An Introduction to Tensors

Refath Bari

July 30, 2020

Abstract

In an effort to familiarize ourselves with Tensors, we must understand the rules of this other-worldly game. We begin with some elementary proofs in Tensor Calculus, which itself was formulated by the Gregorio Ricci-Curbastro, the man behind the famous Ricci Tensor.

1 Introduction

I have constructed a ladder with exactly five rungs. Climbing this ladder will take you to the shoulders of giants – giants in the vein of Albert Einstein, James Clerk Maxwell, Paul Dirac, and Richard Feynman. The language of Tensors is the language of the greatest minds in human history. Indeed, Einstein himself once remarked (in regards to Tensors):

> I admire the elegance of your method of computation. It must be nice to ride through these fields upon the horse of true mathematics while the like of us have to make our way laboriously on foot. **Albert Einstein**

First and foremost, I establish the five ideas we must conquer in our endeavour to colonize Tensors:

1. Vector Spaces
2. Linear Maps
3. Dual Spaces
4. Homomorphisms
5. Cartesian Products

Tackling all of the preceding ideas will set the groundwork; the stage for Tensor Products, from whose vector spaces we can extract actual Tensors. We begin with the idea of Vector Spaces.

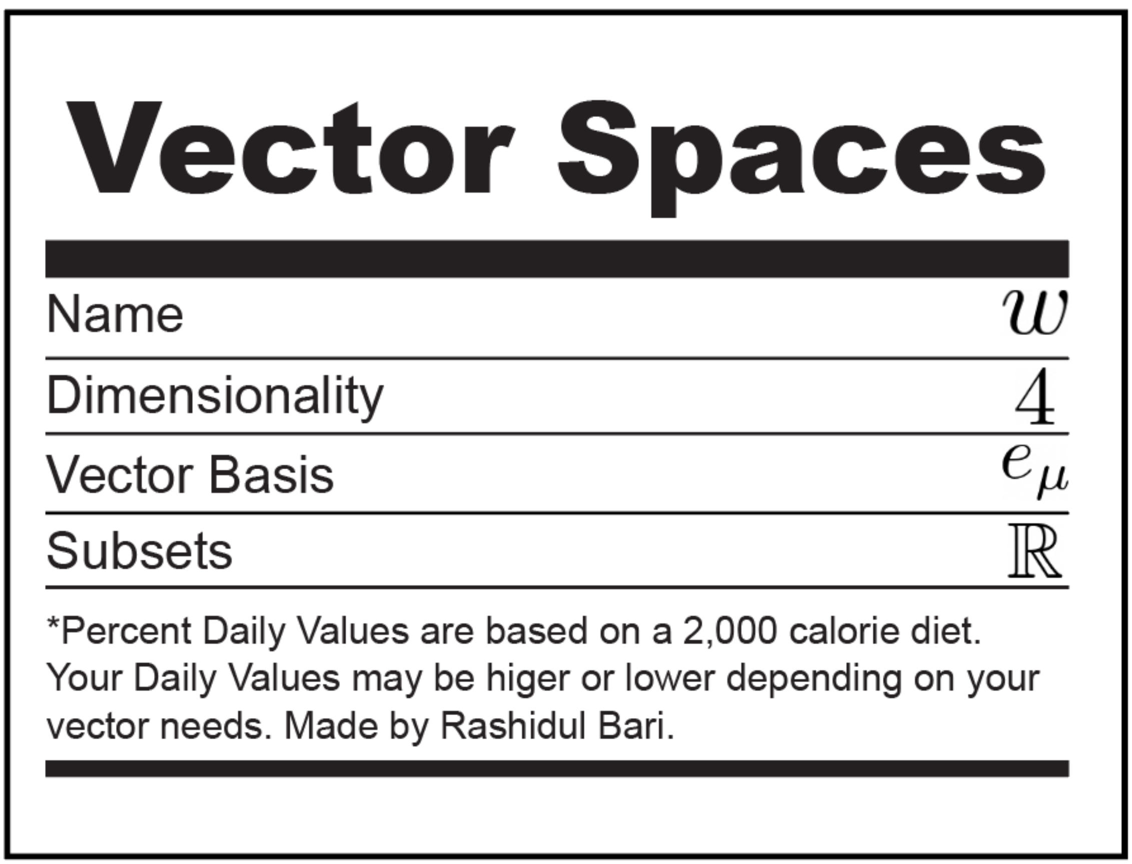
2 A Brief History

First and foremost, I hope to provide a brief history of Tensors, starting with Gauss himself.

1. Tensors originated from Gauss work in Differential Geometry.
2. Tensors were formally introduced by William Hamilton in 1846.
3. It's contemporary usage began in 1898 by Woldemar Voigt.
4. It was developed further in 1890 by Ricci, who invented Tensor Calculus
5. Einstein used Tensors to formulate General Relativity in 1915

3 Vector Spaces

What is a vector space? I like to think of every vector space as a cereal box with only four pieces of Nutritional Information.



