1 Fields

Fields are an abstract structure that describes sets of "numbers" and their operations.

Definition 1.1 (Fields)
A set $\mathcal{F}$ together with two binary operations

- $+ : \mathcal{F} \times \mathcal{F} \to \mathcal{F}$ (Addition)
- $\cdot : \mathcal{F} \times \mathcal{F} \to \mathcal{F}$ (Multiplication)

$\forall x, y, z \in \mathcal{F}$

- $x + (y + z) = (x + y) + z$ (Associativity)
- $x(yz) = (xy)z$ (Associativity)
- $x + y = y + x$ (Commutatitivity)
- $xy = yx$ (Commutatitivity)
- $\exists 0 \in \mathcal{F}$ such that $x + 0 = x \ \forall x \in \mathcal{F}$ (Neutral additive element)
- $\exists 1 \in \mathcal{F}$ such that $x \cdot 1 = x \ \forall x \in \mathcal{F}$ (Neutral scalar multiplication element)
- $\forall x \in \mathcal{F} \ \exists - x \in \mathcal{F} \quad x + (-x) = 0$ (Additive Inverse)
- $\forall y \in \mathcal{F} \backslash \{0\} \exists y^{-1} \exists \mathcal{F} \quad yy^{-1} = 1$ (Multiplicative inverse)
- $x(y + z) = xy + xz$ (Distributivity)

Example of fields: Rational numbers $\mathbb{Q}$, real numbers $\mathbb{R}$ and complex numbers $\mathbb{C}$. Another example is the set $\mathcal{F} = \{0, 1\}$. The set $\mathbb{Q}(\sqrt{2}) = \{a + b\sqrt{2} : a, b \in \mathbb{Q}\}$ is also an example.

Definition 1.2
For any field $\mathcal{F}$ we denote the set of n-tuples by $\mathcal{F}^n$. We define two important operations on $\mathcal{F}^n$.

- Addition: Given two elements $x = (x_1, \ldots, x_n)$ $y = (y_1, \ldots, y_n)$ $x + y = (x_1 + y_1, \ldots, x_n + y_n)$
- Scalar multiplication: Given an element $\lambda \in \mathcal{F}$ and a n-tuple $x \in \mathcal{F}^n$ we define $\lambda x := (\lambda x_1, \ldots, \lambda x_n)$

We often write $0 \in \mathcal{F}^n$ for the n-tuple consisting of n zeros.
Let $x \in \mathcal{F}^n$ we define $-x := (-x_1, \ldots, -x_n)$ and we see that $x + (-x) = 0$.

For any $x \in \mathcal{F}^n$ we have $0 + x = x$

2 Vector Spaces

Definition 2.1 (Vector Space)
Let V be a set and $\mathcal{F}$ be a field.

Let $+ : V \times V \to V$ and $\cdot : \mathcal{F} \times V \to V$ be two binary operations. We say V is a vector space (with respect to these operations) over $\mathcal{F}$, or an $\mathcal{F}$-vector space (VS) if

- Addition is commutative: $\forall u, v \in V \quad u + v = v + u$.
- Addition is associative: $\forall u, v, w \in V \quad u + (v + w) = (u + v) + w$.
- Multiplication is associative: $\forall \lambda, \mu \in \mathcal{F} \forall v \in V \quad (\lambda\mu)v = \lambda(\mu v)$.
- Neutral additive: $\exists 0 \in V$ such that $\forall v \in V \quad 0 + v = v$.
- Inverse addition: $\forall v \in V \exists - v \in V \quad v + (-v) = 0$.
- Neutral scalar multiplication: $1 \in \mathcal{F}$ it holds that $1 \cdot v = v \quad v \in V$.
- Distributivity: $\forall u, v \in V \ \forall \lambda\mu \in \mathcal{F} \quad \lambda(u + v) = \lambda u + \lambda v$ and $(\lambda + \mu)v = \lambda v + \mu v$.

Example 2.1
a) $\mathcal{F}^n$ is an $\mathcal{F}$-VS, it holds that $\forall n \in \mathbb{N}$ especially $\mathcal{F}$ is an $\mathcal{F}$-VS.

b) $V = \{0\} \subseteq \mathcal{F}$ is an $\mathcal{F}$-VS.

c) $\mathcal{F}^\infty := \{(x_1, x_2, \dots) : x_i \in \mathcal{F} \ i \in \mathbb{N}\}$, the set of all infinite sequences is an $\mathcal{F}$-VS

d) Let $V := \{f : S \to \mathcal{F}\}$ be the set of functions from a set S into $\mathcal{F}$ then V is an $\mathcal{F}$-VS with $f, g \in V$ for which $(f + g)(s) := f(s) + g(s) \forall s \in S$. Similiarly $\forall \lambda \in \mathcal{F} \quad (\lambda f)(s) = \lambda(f(s))$. Sometimes you will see this notation:

$$V = \mathcal{F}^S$$

Example. $\mathbb{R}^{[0,1]}$.

Theorem 2.1
Let V be an $\mathcal{F}$-VS. Then the additive neutral element is unique.

Proof 2.1
Suppose there is another additive neutral element: 0 and $0'$ are both neutral. Then

$$\begin{aligned} 0 &= 0 + 0' \text{ since } 0' \text{ is neutral} \\ &= 0' \text{ since } 0 \text{ is neutral} \end{aligned}$$

Hence $0 = 0'$ and there is an unique neutral.

$\blacksquare$

Theorem 2.2
Let V be an $\mathcal{F}$-VS. Then every element in V has a unique additive inverse.

Proof 2.2
Let $v \in V$ and suppose w and w' are both additive inverse for v.

$$w' = 0 + w' = (w + v) + w' = w + (v + w') = w + 0 = w$$

We will from now on decide the unique inverse of v be $-v$ and write $w + (-v) := w - v$.

Theorem 2.3
Let V be an $\mathcal{F}$-VS. Then $\forall v \in V$

$$0 \in \mathcal{F} \quad v = 0 \in V$$

Proof 2.3
We see that

$$0 \cdot v = (0 + 0)v = 0v + 0v$$

Add $-0v$ on both sides

$$0 = 0v.$$

■

Theorem 2.4
Let V be an $\mathcal{F}$-VS. Then $\forall \lambda \in \mathcal{F}$

$$\lambda \cdot 0 = 0.$$

Proof 2.4

$$\lambda 0 = \lambda(0 + 0) = \lambda 0 + \lambda 0$$

Add $-\lambda 0$ on both sides

$$0 = \lambda 0.$$

■

Theorem 2.5
Let V be an $\mathcal{F}$-VS and $-1 \in \mathcal{F}$ is the additive inverse of the multiplicative neutral in $\mathcal{F}$. Then

$$(-1)v = -v \quad \forall v \in V$$

Proof 2.5

$$1 \cdot v + (-1)v = (1-1) \cdot v = 0v = 0$$

by Theorem 2.3.

■

For any VS V the subset $\{0\}$ is also a VS. We generalise this notion.

Definition 2.2 (Subspaces)
Let V be an $\mathcal{F}$-VS then a subset $U \subseteq V$ is called a subspace if U is also an $\mathcal{F}$-VS with respect to the same operations.

Theorem 2.6 (Proposition)
A subset $U \subseteq V$ of an $\mathcal{F}$-VS V is a subspace iff (=if and only if)

- $0 \in U$
- $\forall u, w \in U \quad u + w \in U$
- $\forall \lambda \in \mathcal{F} \quad \forall u \in U \quad \lambda u \in U$

Proof 2.6
$\Rightarrow$ If U is a VS then all these conditions hold.

$\Leftarrow$ Condition 1 implies neutral additive of VS.

By condition 3 we know that $(-1)u \in U$, $(-1)u = -u$ and thereby implies the additive inverse of VS.

■

Example 2.2
1) For any VS V, $\{0\}$ and V itself are subspaces.

2) The set of all polynomials with coefficients in some field $\mathcal{F}$ is a VS, called $\mathcal{F}[x]$. For every $0 \leq d \in \mathbb{N}_0$ the set of polynomials of degree at most d is a subspace.

3) We have seen that $\mathbb{R}^{[0,1]}$ is a $\mathbb{R}$-VS. The sets of continuous or differentiable functions form subspaces.

4) We can classify all subspaces of $\mathbb{R}^3$ in a hierarchy: $\mathbb{R}^3 >$ planes containing the origin $>$ lines going through the origin $> \{0\}$.

Definition 2.3
Let $U_1, U_2, \dots, U_m$ be subspaces of a VS V. Then we define their sum.

$$U_1 + \cdots + U_m := \{u_1 + \cdots + u_m : u_i \in U_i\}$$

Theorem 2.7 (Proposition)
Finite sums of subspaces are subspaces again.

Proof 2.7
We only need to show this for two subspaces.

Let $U_1, U_2 \subseteq V$ be subspaces. Then since $0 \in U_1$ and $0 \in U_2 \Rightarrow 0 + 0 = 0 \in U_1 + U_2$.
Let $u_1 + u_2, u_1' + u_2' \in U_1 + U_2$ then $u_1 + u_2 + u_1' + u_2' = (u_1 + u_1') + (u_2 + u_2') \in U_1 + U_2$
Let $\lambda \in \mathcal{F}$ then

$$\lambda(u_1 + u_2) = \lambda u_1 + \lambda u_2 \in U_1 + U_2.$$

■

Theorem 2.8 (Proposition)
Let $U_1, U_2 \subseteq V$ be subspaces, then $U_1 + U_2$ is the smallest subspace of V containing both.

Proof 2.8
We see that $U_1 \subseteq U_1 + U_2$ because $\forall u_1 \in U_1 \quad u_1 + 0 = u_1 \in U_1 + U_2$, the same applies to U_2.

Assume there exists $W \subseteq U_1 + U_2$ that contains U_1 and U_2. Then there must exist an element $u_1 + u_2 \notin W$. But $u_1 \in W$ and $u_2 \in W \rightarrow W$ is not a subspace.

EX: Functions and reals can be split into subspaces of even and odd reals.

EX: L_1, L_2 lies in $\mathbb{R}^n$ $L_1 + L_2 = \begin{cases} P \text{ plane} \\ L_1 \text{ if } L_1 = L_2 \end{cases}$

EX: P is a plane in $\mathbb{R}^3$ and L is a line in $\mathbb{R}^3$:

$$P + L = \begin{cases} \mathbb{R}^3 \text{ if } L \subsetneq P \\ P \text{ if } L \subseteq P \end{cases}$$

Definition 2.4 (Direct Sum)
Let $U_1, \ldots, U_m \subseteq V$ be subspaces. Then their sum is called a direct sum if $U_1 + \cdots + U_m$ has a unique representation as a sum $u_1 + \cdots + u_m$. We then write $U_1 \oplus \cdots \oplus U_m$ for this sum.

Theorem 2.9 (Proposition)
The sum $U_1 + \cdots + U_m$ is direct iff there is a unique way to write 0 as a sum $u_1 + \cdots + u_m$.

Proof 2.9
$\Rightarrow$ check

$\Leftarrow$

If the sum is not direct then there exists an element that has two different representations

$$u_1 + \cdots + u_m = u_1' + \cdots + u_m'$$

where not all $u_i = u_i'$. Then

$$(u_1 - u_1') + (u_2 - u_2') + \cdots + (u_m - u_m') = 0$$

and at least one different $u_i - u_i' \neq 0$.

■

Theorem 2.10 (Lemma)
$U + W$ is direct iff

$$U \cap W = \{0\}$$

Proof 2.10
$\Rightarrow$: Let $v \in U \cap W$ and $v \neq 0$ then

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

and hence the sum is not direct.

$\Leftarrow$: $0 = u + w \Rightarrow -u \in U \; -w \in W \Rightarrow u = w = 0$

3 Bases and Dimension

A list is an n-tuple.

Definition 3.1 (2.3 and 2.5)
Let $\mathbf{v}_1, \dots, \mathbf{v}_m$ be a list of vectors in an $\mathcal{F}$-VS. Then for any $\lambda_i \in \mathcal{F}$ we call $\lambda_1\mathbf{v}_1 + \cdots + \lambda_m\mathbf{v}_m$ a linear combination of $\mathbf{v}_1, \dots, \mathbf{v}_m$. (Note that λ_i can be zero)

The set of all linear combinations is called the span of $\mathbf{v}_1, \dots, \mathbf{v}_m$ and denoted $\operatorname{span}(\mathbf{v}_1, \dots, \mathbf{v}_m)$. For consistency we let $\operatorname{span}() = \{0\}$.

Theorem 3.1 (Proposition 2.7)
Let $\mathbf{v}_1, \dots, \mathbf{v}_m$ is a list of vectors. Then $\operatorname{span}(\mathbf{v}_i)$ is a subspace and it is the smallest subspace containing all $\mathbf{v}_i$.

Proof 3.1
We show that span is a subspace.

1. $0 \in \operatorname{span}(\mathbf{v}_i)$, just let $\lambda_i = 0 \ \forall\, i$
2. $\sum_{i=1}^m \lambda_i\mathbf{v}_i + \sum \mu_i\mathbf{v}_i = \sum(\lambda_i + \mu_i)\mathbf{v}_i \in \operatorname{span}(\mathbf{v}_i)$
3. $\sum \lambda_i\mathbf{v}_i = \sum(\mu\lambda_i)\mathbf{v}_i \in \operatorname{span}(\mathbf{v}_i)$

$$\mathbf{v}_i = 0\mathbf{v}_1 + 0\mathbf{v}_2 + \cdots + 1\mathbf{v}_i + \cdots + 0\mathbf{v}_m \in \operatorname{span}(\mathbf{v}_i)$$

Assume $W \subseteq \operatorname{span}(\mathbf{v}_i)$ such that $\mathbf{v}_i \in W \ \forall\, i$. Then $\exists x \in \operatorname{span}(\mathbf{v}_i) \backslash W \quad x = \sum \lambda_i\mathbf{v}_i \in W$ which is a contradiction.

■

Definition 3.2 (2.17)
We say a list $\mathbf{v}_i$ of vectors is linearly independent if

$$0 = \lambda_1\mathbf{v}_1 + \cdots + \lambda_m\mathbf{v}_m \Rightarrow \forall\, \lambda_i = 0.$$

Theorem 3.2 (Lemma 20)
A list $\mathbf{v}_i$ is linearly independent iff every vector in $\operatorname{span}(\mathbf{v}_i)$ has a unique representation as a linear combination.

Proof 3.2
$\Rightarrow$(direct proof) Assume that $\sum \lambda_i \mathbf{v}_i = \sum \mu_i \mathbf{v}_i$ then

$$\sum(\lambda_i - \mu_i)\mathbf{v}_i = 0 \Rightarrow \lambda_i - \mu_i = 0 \Rightarrow \lambda_i = \mu_i$$

because $\mathbf{v}_i$ is linear independent.

$\blacksquare$

Remark:

1. If a list $\mathbf{v}_i$ is linearly dependant then there exist λ_i not all zero, such that $\sum \lambda_i \mathbf{v}_i = 0$
2. A single $\mathbf{v}$ is linearly dependant iff $\mathbf{v} = 0$. Because then $1\mathbf{v} = 1 \cdot 0 = 0$, note that $1 \in \mathcal{F}, v \in V, 0 \in V$.

Definition 3.3 (2.27)
Let V be an $\mathcal{F}$-VS. Then

1. A list $\mathbf{v}_i$ such that $V = \text{span}(\mathbf{v}_i)$ is called a generating set (spanning set). If the list is finite (always assumed here) then we say V is finitely generated.
2. A list $\mathbf{v}_i$ is called a basis for V if it is a linearly independent generating set.

Example 3.1

1. $$\lambda_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \lambda_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Rightarrow \lambda_1 = 0 = \lambda_2$$
2. Let $V = \mathcal{F}^n$ and denote by e_i the vector with a one in the i-th coordinate and zero elsewhere.
3. Let $V = \mathbb{R}[x]^{\leq m}(= \mathcal{P}_m(\mathbb{R}))$ then $1, x, x^2, \ldots, x^m$ are a basis with $m+1$ elements.

The $e_1, \ldots e_n$ are the so-called standard basis vectors.

Theorem 3.3 (Lemma)
Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be linearly dependent. Then $\exists j$ such that $\mathbf{v}_j \in \text{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1})$ and $\mathbf{v}_1, \ldots, \mathbf{v}_m$ without $\mathbf{v}_j$ spans the same space.

Proof 3.3
Since $\mathbf{v}_i$ is linearly dependent $\exists \lambda_i \in \mathcal{F}$, not all zero such that $\sum_{i=1}^m \lambda_i \mathbf{v}_i = 0$. Let j be the max index such that $\lambda_j \neq 0$. Then

$$\sum_{i=1}^{j-1} -\frac{\lambda_i}{\lambda_j}\mathbf{v}_i = \mathbf{v}_j \Rightarrow \mathbf{v}_j \in \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{j-1}). \tag{1}$$

Let $\sum_{i=1}^m \mu_i \mathbf{v}_i \in \operatorname{span}(\mathbf{v}_i)$. Substitute equation (1) for $\mathbf{v}_j$

$$\begin{aligned}\mu_1\mathbf{v}_1 + \cdots + \mu_j\mathbf{v}_j + \cdots + \mu_m\mathbf{v}_m &= \mu_1\mathbf{v}_1 + \cdots + \mu_j\left(\sum_{i=1}^{j-1} -\frac{\lambda_i}{\lambda_j}\mathbf{v}_i\right) + \cdots + \mu_m\mathbf{v}_m \\ &= \left(\mu_1 - \frac{\mu_j\lambda_1}{\lambda_j}\right)\mathbf{v}_1 + \left(\mu_2 - \frac{\mu_j\lambda_2}{\lambda_j}\right)\mathbf{v}_2 + \cdots + \left(\mu_{j-1} - \frac{\mu_j\lambda_{j-1}}{\lambda_j}\right)\mathbf{v}_{j-1} + \mu_{j+}\end{aligned}$$

Theorem 3.4 (Steinitz)
Let V be a finitely generated VS. Then the length of any linear independent list is smaller or equal to the length of any generating list.

Proof 3.4
Let $\mathbf{u}_1, \ldots, \mathbf{u}_m$ be linearly independent and $\mathbf{w}_1, \ldots, \mathbf{w}_n$ a generating set.

$$\operatorname{span}(\mathbf{w}_j) = V, \mathbf{u}_1 \in V.$$

Then $(\mathbf{u}_1\mathbf{w}_1, \ldots, \mathbf{u}_1\mathbf{w}_m)$ is linearly dependent. Then for $\sum \lambda_j \mathbf{w}_j = \mathbf{u}_1 \quad wlog \quad \lambda_1 \neq 0 \Rightarrow \frac{1}{\lambda_1}\mathbf{u}_1 - \frac{\lambda_2}{\lambda_1}\mathbf{w}_2 - \frac{\lambda_m}{\lambda_1} = \mathbf{w}_1$ (without loss of generality) point being

$$\operatorname{span}(\mathbf{u}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n) = \operatorname{span}(\mathbf{w}_1, \ldots, \mathbf{w}_n) = V.$$

The new list $S_1 = (\mathbf{u}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n)$ also spans V. Then $(\mathbf{u}_1, \mathbf{u}_2, \mathbf{w}_2, \ldots, \mathbf{w}_n)$ and

$$\mathbf{u}_2 = \lambda_1\mathbf{u}_1 + \lambda_2\mathbf{w}_2 + \cdots + \lambda_n\mathbf{w}_n$$

assume $\lambda_2 \neq 0$ and thus an element $\mathbf{w}_2$ can be pulled out of the set without loss:

$$\Rightarrow S_2 = (\mathbf{u}_1, \mathbf{u}_2, \mathbf{w}_3, \ldots, \mathbf{w}_n)$$

also spans V and we can keep going.

