What are tensors?

nicola santi

A simple but rigorous introduction for young mathematicians

A few years ago, trying to understand what exactly tensors were, I ran into quite a few difficulties: there are a variety of definitions on the Internet (**[1]** gives as many as seven!) that are seemingly incompatible or unrelated to each other, as well as several tutorials that are not very accessible to the beginner because they focus on the formal way of constructing a tensor space or on merely using these spaces to solve physics or electrodynamics problems without clearly explaining what tensors per se are.

In contrast, there are presentations that are also interesting (e.g., **[2]**) but too popular to be useful to a mathematician looking for a more rigorous definition from which to begin the study of tensors.

This article is intended to position itself between the two extremes, between too much and too little formality, and was born out of a desire to save time for those approaching the study of tensors for the first time by providing a simple, I would say colloquial, introduction, while trying never to trivialize the subject and to maintain, as far as the economy of a short tutorial permits, a certain level of detail.

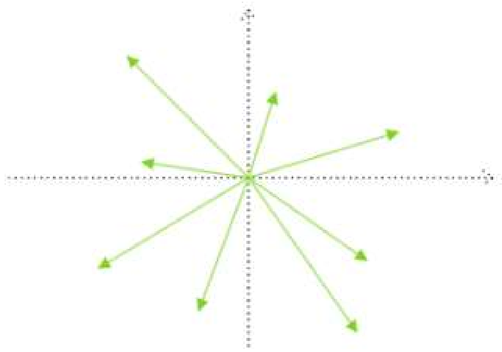
Knowledge of vector spaces V, dual spaces V^* and linear functions is the only prerequisite to reading, so much so that we will start right away by starting with vectors, which are generally easier to understand than tensors. Why is this so?

Vectors

A set V is a vector space, meaning that it contains vectors $\boldsymbol{v}$, if it has two operations, the first to sum the vectors together (vector addition) and the second to multiply them by a real or complex number (scalar product). Vectors obtained by using these two operations on a subset S of V, so summing and scaling all elements of S by every possible scalar, are said to be generated (span) by S.

This is an abstract, axiomatic definition that can be applied to a variety of mathematical entities and not only to lists (arrays) of real numbers in $\mathbb{R}^n$: for example, the set of square matrices $n \times n$ is a vector space as is the set of linear functions from V to $\mathbb{R}$.

Why does the fact that matrices or functionals are (also) vectors not confuse us with respect to the abstract definition of a vector? In my opinion, because vectors have a clear geometric representation: they are all arrows of various lengths coming out of the same point in the plane or space.



This clear representation leads us to say, for example, that matrices are yes matrices, i.e. rectangular table of $n \times m$ numbers, but they are also vectors, that is they also behave like the famous arrows mentioned above. The same goes for functionals which are functions, yes, but they also behave like the arrows, they have the sum and scalar product and therefore are (also) vectors.

Unfortunately, tensors do not have such an immediate geometric representation so stating that matrices or the set of bilinear forms are tensors does not help much because we do not have an intuitive idea, such as arrows, to compare with.

On this issue, Dodson and Poston in **[3]** also point out:
The reader may have noticed a scarcity of pictures in this chapter [about tensors]. This is not because tensors are un-geometric. It is because they are so geometrically various. They include vectors, linear functionals, metric tensors, the "volume" form 1.02 ... and nearly everything else we have looked at so far. That all of these wrap up in the same algebraic parcel is a great convenience, but it does mean that geometrical interpretations must attach to articular types of tensor, not to the tensor concept.