It's worth trying to understand this for a minute. Every vector space really has four major components, as seen above. A vector space is simply a kind of algebraic structure, akin to a ring or a field. First and foremost is its name, a simple letter used to refer to the vector space. Then comes dimensionality, which is the cardinality of the set of basis vectors used to represent the entire vector space. In this instance, our vector space v has a dimensionality 4 (you can think of it as akin to the three dimensions of space and one of time). Then of course, comes the basis vectors themselves. Lets first pause for a moment and ask ourselves what a set of basis vectors even is.

To understand the set of basis vectors, we must first understand the court of law. Marc comitted a crime. He will be tried in a fair and just judicial system. In said system, there must exist a jury. The jury is a set of five people who are representative of the entire population. Much the same way, a set of basis vectors are representative of the entire vector space. Pick out any vector from the vector space, and it should be representable by the basis vectors. Some vectors may be represented by just one basis vector, but those vectors are *special* vectors. In general, all the vectors from the vector space should be and indeed are representable by the set of basis vectors. Let it be known, however, that the set of basis vectors is not unique. A vector space may have any number of sets of basis vectors. What is unique is the *cardinality* of the set of basis vectors, which would be the minimum number of fundamental vectors used to represent the entire span of a vector space.

OK. Enough about Basis Vectors. What about the subsets row? Well, any vector space (at least the ones we're concerned with) involve the subset of real numbers, $\mathbb{R}$. Some vector spaces involve the subset of Complex numbers (which of course, include the Real Numbers if their imaginary component is zero), $\mathbb{Q}$. That right there encompasses all we need to know about the notation of Vector Spaces.

Now is as good a time as any to examine vector spaces from a much broader, holistic perspective – the set of all Algebraic Structures in general. So far as I know, no such hierarchy of this kind exists. Let's take a look at **Bari's Hierarchy of Algebraic Structures**

Let's take a look at each of these algebraic structures and see where Tensors fit in into all of this:

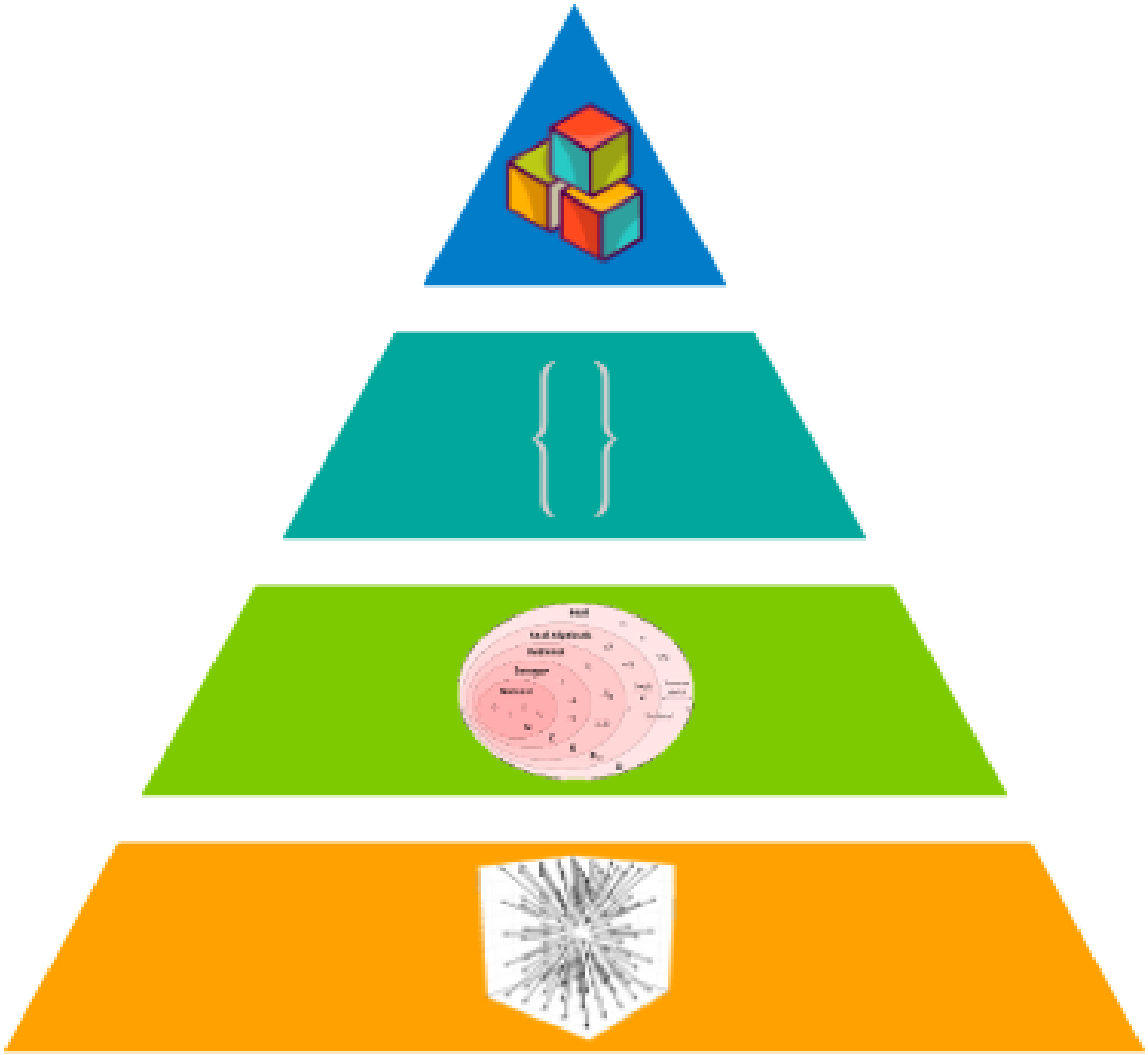
1. **Structures:** Algebraic Structures are the consequence of grouping things, whether they be topological spaces, vector spaces, or integers.

2. **Sets:** Exactly 4 operators can be executed on sets, whether they be unions, intersections, complements or cartesian products.

3. **Fields:** An algebraic structure closed under +, -, *, and / such that distributivity, commutativity, and the zero element hold. Examples of fields include the set of Real Numbers, Rational Numbers, and even Complex Numbers, each denoted by $\mathbb{R}$, $\mathbb{Q}$, and $\mathbb{C}$, respectively.

4. **Vector Space:** Every vector space has a scalar field F associated with it. Addition and Scalar multiplication are closed under vector spaces. This is where Tensors fit in: Tensor Products are an operation that give rise to a different type of vector, known as a Tensor, which is an element in the Vector Space of Tensors.

