$

since $T(v_i) = 0$ for all $v_i \in \beta_N$. For the linear independence part, consider

$$\sum_{k+1}^{n} c_i T v_i = 0$$

for some $c_i \in \mathbb{F}$. Since T is linear,

$$T \sum_{k+1}^{n} c_i v_i = 0$$

so $\sum_{k+1}^{n} c_i v_i \in \ker(T)$. Therefore, there exist $b_1, b_2, \ldots, b_k \in \mathbb{F}$ such that

$$\sum_{k+1}^{n} c_i v_i = \sum_{1}^{k} b_i v_i$$

which can be also written as

$$\sum_{k+1}^{n} c_i v_i + \sum_{1}^{k} b_1 v_1 = 0.$$

But since β is linearly independent, $(b_1, b_2, \ldots, b_n), (c_1, c_2, \ldots, c_n) = 0$. So β_R is linearly independent. Thus

$$\dim(\text{image}(T)) = \text{rank}(T) = n - k$$

as desired. ♠

Recall. Injective, Surjective, Bijective Function
Let A, B sets and let $f : A \to B$ a function. We say f is ***injective*** if for all $y \in B$, there exist a unique $x \in A$ such that $f(x) = y$. We say f is ***surjective*** if there exist $x \in A$ such that $f(x) = y$ for all $y \in B$. We say f is ***bijective*** if f is injective and surjective.

Proposition 2.5.
T Is Surjective If and Only If T Is Injective

Let V, W be vector spaces over F and let $T : V \to W$ be linear. Then T is surjective if and only if $\ker(T) = \{0\}$.

Proof. For the forward direction, suppose that T is surjective and let $x \in \ker(T)$ be arbitrary. Then $Tx = 0 = T(0)$, so by definition of surjection, $x = 0$. For the reverse direction, suppose $\ker(T) = \{0\}$ and $Tx = Ty$. Then

$$Tx - Ty = T(x - y) = 0$$

so it must be that $x - y \in \{0\}$. That is, $x = y$, which means T is surjective, as desired. ♠

Proposition 2.6. *Let V and W be vector spaces over $\mathbb{F}$ of equal finite dimension, and let $T: V \to W$ be linear. Then the following are equivalent.*

(a) *T is injective.*

(b) *T is surjective.*

(c) $\operatorname{rank}(T) = \dim(V)$.

Proof. Proposition 2.5 supplies (a) $\iff$ (b). Observe that

$$T \text{ is injective} \iff \ker(T) = \{0\} \iff \operatorname{nullity}(T) = 0 \iff \operatorname{rank}(T) = \dim(V). \quad \spadesuit$$

Remark 2.6. One of the most important properties of a linear transformation is that it is completely determined by its action on a basis. This result follows from the next theorem and corollary.

Theorem 2.7. Characterization of a Linear Transformation *Let V and W be vector spaces over $\mathbb{F}$ and let $\{v_1, v_2, \dots, v_n\}$ be a basis for V. Moreover, let $w_1, w_2, \dots, w_n \in W$. Then there exists a unique linear $T: V \to W$ such that*

$$Tv_i = w_i$$

for all $i \in \{1, 2, \dots, n\}$.

Proof. Define $T: V \to W$ by

$$v = \sum_{i=1}^{n} c_i v_i \mapsto \sum_{i=1}^{n} c_i w_i.$$

We claim that T is the desired linear transformation. To verify this, let $x, y \in V$ and $c \in \mathbb{F}$. Then $x = \sum_{i=1}^{n} a_i v_i$ and $y = \sum_{j=1}^{n} b_i v_i$ for some $a_1, a_2, \dots, a_n, b_1, \dots, b_n \in \mathbb{F}$, and

$$\begin{aligned} T(cx+y) &= T\left(c\sum_{i=1}^{n} a_i v_i + \sum_{i=1}^{n} b_i v_i\right) = T\sum_{i=1}^{n}(ca_i + b_i)v_i \\ &= \sum_{i=1}^{n}(ca_i + b_i)w_i = c\sum_{i=1}^{n} a_i w_i + \sum_{i=1}^{n} b_i w_i = cTx + Ty. \end{aligned}$$

Moreover, clearly

$$Tv_i = w_i.$$

To verify the uniqueness, let $S: V \to W$ be an arbitrary linear transformation satisfying $Sv_i = w_i$. Then

$$Sx = S\sum_{i=1}^{n} a_i v_i = \sum_{i=1}^{n} a_i S v_i = \sum_{i=1}^{n} a_i w_i = Tx,$$

as desired. $\spadesuit$

Corollary 2.7.1. *Let V, W be a finite-dimensional vector spaces, $\{v_1, v_2, \dots, v_n\}$ be a basis for V, and $T, U: V \to W$ be linear transformations satisfying $Tv_i = Uv_i$ for all $i \in \{1, 2, \dots, n\}$. Then $T = U$.*

Matrix Representations of Linear Transformations

Remark 2.7. Until now, every linear transformation is described by examining its range and null space. In this section, we begin another yet useful approach to describe linear transformation over a vector space: matrix representation of a linear transformation. In fact, we are going to show that there is a special kind of bijection (called isomorphism) between matrices and linear transformations.

Def'n. Orderd Basis of a Vector Space
Let V be a finite-dimensional vector space. An ***ordered basis*** for V is a basis for V endowed with a specific order.

Remark 2.8. Given any ordered basis, we may describe any vector in the vector space by using coordinate vectors, n-tuples ($n = \dim(V)$) that identify abstract vectors in V. For, if $\beta = \{v_1, v_2, \ldots, v_n\}$ is an ordered basis for a vector space V over $\mathbb{F}$, then there exists a unique representation of any $v \in V$ as a linear combination of $v_1, v_2, \ldots, v_n$,

$$v = \sum_{i=1}^{n} a_i v_i$$

for some $a_1, a_2, \ldots, a_n$.

Def'n. Coordinate Vector in an Orderd Basis
Let $\beta = \{v_1, v_2, \ldots, v_n\}$ be an ordered basis for a vector space V over $\mathbb{F}$. A unique n-tuple $(a_1, a_2, \ldots, a_n)$ such that

$$v = \sum_{i=1}^{n} a_i v_i$$

is called the ***coordinate vector*** of v relative to β and denoted by

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

Remark 2.9. We shall show later that the function $[\cdot]_\beta : V \to \mathbb{F}^n$ for any ordered basis β for V is linear.

Def'n. Matrix over a Field
We say an $m \times n$ rectangular array A

$$\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mn} \end{bmatrix}$$

of scalars $A_{ij} \in \mathbb{F}$ a ***matrix*** over $\mathbb{F}$.

Remark 2.10. We shall consistently denote the set of $m \times n$ matrices over $\mathbb{F}$ by $M_{m\times n}(\mathbb{F})$ and the entry on ith row and jth column of a matrix $A \in M_{m\times n}(\mathbb{F})$ by A_{ij}.

Remark 2.11. Similar to how a vector is represented as a coordinate vector in an ordered basis, we may represent a linear transformation as a matrix in an ordered basis. That is, if V, W are finite-dimensional vector spaces with ordered basis $\beta = \{v_1, v_2, \ldots, v_n\}$ and $\gamma = \{w_1, w_2, \ldots, w_m\}$. Then for each $j \in \{1, 2, \ldots, n\}$ there exists unique scalars $a_{ij} \in \mathbb{F}$ such that

$$Tv_j = \sum_{i=1}^{m} a_{ij} w_i.$$

Def'n. Matrix Representation of a Linear Transformation
Consider Remark 2.11. We say the matrix $A \in M_{m\times n}(\mathbb{F})$ defined by $A_{ij} = a_{ij}$ the ***matrix representation***

of T in the ordered bases β and γ. We denote the matrix representation of T by

$$A = [T]_{\beta}^{\gamma}.$$

Remark 2.12. When $V = W$ and $\beta = \gamma$, we write

$$[T]_{\beta} = [T]_{\beta}^{\gamma}$$

for convenience.

Remark 2.13. By Corollary 2.7.1, if $T, U : V \to W$ are linear transformations satisfying $[T]_{\beta}^{\gamma} = [U]_{\beta}^{\gamma}$ then $T = U$. Moreover, observe that jth column of $[T]_{\beta}^{\gamma}$ is $[Tv_j]_{\gamma}$, the coordinate vector of Tv_j relative to γ. In other words,

$$[T]_{\beta}^{\gamma} = \left[[Tv_1]_{\gamma} [Tv_2]_{\gamma} \cdots [Tv_n]_{\gamma} \right].$$

Remark 2.14. Now that we have defined an association from linear transformations to matrices, we are going to prove that

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

for any ordered basis β for V and γ for W is linear, where V, W are finite-dimensional vector spaces. But to do so, we first show that $\mathcal{L}(V, W)$ is a vector space under the following operations: Define addition and scalar multiplications such that

$$(cT + U)v = cTv + Uv$$

for any $T, U \in \mathcal{L}(V, W)$, $c \in \mathbb{F}$, and $v \in V$. Then it is a routine computation to show that $cT + U : V \to W$ is linear (and hence $\mathcal{L}(V, W)$ is closed under the provided operations) and that $\mathcal{L}(V, W)$ is a vector space.

Proposition 2.8. $[\cdot]_{\beta}^{\gamma} : \mathcal{L}(V, W) \to M_{m \times n}(\mathbb{F})$ **Is Linear**

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

$$[cT + U]_{\beta}^{\gamma} = c[T]_{\beta}^{\gamma} + [U]_{\beta}^{\gamma}.$$

Proof. Consider writing $\beta = \{v_1, v_2, \ldots, v_n\}$ and $\gamma = \{w_1, w_2, \ldots, w_m\}$. Then there exist unique scalars a_{ij}, b_{ij} such that $Tv_j = \sum_{i=1}^{m} a_{ij} w_i$ and $Uv_j = \sum_{i=1}^{m} b_{ij} w_i$ for all $j \in \{1, 2, \ldots, n\}$. Hence

$$(cT + U)v_j = cTv_j + Uv_j = c\sum_{i=1}^{m} a_{ij} w_i + \sum_{i=1}^{m} b_{ij} w_i = \sum_{i=1}^{m} (ca_{ij} + b_{ij}) w_i.$$

Thus for all $i \in \{1, 2, \ldots, m\}$ and $j \in \{1, 2, \ldots, n\}$,

$$\left([cT + U]_{\beta}^{\gamma}\right)_{ij} = ca_{ij} + b_{ij} = c\left([T]_{\beta}^{\gamma}\right)_{ij} + \left([U]_{\beta}^{\gamma}\right)_{ij}.$$

as desired. ♠

Compositions of Linear Transformations

Remark 2.15. For simplicity, we shall write TU to denote the composition of linear transformations T and U.

Proposition 2.9. **The Composition of Linear Transformations Is Linear**

Let V, W, Z be vector spaces over $\mathbb{F}$ and let $T : V \to W$ and $U : W \to Z$ be linear. Then $UT : V \to Z$ is linear.

Proof. Let $x, y \in V$ and $c \in \mathbb{F}$ be arbitrary. Then observe that

$$UT(cx+y) = U(T(cx+y)) = U(cTx+Ty) = cU(Tx)+U(Ty) = cUTx+UTy. \qquad \spadesuit$$

Proposition 2.10. **Properties of the Composition of Linear Transformations**

Let V be a vector space and let $T, U_1, U_2 : V \to V$ be linear operators.

(a) $T(U_1+U_2) = TU_1+TU_2$ and $(U_1+U_2)T = U_1T+U_2T$.

(b) $(TU_1)U_2 = T(U_1U_2)$.

(c) $TI = IT = T$, where I is the identity operator. That is, $Iv = v$ for all $v \in V$.

(d) $a(U_1U_2) = (aU_1)U_2 = U_1(aU_2)$ for all $a \in \mathbb{F}$.

Remark 2.16. Although we have stated Proposition 2.10 in terms of linear operators for simplicity, (a), (b), and (d) are valid for any linear transformations T, U_1, U_2 such that the provided compositions are well-defined.

