Linear Algebra II

Hrishikesh Belagali

August 2025

1 Basics

Example

For a set $\mathcal{S}$, let $\mathbb{F}^{\mathcal{S}}$ be the set of all functions from $\mathcal{S}$ to $\mathbb{F}$. Then, defined over canonical addition and scalar multiplications, $\mathbb{F}^{\mathcal{S}}$ is a vector space. The additive identity is the zero function 0, defined as $0(x) = 0$. The additive inverse can be defined as $-f : \mathcal{S} \to \mathbb{F}$ defined as $-f(x) = -(f(x)) \forall x \in \mathcal{S}$.

Note that $\mathbb{F}^n$ and $\mathbb{F}^\infty$ are special cases of $\mathbb{F}^{\mathcal{S}}$, where $\mathcal{S}$ is a finite set of size n or an infinite set, respectively.

Note that the empty set ϕ is not a vector space, nor is it a subspace of any vector space.

Example

The set of differentiable real-valued functions is a subspace of $\mathbb{R}^{\mathbb{R}}$. Note that in calculus, the sum of two continuous functions is continuous, and the sum of two differentiable functions is differentiable. Also, scalar multiples of continuous and differentiable functions are continuous and differentiable, respectively.

Definition

Let $V_1, \cdots V_n$ be subspaces of a vector space $\mathbf{V}$. Then, the sum of these subspaces is defined as

$$V_1 + V_2 + \cdots + V_n = \{v_1 + v_2 + \cdots + v_n \mid v_i \in V_i \text{ for all } i\}$$

Example

Let

$$V_1 = \left\{(w, w, x, x) \in \mathbb{F}^4 | w, x \in \mathbb{F}\right\}$$

$$V_2 = \left\{(y, y, y, z) \in \mathbb{F}^4 | y, z \in \mathbb{F}\right\}$$

Now, let $v_1 \in V_1$ and $v_2 \in V_2$. Then, we can write

$$v_1 = (w_1, w_1, x_1, x_1)$$

$$v_2 = (y_2, y_2, y_2, z_2)$$

for some $w_1, x_1, y_2, z_2 \in \mathbb{F}$. Then, we have

$$v_1 + v_2 = (w_1 + y_2, w_1 + y_2, x_1 + y_2, x_1 + z_2) \in V_1 + V_2$$

Let W be defined as

$$W = \left\{(x, x, y, z) \in \mathbb{F}^4 | x, y, z, \in \mathbb{F}\right\}$$

Then, $v_1 + v_2 \in W$ so $V_1 + V_2 \subseteq W$.
Let $w \in W$. Then, we can write

$$w = (x_w, x_w, y_w, z_w)$$

for some $x_w, y_w, z_w \in \mathbb{F}$. Then, we have

$$w = (x_w, x_w, y_w, z_w) = (x_w, x_w, y_w, y_w) + (0, 0, 0, z_w - y_w) \in V_1 + V_2$$

$\therefore W = V_1 + V_2$

Lemma 1.1

For any subspaces $V_1, \cdots V_n$ of a vector space $\mathbf{V}$, $V_1 + \cdots + V_n$ is a subspace of $\mathbf{V}$. It is also the smallest subspace of V that contains all elements of the form $v_1 + \cdots + v_n$ where $v_i \in V_i$ for all i.

Proof. From the definition and that $V_1, \cdots V_n$ are subspaces, Since the subspaces themselves are closed under addition and scalar multiplication, $V_1 + \cdots + V_n$ is also closed under addition and scalar multiplication. Also, the zero vector $\mathbf{0}$ is in each of the subspaces, so $\mathbf{0} \in V_1 + \cdots + V_n$. Thus, $V_1 + \cdots + V_n$ is a subspace of $\mathbf{V}$.

Note: Generally, the set theoretic union is rarely a subspace, except for trivial cases where one space is a subspace of the other. However, intersections of subspaces are generally subspaces.

Definition: Direct sum

Let $V_1, \cdots V_n$ be subspaces of a vector space $\mathbf{V}$. Then, the sum $V_1 + \cdots + V_n$ is called a direct sum if each element of $V_1 + \cdots + V_n$ can be written in one and only one way as $v_1 + \cdots + v_n$ where $v_i \in V_i$ for all i. In this case, we say that the sum is a direct sum, denoted by

$$W = V_1 \oplus V_2 \oplus \cdots \oplus V_n$$

Example

Let

$$U = \left\{(x, x, y) \in \mathbb{F}^3 | x, y \in \mathbb{F}\right\}$$

Let

$$W = \left\{(x, 0, 0) \in \mathbb{F}^3 | x \in \mathbb{F}\right\}$$

Then, U and W are subspaces of $\mathbb{F}^3$. Any arbitrary vector in $\mathbb{F}^3$ can be written as

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} b \\ b \\ c \end{pmatrix} + \begin{pmatrix} a-b \\ 0 \\ 0 \end{pmatrix}$$

Since this is a unique representation, $U \oplus W = \mathbb{F}^3$.

Theorem 1.2

Let $V_1, \cdots V_n$ be subspaces of a vector space $\mathbf{V}$. Then, $V_1 + \cdots V_n = V_1 \oplus \cdots \oplus V_n$ if and only if the only way to write $\mathbf{0}$ as $v_1 + \cdots + v_n$ where $v_i \in V_i$ for all i is to take each $v_i = \mathbf{0}$. In other words, if $v_1 + \cdots + v_n = \mathbf{0}$ implies that each $v_i = \mathbf{0}$, then the sum is a direct sum.

Proof. Suppose that $V_1 + \cdots V_n$ is a direct sum. Then, the additive identity can be written as the sum of additive identities from each subspace. By definition of a direct sum, this is the *only* way to write the additive identity as a sum.
Suppose that the only way to write zero is as the sum of additive identities from each subspace. Consider an arbitrary vector $v \in V$. Suppose that there are two different ways of writing the sum,

$$v = u_1 + \cdots u_n; u_k \in V_k$$

$$v = v_1 + \cdots v_n; v_k \in V_k$$

Then, we can subtract these two equations

$$0 = (u_1 - v_1) + \cdots + (u_n - v_n); (u_k - v_k) \in V_k$$

Since the only way to write zero is as the sum of additive identities from each subspace, we must have $u_k - v_k = 0$ for all k. Thus, $u_k = v_k$ for all k, and the representation is unique. Therefore, by definition, the sum $V_1 + \cdots V_n$ is a direct sum.

Theorem 1.3

Let U and W be subspaces of a vector space $\mathbf{V}$. Then, the sum $U + W$ is a direct sum if and only if $U \cap W = \{0\}$.