Let's delve into Vector Spaces. You may recall that any vector space must satisfy the following conditions:



1. **Includes Zero Element:** As a consequence of the succeeding two principles, every vector space must include the zero element. I say element because the zero vector is analogous to a different entity in every space.

2. **Closed under Addition:** This means if $a, b \in W$, so should $a + b \in W$

3. **Closed under Multiplication:** Every vector space must be closed under multiplication – specifically *scalar* multiplication such that if $a \in W$ and $c \in \mathbb{R}$, then so should $ca \in W$

The crux of vector spaces comes with Linear Combinations and Linearity. Linear Combinations come in the form of $a\vec{w} + b\vec{z} = \vec{k}$ and Linearity tells us that the Distributive property must hold: $a\vec{w} + a\vec{z} = a(\vec{k} + \vec{z})$. From these two, we say that the Zero Element must exist because for any $\vec{v} \in W$, so is $-\vec{v} \in W$ (by the Scalar Multiplication principle). Add them up and we get $\vec{v} - \vec{v} \in W \Rightarrow \vec{0} \in W$

The cereal analogy still holds. You don't mix two cereals in the same milk. Much the same way, you can't add vectors from two different vector spaces, simply because addition is defined differently – distinctively – for each of the spaces, and there is simply no defined way to add vectors from different spaces. Thus, if $\vec{v} \in W$ and $\vec{j} \in K$, we say $\nexists \vec{v} + \vec{j}$. That right there is a problem; it becomes a pivotal issue when we seek to reconcile the perspectives of two observers in relative motion to each other.

4 Linear Maps

The question now becomes this: how do we establish a relationship between two vector spaces, if we can't even add vectors between them? Here comes the kicker: two vector spaces are isomorphic if there exists a one-to-one and onto relationship between them and if operations in both vector spaces are identical. By extension, then, the dimensionality of both vector spaces must be the same.

Let's talk basis vectors. Say we have some vector space v and there exists some set of basis vectors e_μ where

$$e_\mu = \{e_0, e_1, e_2, e_3\}$$

and $\mu = 0, 1, 2, 3$. Thus, any given vector A from my space v can be written as some linear combination of e_μ:

$$A = ae_0 + be_1 + ce_2 + de_3$$

Indeed, we can simplify the notation even further to

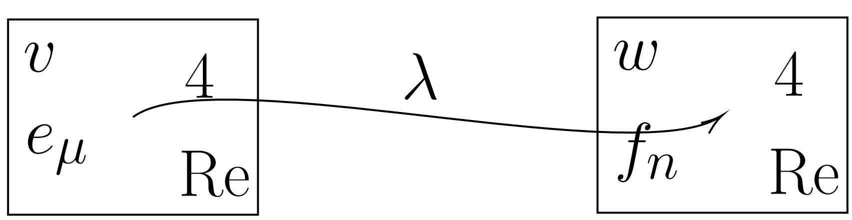
$$A = A^0e_0 + A^1e_1 + A^2e_2 + A^3e_3$$

If that looks a lot like the inner product, that's because it is! Indeed, we have $\sum_{\mu=0}^{3} A^\mu e_\mu$ which can be condensed even further by the Einstein Summation Convention to just

$$\sum_{\mu=0}^{3} A^\mu e_\mu = A^\mu e_\mu$$

.

Armed with a suitable understanding of basis vectors, we are ready to charge forward in our understanding of linear maps. A linear map behaves just like a function (although one can discern it as an operator as well). A linear map is a bridge between worlds: a connection between otherwise mutually-exclusive vector spaces. It is, in many ways, the holy grail.