Remark 2.17. Aside from the addition of matrices and scalar multiplication, an important operation that is not yet introduced is matrix product. Suppose $T : V \to W$ and $U : W \to Z$ are linear, where $\alpha = \{v_1, v_2, \dots, v_n\}, \beta = \{w_1, w_2, \dots, w_m\}, \gamma = \{z_1, z_2, \dots, z_p\}$ are bases of V, W, Z, respectively. Given this, the motivation is to define the product of $[T]_\alpha^\beta$ and $[U]_\beta^\gamma$ to be

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

Def'n. Product of Matrices

Let $A \in M_{m\times n}(\mathbb{F})$ and $B \in M_{n\times p}(\mathbb{F})$. We define the ***product*** of A and B, denoted by AB, to be the $m \times p$ matrix with entries

$$(AB)_{ij} = \sum\nolimits_{k=1}^{n} A_{ik}B_{kj}.$$

Remark 2.18. We verify that the above definition of matrix multiplication is consistent with our motivation.

Proposition 2.11.

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

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

Proof. Write $\alpha = \{v_1, v_2, \dots, v_m\}, \beta = \{w_1, w_2, \dots, w_n\}, \gamma = \{z_1, z_2, \dots, z_p\}$ for convenience. Then

$$Tv_i = \sum\nolimits_{j=1}^{m} \left([T]_\alpha^\beta\right)_{ji} w_j$$

and

$$Tw_j = \sum\nolimits_{k=1}^{p} \left([U]_\beta^\gamma\right)_{kj} z_k$$

for all $i \in \{1,2,\dots,m\}$ and $j \in \{1,2,\dots,n\}$. That is,

$$\begin{aligned} UTv_i &= U\sum_{j=1}^{n}\left([T]_\alpha^\beta\right)_{ji} w_j = \sum_{j=1}^{n}\left([T]_\alpha^\beta\right)_{ji}\sum_{k=1}^{p}\left([U]_\beta^\gamma\right)_{kj} z_k \\ &= \sum_{j=1}^{n}\sum_{k=1}^{p}\left([U]_\beta^\gamma\right)_{kj}\left([T]_\alpha^\beta\right)_{ji} z_k = \sum_{k=1}^{p}\left([U]_\beta^\gamma [T]_\alpha^\beta\right)_{ki} z_k \end{aligned}$$

for each $i \in \{1,2,\dots,m\}$, as desired. ♠

Remark 2.19. Matrix multiplication is not commutative nor cancellative. That is, given $A, B \in M_{n\times n}(\mathbb{F})$, it need not be the case which $AB = BA$, and there exist some nonzero $C \in M_{m\times n}(\mathbb{F})$ and $D \in M_{n\times p}(\mathbb{F})$ such that $CD = 0$, the $m \times p$ zero matrix.

Def'n. Transpose of a Matrix

Let $A \in M_{m\times n}(\mathbb{F})$. We define the ***transpose*** of A, denoted as $A^T \in M_{n\times m}(\mathbb{F})$, by

$$A^T_{ij} = A_{ji}.$$

Remark 2.20. Here are some remarks about the transpose operation. Let $A, B \in M_{m\times n}(\mathbb{F})$ and $c \in \mathbb{F}$. Then

$$(cA+B)^T_{ij} = (cA+B)_{ji} = cA_{ji} + B_{ji} = cA^T_{ij} + B^T_{ij}.$$

That is,

$$(cA+B)^T = cA^T + B^T.$$

Moreover, let $C \in M_{m\times n}(\mathbb{F})$ and $D \in M_{n\times p}(\mathbb{F})$. Then observe that

$$(CD)^T_{ij} = CD_{ji} = \sum_{k=1}^{n} C_{jk}D_{ki} = \sum_{k=1}^{n} D_{ki}C_{jk} = \sum_{k=1}^{n} D^T_{ik}C^T_{kj} = \left(D^T C^T\right)_{ij}.$$

That is, $(CD)^T = D^T C^T$.

Def'n. Kronecker Delta

The ***Kronecker delta*** δ_{ij} is defined as

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}.$$

Remark 2.21. We have a natural analogue of an identity operator. That is, if we define $I \in M_{n\times n}(\mathbb{F})$ by

$$I_{ij} = \delta_{ij},$$

then it is an easy computation to verify that

$$IA = AI = A$$

for any $A \in M_{n\times n}(\mathbb{F})$.

Proposition 2.12. Properties of Matrix Product

Let $A \in M_{m\times n}(\mathbb{F}), B, C \in M_{n\times p}(\mathbb{F})$, and $D, E \in M_{q\times m}(\mathbb{F})$.

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

(b) $a(AB) = (aA)B = A(aB)$ for any $a \in \mathbb{F}$.

(c) $I_m A = AI_n = A$, where $I_k \in M_{k\times k}(\mathbb{F})$ is the $k \times k$ identity matrix.

(d) *Let V be a finite-dimensional vector space, $I : V \to V$ be the identity operator, and let β be an orderd basis for V. Then $[I]_\beta = I$.*

Proposition 2.13. *Let V and W be vector spaces with ordered bases β and γ, respectively, and let $T : V \to W$ be linear. Then,*

$$[Tv]_\gamma = [T]_\beta^\gamma [v]_\beta .$$

Proof. Write $\beta = \{v_1, v_2, \ldots, v_n\}$ and $\gamma = \{w_1, w_2, \ldots, w_m\}$. Let

$$v = \sum_{i=1}^{n} a_i v_i$$

for some $a_1, a_2, \ldots, a_n \in \mathbb{F}$. Then $\left([v]_\beta\right)_i = a_i$ and

$$\begin{aligned} Tv = T\sum_{i=1}^{n} a_i v_i &= \sum_{i=1}^{n} \left([v]_\beta\right)_i \sum_{j=1}^{m} \left([T]_\beta^\gamma\right) w_j \\ &= \sum_{j=1}^{m} \sum_{i=1}^{n} \left([T]_\beta^\gamma\right)_{ji} \left([v]_\beta\right)_i w_j = \sum_{j=1}^{m} \left([T]_\beta^\gamma [v]_\beta\right)_j w_j, \end{aligned}$$

as desired. ♠

Remark 2.22. Let $A \in M_{m\times n}(\mathbb{F})$. Then Proposition 2.13 suggests that there exist a function $L_A : \mathbb{F}^n \to \mathbb{F}^m$ defined by

$$v \mapsto Av.$$

Def'n. Left Multiplication Transformation of a Matrix
Consider Remark 2.22. We say L_A the ***left multiplication transformation*** of A.

Remark 2.23. We shall consistentely write L_A to denote the left multiplication transformation of $A \in M_{m\times n}(\mathbb{F})$.

Proposition 2.14. *Let $A \in M_{m\times n}(\mathbb{F})$ and $L_A : \mathbb{F}^n \to \mathbb{F}^m$ be the left multiplication transformation of A. Then L_A is linear. Moreover, let $B \in M_{m\times n}(\mathbb{F})$ and let $L_B : \mathbb{F}^n \to \mathbb{F}^m$.*

(a) *$[L_A]_\beta = A$.*

(b) *$L_A = L_B$ if and only if $A = B$.*

(c) *$L_{A+B} = L_A + L_B$ and $L_{aA} = aL_A$ for all $a \in \mathbb{F}$.*

(d) *If $T : \mathbb{F}^n \to \mathbb{F}^m$ is linear, then there exists a unique $C \in M_{m\times n}(\mathbb{F})$ such that $T = L_C$. In fact, $C = [T]_\beta^\gamma$.*

(e) *If $E \in M_{n\times p}(\mathbb{F})$, then $L_{AE} = L_A L_E$.*

(f) *$L_I = I$.*

Here, β and γ denotes the standard ordered bases for $\mathbb{F}^n$ and $\mathbb{F}^m$, respectively.

Proposition 2.15.
Matrix Multiplication Is Associative

Let A, B, C be matrices such that $A(BC)$ is defined. Then $(AB)C$ is also defined and $A(BC) = (AB)C$. That is, matrix product is associative.

Proof. Let $A \in M_{m\times n}(\mathbb{F})$ then for $A(BC)$ to be defined, $(BC) \in M_{n\times q}(\mathbb{F})$. That is, $B \in M_{n\times p}(\mathbb{F})$ and $C \in M_{p\times q}(\mathbb{F})$. So AB is defined, and $AB \in M_{m\times p}$, so $(AB)C$ is defined as well. From (e) of Proposition 2.14,

$$L_A(L_{BC}) = L_A(L_B L_C) = (L_A L_B)L_C = (L_{AB})L_C.$$

Then the result follows from (b) of Proposition 2.14. ♠

Invertibility and Isomorphisms

Remark 2.24. By utilizing invertibility of functions, we may investigate inverse of a matrix by using linear transformation $L_A = A$. In particular, if T is linear, then T^{-1} is linear, so L_A^{-1} can be used to determine properties of A^{-1}. It turns out that the concept of isomorphism is related to invertibility, which is discussed in this section as well. Consider the following definition.

> **Def'n. Inverse** of a Linear Transformation
> Let V, W be vector spaces and let $T : V \to W$ be linear. We say a function $U : W \to V$ is the ***inverse*** of T if $TU = I : W \to W$ and $UT = I : V \to V$. If such U exists, then we say T is ***invertible***, and denote U by T^{-1}.

Remark 2.25. Suppose that f and g are invertible functions such that fg is well-defined. Then

(a) $(fg)^{-1} = g^{-1}f^{-1}$ and

(b) $\left(f^{-1}\right)^{-1} = f$. In particular, f^{-1} is invertible.

An important property of invertible function is that f is invertible if and only if f is bijective. Combining this with Proposition 2.6 shows that, a linear transformation $T : V \to W$ is invertible if and only if $\dim(V) = \operatorname{rank}(T)$, where V, W are some vector spaces. Moreover, it turns out that $T^{-1} : W \to V$ is linear as well.

Proposition 2.16.
T^{-1} Is Linear

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

Proof. Let $v, u \in V$ and $c \in \mathbb{F}$ be arbitrary. Then

$$T^{-1}(cTv + Tu) = T^{-1}(T(cv+u)) = cv + u = cT^{-1}(Tv) + T^{-1}(Tu).$$ ♠

Remark 2.26. We also have a natural matrix analogue of the inverse of a linear transformation.

> **Def'n. Inverse** of a Matrix
> Let $A \in M_{n\times n}(\mathbb{F})$. Then A is ***invertible*** if there exists $B \in M_{n\times n}(\mathbb{F})$ such that $AB + BA = I$. Such B is unique and we denote $B = A^{-1}$.

Proposition 2.17.

Let V, W be finite-dimensional vector spaces. Then there exists an invertible $T : V \to W$ if and only if $\dim(V) = \dim(W)$.

Proof. From Remark 2.25, T is invertible if and only if $\dim(V) = \operatorname{rank}(T)$. But by Proposition 2.6, T is invertible if and only if T is surjective, which exactly means $\operatorname{rank}(T) = \dim(W)$. For the reverse direction, if $\dim(V) = \dim(W)$, and $\{v_1, v_2, \ldots, v_n\}$ and $\{w_1, w_2, \ldots, w_n\}$ are bases for V and W respectively, then $T : V \to W$ defined by

$$v_i \mapsto w_i$$

is clearly invertible. ♠

Proposition 2.18. T **Is Invetible If and Only If** $[T]_\beta^\gamma$ **Is Invertible**

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

Proof. For the forward direction, suppose that T is invertible. Then by Proposition 2.17, $\dim(V) = \dim(W) = n$, so $[T]_\beta^\gamma \in M_{n\times n}(\mathbb{F})$. Since $T^{-1} : W \to V$ satisfies $TT^{-1} = I : W \to W$,

$$[T]_\beta^\gamma [T^{-1}]_\gamma^\beta = [TT^{-1}]_\gamma^\gamma = [I]_\gamma^\gamma$$

so $[T]_\beta^\gamma$ is invertible and $\left([T]_\beta^\gamma\right)^{-1} = \left[T^{-1}\right]_\gamma^\beta$. For the reverse direction, suppose $A = \left([T]_\beta^\gamma\right)$ is invertible. Then there is $B \in M_{n\times n}(\mathbb{F})$ such that $AB = BA = I$. Then there exists a unique $U : W \to V$ such that

$$Uw_j = \sum_{i=1}^n B_{ij} v_i,$$

for all $j \in \{1, 2, \ldots, n\}$, where $\gamma = \{w_1, w_2, \ldots, w_n\}$ and $\beta = \{v_1, v_2, \ldots, v_n\}$. By definition, $B = [U]_\gamma^\beta$, and

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

so $UT = I$ and, similarly, $TU = I$. ♠

Corollary 2.18.1.

Let $A \in M_{n\times n}(\mathbb{F})$. Then A is invertible if and only if L_A is invertible. Furthermore, $L_A^{-1} = L_{A^{-1}}$.