Proof. Suppose that $U + W$ is a direct sum. Let $v \in U \cap W$. Then, $v \in U$, and $-v \in W$

$$0 = v + (-v)$$

Since the representation is unique, we must have $v = 0$. Thus, $U \cap W = \{0\}$.

Conversely, suppose that $U \cap W = \{0\}$. Let $u \in U$ and $w \in W$. Then, we can write

$$u + w = 0$$

From the previous result, it suffices to show that $u = w = 0$. This implies that w is the additive inverse of u, meaning $u, w \in U \cap W = \{0\}$. Therefore, $u = w = 0$ and $U + W = U \oplus W$.

Lemma 1.4

Suppose $v_1, \cdots v_m$ is a linearly dependent list in V. Then there exists $j \in \{1, \cdots m\}$ such that

$$v_j \in \text{span}(v_1, \cdots, v_{j-1})$$

If this condition holds, then

$$\text{span}(v_1, \cdots v_m) = \text{span}(v_1, \cdots, v_{j-1}, v_{j+1}, \cdots v_m)$$

Proof. Since $v_1, \cdots v_m$ is linearly dependent, there exist $a_1, \cdots a_m \in \mathbb{F}$ not all zero such that

$$a_1 v_1 + \cdots + a_m v_m = 0$$

Let k be the largest index such that $a_k \neq 0$. Then,

$$v_k = -\frac{a_1}{a_k} v_1 - \cdots - \frac{a_{k-1}}{a_k} v_{k-1}$$

Therefore, $v_k \in \text{span}(v_1, \cdots, v_{k-1})$.

Suppose $v_k \in \text{span}(v_1, \cdots, v_{k-1})$. Then, we can write

$$v_k = b_1 v_1 + \cdots + b_{k-1} v_{k-1}$$

for some $b_1, \cdots, b_{k-1} \in \mathbb{F}$. Let $u \in \text{span}(v_1, \cdots v_m)$. Then,

$$u = c_1 v_1 + \cdots c_m v_m \text{ for some } c_1, \cdots c_m \in \mathbb{F}$$

Substituting for v_k, we get

$$u = c_1 v_1 + \cdots + c_{k-1} v_{k-1} + c_k (b_1 v_1 + \cdots + b_{k-1} v_{k-1}) + c_{k+1} v_{k+1} + \cdots + c_m v_m$$

Therefore, $u \in \text{span}(v_1, \cdots, v_{k-1}, v_{k+1}, \cdots v_m)$. Proving the other direction is trivial, you can set the coefficients of higher indices than k to zero.

Theorem 1.5

Let V be a finite-dimensional vector space. Suppose that $u_1, \cdots u_m$ is linearly independent in V and $w_1, \cdots w_n$. Then, $m \leq n$. In other words, any linearly independent list is smaller or the same size as any spanning list.

Theorem 1.6

Suppose U is a subspace of a finite-dimensional vector space V. Then there exists a subspace W of V such that $V = U \oplus W$.

Proof. Let $u_1, \cdots u_m$ be a basis of U. Since U is a subspace of V, we can extend this basis to a basis of V, say $u_1, \cdots u_m, w_1, \cdots w_n$. Let $W = \text{span}(w_1, \cdots w_n)$. Then, we have

$$V = U + W$$

since any vector in V can be written as a linear combination of the basis vectors.
Now, we need to show that the sum is direct. Suppose that

$$u + w = 0; u \in U, w \in W$$

Then, we can write

$$u = a_1 u_1 + \cdots + a_m u_m; a_i \in \mathbb{F}$$

$$w = b_1 w_1 + \cdots + b_n w_n; b_i \in \mathbb{F}$$

Therefore, we have

$$a_1 u_1 + \cdots + a_m u_m + b_1 w_1 + \cdots + b_n w_n = 0$$

Since the basis vectors are linearly independent, all coefficients must be zero. Thus, $u = w = 0$, and the sum is direct.

Theorem 1.7

If V_1 and V_2 are finite-dimensional subspaces of V, then

$$\text{dim}(V_1 + V_2) = \text{dim}(V_1) + \text{dim}(V_2) - \text{dim}(V_1 \cap V_2)$$

Proof. Let $u_1, \cdots u_m$ be a basis of $V_1 \cap V_2$. We can extend this basis to a basis of V_1, say $u_1, \cdots u_m, v_1, \cdots v_k$. Similarly, we can extend the basis of $V_1 \cap V_2$ to a basis of V_2, say $u_1, \cdots u_m, w_1, \cdots w_l$. We claim that the list

$$u_1, \cdots u_m, v_1, \cdots v_k, w_1, \cdots w_l$$

is a basis of $V_1 + V_2$.
Since the dimension of V_1 is $m + k$ and the dimension of V_2 is $m + l$, the dimension of $V_1 + V_2$ is at most $m + k + l$ Therefore,

$$\dim(V_1 + V_2) = m + k + l = (m + k) + (m + l) - m = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2)$$

It remains to show that the list spans $V_1 + V_2$ and is linearly independent.
Let $v \in V_1 + V_2$. Then, we can write

$$v = v_1 + v_2; v_1 \in V_1, v_2 \in V_2$$

Since $u_1, \cdots u_m, v_1, \cdots v_k$ is a basis of V_1, we can write

$$v_1 = a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \cdots + b_k v_k; a_i, b_i \in \mathbb{F}$$

Similarly, we can write

$$v_2 = c_1 u_1 + \cdots + c_m u_m + d_1 v_1 + \cdots + d_k v_k; c_i, d_i \in \mathbb{F}$$

Therefore, we have

$$v = (a_1 + c_1)u_1 + \cdots + (a_m + c_m)u_m + b_1 v_1 + \cdots + b_k v_k + d_1 w_1 + \cdots + d_l w_l$$

Thus, the list spans $V_1 + V_2$.
To show that the list is linearly independent, suppose we have a linear combination

$$\sum_{i=1}^{m} \alpha_i u_i + \sum_{j=1}^{k} \beta_j v_j + \sum_{l=1}^{l} \gamma_l w_l = 0$$

for some scalars $\alpha_i, \beta_j, \gamma_l \in \mathbb{F}$. We need to show that all coefficients must be zero.
Since $u_1, \cdots u_m, v_1, \cdots v_k$ is a basis of V_1, we can write

$$v_1 = a_1 u_1 + \cdots + a_m u_m + b_1 v_1 + \cdots + b_k v_k; a_i, b_i \in \mathbb{F}$$

Similarly, we can write

$$v_2 = c_1 u_1 + \cdots + c_m u_m + d_1 v_1 + \cdots + d_k v_k; c_i, d_i \in \mathbb{F}$$

Complete ts later

2 Eigenvectors and eigenvalues

2.1 Invariant subspaces

Definition: Invariant subspace

Let $T \in \mathcal{L}(\mathbf{V})$. A subspace U of $\mathbf{V}$ is called invariant under T if

$$u \in U \implies T(u) \in U$$

In other words, U is invariant under T if $T(U) \subseteq U$.