Remark: This shows that no list that is bigger than a generating set can be linearly independent. Also any list that is shorter than a linearly independent list can not generate the whole space.

Theorem 3.5 (Basis)
A list of vectors is a basis for V iff every $\mathbf{v} \in V$ can be uniquely be written as a linear combination.

Proof 3.5
Lemma 20. If you can write every element uniquely then you can write zero uniquely.

Theorem 3.6
Let $\operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_n) = V$. Then there is a subset of $\mathbf{v}_i$ that is a basis.

Proof 3.6
We construct the basis in n-steps.

We add a vector $\mathbf{v}_i$ to our basis if $\mathbf{v}_i \notin \operatorname{span}(\mathbf{v}_1, \ldots, \mathbf{v}_{i-1})$. Let $\mathbf{w}_1, \ldots, \mathbf{w}_m$ be the basis acquired this way. Assume

$$\sum \lambda_i \mathbf{w}_i = 0.$$

Let j be max such that $\lambda_j \neq 0$ then $\sum_{i=1}^{j-1} \lambda_i \mathbf{w}_i = \lambda_j \mathbf{w}_j$, contradiction.

Therefore $\mathbf{w}_i$ is linearly independent and it still spans V.

Theorem 3.7 (Corollary)
Every finitely generated VS has a basis.

Theorem 3.8 (Corollary)
Every linearly independent set can be extended to a basis.

Proof 3.7
Let $(\mathbf{u}_1, \ldots, \mathbf{u}_m)$ be linearly independent and let $(\mathbf{w}_1, \ldots, \mathbf{w}_n)$ be a generated set. Then $(\mathbf{u}_1, \ldots, \mathbf{u}_m, \mathbf{w}_1, \ldots, \mathbf{w}_n)$ is a geenerating set. Use Theorem 3.6 to acquire a basis.

Theorem 3.9 (2.35)
Every basis of a finitely generated VS has the same length.

Proof 3.8
Let B_1 and B_2 be two bases. Since B_1 is linearly independent and B_2 generates V.

$$\begin{aligned} |B_1| &\leq |B_2| \\ |B_2| &\leq |B_1| \\ \Rightarrow |B_1| &= |B_2| \end{aligned}$$

Definition 3.4 (Dimension)
Let V be an $\mathcal{F}$-VS. Then we define dimension as

$$\dim_{\mathcal{F}}(V) = \begin{cases} \text{length of the basis if } V \text{ is finitely generated} \\ \infty \quad \text{otherwise} \end{cases}$$

Theorem 3.10 (Corollary)
Let $U \subseteq V$ be a subspace. Then $\dim(U) \leq \dim(V)$.

Proof 3.9
A basis of U is a linear set in V. Hence it is shorter or equal in length to any generating set of V, especially a basis of V.

Theorem 3.11 (Corollary 2.39)
A linearly independent list of size $\dim(V)$ is already a basis.

Proof 3.10
We can extend the list to a basis. But it is already of length $\dim(V)$ hence nothing is added.

Theorem 3.12 (Corollary 2.42)
Let $\dim(V) = n$ then every generating set of length n is already a basis.

Two sets A, B with size $|A|, |B|$. The union has size: $|A \cup B| = |A| + |B| - |A \cap B|$

Theorem 3.13
Let A, B be subspaces of a finite dimensional space V. Then $\dim(A+B) = \dim(A) + \dim(B) - \dim(A \cap B)$.

Proof 3.11
Let $c_1, \ldots, c_l$ be a basis for $A \cap B$. We extend to a basis $c_1, \ldots, c_l, a_1, \ldots, a_m$ of A and to a basis $c_1, \ldots, c_l, b_1, \ldots, b_n$ of B.

We want to show that $c_i a_j b_k$ is a basis for $A + B$. This is a generating set, now we need to check that it is linearly independent.

Now let

$$\begin{aligned}
0 &= \sum \alpha_i a_i + \sum \beta_j b_j + \sum \mu_k c_k \\
-\sum \alpha_i a_i &= \sum \beta_j b_j + \sum \mu_k c_k \in A \cap B \\
-\sum \alpha_i a_i &= \sum \delta_k c_k \\
0 &= \sum \alpha_i a_i + \sum \delta_k c_k \\
\Rightarrow \alpha_i &= 0 \quad \delta_k = 0 \\
\Rightarrow 0 &= \sum (\beta_j b_j + \sum \gamma_k c_k) \\
\Rightarrow \beta_j &= 0 \quad \gamma_k = 0
\end{aligned}$$

4 Maps

Definition 4.1 (3.2/3.8)
Let $V,\ W$ be two $\mathcal{F}$-VS. A map $T : V \to W$ is called linear if

$$T(\mathbf{v}+\mathbf{v}') = T(\mathbf{v}) + T(\mathbf{v}') \quad \forall \mathbf{v}, \mathbf{v}' \in V$$
$$T(\lambda\mathbf{v}) = \lambda T(\mathbf{v}) \quad \forall \lambda \in \mathcal{F}\ \mathbf{v} \in V.$$

The set of all linear maps from V into W is denoted $\mathrm{Hom}_{\mathcal{F}}(V,\ W)$ meaning homomorphism (in the book: $\mathcal{L}(V,\ W)$). If $V = W$ we also write $\mathrm{End}_{\mathcal{F}}(\mathbf{v}) = \mathrm{Hom}(V,\ V)$.

Example 4.1

$$0 \in \mathrm{Hom}(V,\ W) \quad 0 \in \mathcal{F}\mathbf{v} = 0 \in W$$

Another example is the identity (id):

$$id \in \mathrm{End}(V) \quad id\ \mathbf{v} = \mathbf{v}.$$

Differentiating a polynomial is a linear map. The same applies to integration.
Multiplication by x^2 is a linear map in $\mathrm{Hom}(\mathbb{R}[x],\ \mathbb{R}[x])$.
Most commonly though:

$$T(x,\ y,\ z) = (2x - y, 3y + z)$$

Theorem 4.1
Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis for V and $\mathbf{w}_1, \ldots, \mathbf{w}_n$ some vectors in W. Then there exists a unique linear map T such that $T(\mathbf{v}_i) = \mathbf{w}_i$.

Proof 4.1
We show uniqueness and existence by explicitly calculating images of T. Let $\mathbf{v} \in V$ then exists unique $\lambda \in \mathcal{F}$ such that

$$\mathbf{v} = \sum \lambda_i \mathbf{v}_i.$$

Now

$$T(\mathbf{v}) = T\Big(\sum \lambda_i \mathbf{v}_i\Big) = \sum T(\lambda_i \mathbf{v}_i) = \sum \lambda_i T(\mathbf{v}_i)$$

Theorem 4.2 (Proposition 3.7)
$\mathrm{Hom}(V, W)$ is itself a $\mathcal{F}$-VS with usual addition and scalar multiplication

$$\forall S, T \in \mathrm{Hom}(V, W)$$
$$(S + T)(\mathbf{v}) = S(\mathbf{v}) + T(\mathbf{v})$$
$$\forall \lambda \in \mathcal{F} \quad (\lambda \cdot T)(\mathbf{v}) = \lambda(T(\mathbf{v}))$$

Proof 4.2

1. $0 \in \text{Hom}(V, W)$
2. $S, T \in \text{Hom}(V, W) \Rightarrow S + T \in \text{Hom}(V, W)$
3. $T \in \text{Hom} \Rightarrow \lambda T \in \text{Hom}$

Definition 4.2 (3.8)
Let $T \in \text{Hom}(U, V)$ and $S \in \text{Hom}(V, W)$. Then we define $ST \in \text{Hom}(U, W)$ $(U \underbrace{\rightarrow}_{T} V \underbrace{\rightarrow}_{S} W)$. As $ST(\mathbf{u}) = S(T(\mathbf{u})) = S \circ T(\mathbf{u})$. We see that for three suitable maps

$$\begin{aligned} (ST)U &= S(TU) \\ id\,T &= T\,id = T \\ (S+T)U &= SU + TU \\ S(T+U) &= ST + SU \end{aligned}$$

Note! Composition of linear maps is not commutative: $T, D \in \text{End}(\mathbb{R}[x])$

$$T(p) = x^2 p \quad D(p) = p' \quad TD(p) = x^2 p' \quad DT(p) = x^2 p' + 2xp.$$

Definition 4.3 (3.12 / 3.17)
Let $T \in \text{Hom}(V, W)$. We define the image(range) of T as $\text{im}(T) = \{T\mathbf{v} : \mathbf{v} \in V\} \subseteq W$ and its kernal(nullspace) as $\ker(T) = \{\mathbf{v} \in V : T\mathbf{v} = 0\} \subseteq V$.

Theorem 4.3
Image and kernel are subspaces.

Proof 4.3
We start with the image:

1. $0 \in \text{im}(T) \quad T(0) = T(0+0) = T(0) + T(0) \Rightarrow T(0) = 0$ and bonus $T(0) = T(0 \cdot \mathbf{v}) = 0 \cdot T(\mathbf{v}) = 0$
2. $\mathbf{w}, \mathbf{w}' \in \text{im}(T) \Rightarrow T(\mathbf{v}) = \mathbf{w},\ T(\mathbf{v}') = \mathbf{w}'$

$$\mathbf{w} + \mathbf{w}' = T(\mathbf{v}) + T(\mathbf{v}') = T(\mathbf{v} + \mathbf{v}')$$

3. $\mathbf{w} \in \mathrm{im}(T) \quad \lambda \in \mathcal{F}$
$$T(\lambda \mathbf{v}) = \lambda T(\mathbf{v}) = \lambda \mathbf{w}$$

Now the kernal:

1. By (*) $0 \in \ker$
2. $\mathbf{v}, \mathbf{v}' \in \ker \quad T(\mathbf{v} + \mathbf{v}') = T(\mathbf{v}) + T(\mathbf{v}') = 0 + 0 = 0$
3. $\mathbf{v} \in \ker \quad T(\lambda \mathbf{v}) = \lambda T(\mathbf{v}) = \lambda 0 = 0$

Definition 4.4
Let $f : A \to B$. We say f is injective if $f(a) = f(b) \Rightarrow a = b$ and surjective if $\forall b \; \exists a$ such that $f(a) = b$

Theorem 4.4 (Proposition)
$T \in \mathrm{Hom}(V, U)$ is injective iff $\ker(T) = \{0\}$ and surjective if $\mathrm{im}(T) = W$.

Proof 4.4
Injective: $\Rightarrow$ proof. Assume T is injective. Let $\mathbf{v} \in \ker(T)$ then
$$T(\mathbf{v}) = 0 = T(0) \Rightarrow \mathbf{v} = 0$$
by injectivity.

$\Leftarrow$ proof. Assume $\ker(T) = \{0\}$ and
$$T(a) = T(b) \Rightarrow T(a) - T(b) = 0 \Rightarrow T(a - b) = 0 \Rightarrow a - b = 0 \Rightarrow a = b$$

Surjective is automatically done as it literally means it is the whole thing.

Theorem 4.5
Let V be a finite dimensional VS and $T \in \mathrm{Hom}(V, W)$. Then $\mathrm{im}(T)$ is also finite dimensional and
$$\dim(V) = \dim(\ker(T)) + \dim(\mathrm{im}(T))$$

Proof 4.5
Let $\mathbf{u}_1, \dots, \mathbf{u}_m$ be a basis of $\ker(T)$ which is a subspace of V and we can extend this to a basis of V by adding $\mathbf{v}_1, \dots, \mathbf{v}_n$. The $\dim(V) = m + n \quad \dim(\ker(T)) = m$. Need to show that $\dim(\mathrm{im}(T)) = n$.

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

$$\mathbf{v} = \sum \lambda_i \mathbf{u}_i + \lambda \mu_j \mathbf{v}_j$$

and

$$T(\mathbf{v}) = T\Big(\sum \lambda_i \mathbf{u}_i + \sum \mu_j \mathbf{v}_j\Big) = T\Big(\sum \mu_j \mathbf{v}_j\Big) = \sum \mu_j T(\mathbf{v}_j)$$

which implies ($\Rightarrow$) the set of vectors $T(\mathbf{v}_j)$ generates/spans the image of T. Now we need to show that they are linear independent.

Assume

$$\sum \alpha_j T(\mathbf{v}_j) = 0$$

if they are linear independent then all $\alpha_j = 0$:

$$\begin{aligned} \sum T(\alpha \mathbf{v}_j) &= T\left(\underbrace{\sum \alpha_j \mathbf{v}_j}_{\in \ker(T)}\right) \\ \sum \alpha_j \mathbf{v}_j &= \sum \beta_i \mathbf{u}_i \\ \sum \alpha_j \mathbf{v}_j + \sum (-\beta_i) \mathbf{u}_i &= 0 \\ &\Rightarrow \alpha_j = 0 \end{aligned}$$

because $\mathbf{u}_1, \ldots, \mathbf{u}_m, \mathbf{v}_1, \ldots, \mathbf{v}_n$ is a basis and hence linear independent.

■

Theorem 4.6 (Corollary)
If $\dim(V) > \dim(W)$ then no $T \in \mathrm{Hom}(V, W)$ is injective.

Proof 4.6

$$\begin{aligned} \dim(V) &= \dim(\ker(T)) + \dim(\mathrm{im}(T)) \\ \Rightarrow \dim(V) - \dim(\mathrm{im}(T)) &= \dim(\ker(T)) \\ \mathrm{im}(T) &\leq W \\ \dim(\mathrm{im}(T)) &\leq \dim(W) < \dim(V) \\ &\Rightarrow 1 \leq \dim(\ker(T)) \\ &\Rightarrow \ker(T) \neq \{0\} \end{aligned}$$

which implies T is not injective.

Theorem 4.7 (Corollary)
If $\dim(V) < \dim(W)$ no $T \in \mathrm{Hom}(V, W)$ is surjective.

Proof 4.7

$$\begin{aligned}\dim(V) &= \dim(\ker(T)) + \dim(\operatorname{im}(T))\\ \dim(W) > \dim(V) &\geq \dim(\operatorname{im}(T))\\ \operatorname{im}(T) &\subsetneq W\end{aligned}$$

Definition 4.5
Let $T \in \operatorname{Hom}(V, W)$ and $\mathbf{v}_1, \ldots, \mathbf{v}_m$ basis of V and $\mathbf{w}_1, \ldots, \mathbf{w}_n$ basis of W. Then the matrix of T with respect to these bases is given by the entries A_{jk} defined by

$$T\mathbf{v}_k = \sum A_{jk}\mathbf{w}_j.$$

$A = (A_{jk}) = \mathcal{M}(T)$.

Example 4.2
$\mathbb{R}[x]^{<4}$ with the differentiation mapping, $D \in \operatorname{Hom}(\mathbb{R}[x]^{<4}, \mathbb{R}[x]^{<3})$, the basis of $\mathbb{R}[x]^{<4}$ is $1, x, x^2, x^3$ and for $\mathbb{R}[x]^{<3}$ it is $1, x, x^2$. We get the entries of the matrix by

$$\begin{aligned}D(1) &= 0\\ D(x) &= 1 = 1 \cdot 1 + 0x + 0x^2\\ D(x^2) &= 2x = 0 \cdot 1 + 2 \cdot x + 0x^2\\ D(x^3) &= 3x^2 = 0 \cdot 1 + 0 \cdot x + 3 \cdot x^2\\ &\begin{pmatrix}0 & 1 & 0 & 0\\ 0 & 0 & 2 & 0\\ 0 & 0 & 0 & 3\end{pmatrix}\end{aligned}$$

Another example changes the first basis to $1 + x, x + x^2, x^2 + x^3, x^3$, and the entries of the matrix are

$$\begin{aligned}D(1 + x) &= 1\\ D(x + x^2) &= 1 + 2x\\ D(x^2 + x^3) &= 2x + 3x^2\\ D(x^3) &= 3x^2\\ &\begin{pmatrix}1 & 1 & 0 & 0\\ 0 & 2 & 2 & 0\\ 0 & 0 & 3 & 3\end{pmatrix}\end{aligned}$$

With two linear maps T, S such that $S \circ T$ makes sense, then $\mathbf{u} \to \mathbf{v} \to \mathbf{w}$ and the matrix of the combined bases is $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$.

5 Invertibility and Isomorphisms

Definition 5.1
Let $T \in \mathrm{Hom}(V, W)$, then we say T is invertible if there exists an $S \in \mathrm{Hom}(W, V)$ such that $ST = \mathrm{id}_V$ and $TS = \mathrm{id}_W$. We call S the inverse of T.

Theorem 5.1 (Proposition)
The inverse of an invertible map is unique.

Proof 5.1
Suppose $T \in \mathrm{Hom}(V, W)$ is invertible and S and S' are both inverses. then

$$S = S\mathrm{id}_W = S(TS') = (ST)S' = \mathrm{id}_V S' = S'.$$

We decide $T^{-1} = S$ from now on.

Theorem 5.2
A linear map T is invertible iff it is injective and surjective.

Proof 5.2
Direct proof "$\Rightarrow$".

We want to show it is injective: Assume $T\mathbf{v} = T\mathbf{v}'$. Since it is invertible it has an inverse $T^{-1}T\mathbf{v} = T^{-1}T\mathbf{v}'$ and thus $\mathbf{v} = \mathbf{v}'$.

To show it is surjective we have $\mathbf{w} \in W$ and $T^{-1}\mathbf{w}$ is a method to get it back into V. We do this by $TT^{-1}\mathbf{w} = \mathbf{w}$.

Now indirect proof "$\Leftarrow$".

We construct inverse $S : W \to V$ by defining $S\mathbf{w} = \mathbf{v}$ where $T\mathbf{v} = \mathbf{w}$. This $\mathbf{v}$ exosts because T is surjective and $\mathbf{v}$ is unique because T is injective. Obviously $TS = \mathrm{id}_W$. Consider

$$T(ST) = (TS)T = T$$

and now we want to show that ST is the identity of V:

$$T(ST)\mathbf{v} = (TS)T\mathbf{v} = T\mathbf{v} \Rightarrow ST = \mathrm{id}_V$$

because T is injective.