Bilinear maps

Bilinear maps (or bilinear functions) F are functions of two parameters (two vectors belonging respectively to the vector spaces V and U) that behave linearly with respect to each of these. To mean (vectors in bold):

$$F(\lambda \boldsymbol{x} + \beta \boldsymbol{y},\ \boldsymbol{u}) = \lambda F(\boldsymbol{x}, \boldsymbol{u}) + \beta F(\boldsymbol{y},\ \boldsymbol{u})$$

These bilinear maps have as their codomain a vector space, say W. When the codomain is simply $\mathbb{R}$, we talk about bilinear forms, which are basically the same thing but we will denote them by the letter B.
It is worth remembering that bilinear forms are representable as $n \times m$ matrices (where n is the dimension of V, m that of U) and their values derivable by the usual rules of matrix

calculus.

$$B(\boldsymbol{v}, \boldsymbol{u}) = \boldsymbol{v}^t [B] \boldsymbol{u}$$

Example:

$$v_1 u_1 + 2v_1 u_2 + 3v_2 u_1 + 4v_2 u_2 = \begin{bmatrix} v_1 & v_2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

Linear functions f, when they have values in $\mathbb{R}$, are also expressible by matrix calculus using $1 \times n$ matrix like in the follow example:

$$f(x) = f\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = 3x_1 + 5x_2 = \begin{bmatrix} 3 & 5 \end{bmatrix} \cdot \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Although similar, there are two very important differences between linear functions f and bilinear functions F . The first is that $f(\lambda \boldsymbol{v}) = \lambda f(\boldsymbol{v})$ while $F(\lambda \boldsymbol{v}, \lambda \boldsymbol{u}) = \lambda^2 F(\boldsymbol{v}, \boldsymbol{u})$, please note the power squared.

The second difference concerns the image of bilinear functions, which may or may not be a vector subspace while the images of linear functions are always a vector subspace of W. This last difference is fundamental to the definition of tensor spaces, so it is worth dwelling on it with an example: consider the bilinear map that associates two linear functions, thus belonging to the dual V^* and U^* , with an $x \times m$ matrix in the following way:

$$F(\boldsymbol{v}, \boldsymbol{u}) = \boldsymbol{v}\boldsymbol{u}^t$$

Or, in a more intuitive notation:

$$F(f, g) = f(\cdot)g(\cdot) = fg^t$$

Let us take an example with $n = m$:

$$F\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 4 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 3 & 4 \end{bmatrix} = \begin{bmatrix} 3 & 4 \\ 6 & 8 \end{bmatrix}$$

Recalling that the set of $n \times n$ matrices is also a vector space with dimension $n \cdot n$, we note that the matrices produced in this way are only a subset of that space: in particular, they are only those with the null determinant, which is easy to verify because the columns of our matrix are the same vector u scaled for different values.

More generally, when $n \neq m$, the bilinears function F produce $n \times m$ matrices but all of $rank = 1$.

This examples confirms that the image of bilinear maps can only be a subset of its codomain, in this case the space W of $n \times m$ matrices.

A curiosity might arise as to what is the maximum dimension that can be generated by such bilinear map between two finite vector spaces V and U: the answer is

$$\dim(Im\, F) \leqslant \dim U \cdot \dim V$$

In the next section we will see that bilinear functions of this type can take the name of tensor product because they could construct and thus define a tensor space.

Tensors in a nutshell

A set W is a tensor space if it has two essential characteristics:

1. The first is that it is a vector space (and therefore tensors are also vectors);
2. The second concerns the generation (span) of W: it must be possible to generate all W using the image of a bilinear map F;

All here.

It is required that the bilinear map has as its domain the cartesian of two vector spaces $V \times U$ and as its codomain, of course, W. This map, I remind you, must generate only a subset of W, capable in turn of generating all W. In other words, F is not suriective, $Im\, F \subset W$ just as we saw in the previous paragraph.

A simple example with which to become familiar is the tensor space W given by the bilinear forms B, which I recall are real-valued bilinear functions. W can be generated from the dual spaces V^* and U^* just as we saw in the previous section. In particular, if f and g are functions in V^* and U^* respectively, then their product $f(\boldsymbol{u})g(\boldsymbol{v})$ is a bilinear form.

Representing f as a column vector and g as a row vector, from their product we obtain precisely a matrix representing a bilinear form equal to $f(\boldsymbol{u})g(\boldsymbol{v})$.

