LITERALLY EVERYTHING I KNOW ABOUT

Linear Algebra

Warren Alphonso

A very reductionist summary of LINEAR ALGEBRA AND ITS APPLICATIONS by Lay, Lay, and McDonald, as well as LINEAR ALGEBRA DONE WRONG by Treil.

June 2, 2019

Contents

Chapter 1

Spectral Theory

Spectral theory will be our main tool for analyzing linear operators. In this chapter, we only consider transformations $A : V \to V$ ($n \times n$ matrices).

1.1 Definitions

Definition 1.1.1. A scalar λ is called an **eigenvalue** of operator $A : V \to V$ if there exists a *nonzero* vector $v \in V$ such that

$$Av = \lambda v$$

The vector v is called an **eigenvector** of A (corresponding to the eigenvalue of λ).

Once we know the eigenvalues, finding the eigenvectors is equivalent to solving

$$(A - \lambda I)v = 0$$

$Null(A - \lambda I)$, the set of all eigenvectors and 0, is called the **eigenspace**.

The set of all eigenvalues of an operator is called the **spectrum** of A, denoted $\sigma(A)$.

Since the matrix A is square, $A - \lambda I$ has a nontrivial null space if and only if it is not invertible, which means its determinant will be 0. Thus, for any eigenvalue λ of A,

$$det(A - \lambda I) = 0$$

Definition 1.1.2. If A is an $n \times n$ matrix, $det(A - \lambda I)$ is a degree-n polynomial of variable λ. This is called the **characteristic polynomial** of A. Finding the spectrum of A requires finding the roots to the characteristic polynomial.

This means any operator in $\mathbb{C}^n$ has n eigenvalues, though some may be repeated.

Theorem 1.1.1. *An $n \times n$ matrix A is invertible if and only if it doesn't have an eigenvalue of 0.*

Proof. Proving if:

If A doesn't have an eigenvalue of 0, then $det(A - 0I) \neq 0 \to det(A) \neq 0$, which implies A is invertible.

Proving only if:

If A is invertible, then $det(A) \neq 0$, which implies $det(A - 0I) \neq 0$. ∎

Theorem 1.1.2. *Let A be an $n \times n$ matrix, and let $\lambda_1, \cdots, \lambda_n$ be its complex eigenvalues (counting multiplicities). Then*

$$det(A) = \lambda_1 \cdots \lambda_n$$

Proof. Since $det(A - \lambda I)$ is a degree-n polynomial of variable λ and we know A will have n eigenvalues, we can write

$$det(A) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$$

Plugging in $\lambda = 0$ gives us

$$det(A) = \lambda_1 \cdots \lambda_n$$

∎

Theorem 1.1.3. *Let A be an $n \times n$ matrix, and let $\lambda_1, \cdots, \lambda_n$ be its complex eigenvalues (counting multiplicities). Then*

$$trace(A) = \lambda_1 + \cdots + \lambda_n$$

Proof. Let us begin by analyzing $det(A - \lambda I)$. Notice that in any cofactor expansion, if we pick any element $a_{i,j}$, such that $j \neq k$, then the highest degree of the resulting cofactor will be $n - 2$. This is because cofactoring removes the row and column the chosen entry is on, and since $j \neq k$, we remove the variables $a_{j,j} - \lambda$ and $a_{k,k} - \lambda$. After cofactor expansion, the λ^{n-1} term will be formed by only this equation

$$(a_{1,1} - \lambda) \cdots (a_{n,n} - \lambda) = (-1)^n (\lambda - a_{1,1}) \cdots (\lambda - a_{n,n})$$

so the coefficient of λ^{n-1} amounts to choosing the λ variable $n - 1$ times and choosing one of the other coefficients to get

$$(1)^n (a_{1,1} \lambda^{n-1}) \cdots (a_{n,n} \lambda^{n-1}) = (-1)^n (a_{1,1} + \cdots + a_{n,n}) \lambda^{n-1} \tag{1.1}$$