Definition: Eigenvalues and eigenvectors

A number $\lambda \in \mathbb{F}$ is called an eigenvalue of $T \in \mathcal{L}(\mathbf{V})$ if there exists a non-zero vector $v \in \mathbf{V}$ such that

$$T(v) = \lambda v$$

Such a vector v is called an eigenvector corresponding to the eigenvalue λ.

Definition: Polynomials of linear operators

Let m be a positive integer. Define T^m as $T \circ T \circ \cdots \circ T$ (m times). Define $T^0 = I$. Define $T^{-m} = (T^{-1})^m$ if T is invertible. Then, let $p(z) = a_m z^m + a_{m-1} z^{m-1} + \cdots + a_1 z + a_0$ be a polynomial with coefficients in $\mathbb{F}$. Then, we can define the polynomial of the operator T as

$$p(T) = a_m T^m + a_{m-1} T^{m-1} + \cdots + a_1 T + a_0 I$$

Lemma 1.8

If p, q are polynomials and $T \in \mathcal{L}(\mathbf{V})$, then $(pq)(T) = p(T)q(T) = q(T)p(T)$.

Proof. Let $p(z) = a_m z^m + a_{m-1} z^{m-1} + \cdots + a_1 z + a_0$ and $q(z) = b_n z^n + b_{n-1} z^{n-1} + \cdots + b_1 z + b_0$.
Then, we have

$$p(T) = a_m T^m + a_{m-1} T^{m-1} + \cdots + a_1 T + a_0 I$$
$$q(T) = b_n T^n + b_{n-1} T^{n-1} + \cdots + b_1 T + b_0 I$$

Now, we can compute the product $p(T)q(T)$ as follows:

$$\begin{aligned} p(T)q(T) &= (a_m T^m + a_{m-1} T^{m-1} + \cdots + a_1 T + a_0 I)(b_n T^n + b_{n-1} T^{n-1} + \cdots + b_1 T + b_0 I) \\ &= a_m b_n T^{m+n} + (a_m b_{n-1} + a_{m-1} b_n) T^{m+n-1} + \cdots + (a_0 b_0) I \end{aligned}$$

This is exactly the polynomial $(pq)(T)$, where $(pq)(z) = p(z)q(z)$. Since multiplication of polynomials is commutative, we also have $p(T)q(T) = q(T)p(T)$.

Theorem 1.9

Let $T \in \mathcal{L}(\mathbf{V})$ and $p \in \mathcal{P}(\mathbb{F})$. Then, null $p(T)$ and range $p(T)$ are invariant under T.

Proof. Let $u \in$ null $p(T)$. Then, we have

$$p(T)(u) = 0$$

Applying T to both sides, we get

$$p(T)(Tu) = (p(T)T)u = (Tp(T))u = T(p(T)u) = T(0) = 0$$

Thus, $T(u) \in$ null $p(T)$, showing that null $p(T)$ is invariant under T.

Now, let $v \in$ range $p(T)$. Then, we can write

$$v = p(T)(w)$$

for some $w \in \mathbf{V}$. Applying T, we get

$$T(v) = T(p(T)(w)) = p(T)(T(w))$$

Thus, $T(v) \in$ range $p(T)$, showing that range $p(T)$ is invariant under T.

2.2 Characteristic polynomials

Definition: Monic polynomial

A polynomial $p \in \mathcal{P}(\mathbb{F})$ is called monic if the leading coefficient is 1.

Definition: Minimal polynomial

Let $T \in \mathcal{L}(\mathbf{V})$. A monic polynomial p of smallest degree such that $p(T) = 0$ is called the minimal polynomial of T.

Theorem 1.10

Let $\mathbf{V}$ be finite dimensional. Let $T \in \mathcal{L}(\mathbf{V})$. Then, there is a unique monic polynomial p in $\mathcal{P}(\mathbb{F})$ of smallest degree such that $p(T) = 0$. Furthermore, the degree of $\deg p \leq \dim \mathbf{V}$.

Proof.

Existence: We will prove it by induction on the dimension of $\mathbf{V}$.

Base case:
If $\dim \mathbf{V} = 0$, then $\mathbf{V} = \{0\}$. Thus, T is the zero operator, and we can take $p(z) = 1$.

Inductive step:
Suppose that $\dim \mathbf{V} > 0$ and suppose that the results is true for all operators on all vector spaces of dimension strictly less than $\dim \mathbf{V}$. Let $u \in \mathbf{V}, u \neq 0$ The list $u, Tu, T^2u, \cdots T^n u$ must be linearly dependent for some $n \leq \dim \mathbf{V}$. Then there exists a smallest positive integer m such that $T^m u$ is a linear combination

of the previous $T^k u$'s. Thus, we can write

$$T^m u = -a_{m-1}T^{m-1}u - \cdots - a_1 Tu - a_0 u$$

for some $a_0, \cdots a_{m-1} \in \mathbb{F}$. Let

$$p_1(z) = z^m + a_{m-1}z^{m-1} + \cdots + a_1 z + a_0$$

Then, we have $p_1(T)(u) = 0$.
Since null $p_1(T)$ is invariant under T, $u, Tu, T^2u, \cdots T^{m-1}u \in$ null $p_1(T)$. Note that from the definition of m, the list $u, Tu, T^2u, \cdots T^{m-1}u$ is linearly independent. Therefore, $\dim \text{null } p_1(T) \geq m$. Then,

$$\dim \text{range } p_1(T) = \dim \mathbf{V} - \dim \text{null } p_1(T) \leq \dim \mathbf{V} - m < \dim \mathbf{V} \leq n - m$$

Since range $p_1(T)$ is invariant under T, we can apply the inductive hypothesis to the operator $T|_{\text{range } p_1(T)}$. Thus, there exists a unique monic polynomial $q \in \mathcal{P}(\mathbb{F})$ of degree $\leq n - m$ such that $q(T|_{\text{range } p_1(T)}) = 0$. Then, for all $v \in \mathbf{V}$, we have $p_1(T) \in$ range$q(T)$. Thus, $q(T)(p_1(T)(v)) = 0$. Therefore, we have

$$(qp_1)(T)(v) = q(T)(p_1(T)(v)) = 0$$

for all $v \in \mathbf{V}$. Therefore, $(qp_1)(T) = 0$, and we can take $p = qp_1$ as our minimal polynomial.