Def'n. Isomorphism on Vector Spaces

Let V, W be vector spaces. We say $T : V \to W$ is an ***isomorphism*** if T is linear and invertible. If such T exists, then we say that V and W are ***isomorphic***, denoted as

$$V \cong W.$$

Remark 2.27. Isomorphism is an equivalence relation.

Proposition 2.19. $V \cong W$ **If and Only If** $\dim(V) = \dim(W)$

Let V, W be finite-dimensional vector spaces. Then V and W are isomorphic to each other if and only if $\dim V = \dim W$.

Proof. For the forward direction, suppose that $T : V \to W$ is an isomorphism. Then $\operatorname{nullity}(T) = 0$ so by rank-nullity theorem,

$$\dim(V) = \operatorname{rank}(T) = \dim(W).$$

For the reverse direction, suppose $\dim(V) = \dim(W)$. Then there exists a linear transformation $T : V \to W$ that satisfies

$$Tv_i = w_i,$$

where $\{v_1, v_2, \ldots, v_n\}$ and $\{w_1, w_2, \ldots, w_n\}$ are bases of V and W, respectively. Then T is an isormorphism, since

$$\text{image}(T) = \text{span}\{w_1, w_2, \ldots, w_n\} = W.$$

♠

Corollary 2.19.1. *Let V be a finite-dimensional vector space over $\mathbb{F}$. Then $V \cong \mathbb{F}^n$ if and only if $\dim(V) = n$.*

Remark 2.28. We proceed to discuss the natural isomorphism between $\mathcal{L}(V,W)$ and $M_{m\times n}(\mathbb{F})$, where $\dim(V) = n$ and $\dim(W) = m$, as mentioned in Remark 2.7.

Theorem 2.20. $\mathcal{L}(V,W) \cong M_{m\times n}(\mathbb{F})$

Let V and W be fintie dimensional vector spaces with $\dim(V) = n$ and $\dim(W) = m$, and let β and γ be ordered bases for V and W, respectively. Then the function $[\cdot]_\beta^\gamma : \mathcal{L}(V,W) \to M_{m\times n}(\mathbb{F})$ is an isomorphism.

Proof. The linearity of $[\cdot]_\beta^\gamma$ is supplied by Proposition 2.8. Now the claim is that $[\cdot]_\beta^\gamma$ is surjective. To verify this, let $A \in M_{m\times n}(\mathbb{F})$ be arbitrary. Then there exists a (unique) linear $T : V \to W$ such that

$$Tv_j = \sum\nolimits_{i=1}^m A_{ij} w_i$$

for all $j \in \{1, 2, \ldots, n\}$. It follows from Proposition 2.6 that $[\cdot]_\beta^\gamma$ is bijective. ♠

Corollary 2.20.1. *Let V, W be finite-dimensional vector spaces with $\dim(V) = n$ and $\dim(W) = m$. Then*

$$\dim(\mathcal{L}(V,W)) = mn.$$

Remark 2.29. Similar to how we define a natural isomorphism in terms of ordered bases, $[\cdot]_\beta : V \to \mathbb{F}^n$ for any ordered basis β for a finite-dimensional vector space V is also an isomorphism. The linearity of $[\cdot]_\beta$ follows immediately from the definition of the coordinate vector, and $[\cdot]_\beta$ is surjective, since for any $(c_1, c_2, \ldots, c_n) \in \mathbb{F}$,

$$v = \sum\nolimits_{i=1}^n c_i v_i \in V$$

is the unique representation of v as a linear combination of vectors in an ordered basis $\beta = \{v_1, v_2, \ldots, v_n\}$ for V.

> **Def'n. Standard Representation** of a Vector Space
> Let V be a n-dimensional vector space and let β be an ordered basis for V. The ***standard representation*** of V with respect to β is the isomorphism $[\cdot]_\beta : V \to \mathbb{F}^n$.

Remark 2.30. By using standard representation, we may restate Proposition 2.13 as follows.

Proposition 2.21. *Let V, W be finite-dimensional vector spaces with $\dim(V) = n$ and $\dim(W) = m$, and let β and γ be ordered bases for V and W, respectively. Let $T : V \to W$ be linear and let $A = [T]_\beta^\gamma$. Then*

$$L_A [\cdot]_\beta = [\cdot]_\gamma T.$$

The Change of Basis

Proposition 2.22. *Let V be a finite-dimensional vector space and let β and γ be ordered bases for V. Let $Q = [I]_\gamma^\beta$. Then Q is invertible, and*

$$[v]_\beta = Q[v]_\gamma$$

for all $v \in V$.

Proof. The invertibility of Q is a direct consequence of the invertibility of $I : V \to V$. Moreover, by Proposition 2.13,

$$[v]_\beta = [Iv]_\beta = [I]_\gamma^\beta [v]_\gamma = Q[v]_\gamma.$$ ♠

Def'n. Change of Basis Matrix
Consider Proposition 2.22. We call Q the ***change of basis matrix*** from γ to β.

Remark 2.31. Let V be a finite-dimensional vector space and let $\beta = \{v_1, v_2, \ldots, v_n\}$ and $\gamma = \{u_1, u_2, \ldots, u+n\}$ be ordered bases for V. Let $Q \in M_{n\times n}(\mathbb{F})$ be the change of basis matrix from γ to β. Then

$$u_j = \sum_{i=1}^n Q_{ij} v_i$$

for all $j \in \{1, 2, \ldots, n\}$. That is, the jth column of of Q is $[u_j]_\beta$. Moreover Q^{-1} is the change of basis matrix from β to γ

Proposition 2.23. *Let T be a linear operator on V and β and γ be ordered bases for V. Suppose that Q is the change of basis matrix from γ to β. Then,*

$$[T]_\gamma = Q^{-1}[T]_\beta Q.$$

Proof. Observe that

$$Q[T]_\gamma = [I]_\gamma^\beta [T]_\gamma = [IT]_\gamma^\beta = [TI]_\gamma^\beta = [T]_\beta [T]_\gamma^\beta = [T]_\beta Q.$$

That is, $[T]_\gamma = Q^{-1}[T]_\beta Q$. ♠

Corollary 2.23.1. *Let $A \in M_{n\times n}(\mathbb{F})$ and γ be an ordered basis for $\mathbb{F}^n$. Then $[L_A]_\gamma = Q^{-1}AQ$, where Q is the matrix whose jth column is the jth vector of γ.*

Remark 2.32. Proposition 2.23 and Corollary 2.23.1 motivates the following definition.

Def'n. Similar Matrices
Let $A, B \in M_{n\times n}(\mathbb{F})$. We say A and B are ***similar*** if there exists an invertible $P \in M_{n\times n}(\mathbb{F})$ such that $B = P^{-1}AP$.

3. Linear Equations

Elementary Matrix Operations and Elementary Matrices

Def'n. Elementary Operations on Matrix
Let $A \in M_{m\times n}(\mathbb{F})$. We define the three ***elementary row operations*** on A as follows.

(a) Interchange any two rows of A.

(b) Multiply any row of A by a nonzero scalar $c \in \mathbb{F}$.

(c) Add a scalar multiple of any row of A to another row of A.

We say (a) is a ***type 1*** operation, (b) is a ***type 2*** operation, and (c) is a ***type 3*** operation.

Remark 3.1. We have an analogous definition for the three elementary column operations.

Def'n. Elementary Matrix
We say a matrix $E \in M_{n\times n}(\mathbb{F})$ is an ***elementary matrix*** if E is obtained by applying any elementary operation on I. Depending on which type of operation is used, we call that E is a ***type 1***, ***type 2***, or ***type 3*** elementary matrix, respectively.

Remark 3.2. Any elementary matrix can be obtained by at least two ways. That is, if $E \in M_{n\times n}(\mathbb{F})$ is obtained by applying an elementary row operation on I, then there is an elementary column operation of the same type by applying which E is obtained from I.

Theorem 3.1.

Let $A \in M_{m\times n}(\mathbb{F})$, and suppose that B is obtained from A by performing an elementary column [row] operation on A. Then there exists an $n\times n$ [$m\times m$] elementary matrix E such that $B = AE$ [$B = EA$]. In fact, E is obtained from I by performing the same elementary column [row] operation. Conversely, for any $n\times n$ [$m\times m$] elementary matrix E, $B = AE$ [$B = EA$] is the matrix obtained by performing an elementary column [row] operation on I to obtain E.

Proposition 3.2.
Invertibility of Elementary Matrices

Let $E \in M_{n\times n}(\mathbb{F})$ be elementary. Then E is invertible and E^{-1} is an elementary matrix of the same type.

Proof. Let $E \in M_{n\times n}(\mathbb{F})$ be invertible. Then E can be obtained from I by performing elementary operations, so I can be obtained from E by reversing the operations. Then by Theorem 3.1, there exists an $E' \in M_{n\times n}(\mathbb{F})$ such that $EE' = E'E = I$. Therefore E is invertible and $E' = E^{-1}$. ♠

The Rank and Inverse of a Matrix

Def'n. Rank of a Matrix
Let $A \in M_{n\times n}(\mathbb{F})$. We define the ***rank*** of A, denoted by $\operatorname{rank}(A)$, to be the rank of the left multiplication transformation $L_A : \mathbb{F}^n \to \mathbb{F}^m$.

Theorem 3.3.
Invertible Matrix Theorem I

Let $A \in M_{m\times n}(\mathbb{F})$, and let $P \in M_{m\times m}(\mathbb{F}), Q \in M_{n\times n}(\mathbb{F})$ be invertible.

(a) $\operatorname{rank}(AQ) = \operatorname{rank}(A)$.

(b) $\operatorname{rank}(PA) = \operatorname{rank}(A)$.

(c) $\operatorname{rank}(PAQ) = \operatorname{rank}(A)$.

Proof. For (a), observe that

$$\operatorname{rank}(AQ) = \operatorname{rank}(L_{AQ}) = \operatorname{rank}(L_A L_Q) = \dim(L_A L_Q(\mathbb{F}^n)) = \dim(L_A(\mathbb{F}^n)) = \operatorname{rank}(A),$$

since Q is an invertible matrix so L_Q is an isomorphism and $L_Q(\mathbb{F}^n) = \operatorname{image}(L_Q) = \mathbb{F}^n$. Similar argument can be used for (b). Observe that (c) is an immediate consequence of (a) and (b). ♠

Corollary 3.3.1.

Let $A \in M_{m\times n}(\mathbb{F})$. Then, for all elementary $E \in M_{n\times n}, F \in M_{m\times m}$,

$$\operatorname{rank}(A) = \operatorname{rank}(AE) = \operatorname{rank}(AF).$$

Theorem 3.4.
Invertible Matrix Theorem II

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

Proof. For the forward direction, suppose that A is invertible. Then L_A is an isomorphism, so $\operatorname{rank}(A) = \operatorname{rank}(L_A) = \dim(\mathbb{F}^n) = n$. For the reverse direction, suppose $\operatorname{rank}(A) = n$ then $\operatorname{rank}(A) = n = \dim(\mathbb{F}^n)$ so $L_A : \mathbb{F}^n \to \mathbb{F}^n$ is an isomorpshism, and thus A is invertible. ♠

Proposition 3.5.
Rank of a Matrix Is the Maximal Number of Independent Columns

Let $A \in M_{m\times n}(\mathbb{F})$. Then $\operatorname{rank}(A)$ *is equal to the maximal number of independent columns.*

Proof. Write $A = [A_1\ A_2\ \cdots\ A_n]$ and let $\alpha = \{A_1, A_2, \dots, A_n\}$, the set of columns of A. Moreover, let $\beta = \{e_1, e_2, \dots, e_n\}$ be the standard basis for $\mathbb{F}^n$. Then

$$\begin{aligned}\operatorname{rank}(A) &= \operatorname{rank}(L_A) = \dim(\operatorname{image}(L_A)) = \dim(L_A(\mathbb{F}^n)) = \dim(\operatorname{span}(L_A(\beta)))\\ &= \dim(\operatorname{span}\{Ae_1, Ae_2, \dots, Ae_n\}) = \dim(\operatorname{span}\{A_1, A_2, \dots, A_n\}) = \dim(\operatorname{span}(\alpha)).\end{aligned}$$

But $\dim(\operatorname{span}(\alpha))$ is the maximal number of independent vectors of α, which is the desired result. ♠

Remark 3.3. An important restatement of Proposition 3.5 is that, for any $A \in M_{m\times n}(\mathbb{F})$, $\operatorname{rank}(A)$ is the dimension of the subspace of $\mathbb{F}^m$ that the columns of A span.