Need to check that it is closed under addition and multiplication for it to be linear:

$$TS(x + y) = x + y = TSx + TSy = T(Sx + Sy)$$

By injectivity of T we have that $S(x + y) = Sx + Sy$. Now for multiplication:

$$TS(\lambda x) = \lambda x = \lambda TSx = T(\lambda Sx)$$

and we are good. Thus it is linear.

Definition 5.2
We say two VS are isomorphic if there exists an invertible linear map $T : V \to W$. We write $V \cong W$ and call T a isomorphism.

Theorem 5.3
Any two finite dimensional $\mathcal{F}$-VSs are isomprhic iff they have the same dimension.

Proof 5.3
We have seen that maps between VSs of different dimensions are either not injective or not surjective. Therefore if they are $V \cong W \Rightarrow \dim(V) = \dim(W)$. Now assume $\dim(V) = \dim(W) = n$ and let $\mathbf{v}_1, \dots, \mathbf{v}_n$ and $\mathbf{w}_1, \dots, \mathbf{w}_n$ be bases for V and W respectively. Define $T : V \to W$ by $T(\mathbf{v}) = T(\sum \lambda_i \mathbf{v}_i) = \sum \lambda_i T(\mathbf{v}_i) = \sum \lambda_i \mathbf{w}_i$, then T is an isomorphism. Let $T(\mathbf{v}) = 0$, thus $\mathbf{v} \in \ker(T)$:

$$T(\mathbf{v}) = T(\sum \lambda_i \mathbf{v}_i) = \sum \lambda_i \mathbf{w}_i = 0 \Rightarrow \lambda_i = 0 \ \forall i$$

because the $\mathbf{w}_i$ are linearly independent as they are a basis. This implies $\mathbf{v} = 0 \Rightarrow \ker(T) = \{0\}$.

Say $\mathbf{w} = \sum \mu_i \mathbf{w}_i \in W$. Then

$$T(\sum \mu_i \mathbf{v}_i) = \mathbf{w}$$

$\Rightarrow T$ is surjective.

After fixing bases for V and W we have a map $\mathcal{M} : \mathrm{Hom}_{\mathcal{F}}(V, W) \to \mathcal{F}^{m \times n}$.
One checks that $\mathcal{M}$ is indeed linear.

Theorem 5.4
The map $\mathcal{M}$ is an isomorphism.

Proof 5.4
Need to show it is injective and surjective. We start with showing it is injective:

$$\mathcal{M}(T) = 0$$

each column represents a basis vector of V, and if these are all 0 then $T(\mathbf{v}_i) = 0 \ \forall i$ where $\mathbf{v}_i$ is a basis. Thus $T\mathbf{v} = 0 \ \forall \mathbf{v} \in V$ and thus $T = 0$ is the linear map that maps all vectors to the zerovector. Injectivity is then shown.

Now to show surjectivity we have $A \in \mathcal{F}^{m \times n}$ then we define T such that

$$T\mathbf{v}_k = \sum_{j=1}^{m} A_{jk} \mathbf{w}_j$$

and it follows that $\mathcal{M}(T) = A$.

Theorem 5.5 (Corollary)

$$\dim(\mathrm{Hom}_{\mathcal{F}}(V, W)) = \dim(V) + \dim(W)$$

Proof 5.5
$\mathcal{F}^{m\times n}$ with $E_{i,j}$ which has zeros everywhere except row i and column j where there is a 1. These are a basis.

Definition 5.3

$$\mathrm{End}(V) = \mathrm{Hom}(V, V)$$

is the set of linear maps from V into V, called the endomorphisms.

Theorem 5.6
Let V be a finite dimensional VS and $T \in \mathrm{End}(V)$. Then the following statements are equivalent:

1. T is injective.
2. T is surjective.
3. T is invertible.

Proof 5.6
$\Leftrightarrow$ proof:
The kernal of T is just zero, this implies that $\dim(V) = \dim(\mathrm{im}(T)) + \dim(\ker(T))$ but the dimension of the kernal is zero. Which can only happen if $V = \mathrm{im}(T)$.

No need to check for injectivity and surjectivity if it maps only to zero.

Definition 5.4
Let $V_1, \ldots, V_m$ be $\mathcal{F}$-VS then we define a new VS as

$$V_1 \cdot \cdots \cdot V_m = \{(\mathbf{v}_1, \ldots, \mathbf{v}_m) : \mathbf{v}_i \in V_i\}$$

with obvious addition and scalar multiplication.

Theorem 5.7

$$\dim(V_1 \cdot \cdots \cdot V_m) = \dim(V_1) + \cdots + \dim(V_m).$$

Proof 5.7
We provide a basis. Choose a basis for each component V_i and for each basis vector. Consider the element in $V_1 \cdot \cdots \cdot V_m$ that has $\mathbf{v}_k$ in the ith position and zeros elsewhere. That forms a basis.

Theorem 5.8 (Proposition)
Let $U_1, \ldots, U_m$ are subspaces of V. Define the map $\Gamma : U_1 \cdot \cdots \cdot U_m \to U_1 + \cdots + U_m$. Then $U_1 + \cdots + U_m$ is defined if Γ is injective.

Proof 5.8
Γ is injective: $0 = \Gamma(\mathbf{u}_1, \ldots, \mathbf{u}_m) \Rightarrow \mathbf{v} = (\mathbf{u}_1, \ldots, \mathbf{u}_m) = (0, \ldots, 0)$ which is the same as saying $0 = \sum \mathbf{u}_i$. If the sum is direct then the sum implies $\mathbf{u}_i = 0 \ \forall i$.

Theorem 5.9 (Corollary)
Let $U_1, \ldots, U_m$ be subspaces of a finite dimensional VS V. Then $U_1 + \cdots + U_m$ is direct iff

$$\dim(U_1 + \cdots + U_m) = \dim(U_1) + \cdots + \dim(U_m).$$

Proof 5.9
Γ is obviously surjective. By Theorem 5.6 Γ is an isomorphism iff the dimensions match. Hence if the sum is direct Γ is injective by the Theorem 5.8 and have isomorphism, therefore the dimensions match.

Definition 5.5
Let $U \subset V$ and $\mathbf{v} \in V$ then the set $\mathbf{v} + U := \{\mathbf{v} + \mathbf{u} \ \mathbf{u} \in U\}$ is called an affine subset.

Theorem 5.10 (Lemma)
Let $U \subset V$ then the following are equivalent:

1. $\mathbf{v} - \mathbf{w} \in U$
2. $\mathbf{v} + U = \mathbf{w} + U$
3. $\mathbf{v} + U \cup \mathbf{w} + U \neq \emptyset$

Definition 5.6
Let $U \subset V$ and consider the set of all affine subsets:

$$V/U := \{\mathbf{v} + U : \mathbf{v} \in V\}.$$

This is called the quotient space.

Theorem 5.11
V/U is a VS with additivity given by $(\mathbf{v}+U)+(\mathbf{w}+U) = (\mathbf{v}+\mathbf{w})+U$ and multiplication given by $\lambda(\mathbf{v} + U) = \lambda\mathbf{v} + U$.

Proof 5.10
We need to check the this:
Let $\mathbf{v} + U = \mathbf{v}' + U$ and $\mathbf{w} + U = \mathbf{w}' + U$ then we want

$$(\mathbf{v} + \mathbf{w}) + U = (\mathbf{v}' + \mathbf{w}') + U$$
$$(\mathbf{v} + \mathbf{w}) - (\mathbf{v}' + \mathbf{w}') = (\mathbf{v} - \mathbf{v}') + (\mathbf{w} - \mathbf{w}') \in U$$

since $\mathbf{v} - \mathbf{v}' \in U$ and the same applies to the $\mathbf{w}$'s.

Theorem 5.12

$$\dim(V/U) = \dim(V) - \dim(U).$$

Proof 5.11
Consider the projection map $\pi : V \to V/U$ by $\mathbf{v} \to \mathbf{v} + U$. The kernal $\ker(\pi) = \{\mathbf{v} \in V : \pi(\mathbf{v}) = 0 + U\}$ which is everything in U that is mapped to zero, which is just U. The image is $\mathrm{im}(\pi) = V/U$, and the dimension follow

$$\dim(V) = \dim(V/U) + \dim(U) \Rightarrow \dim(V/U) = \dim(V) - \dim(U).$$

6 Eigenvalues and Eigenspaces

Definition 6.1 (5.5/5.7/5.34)
Let V be an $\mathcal{F}$-VS and $T \in \text{End}(V)$. An element $\lambda \in \mathcal{F}$ is called an eigenvalue if there is a vector $\mathbf{v} \in V$, $\mathbf{v} \neq 0$, such that $T\mathbf{v} = \lambda\mathbf{v}$. Then $\mathbf{v}$ is called an eigenvector for λ, and the subspace $E(\lambda, T) = \ker(T - \lambda I)$ is called the eigenspace.

Example 6.1

- Let $T \in \text{End}(V)$ with $\ker(T) \neq \{0\}$ then 0 is an eigenvalue.
- Let $A = \begin{pmatrix} 1 & -1 \\ -2 & 0 \end{pmatrix}$, $A : \mathbb{R}^2 \to \mathbb{R}^2$. We have

$$A \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \end{pmatrix} = -1 \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \Rightarrow \text{-1 is an eigenvalue.}$$
$$A \cdot \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 2 \end{pmatrix} = 2 \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \Rightarrow \text{2 is an eigenvalue.}$$

Theorem 6.1 (Corollary 5.6)
Let $T \in \text{End}(V)$, $\lambda \in \mathcal{F}$ then the following are equivalent:

1. λ is an eigenvalue.
2. $\ker(T - \lambda I) \neq \{0\} \Leftrightarrow T - \lambda I$ is not injective.

If V is finite dimensional then the following are also equivalent to the above:

1. $T - \lambda I$ is not surjective.
2. $T - \lambda I$ is not invertible.

Proof 6.1

$$T\mathbf{v} = \lambda\mathbf{v} \Leftrightarrow T\mathbf{v} - \lambda\mathbf{v} = (T - \lambda I)\mathbf{v} = 0.$$

Theorem 6.2 (5.10)
Let $T \in \text{End}(V)$ and suppose $\lambda_1, \dots, \lambda_n$ are distinct eigenvalues, with corresponding eigenvectors $\mathbf{v}_1, \dots, \mathbf{v}_n$. Then $\mathbf{v}_1, \dots, \mathbf{v}_n$ are linearly independent.

Proof 6.2
Assume $\exists \mu_i$ such that $\sum_{i=1}^{n} \mu_i \mathbf{v}_i = 0$ and not all $\mu_i = 0$. Wlog let us assume that $\mu_n \neq 0$, then

$$0 = T(\sum \mu_i \mathbf{v}_i) = \sum \mu_i T\mathbf{v}_i = \sum \mu_i \lambda_i \mathbf{v}_i.$$

Subtract $\lambda_n \cdot \sum \mu_i \mathbf{v}_i$ for this

$$0 = \sum \mu_i \lambda_i \mathbf{v}_i - \sum \mu_i \lambda_n \mathbf{v}_i = \sum_{i=1}^{n-1} \mu_i (\lambda_i - \lambda_n) \mathbf{v}_i.$$

Repeating this procedure shows that all $\mu_i = 0$. Hence $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is linearly independent.

Theorem 6.3 (Corollary 5.13)
Let V be finite dimensional, then any endomorphism has at most $\dim(V)$ distinct eigenvalues.

Remark: $T \in \mathrm{End}(V)$ then $T^2 = T \circ T$ and $T^3 = T \circ T^2$. Also $T^\circ = \mathrm{id}$, if T is invertible then $T^{-m} = (T^{-1})^m$. Now we can build polynomials: Let $p(x) \in \mathcal{F}[x]$, $p(T) = c_0 \mathrm{id} + c_1 T + c_2 T^2 + \cdots + c_m T^m$.

Theorem 6.4 (5.21)
Let V be a finite dimensional $\mathbb{C}$-VS then every $T \in \mathrm{End}(V)$ has an eigenvalue.

Proof 6.3
Let $\mathbf{v} \in V$, $\mathbf{v} \neq 0$, then the following list $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \ldots, T^n\mathbf{v}$, where $n = \dim(V)$, is linearly dependent (there is $n+1$ vectors). This implies $\exists c_i \in \mathbb{C}$ such that

$$\sum c_i T^i \mathbf{v} = 0$$

not all $c_i = 0$. Consider

$$p(x) = \sum_{i=0}^{n} c_i x^i = c \prod_{i=1}^{n} (x - \lambda_i).$$

Now we take

$$p(T) = c \cdot \prod_{i=1}^{n} (T - \lambda_i I)$$
$$p(T)\mathbf{v} = c \cdot \prod_{i=1}^{n} (T - \lambda_i I)\mathbf{v} = 0.$$

Somewhere along the way a vector is mapped to zero. Hence at least one of $(T - \lambda_i I)$ is not injective and therefore T has an eigenvalue.

Theorem 6.5 (5.26)
Let $T \in \text{End}(V)$ and $\mathbf{v}_1, \dots, \mathbf{v}_n$ is a basis of V. Then the following are equivalent:

1. The matrix $A = \mathcal{M}(T)$ is upper triangular: $A_{ij} = 0,\ i > j$.
2. $T\mathbf{v}_j \in \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_j)$.
3. $\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_j)$ is invariant under T for each j.

Proof 6.4

1. $$T\mathbf{v}_j = \sum_{i=1}^{n} A_{ij}\mathbf{v}_i = \sum_{i=1}^{j} A_{ij}\mathbf{v}_i$$
   as the image of $\mathbf{v}_j$ is a linear combination of the vectors $\mathbf{v}_1, \dots, \mathbf{v}_j$
2. (3) $\Rightarrow$ (2): $\mathbf{v}_j \in \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_j)$
   $$T\mathbf{v}_j \in \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_j)$$
3. (2) $\Rightarrow$ (3): $T\left(\sum_{i=1}^{j} \lambda_i \mathbf{v}_i\right) = \sum_{i=1}^{j} \lambda_i T(\mathbf{v}_i)$ which sits in $\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_i)$.

Theorem 6.6 (5.27)
Let V be a finite dimensional $\mathbb{C}$-VS. For every $T \in \text{End}V$ there exists a basis such that $\mathcal{M}(T)$ is upper triangular with respect to this basis.

With upper triangular matrices it is ensured that the first basis vector is an eigenvector.

Proof 6.5
We use induction on the dimension of V. For $\dim(V) = 1$ this is true.

Now assume that the theorem holds for all $\mathbb{C}$-VS of dimension lower than $\dim(V)$. T has at least one eigenvalue λ by theorem ?? and we consider the subspace $U = \text{im}(T - \lambda I)$, which is smaller than V. We see that U is invariant under T since

$$\mathbf{u} \in U, \quad T\mathbf{u} = T\mathbf{u} - \lambda\mathbf{u} + \lambda\mathbf{u} = \underbrace{(T - \lambda I)\mathbf{u}}_{\in U} + \underbrace{\lambda\mathbf{u}}_{\in U} \in U.$$

Hence we can consider the restriction of T onto U, $T|_U \in \text{End}U$. Since $\dim(U) < \dim(V)$ there exists a basis $\mathbf{u}_1, \dots, \mathbf{u}_n$ of U such that $T|_U$ is upper triangular. Extend the basis to a basis of V $\mathbf{u}_1, \dots, \mathbf{u}_n, \mathbf{v}_1, \dots, \mathbf{v}_m$. We see that

$$T\mathbf{v}_i = \underbrace{T\mathbf{v}_i - \lambda\mathbf{v}_i}_{\in U} + \lambda\mathbf{v}_i) \in \text{span}(\mathbf{u}_1, \dots, \mathbf{u}_m\mathbf{v}_i) \subseteq \text{span}(\mathbf{u}_1, \dots, \mathbf{u}_n, \mathbf{v}_1, \dots, \mathbf{v}_m).$$

And by Proposition ?? $\mathcal{M}(T)$ is upper triangular with respect to this basis.

Theorem 6.7 (5.30)
Suppose $T \in \text{End}V$ and there is a basis such that $\mathcal{M}(T)$ is upper triangular. Then T is invertible iff all diagonal elements are nonzero.

Proof 6.6
"$\Leftarrow$" proof: Let $\mathcal{M}(T) = A = \begin{pmatrix} \lambda_1 & & \\ & dots & \\ & & \lambda_n \end{pmatrix}$ and all $\lambda_i \neq 0$. We show that T is surjective. Since $T\mathbf{v}_1 = \lambda_1 \mathbf{v}_1 \Rightarrow T\frac{1}{\lambda_1}\mathbf{v}_1 = \mathbf{v}_1$ we see that $\mathbf{v}_1 \in \text{im}(T)$. Also

$$T\mathbf{v}_2 = A_{12}\mathbf{v}_1 + \lambda_2\mathbf{v}_2 \Rightarrow T\left(\mathbf{v}_2 - \frac{A_{12}}{\lambda_1}\mathbf{v}_1\right) = A_{12}\mathbf{v}_1 + \lambda_2\mathbf{v}_2 - A_{12}\mathbf{v}_1 = \lambda_2\mathbf{v}_2.$$

Continuing like this shows that $\mathbf{v}_i \in \text{im}(T)\ \forall i \Rightarrow T$ is surjective which implies that T is invertible.

"$\Rightarrow$" proof: Assume $\lambda_i = 0$ then the subspace $\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_i)$ is mapped by T onto $\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_{i-1})$. Hence T can not be injective as a bigger space is mapped into a smaller space (i space to $i-1$ space). $\exists \mathbf{v} \in \text{span}(\mathbf{v}_1, \dots, \mathbf{v}_i)$ such that $T\mathbf{v} = 0,\ \mathbf{v} \neq 0$.

Theorem 6.8 (5.32)
Let $\mathcal{M}(T)$ be upper triangular then the eigenvalues of T are the diagonal entries.

Proof 6.7
An element $\lambda \in \mathcal{F}$ is an eigenvalue iff $T - \lambda I$ is not invertible. The matrix $\mathcal{M}(T - \lambda I)$ has diagonal elements $(\lambda_i - \lambda)$ where λ_i are the diagonal entries of $\mathcal{M}(T)$. $T - \lambda I$ not invertible iff $\lambda_i - \lambda = 0$ for at least one $i \Leftrightarrow \lambda = \lambda_i$ for some i.

Theorem 6.9 (Proposition)
Let V be a finite dimensional VS and $T \in \text{End}V$. Then $E(\lambda_1, T) + \cdots + E(\lambda_m, T)$ is a direct sum where λ_i are distinct eigenvalues, and

$$\sum \dim(E(\lambda_m, T)) \leq \dim(V).$$

Proof 6.8
Assume there are $\mathbf{u}_i \in E(\lambda_i, T)$ such that not all $\mathbf{u}_i = 0$ and $\sum \mathbf{u}_i = 0$. Every $\mathbf{u}_i$ is an eigenvector to a different eigenvalue(or $\mathbf{u}_i) = 0$) but these are linearly independent. For the sum to be zero, all $\mathbf{u}_i$ must be zero. Hence

$$\dim(E(\lambda_1, T) \oplus \cdots \oplus E(\lambda_m, T)) = \sum \dim(E(\lambda_i, T)) \leq \dim(V).$$

Definition 6.2
We say $T \in \text{End}V$ is diagonalizable if there exists a basis such that $\mathcal{M}(T)$ is diagonal.

