Abstract

This paper will state the Perron-Frobenius Theorem and investigate some of its uses in Markov Chains.

1 Vector Spaces and Fields

To properly state the Perron-Frobenius Theorem, we must first investigate vector spaces. To define a vector space, we must first define a *field* F.

Definition 1 (Field Axioms)**.** A *field* F is a set together with two operations $\cdot : F \times F \to F$ and $+ : F \times F \to F$ such that for any $a, b, c \in F$, we have the following:

- Commutativity of Addition: $a + b = b + a$
- Commutativity of Multiplication $ab = ba$
- Associativity of Additions: $(a + b) + c = a + (b + c)$
- Associativity of Multiplication: $(ab)c = a(bc)$
- Multiplication is Distributive over Addition: $a(b + c) = ab + ac$
- Additive Identity: There exists some $0_F \in F$ such that $a + 0_F = a$
- Multiplicative Identity: There exists some $1_F \in F$ such that $a1_F = a$
- Additive Inverse: There exists some $-a \in F$ such that $a + (-a) = 0_F$
- Multiplicative Inverse: There exists some $a^{-1} \in F$ such that $a(a^{-1}) = 1_F$

These conditions are called that *Field Axioms.*

Some examples of fields are the rational numbers $\mathbb{Q}$, the real numbers $\mathbb{R}$, and the complex numbers $\mathbb{C}$. Now that we have a definition of a field, we can define a *vector space* on a field F.

Definition 2 (Vector Spaces)**.** A vector space V on a field F is a set together with two operations $+ : V \times V \to V$ and $\cdot : F \times V \to V$ satisfying the following:
For any $u, v, w \in V$ and any $a, b \in F$, we have that

- $au = ua$
- $u + v = v + w$
- $(u + v) + w = u + (v + w)$
- $(ab)u = a(bu)$
- V has an element 0 satisfying $u + 0 = 0$
- There exists a $-u \in V$ such that $u + (-u) = 0$
- $u1_F = u$
- $(a + b)u = au + bu$
- $a(u + v) = au + av$ Elements of F are called *scalars* and elements of V are called *vectors.*

We now define the basis of a vector space.

Definition 3 (Bases of Vector Spaces)**.** Suppose we have a vector space V on a field F. A *basis* on V is a set of vectors $a_1, a_2, \ldots, a_n$ satisfying the following:

- Linear Independence: There does not exist $c_1, c_2, \ldots, c_n \in F$ such that $\sum_{i=1}^{n} c_i a_i = 0$ and not all of the c_i are 0 for $1 \leq i \leq n$.
- Spanning: For every $v \in V$ such that $v \neq 0$, there exists $c_1, c_2, \ldots, c_n \in F$ such that $v = \sum_{i=1}^{n} c_i a_i$. If a set of vectors is not spanning, we call the *spanning set* of the set of vectors as any vector that can be expressed as an F-linear combination of the vectors in the set.

We call n the size of the basis.

Remark 4. Every $v \in V$ can be uniquely represented as a linear combination of the elements of a basis. For proof, suppose there are two different representations. Then we can subtract the two resulting equations, and get that some linear combination of the basis is 0, which contradicts linear independence. This implies that every element of a basis must be nonzero.

We call n the *dimension* of V if there exists a basis $a_1, a_2, \ldots, a_n$ of V. However, we do not know that n is the same for every basis. The key idea is the following theorem, which we state without proof.

Theorem 5 (Dimension of a vector space)**.** The dimension of a vector space V on a field F is unique. In other words, every basis of V has the same size.

Now we can safely refer to the *dimension* of a vector space. We will now investigate a specific type of function from one vector space to another, called a *linear transformation*.

2 Linear Transformations

Definition 6 (Linear Transformations)**.** Given to vector spaces V_1, V_2 on a field F, we define a *linear transformation* $T : V_1 \to V_2$ as a function satisfying, for any $a, b \in V_1$ and any $c \in F$, the following hold:

- $T(a + b) = T(a) + T(b)$
- $T(ca) = cT(a)$

Remark 7. Pick a basis $a_1, a_2, \ldots, a_m$ of V_1. We only really care about the values of $T(a_1), T(a_2), \ldots, T(a_m)$. This is because for any $v \in V_1$, we can write

$$v = \sum_{i=1}^{m} c_i a_i \tag{1}$$

where the c_i are in F. Therefore,

$$T(v) = \sum_{i=1}^{m} T(c_i a_i) = \sum_{i=1}^{m} c_i T(a_i) \tag{2}$$

This gives us a way to evaluate $T(v)$ for any $v \in V_1$ in terms of the $T(a_i)$