Theorem 3.6.
Matrix Elimination

Let $A = M_{m\times n}(\mathbb{F})$. Then by finite number of elementary operations $E_1, E_2, \dots, E_p$ on A, A can be transformed into

$$E_p E_{p-1} \cdots A \cdots E_2 E_1 = \begin{bmatrix} I_r & O_1 \\ O_2 & O_3 \end{bmatrix},$$

where O_1, O_2, O_3 are zero matrices and $0 \le r = \operatorname{rank}(A) \le \min(m, n)$.

Proof. Consider spliting the theorem into two cases. First suppose that $A = 0$ then the proof is done. So suppose that $A \neq 0$. Then there must be a nonzero entry, by type 1 operations which can be moved to $(1,1)$ position. Then, by a type 2 operation, it can be turned into 1, and all $(n,1)$ entries can be made zero by type 3 operations. Observe that we have made a matrix of the form

$$\begin{bmatrix} I_1 & O \\ O & A' \end{bmatrix}.$$

So by proceeding on A' inductively, at most $\min(m,n)$ times, we have the desired result. Moreover, since the number of inductive process is finite, we also observe that the number of elementary operations that are used to obtain a matrix of the desired form is finite. Lastly,

$$\operatorname{rank}(A) = \operatorname{rank}(I_r) = r$$

by Lemma 3.5 and invertible matrix theorem I, so $0 \leq r = \operatorname{rank}(A) \leq \min(m,n)$. ♠

Corollary 3.6.1.

For any $A \in M_{m\times n}(\mathbb{F})$, there exist invertible $P \in M_{m\times m}(\mathbb{F})$ and $Q \in M_{n\times n}(\mathbb{F})$ such that PAQ is of the form

$$\begin{bmatrix} I_r & O \\ O & O \end{bmatrix}.$$

Corollary 3.6.2.

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

(a) $\operatorname{rank}(A) = \operatorname{rank}\left(A^T\right)$.

(b) $\operatorname{rank}(A)$ *is the number of maximal independent rows.*

(c) Columns and rows of A span subspaces of $\mathbb{F}^m$ of equal dimension.

Proof. For (a), let $P \in M_{m\times m}(\mathbb{F})$ and $Q \in M_{n\times n}$ be invertible matrices discussed in Corollary 3.6.1. Then,

$$\operatorname{rank}(A) = \operatorname{rank}(PAQ) = \operatorname{rank}\left((PAQ)^T\right) = \operatorname{rank}\left(Q^T A^T P^T\right) = \operatorname{rank}(A^T).$$

Observe that (b) and (c) are immediate consequences of (a) and Lemma 3.5. ♠

Corollary 3.6.3. An Invertible Matrix Is a Product of Elementary Matrices

Let $A \in M_{n\times n}(\mathbb{F})$ be invertible. Then A is a product of elementary matrices.

Proof. Since A is invertible, by invertible matrix theorem II, $\operatorname{rank} A = n$, and by Theorem 3.6 and Corollary 3.6.1, there exist invertible $P = \prod_{i=1}^{p} E_i, Q = \prod_{j_1}^{q} F_i \in M_{n\times n}(\mathbb{F})$ such that

$$PAQ = I.$$

That is,

$$A = P^{-1}IQ^{-1} = P^{-1}Q^{-1} = E_p^{-1}E_{p-1}^{-1}\cdots E_2^{-1}E_1^{-1}F_q^{-1}F_{q-1}^{-1}\cdots F_2^{-1}F_1^{-1},$$

where each E_i^{-1} and F_i^{-1} are elementary matrices by Proposition 3.2. ♠

Proposition 3.7.

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

$$\operatorname{rank}(AB) = \min(\operatorname{rank}(A), \operatorname{rank}(B)).$$

Proof. First, suppose $\operatorname{rank}(A) \leq \operatorname{rank}(B)$. Observe that

$$\operatorname{rank}(AB) = \operatorname{rank}(L_{AB}) = \operatorname{rank}(L_A L_B) = \dim(L_A(\operatorname{image}(L_B)))$$

where $\operatorname{image}(L_B) \subseteq \mathbb{F}^n$ so $L_A(\operatorname{image}(L_B)) \subseteq L_A(\mathbb{F}^n)$. Thus

$$\dim(L_A(\operatorname{image}(L_B))) \leq \dim(L_A(\mathbb{F}^n)) = \dim(\operatorname{image}(L_A)) = \operatorname{rank}(A).$$

When $\operatorname{rank}(B) \leq \operatorname{rank}(A)$, the same argument can be used by taking the transpose of AB, $B^T A^T$. ♠

Four Fundamental Subspaces of a Matrix

Def'n. Column Space, Row Space, Null Space, Left Null Space of a Matrix
Let $A \in M_{m\times n}(\mathbb{F})$. Define the ***column space*** of A, $\text{image}(A)$, by

$$\text{image}(A) = \text{span}\{x \in \mathbb{F}^m : x \text{ column of } A\} \subseteq \mathbb{F}^m.$$

Similarly, define the ***row space*** of A, $\text{image}\left(A^T\right)$, by

$$\text{image}\left(A^T\right) = \text{span}\{x \in \mathbb{F}^n : x \text{ row of } A\} \subseteq \mathbb{F}^n.$$

Moreover, define the ***null space*** of A, $\ker(A)$, by

$$\ker(A) = \{x \in \mathbb{F}^n : Ax = 0\} \subseteq \mathbb{F}^n.$$

Lastly, define the ***left null space*** of A, $\ker\left(A^T\right)$, by

$$\ker(A^T) = \left\{x \in \mathbb{F}^m : A^T x = 0\right\} \subseteq \mathbb{F}^m.$$

$\text{image}(A), \text{image}\left(A^T\right), \ker(A), \ker\left(A^T\right)$ together are called the ***four fundamental subspaces*** of A.

Proposition 3.8. Properties of Four Fundamental Subspaces

Let $A \in M_{m\times n}(\mathbb{F})$. Then the following holds.

(a) $\text{image}(A)$ *and* $\ker(A^T)$ *are subspaces of* $\mathbb{F}^m$ *and* $\text{image}\left(A^T\right)$ *and* $\ker(A)$ *are subspaces of* $\mathbb{F}^n$.

(b) $\text{rank}(A) = \dim(\text{image}\left(A^T\right)) = \dim(\text{image}(A))$.

(c) $\dim(\text{image}(A)) + \dim(\ker(A^T)) = m$ *and* $\dim(\text{image}\left(A^T\right)) + \dim(\ker(A)) = n$.

(d) $\text{image}(A) \oplus \ker(A^T) = \mathbb{F}^m$ *and* $\text{image}\left(A^T\right) \oplus \ker(A) = \mathbb{F}^n$.

Proof. For (a), use Proposition 1.7 for each set. For (b), observe that

$$\begin{aligned}\text{rank}(A) &= \text{the number of linearly independent columns of } A\\ &= \dim(\text{span}\{\text{Col}_1(A), \dots, \text{Col}_n(A)\}) = \dim(\text{Col}(A)),\end{aligned}$$

and similar proof holds for $\text{rank}(A) = \dim\left(\text{image}\left(A^T\right)\right)$. For (c), observe that

$$\dim(\mathbb{F}^n) = \dim(\text{image}(L_A)) + \dim(\ker(L_A)) = \text{rank}(A) + \dim(\ker(A)) = \dim\left(\text{image}\left(A^T\right)\right) + \dim(\ker(A)),$$

by rank-nullity theorem, and similar proof holds for the remaining part of the statement. For (d), observe that

$$\text{image}\left(A^T\right) \cap \ker(A) = \{0\},$$

so $\text{image}\left(A^T\right) \oplus \ker(A)$ is well defined. Then we have

$$\dim\left(\text{image}\left(A^T\right) \oplus \ker(A)\right) = \dim\left(\text{image}\left(A^T\right)\right) + \dim(\ker(A)) - \dim(\text{image}\left(A^T\right) \cup \ker(A)) = n,$$

where $\text{image}\left(A^T\right) \oplus \ker(A) \subseteq \mathbb{F}^n$, so $\text{image}\left(A^T\right) \oplus \ker(A) = \mathbb{F}^n$. Again, similar proof holds for the remaining part of the statement. ♠

Theorem 3.9.
Invertible Matrix Theorem III

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

(a) A is invertible.

(b) Columns of A form a basis for $\mathbb{F}^n$.

(c) Rows of A form a basis for $\mathbb{F}^n$.

(d) A is a product of elementary matrices.

Proof. Consider showing that (a) is equivalent to other statements. For (a) $\iff$ (b), observe that

$$\begin{aligned} A \text{ is invertible} &\iff \operatorname{rank}(A) = n \\ &\iff \text{maximal number of independent columns is } n \\ &\iff \text{columns of } A \text{ form a basis for } \mathbb{F}^n. \end{aligned}$$

Observe that similar proof holds for (a) $\iff$ (c). (a) $\Longrightarrow$ (d) is provided by Corollary 3.6.3. (d) $\Longrightarrow$ (a) is also clear, since each elementary matrix is invertible. ♠

Def'n. Augmented Matrix
A matrix of the form $[A \mid B]$ for some $A \in M_{m\times n}(\mathbb{F})$ and $B \in M_{m\times p}(\mathbb{F})$ is called an ***augmented matrix***.

Proposition 3.10.
Computing the Inverse Matrix

Let $A \in M_{n\times n}(\mathbb{F})$ be invertible. Then the following hold.

(a) There exists finite number of row operations that transforms $[A \mid I]$ into $[I \mid A^{-1}]$.

(b) If there exists $B \in M_{n\times n}(\mathbb{F})$ such that $[I \mid B]$ is obtained from $[A \mid I]$ by a finite number of row operations, then A is invertible and $B = A^{-1}$.

Proof. For (a), write $C = [A \mid I] = \begin{bmatrix} C_1 & C_2 & \cdots & C_m \end{bmatrix}$ for convenience. Then for any $B \in M_{n\times n}(\mathbb{F})$,

$$BC = \begin{bmatrix} BC_1 & BC_2 & \cdots & BC_m \end{bmatrix},$$

so

$$A^{-1}[A \mid I] = [I \mid A^{-1}],$$

where $A^{-1} = E_1E_2\cdots E_p$ for some elementary matrices $E_1,\dots,E_p \in M_{n\times n}(\mathbb{F})$ by Theorem 3.9. That is,

$$E_1\cdots E_p[A \mid I] = [I \mid A^{-1}]$$

which means there exist corresponding row operations to $E_1,\dots,E_p$ that transform $[A \mid I]$ into $[I \mid A^{-1}]$. For (b), observe that there are elementary matrices $G_1,\dots,G_q$ corresponding to row operations that transform $[A \mid I]$ into $[I \mid B]$. That is,

$$G_1\cdots G_q[A \mid I] = [G_1\cdots G_qA \mid G_1\cdots G_qI] = [I \mid B],$$

so $G_1\cdots G_q = A^{-1}$ and thus $B = G_1\cdots G_q = A^{-1}$, as desired. ♠

Systems of Linear Equations

Def'n. System of Linear Equations
A collection of linear equations of the form

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

where $a_{ij}, b_i \in \mathbb{F}$ for all $i \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$ and $x_1, \ldots, x_n$ are n variables taking values in $\mathbb{F}$, is called a ***system of linear equations*** in n unknowns over $\mathbb{F}$.

Def'n. Coefficient Matrix, Augmented Matrix, Solution, Solution Set of a System
Any system can be written as a matrix product $Ax = b$, where

$$Ax = \begin{bmatrix} a_{11} & a_{12} & \cdots & A_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = b.$$

Moreover, the matrix A and $[A \mid b]$ are called the ***coefficient matrix*** and ***augmented matrix*** of the system, respectively. We say $c \in \mathbb{F}^n$ is a ***solution*** of the system if it satisfies $Ac = b$. The set of all solutions $\{c \in \mathbb{F}^n : Ac = b\}$ is called the ***solution set*** of the system.

Def'n. Consistent, Inconsistent, Homogeneous, Inhomogeneous System
A system $Ax = b$ is called ***consistent*** if its solution set is nonempty and ***inconsistent*** otherwise. Moreover, it is called ***homogeneous*** if $b = 0$ and ***inhomogeneous*** otherwise.