Theorem 6.10 (Proposition)
Let V be finite dimensional, $T \in \text{End}V$, and $\lambda_1, \dots, \lambda_m$ are the eigenvalues of T. Then the following are equivalent:

1. T is diagonalizable.
2. V has a basis of eigenvectors.
3. $\dim(V) = \sum \dim(E(\lambda_1, T))$.

Proof 6.9

1. Implies (2). Every vector is mapped to a multiple of itself. And (2) implies (1).
2. Implies (3). Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be a basis of eigenvectors. Each of $\mathbf{v}_i \in E(\lambda_j, T)$ for some j. This implies that
$$\sum \dim(E(\lambda_j, T)) \geq n = \dim(V) \Rightarrow \sum \dim(E(\lambda_j, T)) = \dim(V)$$
by Proposition 6.9.
3. Implies (2). Assume $\sum \dim(E) = \dim(V)$. Choose a basis for each $E(\lambda_j, T)$ then the union of these bases is a basis for V $\mathbf{v}_1, \dots, \mathbf{v}_n$. To show linearly independence we assume that they are not: Assume $\sum \mu_i \mathbf{v}_i = 0$ for not all $\mu_i = 0$. Rearrange the sum by corresponding eigenvalues
$$\sum_{i=1}^{n} \mu_i \mathbf{v}_i = \sum_{j=1}^{m} \mathbf{u}_j = 0,$$
where $\mathbf{u}_j \in E(\lambda_j)$. Each $\mathbf{u}_j$ is either an eigenvector to a distinct eigenvalue or 0. Since eigenvectors to distinct eigenvalues are linearly independent all $\mathbf{u}_j = 0 = \sum_{i \in s_j} \mu_i \mathbf{v}_i$, but $\mathbf{v}_i$ are basis for $E(\lambda_j) \Rightarrow \mu_i = 0$.

Theorem 6.11 (Lemma)
If λ is an eigenvalue then $\dim(E(\lambda, T)) \geq 1$.

Theorem 6.12 (Corollary 5.44)
If $T \in \text{End}V$ has $\dim(V)$ distinct eigenvalues, then T is diagonalizable.

7 Inner Product Spaces

We're going to talk about geometry now with length and angles of vectors.

Definition 7.1 (6.3 Inner Product)
Let V be an $\mathbb{R}$- or $\mathbb{C}$-VS, then a function $\langle\cdot|\cdot\rangle : V \times V \to \mathcal{F}$ is called an inner product if

1. $\langle \mathbf{u}|\mathbf{v}\rangle = \overline{\langle \mathbf{v}|\mathbf{u}\rangle}$
2. $\langle \lambda\mathbf{u} + \mu\mathbf{w}|\mathbf{v}\rangle = \lambda \langle \mathbf{u}|\mathbf{v}\rangle + \mu \langle \mathbf{w}|\mathbf{v}\rangle$
3. $\langle \mathbf{v}|\mathbf{v}\rangle \geq 0$ with equality iff $\mathbf{v} = 0$. Item one ensures that this one makes sense, as it only applies to the reals otherwise.

Example 7.1
The typical Euclidean Spaces: $\mathbb{R}^n, \mathbb{C}^n$:

$$\langle (x_1, \dots, x_n)|(y_1, \dots, y_n)\rangle = \sum_{i=1}^{n} x_i \bar{y}_i = \mathbf{x} \cdot \bar{\mathbf{y}}^\intercal$$

which can also be scaled by a scalar

$$\sum_{i=1}^{n} c_i x_i \bar{y}_i.$$

Another example is a VS of a continous function, real-valued $[-1, 1]$:

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

Theorem 7.1 (Proposition 6.7)
Let V be an inner product space (a VS with an inner product), then

1. $\forall \mathbf{u} \in V, \quad \varphi_{\mathbf{u}} : V \to \mathcal{F},\ \mathbf{v} \to \langle \mathbf{v}|\mathbf{u}\rangle, \quad \varphi_{\mathbf{u}} \in \mathrm{Hom}_{\mathcal{F}}(V, \mathcal{F})$
2. $\langle 0|\mathbf{u}\rangle = \langle \mathbf{u}|0\rangle = 0\ \forall \mathbf{u} \in V$
3. $\langle \mathbf{u}|\mathbf{v} + \mathbf{w}\rangle = \langle \mathbf{u}|\mathbf{v}\rangle + \langle \mathbf{u}|\mathbf{w}\rangle$
4. $\langle \mathbf{u}|\lambda\mathbf{v}\rangle = \bar{\lambda} \langle \mathbf{u}|\mathbf{v}\rangle\ \ \forall \lambda \in \mathcal{F}\ \mathbf{u}, \mathbf{v} \in V$

Definition 7.2
We say $\mathbf{u}$ and $\mathbf{v}$ are orthogonal, in symbols $\mathbf{u} \perp \mathbf{v}$, if $\langle \mathbf{u}|\mathbf{v}\rangle = 0$.

Remark: 0 is orthogonal to everything. Over $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ 0 is the only vector that is orthogonal to itself.

Definition 7.3 (Norm)
Let V be a VS then a function $\|\cdot\| \to \mathbb{R}_{\geq 0}$ is a norm if

1. $\|\mathbf{v}\| = 0$ iff $\mathbf{v} = 0$
2. $\|\lambda \mathbf{v}\| = |\lambda| \|\mathbf{v}\|, \quad \lambda \in \mathcal{F}$
3. $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$.

Definition 7.4 (6.8)
Let V be an inner product space then we can define a norm by $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v} | \mathbf{v} \rangle}$.

Proof 7.1
We will now prove the first two conditions of a norm (Definition 7.3)

1. $\|\mathbf{v}\| = 0 = \sqrt{\langle \mathbf{v} | \mathbf{v} \rangle}$, by condition 3 of inner product.
2. $\|\lambda \mathbf{v}\|^2 = \langle \lambda \mathbf{v} | \lambda \mathbf{v} \rangle = \lambda \bar{\lambda} \langle \mathbf{v} | \mathbf{v} \rangle = |\lambda|^2 \|\mathbf{v}\|^2$.
3. $\sqrt{\langle \mathbf{v} | \mathbf{v} \rangle}$ is a norm condition and we have

$$\begin{aligned}
\|\mathbf{u} + \mathbf{v}\| &\leq \|\mathbf{u}\| + \|\mathbf{v}\| \\
\|\mathbf{u} + \mathbf{v}\|^2 &= \langle \mathbf{u} + \mathbf{v} | \mathbf{u} + \mathbf{v} \rangle \\
&= \|\mathbf{u}\|^2 + \langle \mathbf{u} | \mathbf{v} \rangle + \langle \mathbf{u} \bar{|} \mathbf{v} \rangle + \|\mathbf{v}\|^2 \\
&= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\mathrm{Re}(\langle \mathbf{u} | \mathbf{v} \rangle) \leq \\
&\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2 | \langle \mathbf{u} | \mathbf{v} \rangle | \leq \\
\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| &= (\|\mathbf{u}\| + \|\mathbf{v}\|)^2 \\
\|\mathbf{u} + \mathbf{v}\| &\leq \|\mathbf{u}\| + \|\mathbf{v}\|.
\end{aligned}$$

Theorem 7.2 (Pythagorean Theorem)
Suppose $\mathbf{u} \perp \mathbf{v}$ then

$$\|\mathbf{u} + \mathbf{v}\|^2 = \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.$$

Proof 7.2

$$\begin{aligned}\|\mathbf{u}+\mathbf{v}\|^2 &= \langle \mathbf{u}+\mathbf{v}|\mathbf{u}+\mathbf{v}\rangle \\ &= \langle \mathbf{u}|\mathbf{u}\rangle + \langle \mathbf{u}|\mathbf{v}\rangle + \langle \mathbf{v}|\mathbf{u}\rangle + \langle \mathbf{v}|\mathbf{v}\rangle \\ &= \|\mathbf{u}\|^2 + \|\mathbf{v}\|^2.\end{aligned}$$

Theorem 7.3 (Lemma)
Let $\mathbf{u}, \mathbf{v} \in V,\ \mathbf{v} \neq 0$ we have

$$\begin{aligned} c &= \frac{\langle \mathbf{u}|\mathbf{v}\rangle}{\|\mathbf{v}\|^2} \\ \mathbf{w} &= \mathbf{u} - c\mathbf{v}, \end{aligned}$$

then $\mathbf{u} = \mathbf{w} + c\mathbf{v}$ and $\mathbf{w} \perp \mathbf{v}$.

Proof 7.3
Calculate their inner product to prove that they are orthogonal:

$$\begin{aligned}\langle \mathbf{v}|\mathbf{w}\rangle &= \left\langle \mathbf{v} \middle| \mathbf{u} - \frac{\langle \mathbf{u}|\mathbf{v}\rangle}{\|\mathbf{v}\|^2}\mathbf{v} \right\rangle \\ &= \langle \mathbf{v}|\mathbf{u}\rangle - \left\langle \mathbf{v} \middle| \frac{\langle \mathbf{u}|\mathbf{v}\rangle}{\|\mathbf{v}\|^2}\mathbf{v} \right\rangle \\ &= \langle \mathbf{v}|\mathbf{u}\rangle - \frac{\overline{\langle \mathbf{u}|\mathbf{v}\rangle}}{\|\mathbf{v}\|^2}\|\mathbf{v}\|^2 = \langle \mathbf{v}|\mathbf{u}\rangle - \langle \mathbf{v}|\mathbf{u}\rangle = 0.\end{aligned}$$

We call $c\mathbf{v} = \text{proj}_{\mathbf{v}}(\mathbf{u})$.

Theorem 7.4 (Cauchy-Schwarz Inequality)
Let $\mathbf{u}, \mathbf{v} \in V$ then

$$|\langle \mathbf{u}|\mathbf{v}\rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|.$$

Proof 7.4
If $\mathbf{v} = 0$ then the inequality holds.

Assume now $\mathbf{v} \neq 0$ and consider the orthogonal decomposition

$$\begin{aligned}
\mathbf{u} &= \frac{\langle \mathbf{u}|\mathbf{v}\rangle}{\|\mathbf{v}\|^2}\mathbf{v} + \mathbf{w} \\
\|\mathbf{u}\|^2 &= \left\| \frac{\langle \mathbf{u}|\mathbf{v}\rangle}{\|\mathbf{v}\|^2}\mathbf{v} + \mathbf{w} \right\|^2 = \left\| \frac{\langle \mathbf{u}|\mathbf{v}\rangle}{\|\mathbf{v}\|^2}\mathbf{v} \right\|^2 + \|\mathbf{w}\|^2 \\
&= \frac{\langle \mathbf{u}|\mathbf{v}\rangle^2}{\|\mathbf{v}\|^4}\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 \geq \frac{\langle \mathbf{u}|\mathbf{v}\rangle^2}{\|\mathbf{v}\|^2} \\
\|\mathbf{u}\|^2 &\geq \frac{\langle \mathbf{u}|\mathbf{v}\rangle}{\|\mathbf{v}\|^2} \\
\|\mathbf{v}\|^2\|\mathbf{u}\|^2 &\geq \langle \mathbf{u}|\mathbf{v}\rangle^2 \\
\|\mathbf{v}\|\|\mathbf{u}\| &= |\langle \mathbf{u}|\mathbf{v}\rangle|.
\end{aligned}$$

Here we used Pythagorean theorem and that $\|\mathbf{w}\|^2 \geq 0$.

Theorem 7.5 (Parallellogram)

$$\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = 2\Big(\|\mathbf{u}\|^2 + \|\mathbf{v}\|^2\Big).$$

Definition 7.5 (6.27 / 6.23)
Æet $\mathbf{v}_1, \dots, \mathbf{v}_n \in V$ where V is an inner product space. We say $\mathbf{v}_1, \dots, \mathbf{v}_n$ are orthonormal if

1. $\langle \mathbf{v}_i|\mathbf{v}_j\rangle = 0$ or $\mathbf{v}_i \perp \mathbf{v}_j,\ i \neq j$.
2. $\langle \mathbf{v}_i|\mathbf{v}_i\rangle = 1$ equivalent to $\|\mathbf{v}_i\| = 1$.

(Might also see $\langle \mathbf{v}_i|\mathbf{v}_j\rangle = \delta_{i.j}$ which is the Kronecker delta.)
If $\mathbf{v}_1, \dots, \mathbf{v}_n$ is also a basis we call it an orthonormal basis.

Remark: Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be an orthonormal basis of a VS V. Then $\forall \mathbf{v} \in V$:

$$\mathbf{v} = \langle \mathbf{v}_1|\mathbf{v}_1\rangle \mathbf{v}_1 + \langle \mathbf{v}_1|\mathbf{v}_2\rangle \mathbf{v}_2 + \cdots + \langle \mathbf{v}_1|\mathbf{v}_n\rangle \mathbf{v}_n.$$

Lemma 7.1
Let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be a list of orthonormal vectors. Then $\|\lambda_1\mathbf{v}_1 + \cdots + \lambda_m\mathbf{v}_m\|^2 = |\lambda|^2 + \cdots + |\lambda_m|^2$.

Proof 7.5

$$\|\lambda_1\mathbf{v}_1 + \cdots + \lambda_m\mathbf{v}_m\|^2$$

Note that all vectors are orthogonal to each other as well as any linear combintion of the others. Thus

$$\lambda\langle\mathbf{v}_2|\mathbf{v}_m\rangle + \mu\langle\mathbf{v}_1|\mathbf{v}_m\rangle = 0.$$

The sum is now:

$$\|\lambda_1\mathbf{v}_1 + \cdots + \lambda_{m-1}\mathbf{v}_{m-1}\|^2 + \|\lambda_m\mathbf{v}_m\|^2 = \|\lambda_1\mathbf{v}_1 + \cdots + \lambda_{m-1}\mathbf{v}_{m-1}\|^2 + |\lambda_m|^2\|\mathbf{v}_m\|^2$$

but since the vectors are orthonormal $\|\mathbf{v}_m\|^2 = 1$ and we can continue doing this to obtain

$$\sum|\lambda_i|^2\|\mathbf{v}_i\|^2 = \sum_{i=1}^{m}|\lambda_i|^2.$$

Notice that the standard basis vectors $\mathbf{e}_i$ are an orthonormal basis.

Lemma 7.2 (6.26)
Any list of orthonormals is linear independent.

Proof 7.6
We start by assuming that the linear combination gives zero.

$$\lambda_1\mathbf{v}_1 + \cdots + \lambda_m\mathbf{v}_m = 0$$

which only happens if all $\lambda_i = 0$.

$$\begin{aligned}\|\lambda_1\mathbf{v}_1 + \cdots + \lambda_m\mathbf{v}_m\|^2 &= 0\\ |\lambda_1|^2 + \cdots + |\lambda_m|^2 &= 0 \Rightarrow \lambda_i = 0.\end{aligned}$$

Corollary 7.1 (6.28)
An orthonormal list of length $\dim(V) < \infty$ is a basis.

Gram-Schmidt Orthonormalization

Algorithm for turning a basis into an orthonormal basis: Let $\mathbf{v}_1, \ldots, \mathbf{v}_n$ be a basis of V. We will construct another, orthonormal, basis $\mathbf{w}_1, \ldots, \mathbf{w}_n$.

$$\mathbf{w}_1 = \frac{\mathbf{v}_1}{\|\mathbf{v}_1\|} \text{hence } \|\mathbf{w}_1\| = 1$$

$$\tilde{\mathbf{w}}_2 = \mathbf{v}_2 - \mathrm{proj}_{\mathbf{w}_1}(\mathbf{v}_2) = \mathbf{v}_2 - \langle \mathbf{v}_2 | \mathbf{w}_1 \rangle \, \mathbf{w}_1$$

$$\mathbf{w}_2 = \frac{\tilde{\mathbf{w}}_2}{\|\tilde{\mathbf{w}}_2\|}$$

$$\tilde{\mathbf{w}}_3 = \mathbf{v}_3 - \mathrm{proj}_{\mathbf{w}_1}(\mathbf{v}_3) - \mathrm{proj}_{\mathbf{w}_2}(\mathbf{v}_3) = \mathbf{v}_3 - \langle \mathbf{v}_3 | \mathbf{w}_1 \rangle \, \mathbf{w}_1 - \langle \mathbf{v}_3 | \mathbf{w}_2 \rangle \, \mathbf{w}_2$$

$$\mathbf{w}_3 = \frac{\tilde{\mathbf{w}}_3}{\|\tilde{\mathbf{w}}_3\|}$$

$$\tilde{\mathbf{w}}_i = \mathbf{v}_i - \mathrm{proj}_{\mathbf{w}_1}(\mathbf{v}_i) - \mathrm{proj}_{\mathbf{w}_2}(\mathbf{v}_i) - \cdots - \mathrm{proj}_{\mathbf{w}_{i-1}}(\mathbf{v}_i)$$

$$\mathbf{w}_i = \frac{\tilde{\mathbf{w}}_i}{\|\tilde{\mathbf{w}}_i\|}$$

8 Determinants

We start with change of basis. Remark: We write $\mathcal{M}(T)$ for the matrix representation of a linear map T, implicitly we are assuming that bases have been fixed. Let $T : V \to W$ and $(\mathbf{v}_i)$ is a basis for V and $(\mathbf{w}_j)$ is a basis for W then we will write from now on $\mathcal{M}(T, (\mathbf{v}_i), (\mathbf{w}_j))$ for the matrix representation of T with respect to the bases $(\mathbf{v}_i)$ and $(\mathbf{w}_j)$.

We have seen that $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$, to write it explicitly we write $\mathcal{M}(ST, (\mathbf{u}_i), (\mathbf{w}_j)) = \mathcal{M}(S, (\mathbf{v}_j), (\mathbf{w}_k))\mathcal{M}(T, (\mathbf{u}_i), (\mathbf{v}_j))$, hence the map is $U \to V \to W$ with $T : U \to V$, $S : V \to W$. We will write $\mathcal{M}(T, (\mathbf{v}_i))$ for $\mathcal{M}(T, (\mathbf{w}_i), (\mathbf{v}_i))$ when $T \in \mathrm{End}(V)$.