Uniqueness: Suppose p and q are both monic polynomials of smallest degree such that $p(T) = q(T) = 0$. Then $(p - q)(T) = 0$, and since p and q have the same degree and leading coefficient, $p - q$ has degree strictly less than $\deg p$. Note that $p - q$ is not necessarily monic. However, we can write $p - q = cr$ where c is the leading coefficient of $p - q$ and r is a monic polynomial. Then, we have

$$0 = (p - q)(T) = cr(T)$$

By minimality of $\deg p$, we must have $p - q = 0$, so $p = q$.

Definition

Let $\mathbf{V}$ be finite-dimensional and $T \in \mathcal{L}(\mathbf{V})$. Let $p(z)$ be the minimal polynomial of T. Then,

- The zeros of p are the eigenvalues of T.
- If $\mathbf{V}$ is a complex vector space, then

$$p(z) = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_m)$$

  where $\lambda_1, \cdots \lambda_m$ are the eigenvalues of T (possibly repeated).

Proof. **Forward: If λ is a zero of p, then λ is an eigenvalue of T.**
Suppose that $p(\lambda) = 0$. Then, we can write

$$p(z) = (z - \lambda)q(z)$$

where $q(z)$ is a monic polynomial. Then, we have

$$p(T) = (T - \lambda I)q(T) = 0$$

For some $v \in \mathbf{V}$ such that $q(T)v \neq 0$, we have

$$(T - \lambda I)(q(T)v) = 0$$

Since $\deg q = \deg p - 1 < \deg p$, such a v exists. Thus, $T(q(T)v) = \lambda(q(T)v)$, and $q(T)v$ is an eigenvector corresponding to the eigenvalue λ.

Reverse: If λ is an eigenvalue of T, then λ is a zero of p.
Suppose that λ is an eigenvalue of T. Then, there exists a non-zero vector $v \in \mathbf{V}$ such that

$$T(v) = \lambda v$$

Then, we have

$$T(Tv) = T(\lambda v) = \lambda T(v) = \lambda^2 v$$

Continuing this way, we can show that

$$T^k(v) = \lambda^k v$$

for all non-negative integers k. Then, we have

$$p(T)(v) = (T^m + a_{m-1}T^{m-1} + \cdots + a_1 T + a_0 I)(v) = (\lambda^m + a_{m-1}\lambda^{m-1} + \cdots + a_1\lambda + a_0)v = p(\lambda)v$$

Since $p(T) = 0$, we have $p(\lambda)v = 0$. Since $v \neq 0$, we must have $p(\lambda) = 0$.

By the fundamental theorem of algebra, if $\mathbf{V}$ is a complex vector space, then $p(z)$ can be factored into linear factors. Since the zeros of p are the eigenvalues of T, we can write

$$p(z) = (z - \lambda_1)(z - \lambda_2)\cdots(z - \lambda_m)$$

where $\lambda_1, \cdots \lambda_m$ are the eigenvalues of T (possibly repeated).

Corollary 1.11

Let $\mathbf{V}$ be a nonzero, finite-dimensional, complex vector space and $T \in \mathcal{L}(\mathbf{V})$. Then, T has an eigenvalue.

Note: The theorem is not true for infinite-dimensional vector spaces. For example, consider the vector space $\mathbb{C}^\infty$ and the right shift operator T defined as

$$T(x_1, x_2, x_3, \cdots) = (0, x_1, x_2, x_3, \cdots)$$

Then, T has no eigenvalues. To see this, suppose that λ is an eigenvalue of T. Then, there exists a non-zero vector $v = (x_1, x_2, x_3, \cdots)$ such that

$$T(v) = \lambda v$$

This gives us the equation

$$(0, x_1, x_2, x_3, \cdots) = \lambda(x_1, x_2, x_3, \cdots)$$

Equating components, we find that

$$\begin{aligned} 0 &= \lambda x_1 \\ x_1 &= \lambda x_2 \\ x_2 &= \lambda x_3 \\ &\vdots \end{aligned}$$

If $\lambda \neq 0$, this implies that $x_1 = x_2 = x_3 = \cdots = 0$, contradicting the assumption that v is non-zero. Therefore, we must have $\lambda = 0$. However, if $\lambda = 0$, then we have

$$T(v) = (0, x_1, x_2, x_3, \cdots) = 0$$

which again implies that $v = 0$. Thus, we conclude that T has no eigenvalues.

Recall that $q(z)$ divides $p(z)$ if there exists a polynomial $r(z)$ such that $p(z) = q(z)r(z)$.

Theorem 1.12

Let $\mathbf{V}$ be finite-dimensional and $T \in \mathcal{L}(\mathbf{V})$. Let $p(z)$ be the minimal polynomial of T. Let $q(z) \in \mathcal{P}(\mathbb{F})$. Then, $q(T) = 0$ if and only if $p(z)$ divides $q(z)$.

Proof. If $q = ps$ for some $s \in \mathcal{P}(\mathbb{F})$, then $q(T) = p(T)s(T) = 0$, $s(T) = 0$.
Conversely, suppose that $q(T) = 0$. Using polynomial long division, we can write

$$q(z) = p(z)s(z) + r(z)$$

where $\deg r < \deg p$. Then, we have

$$r(T) = q(T) - p(T)s(T) = 0 - 0 = 0$$

Since $\deg r < \deg p$, we must have $r(z) = 0$. Thus, we conclude that $p(z)$ divides $q(z)$.

Corollary 1.13

Let $\mathbf{V}$ be finite-dimensional and $T \in \mathcal{L}(\mathbf{V})$. and let $U \subseteq \mathbf{V}$ be invariant under T. Let T_U be the restriction of T to U. Thus, the minimal polynomial of T_U divides the minimal polynomial of T.

Theorem 1.14

Let $\mathbf{V}$ be finite dimensional and $T \in \mathcal{L}(\mathbf{V})$. Then, T is not invertible if and only if the minimal polynomial of T does not have a constant term.

Proof. Let p be the minimal polynomial of T. Suppose that T is not invertible. Then, null $T \neq \{0\}$. Then, 0 is an eigenvalue of T. Thus, $p(0) = 0$, so p does not have a constant term. Conversely, suppose that p does not have a constant term. Then, 0 is an eigenvalue of T, so null $T \neq \{0\}$. Thus, T is not invertible.

Theorem 1.15

Every operator on an odd-dimensional real vector space has an eigenvalue.

Proof. later

2.3 Upper-triangular matrices

When discussing operators on $\mathbf{V}$, we'll fix a basis for $\mathbf{V}$. Then, we can represent the operator as a matrix.

Definition: Upper triangular matrix

Let $T \in \mathcal{L}(\mathbf{V})$, where $\mathbf{V}$ is finite-dimensional. The matrix of T with respect to a basis $v_1, \cdots v_n$ of $\mathbf{V}$ is called upper-triangular if

$$Tv_j \in \text{span}(v_1, \cdots v_j)$$

for all $j = 1, \cdots n$. In other words, the matrix of T has all entries below the main diagonal equal to zero.