Proposition 3.11. The Solution Set of a Homogeneous System Is a Subspace

Let $A \in M_{m\times n}(\mathbb{F})$ and consider $Ax = 0$. Then the solution set K_H of the system is a subspace of $\mathbb{F}^n$ and

$$\dim(K_H) = n - \operatorname{rank}(A).$$

Proof. Observe that $A0 = 0$ so $0 \in K_H$. For closure under subtraction, let $v, u \in K_H$ then

$$A(v-u) = Av - Au = 0 - 0 = 0$$

so $(v-u) \in K_H$. For closure under multiplication, let $c \in \mathbb{F}$ then

$$A(cv) = c(Av) = c0 = 0$$

so $cv \in K_H$. Moreover, observe that $K_H = \ker(L_A)$, so by the rank-nullity theorem

$$n = \dim(\mathbb{F}^n) = \operatorname{nullity}(L_A) + \operatorname{rank}(L_A) = \dim(\ker(L_A)) + \operatorname{rank}(A) = \dim(K_H) + \operatorname{rank}(A)$$

rearranging which in terms of $\dim(K_H)$ gives

$$\dim(K_H) = n - \operatorname{rank}(A).$$

♠

Corollary 3.11.1. Properties of the Solution Set of a Homogeneous System

Let K_H be the solution set to $Ax = 0$. Then the following hold.

(a) $K_H \neq \emptyset$. In particular, $0 \in K_H$.

(b) $K_H = \{0\} \iff \operatorname{rank}(A) = n$.

(c) If $m < n$, then the system has a nonzero solution.

Def'n. Full Column Rank of a System
We say a system $Ax = b$ is of ***full column rank*** if it satisfies (b) of Corollary 3.11.1.

Proposition 3.12. $K = K_H + \{c\}$ for any Particular Solution c

Suppose $A \in M_{m\times n}(\mathbb{F})$ and $b \in \mathbb{F}^m$. Let K and K_H be the solution sets for $Ax = b$ and $Ax = 0$, respectively. Then for any solution c to $Ax = b$, we have

$$K = \{c\} + K_H = \{c + k : k \in K_H\}.$$

Proof. Let $x \in \mathbb{F}^n$ such that $Ax = b$. Then for any $k \in K_H$, observe that

$$A(x+k) = Ax + Ak = Ax + 0 = Ax = b,$$

so $(x+k) \in K$. Moreover, if $z \in \mathbb{F}^n$ such that $Az = b$, then

$$A(x-z) = b - b = 0$$

so $(x-z) \in K_H$. In other words, there exists $w = (x-z) \in K_H$ such that

$$x = z + w.$$

♠

Theorem 3.13. Invertible Matrix Theorem IV

Let $A \in M_{n\times n}(\mathbb{F})$. Then A is invertible if and only if $Ax = b$ has a unique solution.

Proof. For the forward direction, suppose that A is invertible. From $Ax = b$, we have

$$A^{-1}b = A^{-1}Ax = Ix = x,$$

so $Ax = b$ has a unique solution $x = A^{-1}b$. For the reverse direction, suppose $Ax = b$ has a unique solution $c \in \mathbb{F}^n$. Let K_H and K be the solution sets for $Ax = 0$ and $Ax = b$, respectively. By Proposition 3.12,

$$K = \{c\} = \{c\} + K_H$$

so $K_H = \{0\}$ and $\dim(K_H) = 0$. Since

$$\dim(K_H) = n - \operatorname{rank}(A),$$

it follows that $\operatorname{rank}(A) = n$, which means A is invertible. ♠

Proposition 3.14. A System Is Consistent If and Only If $\operatorname{rank}(A) = \operatorname{rank}[A \mid b]$

Let $Ax = b$ be a system of linear equations. Then the system is consistent if and only if $\operatorname{rank}(A) = \operatorname{rank}[A \mid b]$.

Proof. Observe that

$$\begin{aligned}
&Ax = b \text{ has a solution} \\
\iff\ & b \in \operatorname{image}(L_A) \\
\iff\ & b \in \operatorname{span}\{\operatorname{Col}_1(A), \dots, \operatorname{Col}_n(A)\} \\
\iff\ & \operatorname{span}\{\operatorname{Col}_1(A), \dots, \operatorname{Col}_n(A)\} = \operatorname{span}\{\operatorname{Col}_1(A), \dots, \operatorname{Col}_n(A), b\} \\
\iff\ & \dim(\operatorname{span}\{\operatorname{Col}_1(A), \dots, \operatorname{Col}_n(A)\}) = \dim(\operatorname{span}\{\operatorname{Col}_1(A), \dots, \operatorname{Col}_n(A), b\}) \\
\iff\ & \operatorname{rank}(A) = \operatorname{rank}[A \mid b],
\end{aligned}$$

which is the desired result. ♠

Def'n. Equivalent Systems

Two systems of linear equations are called ***equivalent*** if they have the same solution set.

Proposition 3.15. **$CAx = Cb$ Is Equivalent to $Ax = b$ Whenever C Is Invertible**

Let $Ax = b$ be a system of m linear equations in n unknowns and let $C \in M_{m\times n}(\mathbb{F})$ be invertible. Then the system $(CA)x = Cb$ is equivalent to $Ax = b$.

Proof. Let K_H be the solution set of $Ax = b$ and suppose that $x \in K_H$. Then

$$(CA)x = C(Ax) = Cb,$$

so x is in the solution set of $CAx = Cb$ as well. Moreover, suppose that x is in the solution set of $(CA)x = Cb$, then

$$CAx = Cb \iff C^{-1}CAx = C^{-1}Cb \iff Ax = b.$$

Thus $x \in K_H$, which is the desired result. ♠

Corollary 3.15.1.

Let $Ax = b$ be a system of m linear equations in n unknowns. If $[A' \mid b']$ is obtained from $[A \mid b]$ by a finite number of elementary row operations, then the system $A'x = b'$ is equivalent to $Ax = b$.

Def'n. Reduced Row Echelon Form (RREF), Leading One

A matrix is said to be in ***reduced row echelon form*** if the following three conditions are satisfied.

(a) Any row containing nonzero entry precedes any row in which all the entries are zero, if any.

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

(c) The first nonzero entry in each row is 1, called the ***leading one*** of the row, and it occurs a column to the right of the first nonzero entry in the preceding row.

Remark 3.4. Given $A \in M_{m\times n}(\mathbb{F})$, we may obtain an RREF from A as follows.

(a) In the leftmost nonzero column, use elementary row operations to get 1 in the first row.

(b) By means of type 3 elementary row operations, use the first row to obtain zeroes in the remaining entries of the leftmost nonzero column.

(c) Consider the submatrix consisting of the columns to the right of the column we just modified and the rows beneath the row that just got a leading one. Use elementary row operations - if necessary - to get a leading one in the top of the first nonzero column of this submatrix.

(d) Use elementary row operations to obtain zeroes below the one created in the preceding step.

(e) Repeat (c) and (d) until no nonzero rows remain.

(f) Work upward, beginning with the last nonzero row and add multiples of each rows above to create zeroes above the first nonzero in each row.

(g) Repeat the precess in (f) for each preceding row until it is performed with the second row, at which time the reduction process is complete.

Def'n. Gaussian Elimination
We call the procedure described in Remark 3.4 the ***Gaussian elimination***.

Proposition 3.16. Gaussian Elimination

Gaussian elimination transforms any matrix to its RREF, and the RREF of a matrix is unique.

Def'n. Free Variable
Let R be the RREF of a coefficient matrix of a system of linear equations $Ax = b$. If the jth column of R does not contain a leading one, then x_j is called a ***free variable***.

Proposition 3.17. Number of Free Variables Is Equal to $n - \operatorname{rank}(A)$

Let $A \in M_{m\times n}(\mathbb{F}), b \in \mathbb{F}$, and B be the RREF of A. Then the following holds.

$$\textit{number of free variables} = n - \textit{number of leading ones} = n - \operatorname{rank}(A).$$

Proof. Observe that

$$\operatorname{rank}(A) = \operatorname{rank}(B) = \text{number of reading ones of } B = \text{number of nonzero rows of } B.$$ ♠

Remark 3.5. Let $A \in M_{m\times n}(\mathbb{F})$ and $b \in \mathbb{F}$. Consider solving the system $Ax = b$ by the following algorithm.

(a) Write the augmented matrix $[A \mid b]$ for the system.

(b) Use elementary row operations (e.g. Gaussian elimination) to transform the augmented matrix into its RREF $[A' \mid b']$.

(c) Write the system of linear equations to the RREF.

(d) If the system contain an equation of the form $0 = 1$, then the system is inconsistent.

(e) Otherwise, assign values $t_1, \ldots, t_{n-r}$ to the free variables and then solve the remaining variables in terms of the free variables. Here $r = \operatorname{rank}(A') = \operatorname{rank}(A)$ is the number of nonzero rows of A'.

(f) Then an arbitrary solution to $Ax = b$ is of the form

$$x = x_0 + \sum_{i=1}^{n-r} t_i u_i$$

where r is the number of nonzero rows in A'.

Then K is given by

$$K = \left\{ x \in \mathbb{F}^n : x = x_0 + \sum_{i=1}^{n-r} t_i u_i \right\}.$$

Def'n. Parametric Value Assigned to Free Variables
The values $t_1, \dots, t_{n-r}$ assigned to free variables in (e) of Remark 3.4 are called ***parametric values***.

Proposition 3.18. *Let $[A \mid b]$ be a consistent system of m linear equations in n variables. Suppose that the RREF of $[A \mid b]$ has r nonzero rows. If the general solution to $Ax = b$ obtained by the procedure described in Remark 3.4 is of the form*

$$x = x_0 + \sum_{i=1}^{n-r} t_i u_i,$$

then $x_0 \in \mathbb{F}^n$ is a solution to $Ax = b$ and $\{u_1, \dots, u_{n-r}\}$ is a basis for the solution set of the corresponding homogeneous system $Ax = 0$.

Proof. To verify that x_0 is a solution to $Ax = b$, observe that, by setting $t_1, \dots, t_{n-r} = 0$,

$$x = x_0 + \sum_{i=1}^{n-r} 0u_i = x_0.$$

So, we may verify that $\{u_1, \dots, u_{n-r}\}$ generates K_H the following. First, write

$$K = \{x_0\} + K_H,$$

where K_H and K are solution sets of $Ax = 0$ and $Ax = b$, respectively. Observe that

$$K = \left\{x : x + x_0 + \sum_{i=1}^{n-r} t_i u_i\right\} = \{x_0\} + \left\{x : x = \sum_{i=1}^{n-r} t_i u_i\right\},$$

so it must be the case which $K_H = \left\{x : x = \sum_{i=1}^{n-r} t_i u_i\right\}$. But this means $K_H = \operatorname{span}\{u_1, \dots, u_{n-r}\}$. Lastly, to verify that $\{u_1, \dots, u_{n-r}\}$ is linearly independent, observe that

$$\dim(K_H) = n - \operatorname{rank}(A) = n - r$$

by Proposition 3.11. ♠

This page intentionally left blank.

4.
Determinants

Determinants

Def'n. Determinant, Cofactor, Cofactor Expansion

Let $A \in M_{n\times n}(\mathbb{F})$. We define the ***determinant*** of A, denoted by $\det A$ or $|A|$, as follows.

(a) If $n = 1$, then

$$\det(A) = A_{11}.$$

(b) For each $n \in \mathbb{N} \setminus \{1\}$, $\det A$ is defined recursively as

$$\det(A) = \sum_{i=1}^{n} (-1)^{1+i} A_{i1} \left(\det \widetilde{A}_{i1}\right),$$

where $\widetilde{A}_{ij}$ is the $(n-1)\times(n-1)$ matrix obtained by removing ith row and jth column.

The scalar $(-1)^{i+j}\det\left(\widetilde{A}_{ij}\right)$ is called the ***cofactor*** of the entry of A in row i and column j. The above determinant equation is also known as the ***cofactor expansion along the first column*** of A.

Remark 4.1. The expression

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

is called the cofactor expansion along the jth column of A.

Remark 4.2. For $A \in M_{2\times 2}(\mathbb{F})$, we have

$$\det(A) = A_{11}A_{22} - A_{12}A_{21}.$$

Remark 4.3. For simplicity, we are going to write

$$A = \left[A_1, A_2, \ldots, A_n\right]$$

to denote $A \in M_{n\times n}(\mathbb{F})$ with rows $A_1, A_2, \ldots, A_n$.