Lemma 8.1 (10.5)
Let $(\mathbf{u}_i), (\mathbf{v}_i)$ be the bases for V. Then

$$\mathcal{M}(\mathrm{id}, (\mathbf{u}_i), (\mathbf{v}_i))^{-1} = \mathcal{M}(\mathrm{id}, (\mathbf{v}_i), (\mathbf{u}_i)).$$

Proof 8.1

$$\mathcal{M}(\mathrm{id}, (\mathbf{u}_1), (\mathbf{v}_i))\mathcal{M}(\mathrm{id}, (\mathbf{v}_i), (\mathbf{u}_i)) = \mathcal{M}(\mathrm{id}, (\mathbf{v}_i), (\mathbf{v}_i)) = I_n.$$

Theorem 8.1 (10.7)
Let $T \in \mathrm{End}(V)$ and $(\mathbf{u}_i), (\mathbf{v}_i)$ bases of V. Then

$$\mathcal{M}(T, (\mathbf{u}_i)) = A^{-1}\mathcal{M}(T, (\mathbf{v}_i))A,$$

where $A = \mathcal{M}(\mathrm{id}, (\mathbf{u}_i), (\mathbf{v}_i))$.

Proof 8.2

$$\begin{aligned} A^{-1}\mathcal{M}(T, \mathbf{v}_i)A &= \mathcal{M}(\mathrm{id}, \mathbf{v}_i, \mathbf{u}_i)\mathcal{M}(T, \mathbf{v}_i)\mathcal{M}(\mathrm{id}, \mathbf{u}_i, \mathbf{v}_i) \\ &= \mathcal{M}(T, \mathbf{v}_i, \mathbf{u}_i)\mathcal{M}(T, \mathbf{u}_i, \mathbf{v}_i) = \mathcal{M}(T, \mathbf{u}_i, \mathbf{u}_i). \end{aligned}$$

Definition 8.1
A map $\det : \mathcal{F}^{n\times n} \to \mathcal{F}$ is called a determinant map if

1. $\det\begin{pmatrix} a_1 \\ \vdots \\ \lambda a_i \\ \vdots \\ a_n \end{pmatrix} = \lambda \det\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$ and $\det\begin{pmatrix} a_1 \\ \vdots \\ a_i + a_i' \\ \vdots \\ a_n \end{pmatrix} = \det\begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} + \det\begin{pmatrix} a_1 \\ \vdots \\ a_i' \\ \vdots \\ a_n \end{pmatrix}$

2. If two rows are identical then $\det = 0$
3. $\det(I_n) = 1$.

Theorem 8.2
For every determinant map it holds that

1. $\det(\lambda A) = \lambda^n \det(A)$
2. If a row of A is zero then $\det(A) = 0$
3. If B results from swappning two rows of A, then $\det(B) = -\det(A)$
4. $\det\begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} = \det\begin{pmatrix} a_1 \\ \vdots \\ a_i + \lambda a_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix}$
5. $\det\begin{pmatrix} \lambda_i & & \\ 0 & \ddots & \\ 0 & 0 & \lambda_n \end{pmatrix} = \prod_{i=1}^n \lambda_i$.

Proof 8.3

1. Apply the first item of Definition 8.1 n-times.
2. $$\det\begin{pmatrix} a_1 \\ \vdots \\ 0 \cdot 0 \\ \vdots \\ a_n \end{pmatrix} = 0\det\begin{pmatrix} a_1 \\ \vdots \\ 0 \\ \vdots \\ a_n \end{pmatrix} = 0.$$
3. $$0 = \det\begin{pmatrix} a_1 \\ \vdots \\ a_i + a_j \\ \vdots \\ a_i + a_j \\ \vdots \\ a_n \end{pmatrix} = \det\begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} + \det\begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} + \det\begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} + \det\begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix}$$
   two of the terms are zero as they have two identical rows.

4.
$$\det\begin{pmatrix} a_1 \\ \vdots \\ a_i + \lambda a_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} = \det\begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix} + \lambda\det\begin{pmatrix} a_1 \\ \vdots \\ a_j \\ \vdots \\ a_j \\ \vdots \\ a_n \end{pmatrix}$$
where the last term is zero as it has two identical rows.

5. If $\lambda_i \neq 0 \ \forall i$ then we can use elementary row operations to transform A into diagonal form in which case we know
$$\det\begin{pmatrix} \lambda & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_n \end{pmatrix} = \det\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix} = \prod_{i=1}^{n} \lambda_i \det(I).$$
If $\lambda_i = 0$ for all i then let i be the largest index such that $\lambda_i = 0$. Then we use row operations to make the ith row all zeroes and the determinant is zero.

Lemma 8.2
$\det(A) \neq 0$ if and only if A is invertible.

Proof 8.4
Using row operations we can transform A into an upper triangular matrix A'. (Equivalently, if $A = \mathcal{M}(T, \mathbf{v}_i)$ then $A' = \mathcal{M}(T, \mathbf{v}_i, \mathbf{u}_i)$). Now we have $\det(A) = \pm\det(A')$ but also the map of such an upper triangular matrix A' is only surjective if all the diagonal values are nonzero and $\det(A') \neq 0$ which is identical to saying that A' is invertible, equivalent to T being invertible and A is invertible: A' is surjective $\Leftrightarrow \lambda_i \neq 0$ $\Leftrightarrow \det(A') \neq 0 \Leftrightarrow A'$ is invertible $\Leftrightarrow T$ is invertible $\Leftrightarrow A$ is invertible.

Corollary 8.1
If there exists a determinant map then it is unique.

Proof 8.5
Let $A \in \mathcal{F}^{n\times n}$ then there are row operations that transform A into an upper diagonal matrix A' then
$$\det(A) = (-1)^k \det(A')$$

where k is the number of row swaps that were performed. Then we know

$$\det(A) = (-1)^k \prod_{i=1}^{n} \lambda_i$$

which only has one set of λ_i.

Theorem 8.3
There is exactly one determinant map for every field $\mathcal{F}$ and integer $n \geq 1$.

Proof 8.6
By induktion on n:

$$n = 1, \quad \det((a_{11})) = a_{11}$$

For $n > 1$ and $A \in \mathcal{F}^{n\times n}$ consider the submatrices $\hat{A}_{ij}$ given by removing the ith row and jth column of A. Then let

$$\det_n(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det_{n-1}(\hat{A}_{ij}).$$

We claim this is a determinant map (and is the same for any j). To do show we show the items of Definition 8.1.

1. Let $A' = \begin{pmatrix} a_1 \\ \vdots \\ \lambda a_k \\ \vdots \\ a_n \end{pmatrix}$ then

$$\begin{aligned}\det_n(A') &= \sum_{i=1, i\neq k}^{n} (-1)^{i+j} a_{ij} \det_{n-1}(\hat{A}_{ij}) + (-1)^{k+j} \lambda a_{kj} \det(\hat{A'}_{kj}) \\ &= \lambda \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det_{n-1}(\hat{A}_{ij}).\end{aligned}$$

2. Let $a_{kj} = a_{lj}\ \forall j$ and assume $k < l$. Then $\det_{n-1}(\hat{A}_{ij}) = 0$ where $i \neq k$ and $i \neq l$. We're left with

$$\det_n(A) = (-1)^{k+j} a_{kj} \det_{n-1}(\hat{A}_{kj}) + (-1)^{l+j} a_{kj} \det_{n-1}(\hat{A}_{lj})$$

We can get $\hat{A}_{lj}$ from $\hat{A}_{kj}$ by swapping rows $l - k - 1$ times. Then

$$\det_{n-1}(\hat{A}_{kj}) = (-1)^{l-k-1} \det(\hat{A}_{lj}).$$

Now we get

$$\begin{aligned}&(-1)^{k+j+l-k-1} a_{kj} \det(\hat{A}_{lj}) + (-1)^{l+j} a_{kj} \det(\hat{A}_{lj}) \\ &(-1)^{l+j-1} a_{kj} \det(\hat{A}_{lj}) + (-1)^{l+j} a_{kj} \det(\hat{A}_{lj}) = 0.\end{aligned}$$

3. s

$$\begin{aligned}\det_n(I) &= \sum_{i=1}^{n}(-1)^{i+j}S_{ij}\det_{n-1}(\hat{I}_{nij}) \\ &= \det_{n-1}(I_{n-1}) = 1.\end{aligned}$$

Corollary 8.2

$$\det(A) = \det(A^\intercal)$$

Proof 8.7
We show that

$$\tilde{\det} : \mathcal{F}^{n\times n} \to \mathcal{F},\ A \to \det(A^\intercal)$$

is a determinant map by checking the conditions of a determinant map and then use uniqueness. This is now an exercise.

Corollary 8.3

$$\det(AB) = \det(A)\det(B)$$

Proof 8.8
If $\det(B) = 0$, then B is not invertible and it follows that AB is not invertible implying that $\det(AB) = 0$. Assume $\det(B) \neq 0$. Define

$$\tilde{\det}(A) = \frac{\det(AB)}{\det(B)}$$

and show that it is a determinant map.

1.

$$\tilde{\det}(\lambda_i I \cdot A) = \tilde{\det}\begin{pmatrix} a_1 \\ \vdots \\ \lambda a_i \\ \vdots \\ a_n \end{pmatrix}$$

$$\tilde{\det}\begin{pmatrix} a_1 \\ \vdots \\ \lambda a_i \\ \vdots \\ a_n \end{pmatrix} = \frac{\det(\lambda_i I \cdot AB)}{\det(B)} = \frac{\det\begin{pmatrix} a_1 b_1 & a_1 b_n \\ \lambda a_i b_1 & \lambda a_i b_n \\ a_n b_1 & a_n b_n \end{pmatrix}}{\det(B)} = \lambda \tilde{\det}(A)$$

jeg gider ikke mere

9 Generalised Eigenspace

Corollary 9.1

$$\det(A{-}1) = \det(A)^{-1}.$$

Proof 9.1

$$1 = \det(I) = \det(AA^{-1}) = \det(A)\det(A^{-1})$$

Definition 9.1
Let $A \in \mathbb{R}^{n\times n}$ or $A \in \mathbb{C}^{n\times n}$, we define its characteristic polynomial as

$$\chi_A(z) = \det(zI - A).$$

We see that the roots of $\chi_A(z)$ are the eigenvalues of A. Over $\mathbb{C}$ $\chi_A(z)$ completely factors:

$$\chi_A(z) = \prod_{i=1}^{n}(z - \lambda_i),$$

where the λ_i are eigenvalues of A. If we plug in $z = 0$ we get $\chi_A(0) = \det(-A))(-1)^n \prod_{i=1}^n \lambda_i$ which implies that $\det(A) = \prod_{i=1}^n \lambda_i$. Futhermore, the coefficient of z^{n-1} is given by

$$-\sum_{i=1}^{n}\lambda_i$$

and we will call $\sum_{i=1}^n \lambda_i$ the trace of A, denoted $\operatorname{tr}(A)$.

Remark: Let $T \in \operatorname{End}(V)$. Then we can define its characteristic polynomial by choosing a basis $(\mathbf{v}_i)$ for V and we calculate $\chi_{\mathcal{M}(T,\mathbf{v}_i)}(z) := \chi_T(z)$. To check let $(\mathbf{u}_i)$ be a different basis and $S = \mathcal{M}(\mathrm{id}, \mathbf{u}_i, \mathbf{v}_i)$, we know now that

$$\begin{aligned}
\mathcal{M}(T,\mathbf{u}_i) &= S^{-1}\mathcal{M}(T,\mathbf{v}_i)S \\
\chi_{\mathcal{M}(T,\mathbf{u}_i)} &= \det(zI - \mathcal{M}(T,\mathbf{u}_i)) = \det(zS^{-1}S - S^{-1}\mathcal{M}(T,\mathbf{v}_i)S) \\
&= \det(S^{-1})\det(zI - \mathcal{M}(T,\mathbf{v}_i))\det(S) = \chi_{\mathcal{M}(T,\mathbf{v}_i)}.
\end{aligned}$$

Theorem 9.1 (Cayley-Hamilton)
Let $T \in \operatorname{End}(V)$ and V is a $\mathbb{C}$-VS with $\chi_T(z)$ as its characteristic polynomial. Then

$$\chi_T(T) = 0.$$

Proof 9.2
We know that there exists a basis such that $\mathcal{M}(T)$ is upper triangular. Assume

$$A = \mathcal{M}(T) = \begin{pmatrix} \lambda_1 & * & * \\ 0 & \ddots & * \\ 0 & 0 & \lambda_n \end{pmatrix}$$

then $\det(zI - A) = \prod(z - \lambda_i)$. We see that

$$\begin{aligned} A\mathbf{e}_k &\in \operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_k) \\ (\lambda_k I - A)\mathbf{e}_k &\in \operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_{k-1}). \end{aligned}$$

Let $\mathbf{v} \in \operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_k)$ such that $(\lambda_k I - A)\mathbf{v} \in \operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_{k-1})$. Now let $\mathbf{v} \in V = \operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_n)$ then $\mathbf{v}_{n-1} = (\lambda_n I - A)\mathbf{v}_n \in \operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_{n-1})$ and we continue: $\mathbf{v}_{n-2} = (\lambda_{n-1} I - A)\mathbf{v}_{n-1} \in \operatorname{span}(\mathbf{e}_1, \dots, \mathbf{e}_{n-2})$ and lastly we get $\mathbf{v}_1 = (\lambda_2 I - A)\mathbf{v}_2 \in \operatorname{span}(\mathbf{e}_1)$ and $\mathbf{v}_1 = (\lambda_1 I - A)\mathbf{v}_1 = 0$. We calculate

$$\begin{aligned} (\lambda_1 I - A)(\lambda_2 I - A) \dots (\lambda_n I - A)\mathbf{v} = 0 \ \forall \mathbf{v} \in V \\ \Rightarrow \chi_T(T) = 0. \end{aligned}$$

Example 9.1

$$\begin{aligned} A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix},\ \chi_A(z) &= (z-2)^2 \\ &= z^2 - 4z + 4 \\ A^2 - 4A + 4I = 0 A^2 & \qquad = 4A - 4I \\ \begin{pmatrix} 4 & 4 \\ 0 & 4 \end{pmatrix} &= \begin{pmatrix} 8 & 4 \\ 0 & 8 \end{pmatrix} - \begin{pmatrix} 4 & 0 \\ 0 & 4 \end{pmatrix}. \end{aligned}$$

Definition 9.2
Consider the set of all permutations of n elements. We call this set the symmetric group S_n.

Example: $\sigma \in S_3$ $\sigma : \{1,2,3\} \to \{1,2,3\}$ and σ is bijective.
Permutations that swaps two elements are called transpositions.
As a matrix we can represent σ by

$$\mathcal{M}_\sigma = \begin{pmatrix} \mathbf{e}_{\sigma(1)} & \mathbf{e}_{\sigma(2)} & \dots & \mathbf{e}_{\sigma(m)} \end{pmatrix}.$$

The determinant is the same as the identity but with exchanged sign as the rows are interchanged:

$$\det(\mathcal{M}_\sigma) = \pm 1$$
$$\text{sign}(\sigma) = \det(\mathcal{M}_\sigma).$$

For example say:

$$\sigma(1) = 3,\ \sigma(2) = 1,\ \sigma(3) = 2$$
$$1 \to 3,\ 2 \to 1,\ 3 \to 2 = 1 \to 1,\ 2 \to 2,\ 3 \to 1 \cdot 1 \to 1,\ 2 \to 3,\ 3 \to 2.$$

Remark: $|S_n| = n!$.

Proposition 9.1 (Leibniz Formula)

$$\det(A) = \sum_{\sigma \in S_n} \text{sign}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)}$$

Proof 9.3
Let

$$A = \begin{pmatrix} *a_1* \\ a_2 \\ a_n* \end{pmatrix}$$

then $a_1 = a_{11}\mathbf{e}_1 + a_{12}\mathbf{e}_2 + \ldots + a_{1n}\mathbf{e}_n$ and

$$\det(A) = \sum_{i=1}^{n} \det \begin{pmatrix} a_{1i}\mathbf{e}_i \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \sum_{i_1=1}^{n} a_{1i} \det \begin{pmatrix} \mathbf{e}_{i_1} \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$$

and we repeat this for all rows. It is then written as

$$\det(A) = \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} a_{1i_1}a_{2i_2}\det\begin{pmatrix}\mathbf{e}_{i_1}\\ \mathbf{e}_{i_2}\\ a_3\\ \vdots\\ a_n\end{pmatrix}$$

$$\vdots$$

$$= \sum_{i_1=1}^{n}\dots\sum_{i_n=1}^{n} a_{1i_1}a_{2i_2}\dots a_{ni_n}\det\begin{pmatrix}\mathbf{e}_{i_1}\\ \vdots\\ \mathbf{e}_{i_n}\end{pmatrix}$$

$$= \sum_{\sigma\in S_n} a_{1,\sigma(1)}\dots a_{n,\sigma(n)}\det\begin{pmatrix}\mathbf{e}_{\sigma(1)}\\ \vdots\\ \mathbf{e}_{\sigma(n)}\end{pmatrix}$$

$$= \sum_{\sigma\in S_n}\operatorname{sign}(\sigma)\prod_{i=1}^{n} a_{i,\sigma(i)}.$$

Proposition 9.2

$$\operatorname{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

Proof 9.4
We define the trace as the negative of the coeficient of z^{n-1} in the characteristic polynomial. Then we need to show that $\chi_A(z) = \prod_{i=1}^{n}(z - a_{ii}) + Q(z)$ where the degree of $Q(z) \leq z^{n-2}$.

Induktion on n:

$$n = 1\ A = (a_{11}),\ \det(zI - a_{11}$$

$$n > 1\ \chi_A(z) = \det\begin{pmatrix} z - a_{11} & -a_{12} & \dots & -a_{1n}\\ -a_{21} & z - a_{22} & \dots & *\\ \vdots & & & \\ -a_{n1} & & & \end{pmatrix} = (z - a_{11})\det((zI \hat{-} A)_{11}) + \sum_{j=2}^{n}(-1)^{j+1}(-a_{j1}) - \det($$

$$\operatorname{degree}((zI \hat{-} A)_{j1}) \leq$$

Example 9.2

$$A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$$

has eigenvalues 2 with multiplicity 2 with eigenvector $\begin{pmatrix} 1 & 0 \end{pmatrix}^\intercal$.

Definition 9.3
Let $A \in \mathcal{F}^{n\times n}$ and λ eigenvalues of A then we define

$$\dim(\ker(A - \lambda I))$$

as the geometrical multiplicity. Also, we define the multiplicity of λ as a root of χ_A as the algebraic multiplicity.

Proposition 9.3
For an eigenvalue λ of A the algebraic multiplicity is always bigger or equal to the geometric multiplicity.

Proof 9.5
Let k be the geometric multiplicity of λ. Then we have k linear independent eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_k$. Extend this to a basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$. With respect to this basis A is on block form with eigenvalues in the first block. Then we have

$$\chi_A = \det(zI_k - \lambda I_k)$$
$$\det(zI_{n-k} - D) = (z-\lambda)^k \det(zI_{n-k} - D)$$

This shows that the algebraic multiplicity is at most k.

10 Adjoints

For this section everything is finite vector spaces over the reals or complex numbers.

