When MAT240 Gets Confusing...

Jack Ceroni

December 2020

Contents

1 Introduction

I do spend a little bit of time discussing concepts that any reader (who took MAT240) will likely already be familiar with (invariant subspaces, eigenspaces, etc.), but I want to keep these notes fairly self-contained.

If you notice any mistakes in these notes, please do not hesitate to send me an email or send me a message on the Math Physics Specialist Discord server.

2 What Are These Notes About?

These notes are meant to sequentially explain how we go about understanding linear operators on vector spaces by "breaking them up" into simpler, more understandable pieces, choosing a nice basis for each of these peices, and then putting them back together.

Formally, these concepts are know as the **decomposition theorem** and **Jordan Canonical Form**.

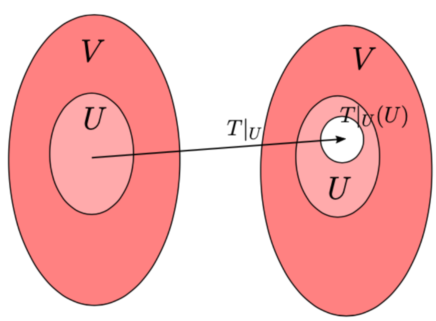
3 Invariant Subspaces

Before we begin our discussion of the decomposition theorem, we first must briefly cover the idea of **invariant subspaces**. Simply put, an invariant subspace U is a subspace of the vector space V that behaves "nicely" under an operator T on V.

More specifically:

Definition 1. *An invariant subspace in defined as a subspace U of V such that if $u \in U$, then $Tu \in U$.*

A consequence of this definition is that the operator $T|_U \in \mathcal{L}(U)$: the restriction of T to the domain U inside of V is well-defined (consider the alternative, if U wasn't invariant then there would exist some $u \in U$ that gets sent somewhere outside of U by $T|_U$, so $T|_U$ wouldn't be a valid map from U to U).



4 Eigenvalues and Eigenvectors

Now that the idea of an invariant subspace has been introduced, we can discuss some simple examples of invariant subspaces.

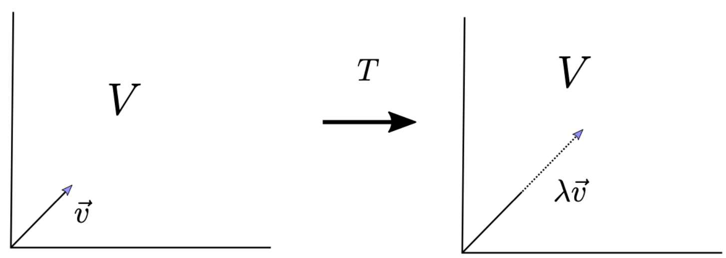
Given an operator T, some fairly obvious examples of spaces that are invariant are Null T and Range T. However, let's consider an even simpler example: a one-dimensional invariant subspace.

Let U be a one-dimensional invariant subspace. Clearly, there is only one linearly independent vector $u \in U$, so U is the space of all vectors of the form cu, for $c \in \mathbb{F}$.

Since T is invariant, we must then have $T(u) = \lambda u$ for some $\lambda \in \mathbb{F}$. It then follows that for any other cu in U, we will have $T(cu) = cT(u) = c\lambda u = \lambda(cu)$.

We call λ an **eigenvalue** of T, and all the vectors in U **eigenvectors** of T corresponding to λ.

Intuitively, eigenvectors of T are vectors such that that T sends them to a scalar multiple of themself, and the eigenvalues are these correspionding scalar multiples.



Before we proceed any further, we stop to prove a very important result about eigenvectors:

Proposition 1. *Given a set of m distinct eigenvalues $\lambda_1, \ ..., \ \lambda_m$, along with a set of corresponding eigenvectors $V = \{v_1, \ ..., \ v_m\}$, the set V is linearly independent.*

Proof. We will prove this proposition by induction. Clearly, this will be true in the case of one eigenvalue, λ. Assume that it holds true given n eigenvalues. We prove it holds true for $n+1$.

Consider the set of eigenvalues $\{\lambda_1, \ ..., \ \lambda_{n+1}\}$ with corresponding eigenvectors $\{v_1, \ ..., \ v_{n+1}\}$. Assume that there is a non-trivial linear combination:

$$a_1 v_1 + \ \cdots \ + a_n v_n + a_{n+1} v_{n+1} = 0$$

Note that since eigenvectors are non-zero, for this non-trivial linear combination to be 0, we must have at least two a_i not equal to 0 otherwise we would have $a_k v_k = 0$, for non-zero a_k, which can't be the case. It follows that at least one a_i with $1 \leq i \leq n$ is non-zero.

We define the linear operator $(T - \lambda_{n+1} I)$. We then have:

$$(T - \lambda_i I)(a_1 v_1 + \ \cdots \ + a_n v_n + a_{n+1} v_{n+1}) = \sum_{k \neq n+1} a_k(\lambda_k - \lambda_{n+1}) v_k = 0$$

But since all eigenvalues are unique, we must have $\lambda_k - \lambda_{n+1} \neq 0$. In addition, it least one a_i in this sum is non-zero. Thus, we have found a non-trivial linear combination of n eigenvectors that yields the zero vector, a contradiction to the inductive hypothesis.