Proposition 4.1. n-Linearity of Determinants

Let $A \in M_{n\times n}(\mathbb{F})$. Then the determinant of A is a linear function of each row, when the remaining rows are held fixed. That is, for each $k \in \{1, \ldots, n\}$, we have

$$\det\left[A_1, \ldots, B_k + \alpha C_k, \ldots, A_n\right] = \det\left[A_1, \ldots, B_k, \ldots, A_n\right] \alpha \det\left[A_1, \ldots, C_k, \ldots, A_n\right].$$

Proof. Write

$$A = \left[A_1, \ldots, B_k + \alpha C_k, \ldots, A_n\right] \quad B = \left[A_1, \ldots, B_k, \ldots, A_n\right] \quad C = \left[A_1, \ldots, C_k, \ldots, A_n\right].$$

Then by definition,

$$\begin{aligned}\det(A) &= \sum_{i=1}^{n} (-1)^{1+i} A_{i1} \det\left(\widetilde{A}_{i1}\right) \\ &= \sum_{i=1, i\neq k}^{n} (-1)^{1+i} A_{i1} \det\left(\widetilde{A}_{i1}\right) + (-1)^{1+k} A_{k1} \det\left(\widetilde{A}_{k1}\right).\end{aligned}$$

Observe that

$$\widetilde{A}_{k1} = \widetilde{B}_{k1} = \widetilde{C}_{k1} \text{ and } A_{k1} = B_{k1} + \alpha C_{k1}.$$

For $i \neq k$, the matrices $\widetilde{A}_{i1}, \widetilde{B}_{i1}, \widetilde{C}_{i1}$ have the same rows, except for one row $i_0 = k-1$ if $i < k$ and $i_0 = k$ if $j > k$. Moreover, row i_0 of $\widetilde{A}_{i1}, \widetilde{B}_{i1}, \widetilde{C}_{i1}$ is $(B_k + \alpha C_k), B_k, C_k$, respectively. So by induction hypothesis,

$$\det\left(\widetilde{A}_{i1}\right) = \det\left(\widetilde{B}_{i1}\right) + \alpha \det\left(\widetilde{C}_{i1}\right)$$

and

$$A_{i1} = B_{i1} = C_{i1}$$

for each $i \neq k$. Plugging these equalities into the cofactor expansion along the first column of A, we get

$$\begin{aligned}
\det(A) &= \sum\nolimits_{i=1}^{n} (-1)^{1+i} A_{i1} \cdot \det\left(\widetilde{A}_{i1}\right) \\
&= \sum\nolimits_{i=1, i\neq k}^{n} (-1)^{1+i} A_{i1} \cdot \det\left(\widetilde{A}_{i1}\right) + (-1)^{1+k} A_{k1} \cdot \det\left(\widetilde{A}_{k1}\right) \\
&= \sum\nolimits_{i=1, i\neq k}^{n} (-1)^{1+i} A_{i1} \cdot \left(\det\left(\widetilde{B}_{i1}\right) + \alpha \det\left(\widetilde{C}_{i1}\right)\right) \\
&\quad + (-1)^{1+k} (B_{k1} + \alpha C_{k1}) \cdot \det\left(\widetilde{A}_{k1}\right) \\
&= \sum\nolimits_{i=1}^{n} (-1)^{1+i} B_{i1} \cdot \det\left(\widetilde{B}_{i1}\right) + \alpha \sum\nolimits_{i=1}^{n} (-1)^{1+i} C_{i1} \cdot \det\left(\widetilde{C}_{i1}\right) \\
&= \det(B) + \alpha \det(C),
\end{aligned}$$

which is the desired result. ♠

Corollary 4.1.1. **Determinant Is Zero A If Has a Zero Row**

Let $A \in M_{n\times n}(\mathbb{F})$. If A has a zero row, then $\det(A) = 0$.

Proof. Write

$$A = \left[R_1, \ldots, R_k, \ldots, R_n\right]$$

where $R_k = 0$. Then,

$$\begin{aligned}
\det A &= \det\left[R_1, \ldots, R_k, \ldots, R_n\right] = \det\left[R_1, \ldots, 2R_k, \ldots, R_n\right] \\
&= 2\det\left[R_1, \ldots, R_k, \ldots, R_n\right] = 2\det(A)
\end{aligned}$$

so $\det(A) = 0$. ♠

Corollary 4.1.2. **Determinants after a Type 2 Elementary Row Operation**

Let $A \in M_{n\times n}(\mathbb{F})$ and $Q \in M_{n\times n}(\mathbb{F})$ be the matrix obtained from A by multiplying a row of A by a scalar $\alpha \in \mathbb{F}$. Then $\det(Q) = \alpha \det(A)$.

Proof. Write

$$A = \left[R_1, \ldots, R_n\right]$$

and suppose that Q is obtained by multiplying row k of A, R_k, by α. Then

$$\det(Q) = \det\left[R_1, \ldots, \alpha R_k, \ldots, R_n\right] = \alpha \det\left[R_1, \ldots, R_k, \ldots, R_n\right] = \alpha \det A,$$

which is the desired result. ♠

Lemma 4.2. $\det(A) = 0$ **If Two Adjacent Rows Are Equal**

Let $A \in M_{n\times n}(\mathbb{F})$. If two adjacent rows of A are equal, then $\det A = 0$.

Proof. We proceed by induction. Let $P(n)$ be the predicate that every $A \in M_{n\times n}(\mathbb{F})$ that has two equal adjacent rows satisfies $\det A = 0$. For $n = 2$, if $A \in M_{n\times n}(\mathbb{F})$ has two equal adjacent rows, then we may write

$$\begin{bmatrix} A_{11} & A_{11} \\ A_{21} & A_{21} \end{bmatrix}$$

so

$$\det(A) = A_{11}A_{21} - A_{11}A_{21} = 0.$$

Now, suppose $P(k)$ for some $k \in \{n \in \mathbb{N} : n > 2\}$. Further suppose that $A \in M_{(k+1)\times(k+1)}(\mathbb{F})$ such that A has two equal adjacent rows, row p and row $p+1$. Then, for each $i \in \{1, \ldots, k+1\} \setminus \{p, p+1\}$, $\widetilde{A}_{i1}$ has two equal adjacent rows, so $\det\left(\widetilde{A}_{i1}\right) = 0$ by induction hypothesis. Moreover, $\widetilde{A}_{p1} = \widetilde{A}_{(p+1)1}$, since $\mathrm{Row}_p(A) = \mathrm{Row}_{p+1}(A)$ and they are adjacent. Thus,

$$\begin{aligned}
\det(A) &= \sum_{i=1}^{k+1} (-1)^{1+i} A_{i1} \det\left(\widetilde{A}_{i1}\right) \\
&= \sum_{i=1, i\neq p, p+1}^{k+1} (-1)^{1+i} A_{i1} \det\left(\widetilde{A}_{i1}\right) + (-1)^{1+p} A_{p1} \det\left(\widetilde{A}_{p1}\right) \\
&\qquad + (-1)^{p+2} A_{(p+1)1} \det\left(\widetilde{A}_{(p+1)1}\right) \\
&= (-1)^{1+p} A_{p1} \det\left(\widetilde{A}_{p1}\right) + (-1)^{p+2} A_{p1} \det\left(\widetilde{A}_{p1}\right) \\
&= \left((-1)^{1+p} + (-1)^{2+p}\right) A_{p1} \det\left(\widetilde{A}_{p1}\right) = 0,
\end{aligned}$$

since $(-1)^{1+p} + (-1)^{2+p} = 0$ for any $p \in \mathbb{Z}$. ♠

Lemma 4.3. $\det(B) = -\det(A)$ **If B Is Obtained by Exchanging Two Adjacent Rows of A**

Let $A \in M_{n\times n}(\mathbb{F})$ and $B \in M_{n\times n}(\mathbb{F})$ be the matrix obtained by exchanging row i and row $i+1$ of A for some $i \in \{1, \ldots, n-1\}$. Then $\det(B) = -\det(A)$.

Proof. Let $R_1, \ldots, R_n$ be the rows of A. Define

$$C = \left[R_1, \cdots, R_i + R_{i+1}, R_i + R_{i+1}, \ldots, R_n\right]$$

then $\det(C) = 0$ by Lemma 4.2. That is,

$$\begin{aligned}
\det(C) &= \det\left[R_1, \ldots, R_i + R_{i+1}, R_i + R_{i+1}, \ldots, R_n\right] \\
&= \det\left[R_1, \ldots, R_i, R_i, \ldots, R_n\right] + \det\left[R_1, \ldots, R_{i+1}, R_{i+1}, \ldots, R_n\right] \\
&\qquad + \det\left[R_1, \ldots, R_i, R_{i+1}, \ldots, R_n\right] + \det\left[R_1, \ldots, R_{i+1}, R_i, \ldots, R_n\right] \\
&= \det\left[R_1, \ldots, R_{i+1}, R_{i+1}, \ldots, R_n\right] + \det\left[R_1, \ldots, R_i, R_{i+1}, \ldots, R_n\right] \\
&= \det(A) + \det(B) = 0,
\end{aligned}$$

which exactly means $\det(B) = -\det(A)$. ♠

Lemma 4.4. Determinant Is Zero If A Has Two Identical Rows

Let $A \in M_{n\times n}(\mathbb{F})$. If A has two identical rows, then $\det(A) = 0$.

Proof. Suppose that A has two identical rows. Then by means of type 1 elementary row operations, A can be transformed into a matrix which has two adjacent identical rows, denote which A'. Suppose n type 1 elementary row operations are used. Then by Lemma 4.2 and Lemma 4.3,

$$\det(A) = (-1)^n \det(A') = 0.$$ ♠

Proposition 4.5.
Determinants after a Type 1 Elementary Row Operation

Let $A \in M_{n\times n}(\mathbb{F})$ and suppose that $B \in M_{n\times n}(\mathbb{F})$ is obtained by exchanging row i and row j of A. Then $\det(B) = -\det(A)$.

Proof. Without loss of generality, suppose $i < k$ and let $R_1, \ldots, R_n$ be the rows of A. Define

$$C = \left[R_1, \ldots, R_i + R_j, \ldots, R_i + R_j, \ldots, R_n\right]$$

then

$$\begin{aligned}
\det(C) &= \det\left[R_1, \ldots, R_i + R_j, \ldots, R_i + R_j, \ldots, R_n\right] \\
&= \det\left[R_1, \ldots, R_i, \ldots, R_i, \ldots R_n\right] + \det\left[R_1, \ldots, R_j, \ldots, R_j, \ldots, R_n\right] \\
&\qquad + \det\left[R_1, \ldots, R_i, \ldots, R_j, \ldots, R_n\right] + \det\left[R_1, \ldots, R_j, \ldots, R_i, \ldots, R_n\right] \\
&= \det\left[R_1, \ldots, R_i, \ldots, R_j, \ldots, R_n\right] + \det\left[R_1, \ldots, R_j, \ldots, R_i, \ldots, R_n\right] \\
&= \det(A) + \det(B) = 0.
\end{aligned}$$

Thus $\det(B) = -\det(A)$. ♠

Proposition 4.6.
Determinants after a Type 3 Elementary Row Operation

Let $A \in M_{n\times n}(\mathbb{F})$. Suppose that B is obtained from A by adding scalar multiple of row j to row i of A. Then $\det(B) = \det(A)$.

Proof. Without loss of generality, suppose $i < j$ and let $R_1, \ldots, R_n$ be the rows of A. Then

$$B = \left[R_1, \ldots, R_i + cR_j, \ldots, R_j, \ldots, R_n\right]$$

for some $c \in \mathbb{F}$. Thus

$$\begin{aligned}
\det(B) &= \det\left[R_1, \ldots, R_i + cR_j, \ldots, R_j, \ldots, R_n\right] \\
&= \det\left[R_1, \ldots, R_i, \ldots, R_j, \ldots, R_n\right] + c\det\left[R_1, \ldots, R_j, \ldots, R_j, \ldots, R_n\right] \\
&= \det\left[R_1, \ldots, R_i, \ldots, R_j, \ldots, R_n\right] = \det(A),
\end{aligned}$$

which is the desired result. ♠

Properties of Determinants

Theorem 4.7.
Characterization of Determinants

As a function of each row, the determinant of a square matrix is a unique function $\mathbb{F}^n \times \mathbb{F}^n \times \cdots \times \mathbb{F}^n \to \mathbb{F}$ that satisfies the following.