Definition 10.1 (7.2)
Let V, W be finite dimensional inner product spaces and $T \in \mathrm{Hom}(V, W)$. Then the adjoint of T is a function $T^* : W \to V$ such that $\langle T\mathbf{v}|\mathbf{w}\rangle = \langle \mathbf{v}|T^*\mathbf{w}\rangle$ for all $\mathbf{v} \in V$, $\mathbf{w} \in W$.

Example 10.1
Let $V = W = \mathbb{C}^2$ then $\langle \mathbf{v}|\mathbf{w}\rangle = \mathbf{v}^\intercal\overline{\mathbf{w}}$. Let also $A \in \mathbb{C}^{2\times 2}$ then

$$\langle A\mathbf{v}|\mathbf{w}\rangle = (A\mathbf{v})^\intercal\overline{\mathbf{w}} = \mathbf{v}^\intercal A^\intercal\overline{\mathbf{w}} = \mathbf{v}^\intercal\overline{\overline{A}^\intercal\mathbf{w}} = \left\langle \mathbf{v}\middle|\overline{A}^\intercal\mathbf{w}\right\rangle$$

and $A^* = \overline{A}^\intercal$.

Proposition 10.1 (7.5)
If $T \in \mathrm{Hom}(V, W)$ is a linear map then $T^* \in \mathrm{Hom}(W, V)$ is also a linear map.

Proof 10.1
We show that $T^*(\mathbf{w}_1 + \mathbf{w}_2) = T^*\mathbf{w}_1 + T^*\mathbf{w}_2$ by showing that for all $\mathbf{v} \in V$ $\langle \mathbf{v}|T^*(\mathbf{w}_1 + \mathbf{w}_2)\rangle = \langle \mathbf{v}|T^*\mathbf{w}_1 + T^*\mathbf{w}_2\rangle$. Hence

$$\begin{aligned}
\langle \mathbf{v}|T^*(\mathbf{w}_1 + \mathbf{w}_2)\rangle &= \langle T\mathbf{v}|\mathbf{w}_1 + \mathbf{w}_2\rangle \\
&= \langle T\mathbf{v}|\mathbf{w}_1\rangle + \langle T\mathbf{v}|\mathbf{w}_2\rangle \\
&= \langle \mathbf{v}|T^*\mathbf{w}_1\rangle + \langle \mathbf{v}|T^*\mathbf{w}_2\rangle = \langle \mathbf{v}|T^*\mathbf{w}_1 + T^*\mathbf{w}_2\rangle \\
\langle \mathbf{v}|T^*\lambda\mathbf{w}\rangle &= \langle T\mathbf{v}|\lambda\mathbf{w}\rangle = \overline{\lambda}\langle T\mathbf{v}|\mathbf{w}\rangle \\
&= \overline{\lambda}\langle \mathbf{v}|T^*\mathbf{w}\rangle = \langle \mathbf{v}|\lambda T^*\mathbf{w}\rangle.
\end{aligned}$$

■

Proposition 10.2 (7.6)
The following hold

1. $(S + T)^* = S^* + T^*$
2. $(\lambda S)^* = \overline{\lambda}S^*$
3. $(T^*)^* = T$
4. $I^* = I$
5. $(ST)^* = T^*S^*$

Proof 10.2
1)

$$\begin{aligned}\langle \mathbf{v}|(S+T)^*\mathbf{w}\rangle &= \langle (S+T)\mathbf{v}|\mathbf{w}\rangle \\ &= \langle S\mathbf{v}|\mathbf{w}\rangle + \langle T\mathbf{v}|\mathbf{w}\rangle \\ &= \langle \mathbf{v}|S^*\mathbf{w}\rangle + \langle \mathbf{v}|T^*\mathbf{w}\rangle \\ &= \langle \mathbf{v}|S^*\mathbf{w} + T^*\mathbf{w}\rangle \\ &= \langle \mathbf{v}|(S^*+T^*)\mathbf{w}\rangle\end{aligned}$$

2)

$$\begin{aligned}\langle \mathbf{v}|(\lambda T)^*\mathbf{w}\rangle &= \langle \lambda T\mathbf{v}|\mathbf{w}\rangle \\ &= \lambda \langle T\mathbf{v}|\mathbf{w}\rangle \\ &= \lambda \langle \mathbf{v}|T^*\mathbf{w}\rangle \\ &= \left\langle \mathbf{v}\middle|\bar{\lambda}T^*\mathbf{w}\right\rangle\end{aligned}$$

3)

$$\begin{aligned}\langle \mathbf{w}|(T^*)^*\mathbf{v}\rangle &= \langle T^*\mathbf{w}|\mathbf{v}\rangle \\ &= \overline{\langle \mathbf{v}|T^*\mathbf{w}\rangle} \\ &= \overline{\langle T\mathbf{v}|\mathbf{w}\rangle} \\ &= \langle \mathbf{w}|T\mathbf{v}\rangle\end{aligned}$$

4)

$$\begin{aligned}\langle \mathbf{v}|I^*\mathbf{u}\rangle &= \langle I\mathbf{v}|\mathbf{u}\rangle \\ &= \langle \mathbf{v}|I\mathbf{u}\rangle\end{aligned}$$

5) $\mathbf{v} \to \mathbf{w} \to \mathbf{u}$

$$\begin{aligned}\langle \mathbf{v}|(ST)^*\mathbf{u}\rangle &= \langle ST\mathbf{v}|\mathbf{u}\rangle \\ &= \langle T\mathbf{v}|S^*\mathbf{u}\rangle \\ &= \langle \mathbf{v}|T^*S^*\mathbf{u}\rangle .\end{aligned}$$

■

Corollary 10.1 (7.7)
Let $T \in \mathrm{Hom}(V, W)$ then the following holds

1. $\ker(T^*) = (\mathrm{im}(T))^\perp$
2. $\mathrm{im}(T^*) = (\ker(T))^\perp$
3. $\ker(T) = (\mathrm{im}(T^*))^\perp$

4. $\mathrm{im}(T) = (\ker(T^*))^{\perp}$.

Proof 10.3
a) Let $\mathbf{w} \in \ker(T^*) \Leftrightarrow T^*\mathbf{w} = 0 \Leftrightarrow \langle \mathbf{v} | T^*\mathbf{w} \rangle = 0 \, \forall \mathbf{v} \in V$ if and only if $\langle T\mathbf{v} | \mathbf{w} \rangle = 0 \Leftrightarrow \mathbf{w} \in (\mathrm{im}(T))^{\perp}$. Note

$$\ker(T^*) = (\mathrm{im}(T))^{\perp} \Rightarrow \ker(T^*)^{\perp} = \mathrm{im}(T)$$

which is d).

c) Use $T = S^*$, then $\ker(S^*)^* = (\mathrm{im}(S^*))^{\perp} \Rightarrow \ker(S) = (\mathrm{im}(S^*))^{\perp}$.

b) is found in analogous.

Proposition 10.3 (7.10)
Let $\mathbf{e}_i$ be an orthonormal basis of V and $\mathbf{f}_j$ be an orthonormal basis of W then

$$B = \mathcal{M}(T^*, \mathbf{f}_j, \mathbf{e}_i) = \overline{\mathcal{M}(T, \mathbf{e}_i, \mathbf{f}_j)}^{\intercal} = \overline{A}^{\intercal}$$

Proof 10.4
The kth column of A is given by the image of $\mathbf{e}_k$ under the map T with respect to the basis $\mathbf{f}_j$: $T\mathbf{e}_k = \sum \lambda_j \mathbf{f}_j$. Since it is an orthonormal basis we know the coefficients and

$$T\mathbf{e}_k = \sum \langle T\mathbf{e}_k | \mathbf{f}_j \rangle \, \mathbf{f}_j.$$

In other words $A_{ji} = \langle T\mathbf{e}_i | \mathbf{f}_j \rangle$.

Likewise $B_{ij} = \langle T^*\mathbf{f}_j | \mathbf{e}_i \rangle$. Now we compare the two

$$\begin{aligned} B_{ij} = \langle T^*\mathbf{f}_j | \mathbf{e}_i \rangle &= \langle \mathbf{f}_j | (T^*)^*\mathbf{e}_i \rangle = \langle \mathbf{f}_j | T\mathbf{e}_i \rangle \\ &= \overline{\langle T\mathbf{e}_i | \mathbf{f}_j \rangle} \\ &= \overline{A_j i} \Rightarrow B = \overline{A}^{\intercal} \end{aligned}$$

T and its adjoint are in different spaces normally.

Definition 10.2 (7.11)
Let $T\mathrm{End}(V)$. We say T is self-adjoint if $T = T^*$.

Theorem 10.1 (7.13)
Let $T \in \mathrm{End}(V)$ be self-adjoint then T only has real eigenvalues.

Proof 10.5
Assume λ is an eigenvalue of T and $T\mathbf{v} = \lambda\mathbf{v}$. Then

$$\begin{aligned}
\lambda\|\mathbf{v}\|^2 &= \langle \lambda\mathbf{v}|\mathbf{v}\rangle \\
&= \langle T\mathbf{v}|\mathbf{v}\rangle \\
&= \langle \mathbf{v}|T\mathbf{v}\rangle \\
&= \langle \mathbf{v}|\lambda\mathbf{v}\rangle \\
&= \overline{\lambda}\,\langle \mathbf{v}|\mathbf{v}\rangle \\
&= \overline{\lambda}\|\mathbf{v}\|^2 \\
&\Rightarrow \lambda = \overline{\lambda} \Rightarrow \lambda \in \mathbb{R}.
\end{aligned}$$

■

Lemma 10.1 (7.14)
Let V be a $\mathbb{C}$-VS, then $\langle T\mathbf{v}|\mathbf{v}\rangle = 0$ for all $\mathbf{v}$ if and only if $T = 0$.

Proof 10.6
$\Leftarrow$ is easy.
$\Rightarrow$: Write

$$\begin{aligned}
4\,\langle T\mathbf{u}|\mathbf{w}\rangle &= \langle T\mathbf{u}|\mathbf{u}\rangle + \langle T\mathbf{u}|\mathbf{w}\rangle + \langle T\mathbf{w}|\mathbf{u}\rangle + \langle T\mathbf{w}|\mathbf{w}\rangle \\
&- \langle T\mathbf{u}|\mathbf{u}\rangle + \langle T\mathbf{u}|\mathbf{w}\rangle + \langle T\mathbf{w}|\mathbf{u}\rangle - \langle T\mathbf{w}|\mathbf{w}\rangle \\
&+ \langle T\mathbf{u}|\mathbf{u}\rangle\, i + \langle T\mathbf{u}|\mathbf{w}\rangle - \langle T\mathbf{w}|\mathbf{u}\rangle + \langle T\mathbf{w}|\mathbf{w}\rangle\, i \\
&- \langle T\mathbf{u}|\mathbf{u}\rangle\, i + \langle T\mathbf{u}|\mathbf{w}\rangle - \langle T\mathbf{w}|\mathbf{u}\rangle - \langle T\mathbf{w}|\mathbf{w}\rangle\, i \\
&= \langle T(\mathbf{u}+\mathbf{w})|\mathbf{u}+\mathbf{w}\rangle - \langle T(\mathbf{w}-\mathbf{v})|\mathbf{w}-\mathbf{u}\rangle + \langle T(\mathbf{u}+\mathbf{w}i)|\mathbf{u}+i\mathbf{w}\rangle\, i \ldots = 0 \\
&\Rightarrow T = 0
\end{aligned}$$

■

Proposition 10.4 (7.15)
Let V be a $\mathbb{C}$-VS then $\langle T\mathbf{v}|\mathbf{v}\rangle \in \mathbb{R}\,\forall \mathbf{v} \in V$ if and only if $T = T^*$.

Proof 10.7
Let $\mathbf{v} \in V$ then

$$\begin{aligned} \langle T\mathbf{v}|\mathbf{v}\rangle - \overline{\langle T\mathbf{v}|\mathbf{v}\rangle} &= \langle T\mathbf{v}|\mathbf{v}\rangle - \langle \mathbf{v}|T\mathbf{v}\rangle \\ &= \langle T\mathbf{v}|\mathbf{v}\rangle - \langle T^*\mathbf{v}|\mathbf{v}\rangle \\ &= \langle (T-T^*)\mathbf{v}|\mathbf{v}\rangle = 0 \end{aligned}$$

If $\langle T\mathbf{v}|\mathbf{v}\rangle \in \mathbb{R} \Rightarrow T = T^*$. If T is self-adjoint read the proof backwards and the result follows. ■

Proposition 10.5 (7.16)
Let $T \in \mathrm{End}(V)$ be self-adjoint and $\langle T\mathbf{v}|\mathbf{v}\rangle = 0$ then $T = 0$.

Proof 10.8
If V is complex then the proposition holds even for not self-adjoint operators by Lemma 10.1. Assume V is a $\mathbb{R}$-VS and $\mathbf{u}, \mathbf{w} \in V$ then

$$\begin{aligned} 4\langle T\mathbf{u}|\mathbf{w}\rangle &= \langle T(\mathbf{u}+\mathbf{w})|\mathbf{u}+\mathbf{w}\rangle - \langle T(\mathbf{u}-\mathbf{w})|\mathbf{u}-\mathbf{w}\rangle \\ &= \langle T\mathbf{u}|\mathbf{u}\rangle + \langle T\mathbf{u}|\mathbf{w}\rangle + \langle T\mathbf{w}|\mathbf{u}\rangle + \langle T\mathbf{w}|\mathbf{w}\rangle \\ &- \langle T\mathbf{u}|\mathbf{u}\rangle + \langle T\mathbf{u}|\mathbf{w}\rangle + \langle T\mathbf{w}|\mathbf{u}\rangle - \langle T\mathbf{w}|\mathbf{w}\rangle \\ &= 0 \\ &\Rightarrow T = 0. \end{aligned}$$

■

Definition 10.3 (7.18)
We say $T \in \mathrm{End}(V)$ is normal if $TT^* = T^*T$.

Example 10.2

1. Self-adjoint operators are normal: $T = T^* \Rightarrow TT = TT$.
2. $V = \mathbb{R}^2$ then look at $\mathcal{M}(T) = \begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}$ then $\mathcal{M}(T^*) = \begin{pmatrix} 2 & 3 \\ -3 & 2 \end{pmatrix}$ and hence is not self-adjoint but the product is the same and it is then normal. Normal is actually bigger than self-adjoint.

Theorem 10.2 (7.20)
T is normal if and only if $\|T\mathbf{v}\| = \|T^*\mathbf{v}\| \; \forall \mathbf{v} \in V$.

Proof 10.9
T is normal is equivalent to saying $T^*T - TT^* = 0$ which implies $\langle (T^*T - TT^*)\mathbf{v} | \mathbf{v} \rangle = 0$.
Side note: $T^*T - TT^*$ is self-adjoint by

$$(T^*T - TT^*)^* = (T^*T)^* - (TT^*)^* = T^*T - TT^*.$$

Now this is equivalent to $\langle T^*T\mathbf{v} | \mathbf{v} \rangle = \langle TT^*\mathbf{v} | \mathbf{v} \rangle$ equivalent to

$$\langle T\mathbf{v} | T\mathbf{v} \rangle = \langle T^*\mathbf{v} | T^*\mathbf{v} \rangle \Leftrightarrow \|T\mathbf{v}\|^2 = \|T^*\mathbf{v}\|^2.$$

Proposition 10.6 (7.21)
Let $T \in \mathrm{End}(V)$ be normal and $T\mathbf{v} = \lambda \mathbf{v}$. Then

$$T^*\mathbf{v} = \overline{\lambda}\mathbf{v}.$$

Proof 10.10
We see that $S = T - \lambda I$ is also normal

$$S^* = T^* - \overline{\lambda} I$$

and then check that it all commutes. Since S is normal we have

$$\|S\mathbf{v}\| = \|S^*\mathbf{v}\|$$
$$\|(T - \lambda I)\mathbf{v}\| = \left\|(T^* - \overline{\lambda} I)\mathbf{v}\right\| = 0.$$

■

Theorem 10.3 (7.22)
Let $T \in \mathrm{End}(V)$ be normal. Then eigenvectors for distinct eigenvalues are orthogonal.

Proof 10.11
Let $\alpha \neq \beta$ be distinct eigenvalues with eigenvectors $\mathbf{u}, \mathbf{v}$. Then we calculate

$$\begin{aligned}
(\alpha - \beta) \langle \mathbf{u} | \mathbf{v} \rangle &= \langle \alpha \mathbf{u} | \mathbf{v} \rangle - \left\langle \mathbf{u} \middle| \overline{\beta} \mathbf{v} \right\rangle \\
&= \langle T\mathbf{u} | \mathbf{v} \rangle - \langle \mathbf{u} | T^* \mathbf{v} \rangle \\
&= \langle T\mathbf{u} | \mathbf{v} \rangle - \langle T\mathbf{u} | \mathbf{v} \rangle = 0 \\
&\Rightarrow \langle \mathbf{u} | \mathbf{v} \rangle = 0.
\end{aligned}$$

■

11 Diagonalizable Matrices

Diagonal matrices are nice so when can we diagonalize a matrix? To do so we want to find a nice orthonormal basis and turn our operator into a diagonal matrix.

Example 11.1

$$\mathcal{M}(T,(e_i)) = \begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix}$$

take another basis (b_i) of $\mathbb{C}^2$, then

$$\mathcal{M}(T,(b_i)) = \underbrace{\mathcal{M}(I,(e_i),(b_i))}_{S}\mathcal{M}(T,(e_i))\mathcal{M}(I,(b_i),(e_i))$$

If (b_i) is orthonormal then $S^{-1} = \overline{S}^{\mathsf{T}}$:

$$\begin{aligned}
b_1 &= \frac{1}{\sqrt{2}}\begin{pmatrix} i \\ 1 \end{pmatrix}, \quad b_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -i \\ 1 \end{pmatrix} \\
\langle b_1|b_1\rangle &= 1 = \langle b_2|b_2\rangle \\
\langle b_1|b_2\rangle &= \frac{1}{2}(i \cdot \overline{(-i)} + 1 \cdot \overline{1}) = 0 \\
S &= \frac{1}{\sqrt{2}}\begin{pmatrix} i & -i \\ 1 & 1 \end{pmatrix} \\
S^{-1} &= \frac{1}{\sqrt{2}}\begin{pmatrix} -i & 1 \\ i & 1 \end{pmatrix} \\
S^{-1}\begin{pmatrix} 2 & -3 \\ 3 & 2 \end{pmatrix} S &= \frac{1}{2}\begin{pmatrix} 3-2i & 2+3i \\ 3+2i & 2-3i \end{pmatrix} S \\
&= \frac{1}{2}\begin{pmatrix} 4+6i & 0+0i \\ 0+0i & 4-6i \end{pmatrix}
\end{aligned}$$

this matrix have two eigenvectors

$$\begin{aligned}
\begin{pmatrix} 1 \\ 0 \end{pmatrix}_{b_i} &= \frac{1}{\sqrt{2}}\begin{pmatrix} i \\ 1 \end{pmatrix}_{e_i} \\
\begin{pmatrix} 0 \\ 1 \end{pmatrix}_{b_i} &= \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}_{e_i}.
\end{aligned}$$

Theorem 11.1 (7.24)
Let V be a $\mathbb{C}$-VS and $T \in \mathrm{End}(V)$. Then the following are equivalent

1. T is normal; i.e. $TT^* = T^*T$.
2. T can be diagonalized with respect to some orthonormal basis of V.

Proof 11.1
want to prove 2. $\implies$ 1..