It follows that the set $\{v_1, \ ..., \ v_{n+1}\}$ is linearly independent and the proof is complete. □

Corollary 1. *Given some finite-dimensional V and some $T \in \mathcal{L}(V)$, T can have at most* dim V *distinct eigenvalues.*

Proof. Assume we had more than dim V eigenvalues. We would then have a set of more than dim V linearly independent eigenvectors, a clear contradiction. □

Corollary 2. *Given a vector space V of dimension n and some $T \in \mathcal{L}(V)$, if there exist n distinct eigenvalues $\lambda_1, \ ..., \ \lambda_n$, then there exists a basis of V composed of eigenvectors.*

4.1 Eigenspaces

Given the fact that we are interested in invariant subspaces, we formalize the notion of a collection of eigenvectors, which we know from above is an invariant subspace:

Definition 2. *Given some operator $T \in \mathcal{L}(V)$ and some eigenvalue λ of T, An **eigenspace** of T is defined to be the subspace $E(\lambda,\ T)$ of V containing the eigenvectors associated to λ.*

Note that from the definition, it is clear that:

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

As it will turn out, eigenspaces are very powerful

5 The Decomposition Theorem

5.1 Motivation

Consider some operator $T \in \mathcal{L}(V)$, where V is finite-dimensional.

In general, it is easier to make sense of linear maps if we can break them up into many pieces, which we can then analyze separately. Intuitively, if we can come up with a collection of subspaces $V_1,\ ...,\ V_n$ such that each has a "simple" structure and:

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

then it is possible that consider T acting on each individual V_k might give some insight into the underlying structure of T.

But, under what circumstances can we do this? Well, if we want to understand how T acts on each individual V_k, then we should consider T restricted to the domain V_k, which we denote $T|_{V_k}$. However, $T|_{V_k}$ isn't always well-defined. As we explained in Section 2, such restrictions only "make sense" when V_k is an **invariant subspace**.

Thus, we want to find subspaces V_k that are **invariant** under T.

5.2 Eigenspaces

Now that we know what we are looking for (invariant subspaces whose direct sum is equal to V), we can start thinking about what these invariant subspaces should actually be.
One of the most familiar examples of an invariant subspace that the reader of these notes is likely familiar

5.3 Operators and Polynomials

5.4 The Decomposition Theorem

Definition 3. *Given some operator $T \in \mathcal{L}(V)$ and some eigenvalue λ of T, An **eigenspace** of T is defined to be the subspace $E(\lambda,\ T)$ of V containing the eigenvectors associated to λ.*

Note that from the definition, it is clear that:

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

Definition 4. *Generalized eigenvectors and eigenspaces*

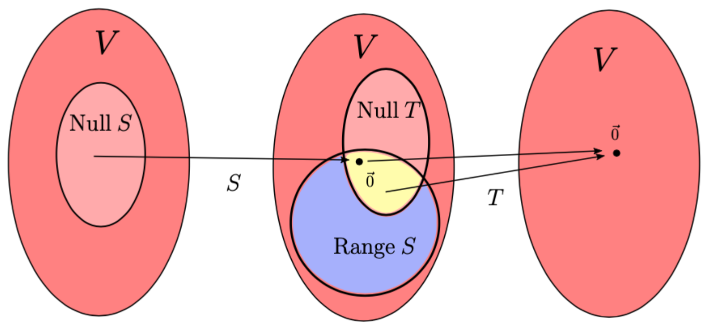
Lemma 1. *V is the null space of the product*

Lemma 2. *The sum of the null spaces is direct*

Lemma 3. *Given two operators S and T in $\mathcal{L}(V)$, where V is finite-dimensional, then:*

$$\dim \text{ Null } ST = \dim \text{ Null } T + \dim \text{ (Range } T \cap \text{Null } S)$$

For some intuition as to why this is true, consider the following diagram:



Clearly, only vectors in the left-most null space and the yellow region get sent to the zero vector, after we "apply" the maps S and then T to V.

If you look closer, you will notice that these two regions are exactly Null T and Range $T \cap$ Null S respectively!

Lemma 4. *Necessary result to prove this.*

Proof. □

Proposition 2. *The null space of the product is equal to the direct sum of all the null spaces*

This allows us to conlude that we can decompose V.

6 Upper Triangular Matrices

Proposition 3. *Every operator over $\mathbb{C}$ has an upper-triangular matrix representation in some basis.*

Lemma 5. *An operator T on V is invertible if and only if all the entries on the diagonal of an upper-triangular representation of T are non-zero.*

Corollary 3. *The eigenvectors of some operator T on V are the diagonal elements of an upper-triangular matrix representation of T*

7 Jordan Canonical Form

7.1 Sending Lemmings Off the Cliff: Finding the Jordan Basis

Now comes arguably the most confusing part of the notes: defining a basis that puts some operator $T \in \mathcal{L}(V)$ in Jordan form. We will attempt to explain each step using diagrams, to help with conceptualizing all of the choices we make.

To briefly recap, we know that given some vector space V and some $T \in \mathcal{L}(V)$, we will have:

$$V = c \prod_{k=1}^{n} (T - \lambda_k I)^{m_k}$$

for sets $\lambda_1, \ldots, \lambda_n$ and $m_1, \ldots, m_n$. Using this fact, we proved the decomposition theorem, showing that:

$$V = (T - \lambda_1 I)^{m_1} \oplus \ \cdots \ \oplus (T - \lambda_n I)^{m_n}$$

We are trying to find a convenient basis. With the above fact, it follows that we only have to find a convenient basis **for each** $(T - \lambda_1 I)^{m_1}$. Then, since the sum of these subspaces is direct, we can combine all of our "nice" bases, giving us a "nice" basis for the whole vector space V.

Hopefully this example illuminates why decomposing V into a collection of simpler pieces is useful in the first place!

7.1.1 Finding a Basis for Each Invariant Subspace