(a) *Determinant is a n-linear function. In other words, if $A = \left[A_1, A_2, \ldots, A_n\right]$*

$$\det(A_1, \ldots, cA_i, \ldots, A_n) = c\det(A_1, \ldots, A_i, \ldots, A_n) = c\det(A)$$

for each $i \in \{1, 2, \ldots, n\}$.

(b) *Whenever there exists $i \in \{1, \ldots, n-1\}$ such that $A_i = A_{i+1}$,* $\det(A) = 0$.

(c) $\det(I) = 1$.

Corollary 4.7.1. Determinant of an Elementary Matrix and Its Transpose

Let E_i be an elementary matrix obtained by type i elementary row operation. Then the following holds.

(a) $\det(E_1) = -1$.

(b) $\det(E_2) = c$*, where* $c \in \mathbb{R} \setminus \{0\}$ *is the coefficient multiplied to a row.*

(c) $\det(E_3) = 1$.

(d) $\det\left(E_i^T\right) = \det(E_i)$.

Proof. Observe that (1), (2), and (3) are direct results of Proposition 5, Corollary 1.2, Proposition 6, respectively, with (3) of Proposition 7. For (4), it is sufficient to recall that E^T is an elementary matrix of the same type as E, provided that E is an elementary matrix. ♠

Corollary 4.7.2.

Let $E \in M_{n\times n}(\mathbb{F})$ *be an elementary matrix. Then* $\det(E) \neq 0$.

Corollary 4.7.3.

Let $A \in M_{n\times n}(\mathbb{F})$ *and* $E \in M_{n\times n}(\mathbb{F})$ *be an elementary matrix. Then* $\det(EA) = \det(E)\det(A)$.

Corollary 4.7.4.

Let $A \in M_{n\times n}(\mathbb{F})$ *and* $E_1, \ldots, E_p \in M_{n\times n}(\mathbb{F})$ *be elementary matrices. Then*

$$\det(E_1E_2\cdots E_pA) = \det(E_1)\det(E_2)\cdots\det(E_p)\det(A).$$

In particular,

$$\det(E_1E_2\cdots E_p) = \det(E_1)\det(E_2)\cdots\det(E_p).$$

Theorem 4.8. Invertible Matrix Theorem V

Let $A \in M_{n\times n}(\mathbb{F})$*. Then A is invertile if and only if* $\det(A) \neq 0$.

Proof. For the forward direction, suppose that A is invertible. Then there exists elementary matrices $E_1, \ldots, E_p \in M_{n\times n}(\mathbb{F})$ such that $A = E_1E_2\cdots E_p$, so $\det(A) = \det(E_1)\det(E_2)\cdots\det(E_p) \neq 0$. For the reverse direction, suppose that $\det(A) \neq 0$. Further suppose that A is not invertible for the sake of contradiction. Then there exists elementary matrices $F_1, F_2, \ldots, F_q \in M_{n\times n}(\mathbb{F})$ such that

$$F_1F_2\cdots F_qA = \begin{bmatrix} I_r & O \\ O & O \end{bmatrix},$$

where $r = \operatorname{rank}(A) < n$. But $\det\begin{bmatrix} I_r & O \\ O & O \end{bmatrix} = 0$, which is a contradiction, since $\det(A) \neq 0$ by assumption and $F_1, F_2, \ldots, F_q$ are elementary matrices. ♠

Corollary 4.8.1. Determinant and Rank

Let $A \in M_{n\times n}(\mathbb{F})$*. Then* $\det(A) = 0$ *if and only if* $\operatorname{rank}(A) < n$.

Proposition 4.9. Determinant of a Matrix Product

Let $A, B \in M_{n\times n}(\mathbb{F})$*. Then* $\det(AB) = \det(A)\det(B)$.

Proof. First suppose that $\det(A) = 0$ or $\det(B) = 0$. Then clearly $\det(AB) = 0$, since AB is not invertible. So suppose that $\det(A), \det(B) \neq 0$. Then A and B are invertible, so there exist elementary $E_1, \ldots, E_p, F_1, \ldots, F_q \in M_{n\times n}(\mathbb{F})$ such that $A = E_1E_2\cdots E_p$ and $B = F_1F_2\cdots F_q$. Therefore,

$$\begin{aligned}\det(AB) &= \det\left(E_1E_2\cdots E_pF_1F_2\cdots F_q\right) \\ &= \det\left(E_1 \det E_2 \cdots \det E_p\right)\left(\det F_1 \det F_2 \cdots \det F_q\right) = \det(A)\det(B),\end{aligned}$$

which is the desired result. ♠

Proposition 4.10.
Determinant of a Matrix and Its Transpose

Let $A \in M_{n\times n}(\mathbb{F})$. *Then* $\det(A) = \det\left(A^T\right)$.

Proof. Suppose that A is not invertible. Then so A^T is not, and we have $\det A = 0 = \det A^T$. So suppose that A is invertible. Then there are elementary matrices $E_1, \ldots, E_p \in M_{n\times n}(\mathbb{F})$ such that

$$\begin{aligned}\det(A^T) &= \det\left(E_1E_2\cdots E_p\right)^T = \det\left(E_p^TE_{p-1}^T\cdots E_1^T\right) \\ &= \det\left(E_p^T\right)\det\left(E_{p-1}^T\right)\cdots\det\left(E_1^T\right) \\ &= \det(E_p)\det(E_{p-1})\cdots\det(E_1) \\ &= \det(E_1)\det(E_2)\cdots\det(E_p) = \det\left(E_1E_2\cdots E_p\right) = \det(A),\end{aligned}$$

as desired. ♠

Theorem 4.11.
Cofactor Expansion Theorem

The determinant of A *can be evaluated by cofactor expansion along any column. That is, for any* $j \in \{1, \ldots, n\}$,

$$\det(A) = \sum_{i=1}^{n}(-1)^{i+j}A_{ij}\det\left(\widetilde{A}_{ij}\right).$$

Proof. Write $A = \left[C_1, C_2, \ldots, C_j, \ldots, C_n\right]$. Observe that it takes $(j-1)$ type 1 elementary column operations to transform A into $A' = \left[C_j, C_1, C_2, \ldots, C_n\right]$. Therefore, $\det(A) = (-1)^{j-1}\det(A')$ and by using cofactor expansion,

$$(-1)^{j-1}\sum_{i=1}^{n}(-1)^{1+i}A'_{i1}\det\left(\widetilde{A}'_{i1}\right) = \sum_{i=1}^{n}(-1)^{i+j}A'_{i1}\det\left(\widetilde{A}'_{i1}\right) = \sum_{i=1}^{n}(-1)^{i+j}A_{ij}\det\left(\widetilde{A}_{ij}\right),$$

since $A'_{i1} = A_{ij}$ and $\widetilde{A}'_{i1} = \widetilde{A}_{ij}$ for all $i \in \{1, \ldots, n\}$ by construction. A more general result involving cofactor expansion along any row can be shown by taking the transpose of A. ♠

This page intentionally left blank.

5. Diagonalization

Eigenvectors and Eigenvalues

Def'n. Eigenvector, Eigenvalue, Eigenspace of a Matrix
Let $A \in M_{n\times n}(\mathbb{F})$. A nonzero $v \in \mathbb{F}^n$ is callend an ***eigenvector*** of A if there exists $c \in \mathbb{F}$ such that

$$Av = cv.$$

Such c is called the ***eigenvalue*** corresponding to v. For any eigenvalue $c \in \mathbb{F}^n$ of A, the set

$$E_c = \{u \in \mathbb{F}^n : Au = cu\} \cup \{0\}$$

is a subspace of $\mathbb{F}^n$ and called the ***eigenspace*** corresponding to c.

Remark 5.1. Let $A \in M_{n\times n}(\mathbb{F})$ and suppose $c \in \mathbb{F}$ is an eigenvalue of A. Then the eigenspace corresponding to c is the null space of $cI - A$, $\ker(cI - A)$. For, any eigenvector $v \in \mathbb{F}^n$ of A corresponding to c satisfies $Av = cv$, which exactly means $(cI - A)v = 0$.

Proposition 5.1.
c Is an Eigenvalue If and Only If $\det(cI - A) = 0$

Let $A \in M_{n\times n}(\mathbb{F})$. Then $c \in \mathbb{F}$ is an eigenvalue of A if and only if $\det(cI - A) = 0$.

Proof. The forward direction is follows easily from Remark 5.1. For the reverse direction, if $\det(cI - A) = 0$, then $cI - A$ is not invertible, so there exists some $v \in \mathbb{F}^n$ such that

$$(cI - A)v = 0.$$

But this exactly means $Av = cv$, so c is an eigenvalue of A. ♠

Remark 5.2. Proposition 5.1 motivates the following.

Def'n. Characteristic Polynomial of a Matrix
Let $A \in M_{n\times n}(\mathbb{F})$. The nth degree polynomial $\det(xI - A)$ in the indeterminate x is called the ***characteristic polynomial*** of A.

Proposition 5.2.
Properties of Characteristic Polynomials

Let $A \in M_{n\times n}(\mathbb{F})$ and let $f \in \mathbb{F}[x]$ be the characteristic polynomial of A.

(a) f is a monic polynomial of $\deg(f) = n$.

(b) A has at most n eigenvalues.

(c) If $B \in M_{n\times n}(\mathbb{F})$ is similar to A, then the characteristic polynomial of B is f.

Proof. For (a), use thee cofactor expansion along any row or column. (b) is a direct consequence of a property of polynomials, that a polynomial of degree n has at most n roots. For (c), since $B = QAQ^{-1}$ for some invertible $A \in M_{n\times n}(\mathbb{F})$,

$$\begin{aligned}\det(cI - B) &= \det\left(cI - QAQ^{-1}\right) = \det\left(QcIQ^{-1} - QAQ^{-1}\right) = \det\left(Q(cI - A)Q^{-1}\right) \\ &= \det(Q)\det(cI - A)\det\left(Q^{-1}\right) = \det(cI - A),\end{aligned}$$

as desired. ♠

Remark 5.3. Let V be a finite-dimensional vector space. Observe that (c) of Proposition 5.2 enables us to define an eigenvalue and eigenvector of a linear operator $T : V \to V$ as an eigenvalue and eigenvector of its matrix representation $[T]_\beta$ for any ordered basis β for V.

Def'n. Eigenvalue, Eigenvector of a Linear Operator
Let $T : V \to V$ be a linear operator on V. A scalar $c \in \mathbb{F}$ is called an ***eigenvalue*** of the linear operator T if there exists a nonzero $v \in V$ such that $Tv = cv$. Such vector v is called an ***eigenvector*** of T corresponding to c.

Def'n. Characteristic Polynomial of a Linear Operator
Let $T : V \to V$ be a linear operator on V. We define the ***characteristic polynomial*** of T to be the characteristic polynomial of any matrix representation $[T]_\beta$ of T, where β is an ordered basis for V.

Remark 5.4. We see that the uniqueness of characteristic polynomial of a linear operator $T : V \to V$ on a finite-dimensional vector space V is guaranteed by (c) of Proposition 5.2. For, if β and γ are ordered bases for V, then $[T]_\beta$ and $[T]_\gamma$ are similar.

Proposition 5.3. *Let $T : V \to V$ be a linear operator on V.*

(a) A scalar $c \in \mathbb{F}$ is an eigenvalue of T if and only if $cI - T$ is not invertible.

(b) Let c be an eigenvalue of T. A vector $v \in V$ is an eigenvector of T corresponding to c if and only if $v \neq 0$ and $v \in \ker(cI - T)$.

Proof. (a) follows easily from Proposition 5.1. For (b), if v is an eigenvector corresponding to c, then $Tv = cv$ or $(cI - T)v = 0$. Conversely, for any nonzero $v \in \ker(cI - T)$, we have $Tv = cv$. ♠

Diagonalization

Def'n. Diagonalizable Linear Operator
Let V be a finite-dimensional vector space. A linear opeartor $T : V \to V$ is called ***diagonalizable*** if there is an ordered basis β for V such that $[T]_\beta$ is an diagonal matrix. Moreover $A \in M_{n\times n}(\mathbb{F})$ is called ***diagonalizable*** if its left multiplication operator $L_A : \mathbb{F}^n \to \mathbb{F}^b$ is diagonalizable.

Remark 5.5. Another way to state the diagonalizability of a (square) matrix is that $A \in M_{n\times n}(\mathbb{F})$ is diagonalizable if and only if A is similar to a diagonal B over $\mathbb{F}$.

