Chapter 1

Cayley Hamilton Theorem

We will build up the machinery needed to prove the following:

Theorem. Cayley Hamilton Theorem

Given a linear operator $T : V \to V$ with $\dim(v) = n$, then

$$P_T(T) = \mathbb{0}_{n\times n}.$$

where P_T is the characteristic polynomial of T.

Proof

1.1 Invariant Subspaces

Def. Invariant Subspace

Given a linear operator $T : V \to V$ and subspace $W \subseteq V$, if $T[W] \subseteq W$ then

$$T|_W : W \to W$$

is a linear operator on W and W is T invariant.

Example 1. Consider some eigenvalue λ of T. Then there is a subspace E_λ of V associated with that eigenvalue. Taking any $v \in E_\lambda$, note that by definition $Tv = \lambda v \in E_\lambda$. Therefore E_λ is an invariant subspace for any eigenvalue λ of T.

Theorem 2.

If $W \subseteq V$ is an invariant subspace under T, then for $T|_W : W \to W$ we have

$$P_{T|_W}(t) \mid P_T(t).$$

Proof. Let $\beta_w = \{w_1, \ldots, w_k\}$ be a basis of W and $\beta = \beta_w \cup \{v_{k+1}, \ldots, v_n\}$ be a basis of V where w_i form a basis of W. Then the matrix form of T is

$$[T]_\beta = \begin{pmatrix} B_1 & B_2 \\ 0 & B_3 \end{pmatrix}$$

where B_1 is $[T|_W]_{\beta_w}$. By subtracting tI from both sides and taking the determinant, we get

$$\det(T - tI) = \det(B_1 - tI)\det(B_3 - tI).$$

But this is just

$$P_T(t) = P_{T|_W}(t) \cdot q(t)$$

with $q(t) = \det(B_3 - tI)$. ◇

1.1.1 Generating Invariant Subspaces

Consider some linear operator T on a finite dimensional space V with $\dim V = n$. Then note for any $v \in V$ that

$$\{0, Tv, T^2 v, \ldots\}$$

must be a linearly dependent set of vectors. If this wasn't the case, then repeated applications of T would produce infinitely many linearly independent vectors within V. Therefore there is some $k \leq n$ such that

$$\{0, Tv, T^2 v, \ldots, T^{k-1} v\}$$

is linearly independent. The span of this set gives a subspace W that is T invariant, something analogous to cyclic groups in group theory. This motivates the following definition.

Def. Cyclic Subspace

Let T be a linear operator on V and $v \in V$. Then the subspace

$$W = \text{span}\left\{0, Tv, T^2v, T^3v, \ldots\right\}$$

is the **T-cyclic subspace of V generated by** v.

Example 2. Consider the operator $T : P(\mathbb{R}) \to P(\mathbb{R})$ with $T(p(x)) = p'(x)$. Starting with x^3, we see that

$$\left\{0, Tx^3, T^2x^3, \ldots\right\} = \left\{0, 3x^2, 6x, 6\right\}.$$

The span of this set then is then $P_3(\mathbb{R})$ which is invariant under T.

Theorem 3.

If $a_0 + a_1Tv + a_2T^2v + \ldots + a_{k-1}T^{k-1}v + T^kv = 0$, then

$$P_{T|_W}(t) = (-1)^k\left(a_0 + a_1t + \ldots a_{k-1}t^{k-1} + t^k\right).$$

Proof. We consider the cyclic invariant subspace W spanned by the basis $\beta = \left\{v, Tv, T^2v, \ldots, T^{k-1}v\right\}$. If $w \in W$, then we know that

$$w = a_0v + a_1Tv + a_2T^2v + \ldots + a_kT^{k-1}v$$

which gives

$$Tw = a_0Tv + a_1T^2v + \ldots + a_kT^kv.$$

◇

1.1.2 The Proof

Proof of Cayley Hamilton Theorem. Let T be a linear operator on V, $v \in V \neq 0$, and W be the cyclic subspace generated by v. ◇