Example:

$$F\left(\begin{bmatrix}1\\2\\5\end{bmatrix},\begin{bmatrix}3\\4\end{bmatrix}\right)=\begin{bmatrix}1\\2\\5\end{bmatrix}\cdot\begin{bmatrix}3 & 4\end{bmatrix}=\begin{bmatrix}3 & 4\\6 & 8\\15 & 20\end{bmatrix}$$
$$rank\begin{bmatrix}3 & 4\\6 & 8\\15 & 20\end{bmatrix}=1$$

Some observations:

- $f(\boldsymbol{u})g(\boldsymbol{v})$ is a bilinear form in f and g, just that F required to define a tensor space;
- The matrix representation of $f(\boldsymbol{u})g(\boldsymbol{v})$ makes it clear to us how it is possible to generate only some matrices, particularly those with $rank\ =1$. This, too, is a requirement of our tensor space because $f(\boldsymbol{u})g(\boldsymbol{v})$ need not produce all of W but only a subset of it that can generate it entirely;

Nomenclature

We introduce the modern nomenclature of tensors, trusting that we understand a little of the underlying mechanism and thus avoid confusion.

The bilinear form F that generates values of W that can span it completely is called the tensor product and denoted $\otimes$. The notation then becomes:

$$\otimes : VxU \rightarrow W$$

The same symbol, also indicates the values produced by this map (i.e., tensors).

$$\boldsymbol{x}\otimes\boldsymbol{y}\ =\ \boldsymbol{w}\in W$$

The tensors generated by this map are, as repeated several times, only a subset of W and are called decomposables because they are derived from the vector product of two specific elements of V and U and thus are decomposable into these two vectors.

Finally, abandoning any attempt at clarity, the same symbol is also used to denote W, the tensor space.

$$W=\ V\otimes U$$

Let us try a second example, in this case of linear functions, using correct notation and terminologies. Let W be the set of linear functions from U to V and f a function in U^*. The vectors in W capable of spanning it completely are provided by the map F, which we correctly denote as a tensor product:

$$(\boldsymbol{v} \otimes f)(\boldsymbol{u}) = f(\boldsymbol{u}) \cdot \boldsymbol{v}$$

Thus W is a tensor space $V \otimes U^*$ of the linear functions from U to V.

Example:

$$\left(\begin{bmatrix} 3 \\ 4 \end{bmatrix} \otimes \begin{bmatrix} 1 & 2 \end{bmatrix}\right)(\boldsymbol{u}) = \underbrace{\left(\begin{bmatrix} 1 & 2 \end{bmatrix} \cdot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}\right)}_{scalar} \cdot \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

Definitions of tensor

We should have all the elements to understand the definition of tensor so let us present it without delay.

Given three vector spaces V, U and W and $\otimes : V \times U \rightarrow W$ a bilinear map, the pair $(W, \otimes)$ is called a tensor space (or product tensor space) over the underlying vector spaces V and U if two properties are satisfied:

- generation: $W = span(Im \otimes)$
- maximum span: $dim\ W = dim\ V \cdot dim\ U$

In the case, the vectors in W are called tensors over V and U, the map x tensor product, and, not infrequently, W is denoted as $V \otimes U$.

The definition reminds us that it is inaccurate to state, for example, that bilinear forms are tensors: we must always add that they are so in conjunction with two specific vector spaces and the tensor product (bilinear function) that maps them to tensor space.

Conclusion

In this article we have tried to provide a friendly and precise introduction of the tensor concept by bringing some examples, including numerical ones. For a more in-depth study we advocate the excellent text by Guo ([1]) that can dispel any doubts about the different definitions of tensors, their historical origin and applications in the main fields of physics, geometry and artificial intelligence.

A future article will attempt to clarify the usefulness of tensors by answering the legitimate question: what are tensors used for?

Bibliography

1. Hongyu Guo,What are tensors exactly, World Scientific
2. The tensor product demystified,https://www.math3ma.com/blog/the-tensor-product-demystified
3. C. T. J. Dodson, T. Poston, Tensor Geometry the Geometric Viewpoint and its Uses, Springer
4. Ray D'Inverno, Introducing Einstein's Relativity, Oxford University Press