Theorem 1.16

Let $T \in \mathcal{L}(\mathbf{V})$, where $\mathbf{V}$ is finite dimensional. Let $v_1, \cdots v_n$ be a basis of $\mathbf{V}$. Then, the following are equivalent:

1. The matrix of T with respect to $v_1, \cdots v_n$ is upper-triangular.
2. $span(v_1, \cdots v_j)$ is invariant under T for $j = 1, \cdots n$.
3. $Tv_j \in \text{span}(v_1, \cdots v_{j-1})$ for $j = 1, \cdots n$.

Proof. $(1) \implies (2)$:
Suppose $k \in \{1, \cdots n\}$ and let $j \in \{1, \cdots k\}$. Then, we can write

$$Tv_j = a_1 v_1 + \cdots + a_j v_j$$

for some $a_1, \cdots a_j \in \mathbb{F}$. Therefore, $Tv_j \in \text{span}(v_1, \cdots v_k)$. Since j was arbitrary, we conclude that $\text{span}(v_1, \cdots v_k)$ is invariant under T.

$(2) \implies (3)$:
From the definition of $\text{span}(v_1, \cdots v_j)$ being invariant under T, we have

$$Tv_k \in \text{span}(v_1, \cdots v_k)$$

(3) $\implies$ (1):
Suppose $k \in \{1, \cdots n\}$ and let $j \in \{1, \cdots k\}$. Then, we can write

$$Tv_j = a_1 v_1 + \cdots + a_j v_j$$

for some $a_1, \cdots a_j \in \mathbb{F}$. Therefore, $Tv_j \in \text{span}(v_1, \cdots v_k)$. Since j was arbitrary, we conclude that $\text{span}(v_1, \cdots v_k)$ is invariant under T.

Theorem 1.17

Let $T \in \mathcal{L}(\mathbf{V})$, where $\mathbf{V}$ is finite dimensional. If $\mathcal{M}(T)$ is upper-triangular with respect to some basis of $\mathbf{V}$, and $\lambda_1, \cdots \lambda_n$ are the entries on the main diagonal of $\mathcal{M}(T)$, then

$$(T - \lambda_1 I)(T - \lambda_2 I) \cdots (T - \lambda_n I) = 0$$

Proof. Let $p(z) = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_n)$. It suffices to show that $(T - \lambda_1 I)(T - \lambda_2 I) \cdots (T - \lambda_n I)$ vanishes on $\text{span}(v_1, \cdots v_k)$ for all $k = 1, \cdots n$. For $k = 1$, this is true since $Tv_1 = \lambda_1 v_1$.
Now note $(T - \lambda_2 I) \in \text{span}(v_1)$, so $(T - \lambda_1 I)(T - \lambda_2 I)v_2 = 0$. Continuing this way, we can show that

$$(T - \lambda_1 I)(T - \lambda_2 I) \cdots (T - \lambda_k I)v_k = 0$$

for all $k = 1, \cdots n$. Thus, we conclude that

$$(T - \lambda_1 I)(T - \lambda_2 I) \cdots (T - \lambda_n I) = 0$$

Theorem 1.18

Let $T \in \mathcal{L}(\mathbf{V})$, where $\mathbf{V}$ is finite-dimensional. Then, the eigenvalues are precisely the entries on the main diagonal of any upper-triangular matrix representing T.

Proof. Let $p(z)$ be the minimal polynomial of T. Let $q(z) = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_n)$, where $\lambda_1, \cdots \lambda_n$ are the entries on the main diagonal of the upper-triangular matrix representing T. Then, $q(T) = 0$. Thus, p divides q. Then, $p(z)$ must be of the form

$$p(z) = (z - \lambda_{i_1})(z - \lambda_{i_2}) \cdots (z - \lambda_{i_m})$$

where $i_1, i_2, \cdots i_m$ are some indices in $\{1, 2, \cdots n\}$.
Let $j \in I = \{1, 2, \cdots n\}$ be arbitrary. Since the jth diagonal of a product of upper-triangular matrix is the product of the jth diagonals of the matrices, the jth element of $p(T)$ is

$$\prod_{i \in I} (\lambda_j - \lambda_i) = 0$$

Thus, there exists $i \in I$ such that $\lambda_j - \lambda_i = 0$, or $\lambda_j = \lambda_i$. Therefore, λ_j is an eigenvalue of T.

Theorem 1.19

Let $\mathbf{V}$ be a finite-dimensional vector space and $T \in \mathcal{L}(\mathbf{V})$. Then, T has an upper-triangular matrix with respect to some basis of $\mathbf{V}$ if and only if the minimal polynomial of T has the form

$$p(z) = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_n)$$

where $\lambda_1, \cdots \lambda_n$ are in $\mathbb{F}$ (not necessarily distinct).

Note: This is always true for $\mathbb{C}$, it becomes an issue for $\mathbb{R}$.

Proof. First, suppose that the matrix of T is upper-triangular with respect to some basis $\mathcal{B} = \{v_1, \cdots v_n\}$. Let $\alpha_1, \cdots \alpha_n$ be the diagonal entries of the matrix of T and $q(z) = (z - \alpha_1)(z - \alpha_2) \cdots (z - \alpha_n)$. Then, we have $q(T) = 0$. Let $p(z)$ be the minimal polynomial of T. Then, p divides q, so $p(z)$ must be of the form

$$p(z) = (z-\alpha_1)(z-\alpha_2)\cdots(z-\alpha_n).$$

Conversely, suppose that the minimal polynomial of T has the form

$$p(z) = (z-\lambda_1)(z-\lambda_2)\cdots(z-\lambda_m)$$

where $\lambda_1, \cdots \lambda_m$ are in $\mathbb{F}$. We prove the result by induction on $m = \deg p$.

Base case: $m = 1$
In this case, we have $p(z) = z - \lambda_1$. Thus, $T = \lambda_1 I$. Therefore, the matrix of T with respect to any basis is upper-triangular.

Inductive step:
Suppose that the result holds for all $k < m$. Let $U = \text{range }(T - \lambda_m I)$. Then, U is invariant under T. Let T_U be the restriction of T to U. Then, the minimal polynomial of T_U divides $p(z)/(z-\lambda_m)$. Then, the minimal polynomial of T_U has the form

$$(z-\lambda_1)(z-\lambda_2)\cdots(z-\lambda_{m-1})$$

By the inductive hypothesis, there exists a basis $u_1, \cdots u_m$ of U such that the matrix of T_U is upper triangular with respect to this basis. Then, for $k \in \{1, \cdots m\}$, we have

$$Tu_k = T|_u(u_k) \in \text{span }(u_1, \cdots u_k)$$

Now, extend this basis of U to a basis of $\mathbf{V}$, say $u_1, \cdots u_m, v_1, \cdots v_n$ where $n = \dim \mathbf{V} - \dim U$. If $k \in \{1, \cdots n\}$, then we have

$$Tv_k = (T - \lambda_m I)v_k + \lambda_m v_k \in \text{span }(u_1, \cdots u_m, v_1, \cdots v_k)$$

Therefore, T is upper-triangular with respect to the basis $u_1, \cdots u_m, v_1, \cdots v_n$.