Assume $\mathcal{M}(T, (\mathbf{b}_i))$ is diagonal and $\mathbf{b}_i$ is an orthonormal basis. First we want to show that T is normal.

$$\mathcal{M}(T^*, (\mathbf{b}_i)) = \overline{\mathcal{M}(T, (\mathbf{b}_i))}^\intercal$$

and clearly $\mathcal{M}(T, (\mathbf{b}_i))$ and $\mathcal{M}(T^*, (\mathbf{b}_i))$ commute as they are both diagonal. Hence T and T^* commute $\Rightarrow$ T is normal.

Now we want to prove 1. $\implies$ 2..

Assume T is normal and then show that it can be diagonalized. We have that there is a basis such that T is upper triangular with respect to this basis and that basis can then be made orthonormal by Gram-Schmidt, thus there is an orthonormal basis.

$$\begin{aligned}
\mathcal{M}(T) &= \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ 0 & \dots & a_{nn} \end{pmatrix} \\
\mathcal{M}(T^*) &= \begin{pmatrix} \overline{a_{11}} & 0 & 0 \\ \vdots & \ddots & 0 \\ \overline{a_{1n}} & \dots & \overline{a_{nn}} \end{pmatrix} \\
\|T\mathbf{b}_1\| &= \|a_{11}\mathbf{b}_1\| = |a_{11}| \\
\|T^*\mathbf{b}_1\| &= \left\| \sum_{i=1}^{n} \overline{a_{1i}}\mathbf{b}_i \right\| = \sum_{i=1}^{n} |a_{1i}| \\
&\Rightarrow a_{1i} = 0 \quad \text{for all } i > 1. \\
\|T\mathbf{b}_2\| &= |a_{22}| = \|T^*\mathbf{b}_2\| = \sum_{i=2}^{n} |a_{2i}| \\
&\Rightarrow a_{2i} = 0 \quad \text{for all } i > 1. \\
&\vdots
\end{aligned}$$

$\mathcal{M}(T)$ is already diagonal. ■

Lemma 11.1
Let $T \in \operatorname{End}(V)$ be self-adjoint and $b, c \in \mathbb{R}$ such that $b^2 < 4c$. Then $T^2 + bT + cI$ is invertible.

Proof 11.2
Let $\mathbf{v} \in V, v \neq 0$, we then want to show that $(T^2 + bT + cI)\mathbf{v}$ is nonzero. Thus

$$\begin{aligned}
\left\langle (T^2 + bT + cI)\mathbf{v} \middle| \mathbf{v} \right\rangle &= \underbrace{\left\langle T^2\mathbf{v} \middle| \mathbf{v} \right\rangle}_{\langle T\mathbf{v}|T\mathbf{v}\rangle} + b\left\langle T\mathbf{v}|\mathbf{v}\right\rangle + c\left\langle \mathbf{v}|\mathbf{v}\right\rangle \\
&= \|T\mathbf{v}\|^2 + ?? + c\|\mathbf{v}\|^2 \\
-|b|\|T\mathbf{v}\|\|\mathbf{v}\| \leq b\left\langle T\mathbf{v}|\mathbf{v}\right\rangle &\leq b\|T\mathbf{v}\|\|\mathbf{v}\| \\
\left\langle (T^2 + bT + cI)\mathbf{v} \middle| \mathbf{v} \right\rangle &\geq \|T\mathbf{v}\|^2 - |b|\|T\mathbf{v}\|\|\mathbf{v}\| + c\|\mathbf{v}\|^2 \\
&= (\|T\mathbf{v}\| - \frac{|b|}{2}\|v\|)^2 + (c - \frac{b}{4})\|\mathbf{v}\|^2 > 0,
\end{aligned}$$

hence $T^2 + bT + cI$ has no kernal is thereby invertible. ■

Proposition 11.1
Let $V \neq \{0\}$ and $T \in \mathrm{End}(V)$, T is self-adjoint. Then T has an eigenvalue.

Proof 11.3
Already shown for $\mathbb{C}$-VS. Assume V is an $\mathbb{R}$-VS. Let $\mathbf{v} \in V, \mathbf{v} \neq 0$ then $\mathbf{v}, T\mathbf{v}, T^2\mathbf{v}, \ldots, T^n\mathbf{v}$ where $n = \dim(V)$, is linear dependent $\Rightarrow \exists a_i \in \mathbb{R} | a_0\mathbf{v} + a_1T\mathbf{v} + \ldots + a_nT^n\mathbf{v} = 0$. The polynomial $a_0 + a_1x + \ldots + a_nx^n$ splits. To show this let $f \in \mathbb{R}[x], z \in \mathbb{C}\backslash\mathbb{R}$ and $f(z) = 0$, then we can write out

$$0 = \overline{\sum_{i=0}^{n} a_i z^i} = \sum_{i=0}^{n} a_i \overline{z^i} = f(\overline{z}).$$

For this proof it means that

$$\begin{aligned}
(x - z)(x - \overline{z}) &= x^2 - 2\mathrm{Re}(z)x + \underbrace{\|z\|^2}_{z\overline{z}} \\
(2\mathrm{Re}(z))^2 = 4\mathrm{Re}(z)^2 &< 4(\mathrm{Re}(z)^2 + \mathrm{Im}(z)^2) = 4c.
\end{aligned}$$

We have that polynomials splits like

$$f(x) = c\prod_{i=1}^{M}(x^2 + b_ix + c_i) \cdot \prod_{j=1}^{m}(x - \lambda_j).$$

and the given polynomial is written as

$$0 = \left(c\prod_{i=1}^{M}(T^2 + bT + cI)\prod_{j=1}^{m}(T - \lambda_jI) \right)\mathbf{v}$$

all of the quadratic parts (the first product operator) are invertible. This shows that $m \geq 1$ and (one of the λ_j is an eigenvalue) $T - \lambda_jI$ is not invertible for some j. ■

Proposition 11.2
Let $T \in \operatorname{End}(V)$, T is self-adjoint and $U \subset V$ is a subspace that is invariant under T then

1. $U^\perp$ is invariant under T
2. $T|_U \in \operatorname{End}(U)$ is self-adjoint
3. $T|_{U^\perp} \in \operatorname{End}(U^\perp)$ is self-adjoint.

Proof 11.4
1. In Exercise 7.A3 we showed that U is invariant under T iff $U^\perp$ is invariant under T^*.
2.

$$\begin{aligned}\left\langle T\Big|_U \mathbf{u}\middle|\mathbf{u}'\right\rangle &= \langle T\mathbf{u}|\mathbf{u}'\rangle \\ &= \langle \mathbf{u}|T\mathbf{u}'\rangle \\ &= \left\langle \mathbf{u}\middle|T\Big|_U \mathbf{u}'\right\rangle.\end{aligned}$$

■

Theorem 11.2
Let V be a $\mathbb{R}$-VS and $T \in \operatorname{End}(V)$. Then the following are equivalent

1. T is self-adjoint
2. T can be diagonalized with respect to some orthonormal basis of V.

Proof 11.5
We want to prove 2. $\implies$ 1..

If T has a diagonal matrix, then

$$\mathcal{M}(T)^\intercal = \mathcal{M}(T).$$

But $\mathcal{M}(T)^\intercal$ is the adjoint hence $T = T^*$.

Now we want to prove 1. $\implies$ 2..

We use induction on the dimension of V. If $\dim(V) = 1$ then all the linear operators are 1×1 matrices which are all diagonal and the theorem holds. Now assume $\dim(V) > 1$, then T has an eigenvector $\mathbf{u}$, with $\|\mathbf{u}\| = 1$ (by Proposition 11.1). We split $V = U \oplus U^\perp$ where $U = \operatorname{span}(\mathbf{u})$. By the previous Theorem $U^\perp$ is invariant under T and $T|_{U^\perp}$ is self-adjoint. Furtheremore $\dim\left(U^\perp\right) = n - 1$. By induction hypothesis there exists an orthonormal basis of $U^\perp$ such that $T|_{U^\perp}$ has a diagonal matrix. By adding $\mathbf{u}$ to this basis gives an orthonormal basis of V of eigenvectors of $T \Rightarrow T$ has a diagonal matrix with respect to this basis.

12 Matrix Factorization/Decompositions

Let

$$\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

be a vector of indeterminants. Then a system of linear equations can be written as

$$A\mathbf{x} = \mathbf{b}$$

where $A \in \mathcal{F}^{n \times n}$ and $\mathbf{b} \in \mathcal{F}^n$. If A is upper (lower) triangular it is easy to calculate a solution, if it exists and is unique.

$$A = \begin{pmatrix} a_{11} & * & * \\ 0 & \ddots & * \\ 0 & 0 & a_{nn} \end{pmatrix}$$

and

$$x_n = \frac{b_n}{a_{nn}},\ a_{nn} \neq 0.$$

Assume all $a_{ii} \neq 0$:

$$x_{n-1} = \frac{(b_{n-1} - a_{n-1}x_n)}{a_{n-1}}$$
$$\vdots$$
$$x_i = \frac{b_i - \sum_{j=i-1}^{n} a_{ij}x_j}{a_{ii}}.$$

Given a matrix $A = A^{(0)}$ we can use Gaussian Elimination to obtain an equivalent system of linear equations with an upper triangular matrix $A^{(n-1)}$.

Let us assume that $a_{11}^{(0)} \neq 0$ and is our pivot element, then we start by subtracting multiples of the first row from the other rows. Let

$$l_{i,1} := \frac{a_{i1}}{a_{11}} \text{ for } i \geq 2.$$

Then we define a new matrix $A^{(1)}$ as

$$A_{ij}^{(1)} = \begin{cases} A_{ij}^{(0)} & \text{if } i = 1 \\ A_{ij}^{(0)} - l_{i1}A_{1j}^{(0)} & \text{otherwise} \end{cases}.$$

Now we calculate the matrix product:

$$L_1A^{(0)} = A^{(1)} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ -l_{21} & 1 & \cdots & 0 \\ -l_{31} & 0 & 1 & 0 \\ \vdots & 0 & \cdots & \cdots \\ -l_{n1} & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ 0 & a_{22}^{(1)} & \\ \vdots & & * \\ 0 & & \end{pmatrix}$$

Assume that $a_{22}^{(1)}$ then $L_2A^{(1)} = A^{(2)}$ where L_2 is similar to L_1 with identity on the diagonal and then $-l_{i2}$ on the second column. Continuing we see that

$$L_{n-1}L_{n-2}\cdots L_2L_1A^{(0)} = A^{(n-1)}$$

is upper triangular.

Remarks regarding the L_i matrices:

1. $L_i^{-1} = \Lambda_i = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & l_{i+1,i} & 1 \end{pmatrix}$ (just flipped sign on the l_{ji})

2. $\prod_{j=1}^{n-1} \Lambda_j = \Lambda_1\Lambda_2\cdots\Lambda_{n-1} = \begin{pmatrix} 1 & 0 & 0 \\ l_{21} & 1 & 0 \\ l_{31} & l_{32} & 1 \end{pmatrix}$ (WLOG: just extend the matrix)

Consider a strictly lower triangular matrix

$$P_k = \begin{pmatrix} A & 0 \\ B & 0 \end{pmatrix}$$

where $A \in \mathcal{F}^{k\times k}$ and $B \in \mathcal{F}^{n-k\times k}$. Take now

$$\Lambda_{k+1} = \begin{pmatrix} I_k & 0 \\ 0 & C \end{pmatrix}$$

then

$$P_k\Lambda_{k+1} = \begin{pmatrix} A & 0 \\ B & 0 \end{pmatrix} = P_k.$$

Now the claim is

$$\prod_{j=1}^{k} \Lambda_j = P_k + I$$

$$\prod_{j+1}^{k+1} \Lambda_j = (P_k + I)\Lambda_{k+1} = P_k + \Lambda_{k+1}$$

and now

$$\prod L_j A = A^{(n-1)} = U$$
$$A = \underbrace{\prod_{j=1}^{n-1} \Lambda_j}_{L} U = LU.$$

Going back to the linear equations we have

$$A\mathbf{x} = LU\mathbf{x} = \mathbf{b},$$

solve $L\mathbf{y} = \mathbf{b}$ via substitution. Then solve $U\mathbf{x} = \mathbf{y}$ via substitution. Then

$$L\mathbf{y} = \mathbf{b} = LU\mathbf{x} = \mathbf{b},$$

which is easier for the computer and for inverses

$$A\mathbf{x}^{(iL)} = A\begin{pmatrix}\mathbf{x}^{(1)} & \cdots & \mathbf{x}^{(L)}\end{pmatrix}.$$

13 Cholesky Decomposition

Proposition 13.1 (3.7 in Arne's notes)
Let A be self-adjoint. Then the following are equivalent:

i) A is positive definite

ii) All eigenvalues of A are strictly positive

iii) All leading principal minors are positive

Proof 13.1

i) Want to show that i) $\Rightarrow$ ii): Assume $\mathbf{v}$ is eigenvector. Then

$$0 < \langle A\mathbf{v}|\mathbf{v}\rangle = \langle \lambda\mathbf{v}|\mathbf{v}\rangle = \lambda\langle\mathbf{v}|\mathbf{v}\rangle = \lambda\|\mathbf{v}\|^2$$

and then λ is positive.

ii) Want to show that ii) $\Rightarrow$ i): Since A is self-adjoint we have an unitary (orthogonal for reals) matrix U such that

$$A = UDU^*$$

where D is diagonal with the eigenvalues on the diagonal. Let $\mathbf{x} \in V$, $\mathbf{x} \neq 0$, then

$$\langle A\mathbf{x}|\mathbf{x}\rangle = \langle UDU^*\mathbf{x}|\mathbf{x}\rangle = \langle DU^*\mathbf{x}|U^*\mathbf{x}\rangle = \left\langle \sqrt{D}^2U^*\mathbf{x}\middle|U^*\mathbf{x}\right\rangle$$

because D is diagonal with positive eigenvalues on the diagonal. Continuing we have

$$\left\langle \sqrt{D}^2U^*\mathbf{x}\middle|U^*\mathbf{x}\right\rangle = \left\langle \sqrt{D}U^*\mathbf{x}\middle|\sqrt{D}U^*\mathbf{x}\right\rangle = \left\|\sqrt{D}U^*\mathbf{x}\right\|^2 > 0.$$

iii) Want to show that i) $\Rightarrow$ iii): Let A_k be the $k \times k$ leading principal submatrix of A. By Proposition 3.5 (Arne's notes) A_k is positive definite. The determinant of A_k is the product of its eigenvalues. Since A is positive definite so is A_k and it implies ii): all eigenvalues of A_k are positive which imply that the product of all eigenvalues of A_k is positive.

iv) Want to show that iii) $\Rightarrow$ i): Induction on the dimension m of A. For $m = 1$; $A = \begin{pmatrix} a_{11} \end{pmatrix}$ and it i), ii) and iii) holds. Let $A \in \mathcal{F}^{m+1\times m+1}$ and A_m is the leading principal $m \times m$ submatrix. Since A is self-adjoint so is A_m. All the leading principal minors of A_m (a subset of the leading minors of A) are positive hence by induction hypothesis we know that A_m is positive definite. $\Rightarrow$ all eigenvalues of A_m are positive. We know there exists an unitary (orthogonal) matrix U such that

$$A_m = UDU^*$$

with D being diagonal with eigenvalues on the diagonal (all of which are positive). Now consider

$$Q = \begin{bmatrix} U & 0_{m\times 1} \\ 0_{1\times m} & 1 \end{bmatrix}$$

and

$$B = QAQ^* = \begin{bmatrix} U & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} A_m & b \\ \bar{b} & b_{m+1} \end{bmatrix} \begin{bmatrix} U^* & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 & 0 & 0 & b_1 \\ 0 & 0 & \lambda_2 & 0 & b_2 \\ 0 & 0 & 0 & \lambda_3 & b_3 \\ \bar{b}_1 & \bar{b}_2 & \bar{b}_3 & b_4 & \end{bmatrix}.$$

The determinant of B is then

$$\begin{vmatrix} \lambda_1 & 0 & 0 & 0 & b_1 \\ 0 & 0 & \lambda_2 & 0 & b_2 \\ 0 & 0 & 0 & \lambda_3 & b_3 \\ 0 & 0 & 0 & 0 & b_4 - \sum_{i=1}^m \frac{b_i \bar{b}_i}{\lambda_i} \end{vmatrix} = \prod_{i=1}^m \lambda_i \left(b_{m+1} - \sum_{i=1}^m \frac{|b_i|^2}{\lambda_i} \right) > 0$$

and hence

$$b_{m+1} - \sum_{i=1}^m \frac{|b_i|^2}{\lambda_i} > 0.$$

Now let $\mathbf{x} \neq 0$ then

$$\begin{aligned} \langle B\mathbf{x} | \mathbf{x} \rangle &= \left(x_1, \dots, x_{m+1} \right) B^\intercal \begin{pmatrix} x_1 \\ \vdots \\ x_{m+1} \end{pmatrix} \\ &= \lambda_1 x_1^2 + \bar{b}_1 x_1 x_{m+1} + \lambda_2 x_2^2 + \bar{b}_2 x_2 x_{m+1} + \ldots + \lambda_m x_m^2 \\ &+ \bar{b}_m x_m x_{m+1} + b_1 x_{m+1} x_1 + b_2 x_{m+1} x_2 + \ldots + b_m x_{m+1} x_m + b_{m+1} x_{m+1}^2 \\ &= \lambda_1 \left(x_1^2 + \frac{2\,\mathrm{Re}(b_1)}{\lambda_1} (x_1 x_{m+1}) + \frac{\mathrm{Re}(b_1)^2 x_{m+1}^2}{\lambda_1^2} \right) \\ &- \frac{\mathrm{Re}(b_1)^2 x_{m+1}^2}{\lambda_1} \cdots \\ &= \sum_{i=1}^m \lambda_i \left(x_i + \frac{\mathrm{Re}(b_i) x_{m+1}}{\lambda_i} \right)^2 \\ &- x_{m+1}^2 \sum_{i=1}^m \frac{\mathrm{Re}(b_i)^2}{\lambda_i} + b_{m+1} x_{m+1}^2. \end{aligned}$$

We want to show that this is positive. The eigenvalues are positive and the square is obviously positive. We then have

$$x_{m+1}^2 \left(b_{m+1} - \sum_{i=1}^m \frac{\mathrm{Re}(b_i)^2}{\lambda_i} \right) \geq x_{m+1}^2 \left(b_{m+1} - \sum_{i=1}^m \frac{|b_i|^2}{\lambda_i} \right) > 0.$$

Theorem 13.1 (3.11 in Arne's notes)
Let $A \in \mathcal{F}^{m \times m}$ be positive definite. Then there exists a unique upper triangular R with positive entries along the diagonal such that $A = R^* R$.