Note we can rewrite the characteristic equation as

$$det(A - \lambda I) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda) = (-1)^n (\lambda - \lambda_1) \cdots (\lambda - \lambda_n)$$

Now let us identify the coefficient of the λ^{n-1} term

$$(-1)^n (\lambda_1 \lambda^{n-1} + \lambda_2 \lambda^{n-1} + \cdots + \lambda_n \lambda^{n-1}) = (-1)^n (\lambda_1 + \cdots + \lambda_n) \lambda^{n-1} \tag{1.2}$$

Comparing coefficients in Equations 1.1 and 1.2, $trace(A) = \lambda_1 + \cdots + \lambda_n$. ∎

1.2   Diagonalization

We can use spectral theory to find the diagonalization of operators, which means that given an operator, we find the basis in which the matrix of the operator is diagonal. This makes powers of an operator much easier to compute.

Theorem 1.2.1. *A matrix A in $\mathbb{F}^n$ can be written as $A = PDP^{-1}$, where D is a diagonal matrix and P is invertible, if and only if there exists a basis in $\mathbb{F}^n$ of eigenvectors of A.*
In this case, the diagonal entries of D are the eigenvalues of A and the columns of P are the corresponding eigenvectors.

Proof. To understand the intuition behind this, note that $P = [I]_{S,B}$, where S is the standard basis and B is the basis for the eigenspace, since each column is the representation of a basis vector written in S. Rewriting $A = PDP^{-1}$ as $D = P^{-1}AP = [I]_{B,S}A[I]_{S,B}$ which means $D = [A]_{B,B}$, which is a diagonal operator if and only if its diagonal entries are eigenvalues whose corresponding eigenvectors are b_k.
A simpler, more direct proof is to rewrite $AP = PD$.

$$AP = \begin{bmatrix} Ab_1 & \cdots & Ab_n \end{bmatrix} = \begin{bmatrix} \lambda_1 b_1 & \cdots & \lambda_n b_n \end{bmatrix}$$

$$PD = \begin{bmatrix} b_1 \lambda_1 & \cdots & b_1 \lambda_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 b_1 & \cdots & \lambda_n b_n \end{bmatrix}$$

∎

Of course, for P to be invertible, the eigenvectors $b_1, \cdots, b_n$ must be linearly independent. Luckily, we can easily check if this is the case with the following theorem.

Theorem 1.2.2. *Let $\lambda_1, \cdots, \lambda_n$ be **distinct** eigenvalues for A, and let $b_1, \cdots, b_n$ be their corresponding eigenvectors. Then $b_1, \cdots, b_n$ are linearly independent.*

Proof. We proceed by induction over the n eigenvectors of A.
Base case: $n = 1$
This is trivial because by definition, an eigenvector is nonzero. Any set consisting of a single nonzero vector is linearly independent.
Inductive Hypothesis:
Assume it holds for $n = k$.
Inductive Step: $n = k + 1$
Suppose there exists a non-trivial solution to

$$\sum_{i=1}^{k+1} c_i b_i = 0$$

We can apply $(A - \lambda_{k+1} I)$ to both sides to get

$$\sum_{i=1}^{k+1} c_i (A - \lambda_{k+1} I) b_i = 0$$

Since $(A - \lambda_{k+1} I) b_{k+1} = 0$, we can write

$$\sum_{i=1}^{k} c_i (A - \lambda_{k+1} I) b_i = \sum_{i=1}^{k} c_i (\lambda_i - \lambda_{k+1}) b_i = 0$$

By the inductive hypothesis, we know the first k eigenvectors are linearly independent, so the coefficient $c_i(\lambda_i - \lambda_{k+1})$ must be 0 for $0 \leq i \leq k$, and since eigenvalues are distinct, $c_i = 0$ for $0 \leq i \leq k$.

Now we can reduce our original summation

$$\sum_{i=1}^{k+1} c_i b_i = c_{k+1} b_{k+1} = 0$$

This means that c_{k+1} must be 0, so the summation only has the trivial solution, which means the eigenvectos are linearly independent. ∎