Chapter 2

Inner Products and Norms

When working in $\mathbb{R}^n$, there is the familar idea of the scalar/dot product. Given two vectors $\mathbb{x}$ and $\mathbb{y}$ then their scalar product is

$$x \cdot y = \sum_{i=1}^{n} x_i y_i.$$

The concept of euclidean length is also captured by scalar products via

$$\sqrt{x \cdot x} = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

This scalar product on $\mathbb{R}^n$ does not generalize to other vector spaces, or it may not be a useful notion of length/product of vectors even when working in $\mathbb{R}^n$. Therefore it is useful to generalize this notion of a scalar product.

Def. Inner Product

A mapping $\langle \cdot, \cdot \rangle : V \times V \to F$ is an inner product if for all $x, y \in V$ and $s \in F$

1. $\langle x + z, y \rangle = \langle x, y \rangle + \langle z, y \rangle$ for all $z \in V$
2. $\langle sx, y \rangle = s\langle x, y \rangle$
3. $\overline{\langle x, y \rangle} = \langle y, x \rangle$
4. $\langle x, x \rangle > 0$ when $x \neq 0$

Example 3. The vector space $M_{n\times n}(\mathbb{R})$ of real n by n matrices can be endowed with an inner product where $\langle A, B \rangle = \operatorname{tr} B^t A$.

Example 4. The vector space $C([0, 2\pi])$ of continuous complex functions on the interval 0 to 2π can be endowed with an inner product where

$$\langle f, g \rangle = \int_0^{2\pi} f(t)\overline{g(t)}\mathrm{d}t.$$

An important concept that can be generalized from $\mathbb{R}^n$ is orthogonality. It is common to compare the scalar product of two vectors to 0 to determine if they are orthogonal or not. This motivates a generalized notion of orthogonality.

Def. Orthogonal Vectors

Let V be a vector space equipped with an inner product $\langle \cdot, \cdot \rangle$. Then $x, y \in V$ are orthogonal if $\langle x, y \rangle = 0$.

Example 5. Consider from 4 the family of functions $f_m(t) = e^{imt}$. Then for any f_m, f_n

$$\begin{aligned}\langle f_m, f_n\rangle &= \int_0^{2\pi} f_m(\tau)\overline{f_n(\tau)}\mathrm{d}\tau \\ &= \int_0^{2\pi} e^{i(m-n)\tau}\mathrm{d}\tau \\ &= \left.\frac{e^{i(m-n)\tau}}{i(m-n)}\right|_0^{2\pi} = 0.\end{aligned}$$

Hence all f_m are orthogonal to each other.

Def. Vector Norm

Let V be a vector space equipped with an inner product $\langle\cdot,\cdot\rangle$. Then the **norm or length** of x is

$$\|x\| = \sqrt{\langle x, x\rangle}.$$

Theorem. Cauchy-Schwarz Inequality

For any vector space V with an inner product $\langle\cdot,\cdot\rangle$ and $x, y \in V$,

$$|\langle x, y\rangle| \leq \|x\|\|y\|.$$

Proof. ◇

The triangle inequality then follows quickly from Cauchy-Schwarz.

Theorem. Triangle Inequality

For any vector space V with an inner product $\langle\cdot,\cdot\rangle$ and $x, y \in V$,

$$\|x + y\| \leq \|x\| + \|y\|.$$

This means that for any inner product on a space, it has its own version of a triangle inequality. This offloads the burden of proving directly that a norm satisfies the triangle inequality to finding some notion of an inner product that gives rise to that norm.

2.1 A General Notion of Norms

It is important to note that while every inner product gives rise to a norm, not every norm can be reverse engineered into a inner product.

Example 6. On the vector space $M_{n\times n}(\mathbb{R})$,

$$\|A\| = \sup_{\substack{x\in\mathbb{R}^n \\ \|x\|\leq 1}} Ax$$

defines a norm, but there exists no inner product that gives rise to it.

