Tensorama

John Kerl

Department of Mathematics, University of Arizona
Graduate Student Colloquium

February 10, 2010

Polysemy: the coexistence of many possible meanings for a word or phrase.

Let us be patient! These severe afflictions
Not from the ground arise;
But oftentimes celestial benedictions
Assume this dark disguise.
— Henry Wadsworth Longfellow (1807-1882).

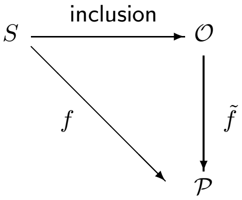
Tensors are encountered throughout math and physics. They are presented in at least four seemingly distinct guises:

(1) Tensor products from abstract algebra.
(2) Tensors as k-linear functions $T : V^k \to \mathbb{R}$.
(3) Tensors as k-dimensional arrays: a 0-tensor is related to a scalar, a 1-tensor is related to an array, a 2-tensor is related to a matrix, a 3-tensor is related to a 3-dimensional array, etc.
(4) Old-fashioned tensors from physics ("transform according to ..."). (In this guise, tensors appear particularly foreign. You will see lots of superscripts, subscripts, Einstein summation, etc. Why bother? Well, this is the way tensors are usually viewed in applications, so this is the language that your scientific collaborators will be speaking.)

Freeness

Definition: An object (group, ring, module, etc.) $\mathcal{O}$ is *free* on a set $S \subset \mathcal{O}$ if for any other respective object P and a morphism $f : S \to P$ there is a unique morphism $\tilde{f}$ from $\mathcal{O}$ to $\mathcal{P}$ such that $\tilde{f}|_S = f$.

A commutative diagram is a bit more intuitive:



Freeness, continued

Still ... *what*? I claim that we all already know this well — at least, if $\mathcal{O}$ and $\mathcal{P}$ are vector spaces, S is a basis for $\mathcal{O}$, and f is a linear transformation. Specifying the images of the basis vectors uniquely specifies the images of *all* vectors, since any vector is a linear combination of basis elements.

We have

$$A(\mathbf{v}) = A\left(\sum_{j=1}^{n} v_j \mathbf{e}_j\right) = \sum_{j=1}^{n} v_j A\left(\mathbf{e}_j\right)$$

where $A(\mathbf{e}_j)$ is the jth column of A.

Example:

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \left(\begin{array}{c|c} 1 & 0 \\ 0 & 1 \end{array} \right) = \left(\begin{array}{c|c} 1 & 3 \\ 2 & 4 \end{array} \right)$$

i.e. $\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} 3 \\ 4 \end{pmatrix}$.

Relations

Definition: a *relation* is some non-trivial expression (containing elements of an object) which is equal to the identity.

Group examples: $aba^{-1}b^{-1} = e$ (commutator relation) or $a^m = e$ (finite-order relation).

Intuition: A set $S \subset \mathcal{O}$ is free on S if the elements of S have only the minimal relations required for $\mathcal{O}$'s category.

Example: If G is a free group on $\{a, b, c\}$, then ab, a^2, $b^{-1}ac$, etc. don't simplify. But we must permit $aa^{-1} = e$ (which we wouldn't permit for a free semigroup) since we want G to *be* a group.

Example: A is free abelian on $\{a, b, c\}$. Sample elements: $3a + 2b$, $-7a + 4b - c$.

A free abelian group is a $\mathbb{Z}$-module.

Modules and vector spaces

If we allow rational instead of integer coefficients, we have a free $\mathbb{Q}$-module. Sample element: $\frac{2}{3}a + \frac{4}{7}b - 10c$.

Free means no non-trivial relations among basis elements: $Ka + Lb + Mc = 0$ implies $K = L = M = 0$. For vector spaces V, a basis set S is linearly independent $\iff$ V is free on S.

Vector spaces (free $\mathbb{R}$-modules) are typically obtained in one of two (very different) ways:

(1) Start with a large V and find a small S inside.

(2) Start with (a smaller) S and form a (larger) free $\mathbb{R}$-module V generated by S.

We say we work with equivalence classes. But what we often mean are coset representatives and transformation rules.

Example: $\mathbb{Z}/5\mathbb{Z}$. Do we really think

$$\{\ldots, -1, 4, 9, \ldots\} + \{\ldots, -2, 3, 8, \ldots\} = \{\ldots, -3, 2, 7, \ldots\}$$

every time we think of addition mod 5?

No! We think

$$4 + 3 = 7 \rightarrow 2.$$

We choose canonical representatives $\{0, 1, 2, 3, 4\}$. When we do modular arithmetic, we lift to $\mathbb{Z}$ and add (e.g. 7). Then we transform the answer, taking it ± 5 repeatedly, to get a canonical representative (e.g. 2).

Tensor products of vector spaces

For vector spaces[1] V and W, the tensor product $V \otimes W$ is the free vector space[2] on $V \oplus W$ — where the S is the (huge!) set of *all* $(\mathbf{v}, \mathbf{w})$ pairs — modulo the relations (for all scalars r and vectors $\mathbf{v}$, $\mathbf{w}$, etc.):

(i) $r\mathbf{v} \otimes \mathbf{w} = \mathbf{v} \otimes r\mathbf{w}$,

(ii) $(\mathbf{v}_1 + \mathbf{v}_2) \otimes \mathbf{w} = \mathbf{v}_1 \otimes \mathbf{w} + \mathbf{v}_2 \otimes \mathbf{w}$,

(iii) $\mathbf{v} \otimes (\mathbf{w}_1 + \mathbf{w}_2) = \mathbf{v} \otimes \mathbf{w}_1 + \mathbf{v} \otimes \mathbf{w}_2$.