Corollary 1.20

Let $\mathbf{V}$ be a finite dimensional complex vector space and $T \in \mathcal{L}(\mathbf{V})$. Then, T has an upper-triangular matrix with respect to some basis of $\mathbf{V}$.

Proof. It follows from the fundamental theorem of algebra.

Definition: Diagonalizable

Let $\mathbf{V}$ be finite dimensional and $T \in \mathcal{L}(\mathbf{V})$. We say that T is diagonalizable if T has a diagonal matrix with respect to some basis of $\mathbf{V}$.

Example

Define $T \in \mathcal{L}(\mathbb{R}^2)$ by

$$T\left(x, y\right) = \left(41x + 7y, -20x + 74y\right)$$

Then, the matrix of T wrt the standard basis is

$$\mathcal{M}(T) = \begin{pmatrix} 41 & 7 \\ -20 & 74 \end{pmatrix}$$

With respect to the basis $v_1 = \begin{pmatrix} 1 \\ 4 \end{pmatrix}, v_2 = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$, the matrix of T is

$$\mathcal{M}(T) = \begin{pmatrix} 69 & 0 \\ 0 & 46 \end{pmatrix}$$

Therefore, T is diagonalizable.

It is convenient to give a name to the set of all eigenvectors (and $\mathbf{0}$).

Definition: Eigenspace

Let $\mathbf{V}$ be finite dimensional and $T \in \mathcal{L}(\mathbf{V})$. Let λ be an eigenvalue of T. The eigenspace corresponding to λ is defined as

$$E(\lambda, T) = \text{null } (T - \lambda I)$$

Note that T restricted to $E(\lambda, T)$ is just multiplication by λ.

Theorem 1.21

If $\lambda_1, \cdots \lambda_m$ are distinct eigenvalues of $T \in \mathcal{L}(\mathbf{V})$, then then the sum of the corresponding eigenspaces is a direct sum, and $\dim E(\lambda_1, T) + \cdots + \dim E(\lambda_m, T) \leq \dim \mathbf{V}$.

Proof. Suppose $v_1, \cdots v_m = 0$, where $v_k \in E(\lambda_k, T)$ for $k = 1, \cdots m$. We proved before that eigenvectors corresponding to distinct eigenvalues are linearly independent. Thus, $v_1 = \cdots = v_m = 0$. Therefore, the sum is direct. The dimension inequality follows from the properties of direct sums.

We will be developing a criterion for diagonalizability.

Theorem 1.22

Let $\mathbf{V}$ be finite dimensional and $T \in \mathcal{L}(\mathbf{V})$. Let $\lambda_1, \cdots \lambda_m$ be the distinct eigenvalues of T. Then, the following are equivalent:

1. T is diagonalizable.
2. $\mathbf{V}$ has a basis consisting of eigenvectors of T.
3. $\mathbf{V} = E(\lambda_1, T) \oplus E(\lambda_2, T) \oplus \cdots \oplus E(\lambda_m, T)$
4. $\dim E(\lambda_1, T) + \dim E(\lambda_2, T) + \cdots + \dim E(\lambda_m, T) = \dim \mathbf{V}$

Proof. $(1) \iff (2)$:
Note that T has a diagonal matrix with respect to some basis $(v_1, \cdots v_n)$ if and only if

$$Tv_j = \lambda_j v_j$$

This means that the basis consists of eigenvectors of T.

$(2) \iff (3)$:
Since $\mathbf{V}$ has a basis of eigenvectors of T, every vector in $\mathbf{V}$ can be written as a linear combination of these eigenvectors. $v = v_1, \cdots v_n$ where $v_k \in E(\lambda_{i_k}, T)$ for some $i_k \in \{1, \cdots m\}$. Thus,

$$\mathbf{V} = E(\lambda_1, T) + E(\lambda_2, T) + \cdots + E(\lambda_m, T)$$

.

$$\mathbf{V} = E(\lambda_1, T) \oplus E(\lambda_2, T) \oplus \cdots \oplus E(\lambda_m, T)$$

$(3) \iff (4)$:
This follows from the dimension formula for direct sums.

$(4) \implies (2)$:
Let $u_1, \cdots u_k$ be a basis of $E(\lambda_1, T)$, $v_1, \cdots v_l$ be a basis of $E(\lambda_2, T)$, and so on. Then, we claim that the list

$$u_1, \cdots u_k, v_1, \cdots v_l, \cdots$$

is linearly independent. Suppose that

$$a_1 u_1 + \cdots + a_k u_k + b_1 v_1 + \cdots + b_l v_l = 0$$

Then, we have

$$a_1 u_1 + \cdots + a_k u_k = -b_1 v_1 - \cdots - b_l v_l$$

Since $u_1, \cdots u_k$ are linearly independent, we must have $a_1 = \cdots = a_k = 0$. Similarly, since $v_1, \cdots v_l$ are linearly independent, we must have $b_1 = \cdots = b_l = 0$. Therefore, the list is linearly independent.

Theorem 1.23

If T has $\dim \mathbf{V}$ distinct eigenvalues, then T is diagonalizable.

Proof. Let $n = \dim \mathbf{V}$. and let $\lambda_1, \cdots \lambda_n$ be the distinct eigenvalues of T. Let v_k be the eigenvector corresponding to λ_k for $k = 1, \cdots n$. Then, the list $v_1, \cdots v_n$ is linearly independent. Since the list has length n, it is a basis of $\mathbf{V}$. Therefore, T is diagonalizable.

Example

Let $T \in \mathcal{L}(\mathbb{R}^3)$ be defined by

$$T\left(x, y, z\right) = \left(2x + y, 5y + 3z, 8z\right)$$

Then, the matrix of T with respect to the standard basis is

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 5 & 3 \\ 0 & 0 & 8 \end{pmatrix}$$

The eigenvalues of T are 2, 5, and 8. The corresponding eigenspaces are

$$E(2, T) = \text{span}\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad E(5, T) = \text{span}\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \quad E(8, T) = \text{span}\begin{pmatrix} 1 \\ 6 \\ 0 \end{pmatrix}$$

Since the eigenvalues are distinct, T is diagonalizable. A basis of $\mathbb{R}^3$ consisting of eigenvectors of T is given by

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 6 \\ 0 \end{pmatrix}$$

With respect to this basis, the matrix of T is

$$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 5 & 0 \\ 0 & 0 & 8 \end{pmatrix}$$

Now, we will find the necessary and sufficien condition for diagonalizability in terms of the minimal polynomial.