Def. Generalized Norm

Let V be a vector space. Then a norm $\|\cdot\| : V \to \mathbb{R}$ satisfies $\forall x, y \in V$ and $s \in C$

1. $\|x\| \geq 0$ and $\|x\| = 0 \Leftrightarrow x = 0$
2. $\|sx\| = |s|\|x\|$
3. $\|x + y\| \leq \|x\| + \|y\|$ $(\star)$

Example 7. The map

$$\|x\|_\infty := \max_{i\in\{1,\ldots,n\}} |x_i|$$

on $\mathbb{R}^n$ is a norm. Consider the requirements to be a norm

1. Since the norm takes the maximum of the absolute value of each component, the norm will be a non negative result, meaning $\|x\|_\infty \geq 0$. If the norm is 0, then the largest term in magnitude was 0, hence $x = 0$. The reverse follows easily.

2. With $s \in C$

$$\begin{aligned} \|sx\|_\infty &= \max_{i\in\{1,\ldots,n\}} |sx_i| \\ &= |s| \max_{i\in\{1,\ldots,n\}} |x_i| \\ &= |s|\|x\|_\infty. \end{aligned}$$

3. The triangle inequality follows from the triangle inequality on the reals and the linearity of the maximum function.

There is a famous and important class of norms defined on euclidean space known as the p-norms. They give rise to L^p spaces which are crucial to functional analysis.

Def. L_p Norm

Given $p \in$, the map

$$L_p(x) := \sum_i \left(|x_i|^p\right)^{\frac{1}{p}}$$

is a norm for any $\mathbb{R}^n$.

Chapter 3

Orthogonality

When a vector space has an inner product, there is a notion of orthgonality as was defined in Inner Products and Norms. Orthogonality of vectors tends to make computations and proofs simpler, hence building and working in an orthogonal basis is advantageous. Imposing normality of the basis further improves the situation.

Def. Orthogonal Basis

A basis $\beta = \{v_1, \ldots, v_n\}$ of a vector space V with inner product $\langle \cdot, \cdot \rangle$ is an **orthonormal basis** if $\|v_i\| = 1$ and $\langle v_i, v_j \rangle = 0$ for $i \neq j$.

Theorem 6.

Suppose $\beta = \{v_1, \ldots, v_n\}$ is an orthonormal basis of some vector space V. Then for any $x \in V$

$$x = \sum_i \langle x, v_i \rangle v_i.$$

Proof. Since β is a basis, $x \in V$ can be written as

$$x = \sum_i a_i v_i$$

for scalars a_i. Then note

$$\begin{aligned}\langle x, v_i\rangle = \left\langle \sum_j a_j v_j, v_i \right\rangle &= \sum_j \langle a_j v_j, v_i\rangle \\ &= \sum_j a_j \langle v_j, v_i\rangle \\ &= a_i \langle v_i, v_i\rangle \\ &= a_i\end{aligned}$$

Substituting the expression for each a_i gives the desired result. ◇

Theorem 7.

Any set of non-zero orthogonal vectors is linearly independent.

Proof. Let $\{v_1, \ldots, v_k\}$ be a set of orthogonal vectors with $v_i \neq 0$. Assume towards contradiction that this set is not linearly independent. Then there exists scalars a_i such that

$$\sum_i a_i v_i = 0.$$

Therefore at least one a_i is non-zero. Note that for any v_j

$$\left\langle \sum_i a_i v_i, v_j \right\rangle = a_j \left\| v_j \right\|^2$$

from the previous proof. But at the same time

$$\left\langle \sum_i a_i v_i, v_j \right\rangle = \langle 0, v_j\rangle = 0$$

meaning $a_j \left\| v_j \right\|^2 = 0$. Since v_j is non-zero, then $a_j = 0$. However, this is true for any j meaning all a_i must be zero, a contradiction. ◇