We will also see the linear map written in various other dialects including:

1. **Domain-Range Notation:** $\lambda : v \to w$

2. **Operator Notation:** $\lambda v \to w$

3. **Function Notation:** $\lambda(v) \to w$

4. **Braket Notation:** $\langle \lambda | v \rangle \to w$

Now to connect any vector A from my vector space v to w, I need to actually *impose* my transmation, my linear map onto my basis vectors. In other words, if I can just find where my linear map λ takes my basis vectors e_μ to, I'm home free! Indeed, this is exactly what I can do, by letting my actual linear map *act* out onto my basis vectors, as such:

$$\lambda e_0 = \langle \lambda | e_0 \rangle = B^n f_n$$
$$\lambda e_1 = \langle \lambda | e_1 \rangle = C^n f_n$$
$$\lambda e_2 = \langle \lambda | e_2 \rangle = D^n f_n$$
$$\lambda e_3 = \langle \lambda | e_3 \rangle = E^n f_n$$

Great! Now we know where all the basis vectors of v map onto. Now we simply have

$$\langle \lambda, A \rangle = \langle \lambda, A^\mu e_\mu \rangle$$

by Einstein Summation Convention (repeated product signifies sum). But we know what all the e_μ map onto! And we can simply extract the constants (since one of the principles of Vector Spaces – scalar multiplication, permits us to do exactly that) as such:

$$\langle \lambda, A^\mu e_\mu \rangle = A^0 \langle \lambda, e_0 \rangle + A^1 \langle \lambda, e_1 \rangle + A^2 \langle \lambda, e_2 \rangle + A^3 \langle \lambda, e_3 \rangle$$

Before we part ways from this section, it's worth taking a look at the fundamental properties of linear maps:

$$\langle \lambda, v + p \rangle = \langle \lambda, v \rangle + \langle \lambda, p \rangle$$
$$\langle \lambda, qv + bp \rangle = q \langle \lambda, v \rangle + b \langle \lambda, p \rangle$$

5 Dual Spaces

Great! We've accomplished our original goal of connecting any vector from the original vector space to the new one. But if we didn't want to create a whole new vector space? What if we wanted to keep as simple an algebraic structure as possible? Enter Dual Spaces. For any vector space V, there exists some Dual Space which is also a vector space. The elements of the dual space are no vectors – they're linear maps. In fact, they're the set of *all* linear maps between two vector spaces and the map between our vector space V and the dual space $V*$ is just one element $T \in V*$. We now have some $V* : V \to \mathbb{R}$. Thus, this $V*$ takes our vector space V to the set of real numbers $\mathbb{R}$. Of course, $\mathbb{R}$ itself is a one-dimensional vector space. We thus denote this as the 'realization' of a 1-D vector space. What magic!

6 Homomorphisms

All this buzz over vector demands that we acknowledge the idea of Homomorphisms. Indeed, the idea of Dual Spaces is a child of the idea of $Hom(v, w)$, which is simply a set of all the linear maps between two vector spaces. I hereby issue a formal definition for $Hom(v, w)$:

Definition 6.1. $Hom(v, w)$ For any vector spaces v, w over scalar field F, $\exists Hom(v, w)$ where all $T \in Hom(v, w)$ such that $T : v \to k$. In other words, $Hom(v, w)$ contains all linear maps T from vector space v to w.

Great. $Hom(v, w)$ exists. Can we impose some algebraic structure onto $Hom(v, w)$? Indeed, what can we even classify it as? Well we know that our homomorphic set has at least two operations: it's closed under addition and scalar multiplication. So yes! $Hom(v, w)$ does indeed have some structure – in fact, it's a vector space! Indeed, the time has come for a formal theorem in regards to the algebraic structure of $Hom(v, w)$:

Theorem 6.1. *$Hom(v, w)$ is a vector space over the scalar field F with respect to the addition of linear maps and scalar multiplication.*

This is a whole new Vector Space and it's not a subspace of anything! This means that $Hom(v, w)$ must satisfy three key tenets:

1. **Commutavity:** Is $s + t = t + s$?
2. **Distributivity:** Is $a(b + c) = ab + ac$?
3. **Zero Element:** Is the analagoue of $0 \in Hom(v, k)$?