Proposition 5.4. **$T : V \to V$ Is Diagonalizable If and Only If a Eigenbasis for V Exists**

Let $T : V \to V$ be a linear operator on an n-dimensional vector space V. Then T is diagonalizable if and only if there exists an ordered basis β for V consisting of eigenvectors of T.

Proof. For the forward direction, suppose that T is diagonalizable. Then there exists ordered basis $\beta = \{v_1, v_2, \ldots, v_n\}$ such that $[T]_\beta$ is diagonal. So $\left([T]_\beta\right)_{ii} = c_i$ for some $c_i \in \mathbb{F}$, and $\left([T]_\beta\right)_{ij} = 0$ whenever $i \neq j$. That is,

$$Tv_i = c_i v_i$$

so v_i is an eigenvector of T for each $i \in \{1, 2, \ldots, n\}$. For the reverse direction, suppose that $\beta = \{v_1, v_2, \ldots, v_n\}$ is consisting of eigenvectors. Then $Tv_i = c_i v_i$ for some $c_i \in \mathbb{F}$, which exactly means

$$[T]_\beta = \begin{bmatrix} c_1 & 0 & \cdots & 0 \\ 0 & c_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & c_n \end{bmatrix}.$$ ♠

Corollary 5.4.1.

Let $A \in M_{n\times n}(\mathbb{F})$. Then A is diagonalizable if and only if there is an ordered basis for $\mathbb{F}^n$ consisting of eigenvectors of A.

Proposition 5.5. Eigenvectors Corresponding to Distinct Eigenvalues Are Linearly Independent

Let $T : V \to V$ be a linear opeartor where $\dim(V) = n$. Let $c_1, c_2, \ldots, c_m$ be distinct eigenvalues of T. If $v_1, v_2, \ldots, v_m$ are eigenvectors of T corresponding to $c_1, c_2, \ldots, c_m$, then $\{v_1, v_2, \ldots, v_m\}$ is linearly independent.

Proof. We proceed by induction on m. When $m = 1$, observe that $\{v_1\}$ is linearly independent, since any eigenvector is nonzero. Moreover, suppose that $\{v_1, v_2, \ldots, v_k\}$ is linearly independent for some $k \le m-1$. For the sake of contradiction, further suppose that $\{v_1, v_2, \ldots, v_{k+1}\}$ is linearly dependent. Then there exists $(a_1, \ldots, a_k) \in \mathbb{F}^k$ such that

$$v_{k+1} = \sum_{i=1}^{k} a_i v_i.$$

Then

$$c_{k+1} v_{k+1} = T v_{k+1} = T \sum_{i=1}^{k} a_i v_i = \sum_{i=1}^{k} a_i T v_i = \sum_{i=1}^{k} a_i c_i v_i,$$

so

$$v_{k+1} = \sum_{i=1}^{k} a_i \frac{c_i}{c_{k+1}} v_i,$$

where each $\frac{c_i}{c_{k+1}} \neq 1$ since each eigenvalue is distinct, so we have two different representation of v_{k+1} as a linear combination of $v_1, v_2, \ldots, v_k$, which contradicts the linear independence of $v_1, v_2, \ldots, v_k$. Thus $\{v_1, v_2, \ldots, v_m\}$ is linearly independent. ♠

Corollary 5.5.1. Diagonalizable If $n = \dim(V)$ Distinct Eigenvalues

Let $T : V \to V$ be a linear operator where $\dim(V) = n$. If T has n distinct eigenvalues, then T is diagonalizable.

Proof. Let $c_1, c_2, \ldots, c_n$ be distinct eigenvalues of T and $v_1, v_2, \ldots, v_n$ be the corresponding eigenvectors. Then $\{v_1, v_2, \ldots, v_n\}$ is linearly independent by Proposition 5.7, so is a basis for V, which exactly means T is diagonalizable by Corollary 5.5.1. ♠

Proposition 5.6.

Let V be a finite-dimensional vector space and let $T : V \to V$ be diagonalizable. Then the characteristic polynomial of T can be written as a product of linear polynomials in $\mathbb{F}$.

Proof. Observe that there exists and ordered basis β for V such that $[T]_\beta$ is diagonal with $\left([T]_\beta\right)_{ii} = c_i$ for some $c_1, c_2, \ldots, c_n \in \mathbb{F}$. That is,

$$\prod_{i=1}^{n} (x - c_i)$$

is the characteristic polynomial of T. ♠

Def'n. Algebraic Multiplicity, Geometric Multiplicity of an Eigenvalue

Let c be an eigenvalue of a linear operator $T : V \to V$ on an n-dimensional vector space V and let $p(t)$ be the characteristic polynomial of T. The ***algebraic multiplicity*** of c is the largest $k \in \mathbb{N}$ such that $(t-c)^k | p(t)$. Moreover, the ***geometric multiplicity*** is $\dim(E_c)$, the dimension of the eigenspace corresponding to c.

Proposition 5.7. *Let $T : V \to V$ be a linear operator where $\dim(V) = n$ and let c be an eigenvalue of T having algebraic multiplicity m. Let $E_c \subseteq V$ be the eigenspace corresponding to c. Then*

$$1 \leq \dim(E_c) \leq m.$$

Proof. For convenience, let $k = \dim(E_c)$ and let $\{v_1, v_2, \dots, v_k\}$ be a basis for E_c. By the basis extansion theorem, add $v_{k+1}, v_{k+2}, \dots, v_n$ to form $\beta = \{v_1, v_2, \dots, v_n\}$, an ordered basis for V. Then,

$$[T]_\beta = \begin{bmatrix} cI & B \\ 0 & C \end{bmatrix}$$

for some $B \in M_{k\times(n-k)}(\mathbb{F})$ and $C \in M_{(n-k)(n-k)}(\mathbb{F})$, since $Tv_i = cv_i$ for each $i \in \{1, 2, \dots, k\}$. Therefore

$$f = \det\left(xI - [T]_\beta\right) = \det \begin{bmatrix} (x-c)I & B \\ 0 & xI - C \end{bmatrix} = \det((x-c)I)\det(xI - C) = (x-c)^k \det(xI - C)$$

is the characteristic polynomial of T. Thus $(x-c)^k | f$, which exactly means $k \leq m$, as desired. ♠

Proposition 5.8. *Let T be a linear operator and let $c_1, c_2, \dots, c_n$ be distinct eigenvalues of T. For each $i \in \{1, 2, \dots, k\}$, let v_i be an eigenvector corresponding to c_i. If*

$$\sum_{i=1}^{k} v_i = 0$$

then $v_i = 0$ for all $i \in \{1, 2, \dots, k\}$.

Proof. It is clear that $v_1, v_2, \dots, v_k$ are linearly independent. Thus, if

$$\sum_{i=1}^{k} v_i = 0,$$

then it must be the case that $v_i = 0 \in E_{c_i}$ for each $i \in \{1, 2, \dots, k\}$, which is the desired result. ♠

Proposition 5.9. *Let $T : V \to V$ be a linear opeartor and let $c_1, c_2, \dots, c_k$ be distinct eigenvalues of T. For each $i \in \{1, 2, \dots, k\}$, let S_i be a finite linearly independent subset of the eigenspace E_{c_i}. Then*

$$S = \bigcup_{i=1}^{k} S_i$$

is linearly independent.

Proof. Let $n_i = |S_i|$ and write $S_i = \{S_{i1}, S_{i2}, \dots, S_{in_i}\}$ for convenience. Then

$$S = \left\{S_{ij} : 1 \leq i \leq k, 1 \leq j \leq n_i\right\}.$$

For the sake of contradiction, suppose that S is linearly dependent. That is, there exist nonzero $a_{ij} \in \mathbb{F}$ such that

$$\sum_{i=1}^{k} \sum_{j=1}^{n_i} a_{ij} S_{ij} = 0.$$

Define $v_i = \sum_{j=1}^{n_i} a_{ij} S_{ij}$ then

$$T(v_i) = T\left(\sum_{j=1}^{n_i} a_{ij} S_{ij}\right) = \sum_{j=1}^{n_i} a_{ij} T(S_{ij}) = \sum_{j=1}^{n_i} a_{ij} c_i S_{ij} = c_i \sum_{j=1}^{n_i} a_{ij} S_{ij} = c_i v_i,$$

so each v_i is an eigenvector corresponding to c_i, and $\{v_1, v_2, \ldots, v_k\}$ is linearly independent. But this means

$$\sum_{i=1}^{k} v_i = \sum_{i=1}^{k} \sum_{j=1}^{n_1} a_{ij} S_{ij} \neq 0,$$

which is a contradiction. Thus S is linearly independent, as desired. ♠

Proposition 5.10. *Let $T : V \to V$ be a linear operator such that the characteristic polynomial f of T is a product of linear polynomials over $\mathbb{F}$ and $\dim(V) = n$. Let $c_1, c_2, \ldots, c_k$ be all distinct eigenvalues of T.*

(a) T is diagonalizable if and only if the algebraic multiplicity of c_i is equal to the geometric multiplicity of c_i for all $i \in \{1, 2, \ldots, k\}$.

(b) If T is diagonalizable and each β_i is an ordered basis for the eigenspace corresponding to c_i, then $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$ is an eigenbasis for V.

Proof. Write

$$f = \prod_{i=1}^{k} (x - c_i)^{m_i}$$

where $\sum_{i=1}^{k} m_i = \deg(f) = n$ by the assumption. For the forward direction of (a), suppose that T is diagonalizable. Let β be an ordered basis for V and let each $E_i \subseteq V$ be the eigenspace corresponding to c_i. Define $\beta_i = \beta \cap E_i$ then $|\beta_i| \leq \dim(E_i)$, since β_i is linearly independent. Moreover, $\dim(E_i) \leq m_i$ by the previous proposition. So we have

$$\sum_{i=1}^{k} |\beta_i| \leq \sum_{i=1}^{k} \dim(E_i) \leq \sum_{i=1}^{k} m_i = n,$$

but clearly $\sum_{i=1}^{k} |\beta_i| = n$ as well, since $\beta_i \cap \beta_j = \emptyset$ whenever $i \neq j$. Therefore

$$\sum_{i=1}^{k} [m_i - \dim(E_i)] = 0,$$

but since $m_i \geq \dim(E_i)$ for each $i \in \{1, 2, \ldots, k\}$, it must be the case that $m_i = \dim(E_i)$. For the reverse direction of (a), suppose $m_i = \dim(E_i)$ for each $i \in \{1, 2, \ldots, k\}$. We simultaneously show that T is diagonalizable and prove (b). For each $i \in \{1, 2, \ldots, k\}$, let β_i be an ordered basis for E_i, and let $\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k$. Then β is linearly independent by Proposition 5.11. Furthermore, since $\dim\{E_i\} = |\beta_i| = m_i$ for each $i \in \{1, 2, \ldots, k\}$,

$$|\beta| = \sum_{i=1}^{k} |\beta_i| = \sum_{i=1}^{k} m_i = n.$$

Therefore β is an ordered basis for V containing eigenvectors of T, which means T is diagonalizable. ♠

Remark 5.6. Proposition 5.10 provides the following way of checking whether a linear operator $T : V \to V$ is diagonalizable or not.

Proposition 5.11.
Diagonalizability Test

Let $T : V \to V$ is diagonalizable if and only if the following conditions hold.

(a) The characteristic polynomial of T is a product of linear factors.

(b) For each eigenvalue c of T, the algebraic multiplicity of c equals $\dim(V) - \operatorname{rank}(cI - T)$.

Proposition 5.12. **Eigendecomposition**

Let $A \in M_{n\times n}(\mathbb{F})$ be diagonalizable, $c_1, c_2, \ldots, c_k$ be all distinct eigenvalues, and $\beta_1, \beta_2, \ldots, \beta_k$ be bases for the corresponding eigenspaces $E_1, E_2, \ldots, E_n$. Let

$$\beta = \beta_1 \cup \beta_2 \cup \cdots \cup \beta_k,$$

let $P \in M_{n\times n}(\mathbb{F})$ be such that $\mathrm{Col}_j(P) \in \beta$ for each $j \in \{1, 2, \ldots, n\}$ and $\mathrm{Col}_j(P) \neq \mathrm{Col}_k(P)$ whenever $j \neq k$, and let $D \in M_{n\times n}(\mathbb{F})$ be a diagonal matrix whoe diagonal entries are eigenvalues corresponding to the columns of P. Then $A = PDP^{-1}$.