Theorem 1.24

Let $\mathbf{V}$ be finite dimensional and $T \in \mathcal{L}(\mathbf{V})$. T is diagonalizable if and only if the minimal polynomial of T has the form

$$p(z) = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_m)$$

where $\lambda_1, \cdots \lambda_m$ are distinct eigenvalues of T.

Proof. First, suppose that T is diagonalizable. Let $v_1, \cdots v_n$ be a basis of $\mathbf{V}$ consisting of eigenvectors of T, Let $\lambda_1, \cdots \lambda_m$ be the distinct eigenvalues of T. Then, for each λ_k there exists a v_j such that $Tv_j = \lambda_k v_j$. It follows that

$$(T - \lambda_1 I)(T - \lambda_2 I) \cdots (T - \lambda_m I)v_j = 0$$

for all $j = 1, \cdots n$. Then, the minimal polynomial of T is

$$(z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_m)$$

Since the minimal polynomial divides this from above, but each eigenvalues is a root of the minimal polynomial, it can't have smaller degree. Therefore, the minimal polynomial has the desired form.

Conversely, suppose the minimal polynomial of T has the form

$$p(z) = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_m)$$

where $\lambda_1, \cdots \lambda_m$ are distinct eigenvalues of T.

Base Case: $m = 1$
In this case, we have $p(z) = z - \lambda_1$. Thus, $T = \lambda_1 I$. Therefore, the matrix of T with respect to any basis is diagonal.

Inductive Step:
Let $m > 1$ and suppose that the result holds for all $k < m$. Let $U =$ range $(T - \lambda_m I)$. Then, U is invariant under T. Let T_U be the restriction of T to U. Then, the minimal polynomial of T_U divides $p(z)/(z - \lambda_m)$. Then, the minimal polynomial of T_U has the form

$$(z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_{m-1})$$

By the inductive hypothesis, T_U is diagonalizable and so U has a basis $u_1, \cdots u_m$ consistings of eigenvectors of T_U. Now, let $u \in$ range $(T - \lambda_m I) \cap$ null $(T - \lambda_m I)$ be arbitrary. Then, we have $Tu = \lambda_m u$ and so

$$(T - \lambda_1 I) \cdots (T - \lambda_{m-1})u = (\lambda_m - \lambda_1) \cdots (\lambda_m - \lambda_{m-1})u = 0$$

Since the eigenvalues are distinct, we must have $u = 0$. Therefore, range $(T - \lambda_m I) \cap$null $(T - \lambda_m I) = \{0\}$. and $U +$ null $(T - \lambda_m I)$ is a direct sum. Let $w_1, \cdots w_n$ be a basis of null $(T - \lambda_m I)$. Then,

$$v_1, \cdots v_m, w_1, \cdots w_n$$

is linearly independent. Finally,

$$m + n = \dim \text{range } (T - \lambda_m I) + \dim \text{null } (T - \lambda_m I) = \dim \mathbf{V}$$

Therefore, $\mathbf{V}$ has a basis of eigenvectors of T and so T is diagonalizable.

Theorem 1.25

Suppose T is diagonalizable and U is a subspace of $\mathbf{V}$ invariant under T. Then, the restriction of T to U is diagonalizable.

Proof. The minimal polynomial of T_U divides the minimal polynomial of T. Since the minimal polynomial of T has the form

$$(z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_m)$$

where $\lambda_1, \cdots \lambda_m$ are distinct eigenvalues of T, the minimal polynomial of T_U also has this form. Therefore, T_U is diagonalizable.

3 Inner Product Spaces

We define the norm of $x = (x_1, x_2, \cdots x_n) \in \mathbb{F}^n$ as

$$\|x\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

This is called the Euclidean norm. We define the dot product of $x = (x_1, x_2, \cdots x_n)$ and $y = (y_1, y_2, \cdots y_n)$ in $\mathbb{F}^n$ as

$$x \cdot y = x_1 y_1 + x_2 y_2 + \cdots + x_n y_n$$

The dot product satisfies the following properties:

- $\|x\|^2 = x \cdot x$.
- $x \cdot x \geq 0$ with equality if and only if $x = 0$.
- $x \cdot y = y \cdot x$.
- $(ax + by) \cdot z = a(x \cdot z) + b(y \cdot z)$ for all $a, b \in \mathbb{F}$.

We define the inner product on $\mathbb{C}^n$ as

$$x \cdot y = x_1\overline{y_1} + x_2\overline{y_2} + \cdots + x_n\overline{y_n}$$

Definition: Inner product

Let $\mathbf{V}$ be a vector space over $\mathbb{F}$. An inner product on $\mathbf{V}$ is a function that takes each ordered pair of vectors $u, v \in \mathbf{V}$ and produces a scalar in $\mathbb{F}$

$$\langle \cdot, \cdot \rangle : \mathbf{V} \times \mathbf{V} \to \mathbb{F}$$

such that for all $u, v, w \in \mathbf{V}$ and $a \in \mathbb{F}$, the following properties hold:

- $\langle u, v \rangle = \overline{\langle v, u \rangle}$
- $\langle au + bw, v \rangle = a\langle u, v \rangle + b\langle w, v \rangle$
- $\langle v, v \rangle \geq 0$ with equality if and only if $v = 0$

Note: If $\mathbb{F} = \mathbb{R}$, then $\langle u, v \rangle = \langle v, u \rangle$, i.e conjugate symmetry boils down to symmetry.

Example

We can define an inner product on the vector space of continuous real-valued functions on $[-1, 1]$ by

$$\langle f, g \rangle = \int_{-1}^{1} f(x)g(x)\, dx$$

Definition: Inner product space

A vector space $\mathbf{V}$ equipped with an inner product $\langle \cdot, \cdot \rangle$ is called an inner product space.

Throughout, we assume that $\mathbf{V}$ and $\mathbf{W}$ are inner product spaces.

Lemma 1.26

1. For each fixed $v \in \mathbf{V}$, the function $\mathbf{V} \to \mathbb{F}, u \mapsto \langle u, v \rangle$ is linear.
2. $\langle 0, v \rangle = \langle v, 0 \rangle = 0$ for all $v \in \mathbf{V}$.
3. $\langle v, 0 \rangle = 0$ for all $v \in \mathbf{V}$.
4. $\langle u, v + w \rangle = \langle u, v \rangle + \langle u, w \rangle$ for all $u, v, w \in \mathbf{V}$.
5. $\langle u, av \rangle = \overline{a}\langle u, v \rangle$ for all $u, v \in \mathbf{V}$ and $a \in \mathbb{F}$.