But wait. What *is* the zero element of $Hom(v, k)$? To answer that question, we must first understand what the actual elements of $Hom(v, k)$ are. They're just linear maps! So the equivalent of the zero vector in $Hom(v, k)$ is quite simply, the zero map – i.e., the map that takes everything to zero.

In fact, linear maps are uncannily like linear transformations. Indeed, when I first encountered them, I treated them identically. Thus, if linear maps are just linear transformations in disguise, there's no reason we can't use a matrix to represent them! The bridge between linear maps and their matrix cousins are the basis vectors, for whom we've devoted an entire section of this paper to. We now have the vocabulary to conjecture that the vector space $Hom(v, w)$, being a set of linear maps, is indeed isomorphic to matrices. Say we have some linear maps $T, S : v \to w$, or perhaps more expressively, $T, S : v \in R^n \to w \in R^m$. Say e_μ serves as the basis for v and f_n for w. Thus,

$$[T + S]_E^F = [T]_E^F + [S]_E^F$$

Basically, $Hom(v, w)$ satisfies the first tenet of a linear map: its closed under addition. Likewise, we have

$$[aT]_E^F = a[T]_E^F$$

Thus, scalar mulitplication is also satisfied. OK! Wow! We thus have some linear map that takes our map to a matrix!

$$K : T \rightarrow [T]_E^F$$

If this does not qualify as magic, I don't know what does. We can thus say for any linear map ϕ that takes our $Hom(v, w)$ to a matrix:

$$\phi : Hom(v, w) \rightarrow M_{m*n}(F)$$

It should be noted that since they're actual linear transformations, v and w are square matrices of order m and n respectively. More precisely

$$dim(v) = m * m$$
$$dim(w) = n * n$$

We can thus say that $\phi(T) = [T]_E^F$. To conclude that $Hom(v, w)$ is indeed isomorphic to a matrix, we must prove that it is one to one and onto with said matrix. In other words, $Hom(v, w)$ is a vector space full of elements that are simply linear maps between vector spaces. $M_{m*n}(F)$ is a vector space of matrices. To prove they're isomorphic, we can either prove they're one to one and onto with each other, or equivalently,

$$dim(Hom(v, k)) = dim(v)dim(w)$$

In the interest of length, we will assume that these axioms are upheld.

7 Cartesian Products

We begin with a formal definition of Cartesian Products:

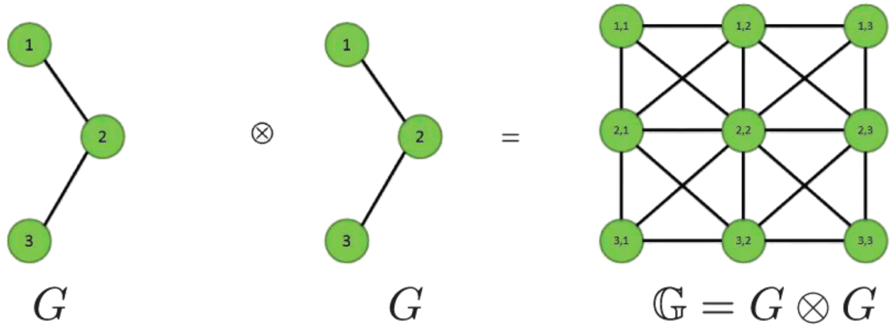
Definition 7.1. Cartesian Products The Cartesian Product $A \times B$ is set of all ordered pairs (a, b) where $a \in A$ and $b \in B$

Examples are in order. We need only look to (5,7) on the coordinate plane, where 5 is my x-component and 7 is my y-component. Thus, we say $(5, 7) \in (R \times R)$

8 Tensor Products

Let me begin with a definition of Tensors that may leave you unsatisfied: they're simply vectors from a special vector space. That vector space – the Tensor Product Space – is the consequence of the Tensor Product, an operation performed on two vector spaces to produce a totally different vector space whose elements are Tensors.