Proof 13.2
Let $m > 1$. We proceed recursively to find R. Let

$$A = \begin{bmatrix} a_{11} & b^*_{1\times m-1} \\ b_{m-1\times 1} & B_{m-1\times m-1} \end{bmatrix}$$

with $a_{11} > 0$. Now

$$\begin{bmatrix} 1 & 0 \\ -\frac{1}{a_{11}} b_{1\times m-1} & I \end{bmatrix} A = \begin{bmatrix} a_{11} & b^*_{m-1\times 1} \\ 0 & \tilde{B} \end{bmatrix}.$$

Then

$$A = \begin{bmatrix} 1 & 0 \\ \frac{1}{a_{11}} b & I \end{bmatrix} \begin{bmatrix} a_{11} & b^* \\ 0 & B - \frac{1}{a_{11}} b \cdot b^* \end{bmatrix}$$

resulting in

$$A = \begin{bmatrix} 1 & 0 \\ \frac{1}{a_{11}} b & I \end{bmatrix} \begin{bmatrix} a_{11} & 0 \\ 0 & B - \frac{1}{a_{11}} b \cdot b^* \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{a_{11}} b^* \\ 0 & I \end{bmatrix}.$$

To get $A = LU$ the diagonal matrix in the middle has to be the identity and therefore the square root of a_{11} is taken:

$$A = \begin{bmatrix} 1 & 0 \\ \frac{1}{\sqrt{a_{11}}} b & I \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & B - \frac{1}{\sqrt{a_{11}}} b \cdot b^* \end{bmatrix} \begin{bmatrix} 1 & \frac{1}{\sqrt{a_{11}}} b^* \\ 0 & I \end{bmatrix}.$$

We then call the first matrix (the lower triangular one) for R_1^* and the last matrix (the upper triangular one) for R_1. The diagonal matrix in the middle is called A_2. We now have

$$(R_1^{-1})^* A R_1^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & B - \frac{1}{a_{11}} bb^* \end{bmatrix}$$

which is positive definite. Therefore $B - \frac{1}{a_{11}} bb^*$ is positive definite. Repeat the process for A_2 (or $B - \frac{1}{a_{11}} bb^*$) to get matrices

$$R_2 = \begin{bmatrix} 1 & 0 \\ 0 & \tilde{R}_2 \end{bmatrix}$$

such that

$$A_2 = R_2^* \begin{bmatrix} 1 & & 0 \\ 0 & 1 & 0 \\ 0 & \tilde{A}_3 & \end{bmatrix} R_2.$$

After m steps we get

$$A = (R_1^* R_2^* R_3^* \ldots R_m^*) I (R_m \ldots R_3 R_2 R_1) = R^* R.$$

14 Singular Value Decomposition

From now on V and W are finite dimensional inner product spaces.

Proposition 14.1 (5.1 in Arne's notes)
$A \in \text{Hom}(V, W)$. Then A^*A and AA^* have the same non-zero eigenvalues with the same (geometric) multiplicities. If $\mathbf{v}$ is an eigenvector of A^*A corresponding to a non-zero eigenvalue λ then $A\mathbf{v}$ is an eigenvector of AA^* for the same eigenvalue.

Proof 14.1
Let $\lambda \neq 0$ be an eigenvalue of A^*A and $\mathbf{v} \neq 0$ the corresponding eigenvector. Then

$$A^*A\mathbf{v} = \lambda\mathbf{v} \Rightarrow AA^*(A\mathbf{v}) = \lambda(A\mathbf{v})$$

and $A\mathbf{v} \neq 0$ since then that would imply that $\lambda = 0$ which would be a contradiction since we assumed that it is non-zero. From the equation we see that A^*A and AA^* have the same (non-zero) eigenvalue.

Assume the multiplicity of λ is m, i.e. there is a basis $\mathbf{v}_1, \ldots, \mathbf{v}_m$ of the eigenspace. We then have

$$\ker(A^*A - \lambda I)$$

where $A\mathbf{v}_1, \ldots, A\mathbf{v}_m$ is a basis for $\ker(AA^* - \lambda I)$. To show linear independence assume

$$\begin{aligned} c_1 A\mathbf{v}_1 + \ldots + c_m A\mathbf{v}_m &= 0 \\ \Rightarrow c_1 A^*A\mathbf{v}_1 + \ldots + c_m A^*A\mathbf{v}_m &= 0 \\ \Leftrightarrow \lambda(c_1\mathbf{v}_1 + \ldots + c_m\mathbf{v}_m) &= 0 \\ \lambda \neq 0 \Rightarrow c_1\mathbf{v}_1 + \ldots + c_m\mathbf{v}_m &= 0 \\ \Rightarrow c_i &= 0 \; \forall i \end{aligned}$$

because $\mathbf{v}_i$ are a basis. This all implies that $\text{span}(A\mathbf{v}_1, \ldots, A\mathbf{v}_m) \subseteq \ker(AA^* - \lambda I)$, but V and W have different dimensions, however, by symmetry every eigenvector to λ of AA^* induces an eigenvector $A^*\mathbf{w}$ of A^*A. Let $\mathbf{w}_1, \ldots, \mathbf{w}_n$ be a basis for $\ker(AA^* - \lambda I)$ implies $\text{span}(A^*\mathbf{w}_1, \ldots, A^*\mathbf{w}_n) \subseteq \ker(A^*A - \lambda I)$ and so they sit in each other. We have

$$\begin{aligned} \dim(\ker(A^*A - \lambda I)) &\leq \dim(\ker(AA^* - \lambda I)) \\ &\geq \\ &= \end{aligned}$$

completing the proof.

Proposition 14.2 (5.3 in Arne's notes)
Let $A \in \text{Hom}(V, W)$.

i) The eigenvalues of A^*A are non-negative (≥ 0)

ii) $\ker(A^*A) = \ker(A)$ and $\ker(AA^*) = \ker(A^*)$

Proof 14.2
Assume that $A^*A\mathbf{v} = \lambda\mathbf{v},\ \mathbf{v} \neq 0$ then

$$\begin{aligned}\lambda\|\mathbf{v}\|^2 = \lambda\langle\mathbf{v}|\mathbf{v}\rangle = \langle\lambda\mathbf{v}|\mathbf{v}\rangle = \langle A^*A\mathbf{v}|\mathbf{v}\rangle \\ = \langle A\mathbf{v}|A\mathbf{v}\rangle = \|A\mathbf{v}\|^2 \geq 0 \Rightarrow \lambda \geq 0.\end{aligned}$$

ii): We see that $\ker(A) \subseteq \ker(A^*A)$. Let $\mathbf{v} \in \ker(A^*A)$ then

$$\begin{aligned}0 &= \langle A^*A\mathbf{v}|\mathbf{v}\rangle = \langle A\mathbf{v}|A\mathbf{v}\rangle = \|A\mathbf{v}\|^2 \\ \Rightarrow &A\mathbf{v} = 0 \quad \ker(A^*A) \subseteq \ker(A).\end{aligned}$$

Proposition 14.3 (5.6 in Arne's Notes)
Let $A \in \mathrm{Hom}(V, W)$. Then $\mathrm{rank}(A) = \dim(\mathrm{im}(A)) = \mathrm{rank}(A^*A)$, and $\mathrm{rank}(A) = \mathrm{rank}(A^*)$.

Proof 14.3

$$\begin{aligned}\dim(\mathrm{im}(A)) &= \dim(V) - \dim(\ker(A)) \\ &= \dim(V) - \dim(\ker(A^*A)) \\ &= \dim(\mathrm{im}(A^*A)) = \mathrm{rank}(A^*A).\end{aligned}$$

Analogous:

$$\mathrm{rank}(A^*) = \mathrm{rank}(AA^*).$$

Since AA^* is self-adjoint there exists an orthonormal basis of W of eigenvectors of AA^*. Since AA^* and A^*A have the same non-zero eigenvalues with the same multiplicity. Hence their images have the same dimension, which implies that

$$\mathrm{rank}(AA^*) = \mathrm{rank}(A^*A) \Rightarrow \mathrm{rank}(A) = \mathrm{rank}(A^*).$$

Proposition 14.4 (5.7 in Arne's Notes)
Let $A \in \mathrm{Hom}(V, W)$. Then

$$(\mathrm{A})^{\perp} = \ker(A^*).$$

Proof 14.4
Let $\mathbf{w} \in \ker(A^*)$ and $\mathbf{v} \in V$.

$$\begin{aligned}&\langle A\mathbf{v}|\mathbf{w}\rangle = \langle \mathbf{v}|A^*\mathbf{w}\rangle = \langle \mathbf{v}|0\rangle = 0 \\ &\ker(A^*) \subseteq (\mathrm{im}(A))^{\perp}\end{aligned}$$

but this is not necessarily the only one doing this. To show uniqueness let $\mathbf{w}' \in (\mathrm{im}\, A)^{\perp}$ then we have

$$\forall \mathbf{v} \in V : \langle A\mathbf{v}|\mathbf{w}'\rangle = 0 = \langle \mathbf{v}|A^*\mathbf{w}'\rangle$$

and if it is orthogonal to everything then it must be zero:

$$\Rightarrow A^*\mathbf{w}' = 0 \Rightarrow \mathbf{w}' \in \ker(A^*) \Rightarrow \mathrm{im}(A)^{\perp} \subseteq \ker(A^*).$$

Definition 14.1 (Singular Values: 5.4 in Arne's Notes)
Let $A \in \text{Hom}(V, W)$. We call the positive square root of the non-zero eigenvalues of A^*A the singular values of A. They are denoted by $\sigma_i(A)$ where $\sigma_i(A) \geq \sigma_2(A) \geq \ldots$ and every singular value appears with the corresponding multiplicity of the eigenvalue.

Theorem 14.1 (5.8 in Arne's Notes)
Let $A \in \text{Hom}(V, W)$ and $\text{rank}(A) = r$, and $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$ are the singular valies of A. Then there exists an orthonormal basis $\mathbf{v}_1, \ldots, \mathbf{v}_n$ of V and orthonormal vectors $\mathbf{w}_1, \ldots, \mathbf{w}_r$ in W such that

$$A\mathbf{x} = \sum_{j=1}^{r} \sigma_j \langle \mathbf{x}|\mathbf{v}_j \rangle \mathbf{w}_j \; \forall \mathbf{x} \in V.$$

Proof 14.5
Since A^*A is self-adjoint (which is awesome) then there exists a basis of orthonormal eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ where the $\mathbf{v}_j$ are sorted by descending size of the corresponding eigenvalue i.e.

$$A^*A\mathbf{v}_j = \begin{cases} \sigma_j^2 \mathbf{v}_j & \text{for} 1 \leq j \leq r \\ 0 & \text{for} j > r. \end{cases}$$

Now let $vbw_j := \frac{1}{\sigma_j} A\mathbf{v}_j$ for $1 \leq j \leq r$. We see that

$$\begin{aligned} \langle \mathbf{w}_j|\mathbf{w}_k \rangle &= \frac{1}{\sigma_j\sigma_k} \langle A\mathbf{v}_j|A\mathbf{v}_k \rangle \\ &= \frac{1}{\sigma_j\sigma_k} \langle \mathbf{v}_j|A^*A\mathbf{v}_k \rangle \\ &= \frac{\sigma_k^2}{\sigma_j\sigma_k} \langle \mathbf{v}_j|\mathbf{v}_k \rangle \\ &= \begin{cases} 1 & \text{if} k = j \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Hence the vectors $\mathbf{w}_j$ are orthonormal.

$$\forall \mathbf{x} \in V : \mathbf{x} = \sum_{j=1}^{n} \langle \mathbf{x}|\mathbf{v}_j \rangle \mathbf{v}_j$$

because the basis is orthonormal. Now remember that $A^*A\mathbf{v} = 0 \Rightarrow A\mathbf{v} = 0$, then

$$\begin{aligned} A\mathbf{x} &= \sum_{j=1}^{n} \langle \mathbf{x}|\mathbf{v}_j \rangle A\mathbf{v}_j \\ &= \sum_{j=1}^{r} \langle \mathbf{x}|\mathbf{v}_j \rangle A\mathbf{v}_j \\ &= \sum_{j=1}^{r} \sigma_j \langle \mathbf{x}|\mathbf{v}_j \rangle \mathbf{w}_j. \end{aligned}$$

Remark: Let us consider $\mathcal{M}(A, \mathbf{v}_1, \ldots, \mathbf{v}_n, \mathbf{w}_1, \ldots, \mathbf{w}_m)$: A has zeros for rows after row r and for columns after column r. Otherwise it has the singular values on the diagonal (of the $r \times r$) block of A). We have

$$\begin{aligned}
A\mathbf{v}_j &= \sigma_j \mathbf{w}_j \; 1 \leq j \leq r \\
A\mathbf{x} &= \sum_{j=1}^{r} \sigma_j \langle \mathbf{x} | \mathbf{v}_j \rangle = \begin{bmatrix} \mathbf{w}_1 & \ldots & \mathbf{w}_r \end{bmatrix} \begin{bmatrix} \sigma_1 \langle \mathbf{x} | \mathbf{v}_1 \rangle \\ \vdots \\ \sigma_r \langle \mathbf{x} | \mathbf{v}_r \rangle \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{w}_1 & \ldots & \mathbf{w}_r \end{bmatrix} \begin{bmatrix} \sigma_1 & \ldots & 0 \\ 0 & \ddots & \vdots \\ 0 & \ldots & \sigma_r \end{bmatrix} \begin{bmatrix} \langle \mathbf{x} | \mathbf{v}_1 \rangle \\ \vdots \\ \langle \mathbf{x} | \mathbf{v}_r \rangle \end{bmatrix} \\
&= \begin{bmatrix} \mathbf{w}_1 & \ldots & \mathbf{w}_r \end{bmatrix} \operatorname{diag}(\sigma_j) \begin{bmatrix} \mathbf{v}_1 & \ldots & \mathbf{v}_r \end{bmatrix}^* \mathbf{x} \\
&= U \Sigma V \\
A &= U \Sigma V^*.
\end{aligned}$$

To determine U, Σ and V^* we need to find the eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_r$ of A^*A. Now let us look at sizes: A is an $m \times n$ matrix. Then U is an $m \times r$ matrix, Σ is an $r \times r$ matrix, and V^* is an $r \times n$ matrix. This can also be done by having U' as an $m \times m$ matrix, Σ' as an $m \times n$ matrix, and V'^* as an $n \times n$ matrix. This is regarded as the full SVD. The full SVD is achieved by using all the eigenvectors. The diagonal matrix becomes the earlier diagonal matrix but extended with zeros.

Example 14.1

$$A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}$$

Before checking the eigenvalues the size of A^*A is determined, which is 3×3, but A has at most rank $= 2$. Instead the eigenvalues of AA^* are determined.

$$AA^* = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \begin{bmatrix} 3 & 2 & & \\ 2 & 3 & 2 & -2 \end{bmatrix} = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$$

$$\begin{aligned}
\chi_{AA^*}(x) = (17 - x)^2 - 64 &= \begin{vmatrix} 17 - x & 8 \\ 8 & 17 - x \end{vmatrix} \\
&= x^2 - 34x + 225 = (x - 25)(x - 9) \\
\sigma_1 = 5 \; \sigma_2 &= 3 \\
\Sigma &= \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix}
\end{aligned}$$

and now we need the eigenvectors.

$$
\begin{aligned}
A^*A &= \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix} \\
A^*A\mathbf{v} &= 25\mathbf{v} \\
(A^*A - 25I)\mathbf{v} &= 0 \\
\begin{bmatrix} -12 & 12 & 2 \\ 12 & -12 & -2 \\ 2 & -2 & -17 \end{bmatrix} \underbrace{\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}}_{\tilde{\mathbf{v}}_1} &= 0 \\
\mathbf{v}_1 &= \frac{1}{\sqrt{2}}\tilde{\mathbf{v}}_1.
\end{aligned}
$$

$$
\begin{aligned}
\begin{bmatrix} 4 & 12 & 2 \\ 12 & 4 & -2 \\ 2 & -2 & -1 \end{bmatrix} \to \begin{bmatrix} 0 & 16 & 4 \\ 0 & 16 & 4 \\ 2 & -2 & -1 \end{bmatrix} \underbrace{\begin{bmatrix} -1 \\ 1 \\ -4 \end{bmatrix}}_{\tilde{\mathbf{v}}_2} &= 0 \\
\mathbf{v}_2 &= \frac{1}{\sqrt{16}}\tilde{\mathbf{v}}_2.
\end{aligned}
$$

Now

$$
V = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{18}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{18}} \\ 0 & -\frac{4}{\sqrt{12}} \end{bmatrix}.
$$

$$
\begin{aligned}
\tilde{\mathbf{w}}_1 &= \frac{1}{\sigma_1}A\tilde{\mathbf{v}}_1 = \frac{1}{5}\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
\mathbf{w}_1 &= \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \\
\tilde{\mathbf{w}}_2 &= \frac{1}{\sigma_2}A\tilde{\mathbf{v}}_2 = \frac{1}{3}\begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \begin{bmatrix} -1 & 1 & -4 \end{bmatrix} = \begin{bmatrix} -9 \\ 9 \end{bmatrix} \\
\mathbf{w}_2 &= \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix},
\end{aligned}
$$

and the vectors $\mathbf{w}_1$ and $\mathbf{w}_2$ are orthogonal.

$$
U = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix},
$$

now check if you get A:

$$\begin{aligned} U\Sigma V^* &= \frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\1 & 1\end{bmatrix}\begin{bmatrix}5 & 0\\0 & 3\end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & -\frac{1}{3}\\1 & \frac{1}{3}\\0 & -\frac{4}{3}\end{bmatrix}^* \\ &= \frac{1}{2}\begin{bmatrix}1 & -1\\1 & 1\end{bmatrix}\begin{bmatrix}5 & 5 & 0\\-1 & 3 & 4\end{bmatrix} \\ &= \frac{1}{2}\begin{bmatrix}6 & 4 & 4\\4 & 6 & -4\end{bmatrix} \\ &= \begin{bmatrix}3 & 2 & 2\\2 & 3 & -2\end{bmatrix} = A. \end{aligned}$$

The full SVD is:

$$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\\1 & 1\end{bmatrix}\begin{bmatrix}5 & 0 & 0\\0 & 3 & 0\end{bmatrix}\begin{bmatrix}\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & \\ -\frac{1}{\sqrt{18}} & \frac{1}{\sqrt{10}} & -\frac{4}{\sqrt{18}}\\ \frac{2}{3} & -\frac{2}{3} & -\frac{1}{3}\end{bmatrix}.$$

To create examples yourself make Σ and choose sensible singular values.