Proof.

1. This follows from the definitions of linearity and homogeneity of the inner product.
2. Since any linear map takes 0 to 0, we have $\langle 0, v \rangle = 0$.
3. This follows from conjugate symmetry and part (2).
4. This follows from linearity in the first argument.
5. This follows from conjugate symmetry.

Each inner product determines the norm of a vector.

Definition: Norm

For a vector $v \in \mathbf{V}$, the norm of v is defined as

$$\|v\| = \sqrt{\langle v, v \rangle}$$

Lemma 1.27

Let $v \in \mathbf{V}$.

1. $\|v\| = 0$ iff $v = 0$.
2. $\|av\| = |a|\|v\|$ for all $a \in \mathbb{F}$.

Proof.

1. This follows from the properties of inner products.
2. We have

$$\|av\| = \sqrt{\langle av, av\rangle} = \sqrt{a\bar{a}\langle v, v\rangle} = \sqrt{|a|^2\langle v, v\rangle} = |a|\|v\|$$

Definition

Two vectors $u, v \in \mathbf{V}$ are said to be orthogonal if $\langle u, v\rangle = 0$.

Note that $\langle u, v\rangle = 0$ iff $\langle v, u\rangle = 0$, so the order doesn't matter here. For $u, v \in \mathbb{R}^2$, the inner product

$$\langle u, v\rangle = \|u\|\|v\| \cos\theta$$

where θ is the angle between u and v. This also holds more generally in $\mathbb{R}^n$. With the usual Euclidean inner product, we have

$$\langle u, v\rangle = 0 \iff \cos\theta = 0 \iff \theta = \pm\frac{\pi}{2}$$

Thus, orthogonal vectors are perpendicular to each other.

Definition: Orthonormal list

A list of vectors $v_1, \cdots v_n$ in $\mathbf{V}$ is called orthonormal if $\|v_j\| = 1$ for all $j = 1, \cdots n$ and $\langle v_j, v_k\rangle = 0$ for all $j \neq k$.

Theorem 1.28

Suppose $e_1, \cdots e_n$ is an orthonormal list in $\mathbf{V}$. Then,

$$\|a_1e_1 + a_2e_2 + \cdots + a_ne_n\|^2 = |a_1|^2 + |a_2|^2 + \cdots + |a_n|^2$$

for all $a_1, a_2, \cdots a_n$ in $\mathbb{F}$.

Proof. Since each e_k has norm 1, this is immediate from the Pythagorean theorem applied repeatedly.

Theorem 1.29

Every orthonormal list of vectors in $\mathbf{V}$ is linearly independent.

Proof. Suppose $a_1, a_2, \cdots a_n$ in $\mathbb{F}$ such that

$$a_1e_1 + a_2e_2 + \cdots + a_ne_n = 0$$

Then, we have

$$0 = \|a_1e_1 + a_2e_2 + \cdots + a_ne_n\|^2 = |a_1|^2 + |a_2|^2 + \cdots + |a_n|^2$$

Therefore, $a_1 = a_2 = \cdots = a_n = 0$. Thus, the list is linearly independent.

Theorem 1.30: Bessel's Inequality

Suppose $e_1, \cdots e_n$ is an orthonormal list in $\mathbf{V}$. Then, for every $v \in \mathbf{V}$, we have

$$\sum_{j=1}^{n} |\langle v, e_j \rangle|^2 \leq \|v\|^2$$

Proof. Let $v \in \mathbf{V}$. We breakup v into a component in span$(e_1, \cdots e_n)$ and a component orthogonal to this subspace. Write

$$v = \langle v, e_1 \rangle e_1 + \langle v, e_2 \rangle e_2 + \cdots + \langle v, e_n \rangle e_n + w$$

Let u be the first part of this sum. Then, we have

$$\langle w, e_k, = \rangle \langle v, e_k, - \rangle \langle v, e_k, \langle \rangle, e_k, e_k \rangle = 0$$

for all $k = 1, \cdots n$. This implies $\langle u, w \rangle = 0$. Therefore, by the Pythagorean theorem, we have

$$\|v\|^2 = \|u\|^2 + \|w\|^2 \geq \|u\|^2$$

Definition: Orthonormal basis

A basis of $\mathbf{V}$ that is an orthonormal list is called an orthonormal basis.

Theorem 1.31

Let $\mathbf{V}$ be a finite-dimensional inner product space. Then, every orthonormal list of vectors in $\mathbf{V}$ of length equal to $\dim \mathbf{V}$ isan orthonormal basis of $\mathbf{V}$.

Proof. Every orthonormal list is linearly independent. Since the list has the right number of elements, it's a basis for $\mathbf{V}$.

Theorem 1.32: Parseval's Identity

Let $e_1, \cdots e_n$ be an orthonormal basis of $\mathbf{V}$. Then, for every $v, u \in \mathbf{V}$, we have

1. $v = \sum_{j=1}^{n} \langle v, e_j \rangle e_j$
2. $\|v\|^2 = \sum_{j=1}^{n} |\langle v, e_j \rangle|^2$
3. $\langle u, v \rangle = \sum_{j=1}^{n} \langle u, e_j \rangle \overline{\langle v, e_j \rangle}$

Proof.

1. Since $e_1, \cdots e_n$ is a basis of $\mathbf{V}$, we can write v as a linear combination of these vectors.

   $$v = a_1 e_1 + a_2 e_2 + \cdots + a_n e_n$$

   for some scalars $a_1, a_2, \cdots a_n$ in $\mathbb{F}$. Since the basis is orthonormal, we have

   $$\langle v, e_k \rangle = a_k$$

   for all $k = 1, \cdots n$. Therefore, we have

   $$v = \sum_{j=1}^{n} \langle v, e_j \rangle e_j$$

2. This follows from part (1) and the previous theorem.

3. Using part (1), we have

$$\langle u, v\rangle = \sum_{j=1}^{n} \langle u, e_j\rangle \overline{\langle v, e_j\rangle}$$

Definition: Gram-Schmidt Procedure

Let $\mathbf{V}$ be an inner product space. The Gram-Schmidt Procedure is a method for converting a linearly independent list of vectors $v_1, \cdots v_n$ in $\mathbf{V}$ into an orthonormal list $e_1, \cdots e_n$ such that

$$\text{span}(v_1, \cdots v_k) = \text{span}(e_1, \cdots e_k)$$

for each $k = 1, \cdots n$. The procedure is as follows:

1. Let $e_1 = \frac{v_1}{\|v_1\|}$.
2. For $k = 2, \cdots n$, let

$$w_k = v_k - \sum_{j=1}^{k-1} \langle v_k, e_j\rangle e_j$$

and let

$$e_k = \frac{w_k}{\|w_k\|}$$