To begin our discussion of Tensor Products, let's begin with the Direct Product. For finite-dimensional vector spaces, the Direct Product $\times$ and Direct Sum $\bigoplus$ are the same, so we can begin our discussion by either one. Here comes the formal definition:



Definition 8.1. Product of Vector Spaces Given some finite number of vector spaces $V_1, V_2, V_3, ..., V_m$, we define their direct product as follows:

$$V_1 \times V_2 \times V_3 \times ... \times V_m = \{u_1, u_2, u_3, ..., u_m\}$$

where $u_1 \in V_1, u_2 \in V_2, u_3 \in V_3, ..., u_m \in V_m$We proceed to define the sum of vector spaces as component by component addition. Take any two group of vectors $\{u_1, u_2, u_3, ..., u_m\}$ and $\{w_1, w_2, w_3, ..., w_m\}$ from the vector spaces $V_1, V_2, V_3, ..., V_m$ and we define their sums as follows:

$$(u_1, u_2, u_3, ..., u_m) + (w_1, w_2, w_3, ..., w_m) = (u_1 + w_1, u_2 + w_2, u_3 + w_3, ..., u_m + w_m)$$

Finally, we define Scalar Multiplication as expected:

$$\lambda(u_1, u_2, u_3, ..., u_m) = (\lambda u_1, \lambda u_2, \lambda u_3, ..., \lambda u_m)$$

We also claim that the product of vector spaces is itself a vector space. This can be easily verified by the three tenets of a vector space, as outlined above.

$$\mathbf{v} \otimes \mathbf{w} = \begin{bmatrix} v_1 w_1 & v_1 w_2 & \cdots & v_1 w_m \\ v_2 w_1 & v_2 w_2 & \cdots & v_2 w_m \\ \vdots & \vdots & \ddots & \vdots \\ v_n w_1 & v_n w_2 & \cdots & v_n w_m \end{bmatrix}$$

From this definition of the direct product, we can get the Tensor Product. First and foremost, let us redefine scalar multiplication over the Field F, which can either be $\mathbb{R}$ or $\mathbb{C}$. Now, when we multiply a set of vectors from our vector spaces, we can simply multiply the scalar by any component we choose:

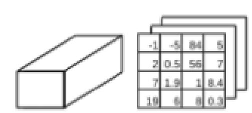
$$\lambda(v, w) = (\lambda v, w) = (v, \lambda w)$$

for all $\lambda \in F$. OK, we now redefine the addition operation over this space such that two ordered pairs of vectors from different spaces can be added only if one of their components is the same:

$$(u_1, u_2) + (u_1', u_2) = (u_1 + u_1', u_2)$$

The above sum only works because both ordered pairs share the second component w. All other additions are simply themselves. You can almost think of this as akin to like term addition. $v + w$ is totally valid addition, but we can't simplify it any further. In much the same vein, we say that $(x, y) + (v, w)$ is still valid addition, but can't be simplified any further (unless, of course, $x = v$ and $y = w$)

And that's it! That right there is the definition of the Tensor Product. All we need to do now is give it some new notation, which we can surely afford in the form of: $v \otimes w$.

Tensors			
Scalar	Vector	Matrix	Cube
Mass, 5 kg, 1, 7, 3.14, Temperature, Distance, 10m	$\begin{pmatrix} 1 \\ 2 \end{pmatrix}$ $\begin{pmatrix} 3 & 1 \end{pmatrix}$	$\begin{pmatrix} 3 & 1 \\ 4 & 1 \end{pmatrix}$ $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$	
Magnitude	Magnitude + Direction	Fields, Transformations	Motion Capture Data, Facial Recognition
Rank 0	Rank 1	Rank 2	Rank 3

And thus we conclude our exploration into Tensor Products. For the sake of brevity, I've neglected any discussion of bi-linear or multi-linear maps, or even the applications of Tensor Products, which is the ball really gets rolling.

References

[1] Pavel Grinfeld, *Introduction to Tensor Analysis and the Calculus of Moving Surfaces*, (Springer-Verlag, New York, 2013).