We can go further to write the $T(a_i)$ in terms of a basis $b_1, b_2, \ldots, b_n$ of V_2. We can write

$$T(v) = \sum_{i=1}^{m} c_i \sum_{j=1}^{n} d_i j b_j \tag{3}$$

Clearly, the d_{ij} are fixed. Therefore, we can represent the transformation T with a $m \times n$ matrix, where the ij entry is d_{ij}. Our matrix T would look like

$$\begin{pmatrix} d_{11} & d_{12} & \ldots & d_{1n} \\ d_{21} & d_{22} & \ldots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{m1} & d_{m2} & \ldots & d_{mn} \end{pmatrix} \tag{4}$$

3 Square Matrices

Definition 8 (Linear Operators)**.** Given a linear transformation $T : V_1 \to V_2$, we say that T is a *linear operator* if $V_1 = V_2$.

Remark 9. If T is a linear operator, then the matrix T is a *square matrix*. Equivalently, the number of rows equals the number of columns.

Definition 10 (Eigenvectors and Eigenvalues)**.** Given a linear transformation $T : V \to V$ where V is a vector space on a field F, we say that $v \in V$ is an *eigenvector* of T if $T(v) = \lambda v$ for some $\lambda \in F$. λ is called an *eigenvalue* of T. The set of all eigenvectors v with eigenvalue λ plus 0 is called the *eigenspace* for λ.

The following is a very important theorem about eigenvectors, which we state without proof.

Theorem 11 (Eigenvectors are Linearly Independent)**.** Suppose we have a linear transformation T on a n-dimensional vector space V yielding n different eigenvalues, with each eigenvalue yielding one eigenvector. Then these k eigenvectors are linearly independent.

4 Perron-Frobenius Theorem

We can now state the Perron-Frobenius Theorem for positive matrices, the main focus of this paper.

Theorem 12 (Perron-Frobenius)**.** Given an $n \times n$ matrix A with positive entries, we have the following:

- There exists a positive real eigenvalue r of A such that all other eigenvalues of A have absolute value less than r.
- r has a one dimensional eigenspace.
- There exists an eigenvector v of A known as the Perron-Frobenius eigenvector that has all positive entries.
- All positive eigenvectors of A are given by positive real multiples of v.

We will not prove this theorem, but we will investigate its use in Ergodic Markov Chains. Suppose we have a $\triangle ABC$ and a particle is at vertex A. Each second, from vertex A it moves to B with probability 0.6, C with probability 0.4. From B it moves to C with probability 0.6, and A with probability 0.4. From C it moves to A with probability 0.6, and B with probability 0.4. We have that the transition matrix T for the associated Markov Chain is

$$\begin{pmatrix} 0 & 0.6 & 0.4 \\ 0.4 & 0 & 0.6 \\ 0.6 & 0.4 & 0 \end{pmatrix}$$

This is just a matrix where each entry represents the probability of transitioning from one vertex to another. The question now is, after an infinite amount of time, where is the particle likely to be? Does its starting position matter? The symmetry of this suggests that it will be equally likely for the particle to be at any vertex after an infinite amount of time. However, because the sum of each row of this matrix is 1, we have that the vector

$$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

is an eigenvector of the matrix T, with Perron-Frobenius eigenvalue 1 due to the Perron-Frobenius theorem. Here, the Perron-Frobenius theorem is telling us that multiples of the stated vector are the only positive eigenvectors of this transition matrix. Even though all of the entries of this matrix are not positive, this argument still holds because the matrix is ergodic, which means that with a sufficiently large amount of time, we can get from one vertex to another with positive probability. This would not hold if, for example, we had a quadrilateral, because we would only be able to reach our original vertex after an even number of seconds. Therefore, the Perron-Frobenius theorem implies that there exists a unique positive eigenvector π such that

$$\pi A = \pi$$

where A is the transition matrix. We will not prove this rigorously, but this implies that the distribution of where the particle is would naturally converge to this eigenvector, as it is unchanged when multiplied to the matrix. This is the intuition behind the Ergodic Theorem, which states that for an ergodic square stochastic matrix A, the limit as n approaches infinity of A^n is a matrix with row vectors that are the same, because the starting point will not matter.

References

[1] https://mathworld.wolfram.com/FieldAxioms.html

[2] https://www.overleaf.com/project/5e8b89c30157870001f2dd3b

[3] https://www.overleaf.com/project/5eaaf3d9cce9c50001bc56ee