So, elements $\mathbf{v} \otimes \mathbf{w}$ are *equivalence classes* of $(\mathbf{v}, \mathbf{w})$ pairs with these transformation rules. But here, unlike modular arithmetic, there is no clear favorite representative for each equivalence class. And, the generating set S is already large.

Example: in $\mathbb{R}^2 \otimes \mathbb{R}^2$,

$$\begin{pmatrix} 2 \\ 4 \end{pmatrix} \otimes \begin{pmatrix} 3 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \end{pmatrix} \otimes \begin{pmatrix} 6 \\ 10 \end{pmatrix}.$$

[1] In this talk, for simplicity, I take them to be finite-dimensional.
[2] A free abelian group along with the scalar operations is a vector space.

Tensor products of vector spaces in coordinates

Theorem: $\mathbb{R} \otimes V \cong V$ with $r \otimes \mathbf{v} \mapsto 1 \otimes r\mathbf{v}$.
Theorem: $(V_1 \oplus V_2) \otimes W \cong (V_1 \otimes W) \oplus (V_2 \otimes W)$ and likewise on the right.
Consequence: $\mathbb{R}^m \otimes \mathbb{R}^n \cong \mathbb{R}^{mn}$ with basis elements $\mathbf{e}_i \otimes \mathbf{f}_j$. Specifically,

$$\mathbf{v} \otimes \mathbf{w} = \left(\sum_{i=1}^{m} v_i \mathbf{e}_i \right) \otimes \left(\sum_{j=1}^{n} w_i \mathbf{f}_j \right) = \sum_{i=1}^{m} \sum_{j=1}^{n} v_i w_j (\mathbf{e}_i \otimes \mathbf{f}_j).$$

(Not all elements of $V \otimes W$ are of the pure form $\mathbf{v} \otimes \mathbf{w}$, but all elements are linear combinations of such.) Coefficients of $\mathbf{v} \otimes \mathbf{w}$ are $v_i w_j$. We can write these in a tableau:

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \otimes \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{array}{c|cc} & w_1 & w_2 \\ \hline v_1 & v_1 w_1 & v_1 w_2 \\ v_2 & v_2 w_1 & v_2 w_2 \end{array}; \quad \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{array}{c|cc} & 0 & 1 \\ \hline 1 & 0 & 1 \\ 0 & 0 & 0 \end{array}.$$

Then

$$\begin{pmatrix} 2 \\ 4 \end{pmatrix} \otimes \begin{pmatrix} 3 \\ 5 \end{pmatrix} = \begin{array}{c|cc} & 3 & 5 \\ \hline 2 & 6 & 10 \\ 4 & 12 & 20 \end{array} \quad \text{and} \quad \begin{pmatrix} 1 \\ 2 \end{pmatrix} \otimes \begin{pmatrix} 6 \\ 10 \end{pmatrix} = \begin{array}{c|cc} & 6 & 10 \\ \hline 1 & 6 & 10 \\ 2 & 12 & 20 \end{array}.$$

Much nicer! Now we have 2D arrays in coordinates: the link between guises (1) and (3) from the beginning of the talk is now becoming clear. Next, guise (2).

Linear functionals

Definition: the dual space of V is $V^* = \{\boldsymbol{\lambda} : V \to \mathbb{R} : \boldsymbol{\lambda} \text{ linear}\}$. It's the vector space of linear functionals on V.

Representation theorem (trivial for finite-dimensional inner-product spaces; Riesz-Fréchet for infinite-dimensional inner-product spaces): Each linear functional $\boldsymbol{\lambda} \in V^*$ has a unique $\mathbf{u}$ in V so that $\boldsymbol{\lambda}(\mathbf{v}) = \mathbf{u}^T\mathbf{v} = \mathbf{u}^*\mathbf{v}$ for all $\mathbf{v}$. Specifically,

$$\boldsymbol{\lambda}(\mathbf{v}) = \boldsymbol{\lambda}\left(\sum_{j=1}^{n} v_j \mathbf{e}_j\right) = \sum_{j=1}^{n} \boldsymbol{\lambda}(\mathbf{e}_j)\, v_j$$

defines the vector $\mathbf{u}$ with

$$u_j = \boldsymbol{\lambda}(\mathbf{e}_j).$$

Then

$$\boldsymbol{\lambda}(\mathbf{v}) = \mathbf{u}^*\mathbf{v} = \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = u_1v_1 + u_2v_2 + u_3v_3.$$

A linear functional (in V^*) is a row vector. A vector (in V) is a column vector.

Bases for linear functionals

Given a basis $\{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ for V, the dual basis $\{\mathbf{b}_1^*, \ldots, \mathbf{b}_n^*\}$ for V^* is such that

$$\mathbf{b}_i^*(\mathbf{b}_j) = \delta_{ij}.$$

How to compute them? Notation for matrices of row or column vectors (e.g. $n = 3$):

$$\left(\mathbf{u} \,\middle|\, \mathbf{v} \,\middle|\, \mathbf{w}\right) = \left(\begin{array}{c|c|c} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{array}\right) \quad \text{and} \quad \left(\begin{array}{c} \boldsymbol{\lambda} \\ \hline \boldsymbol{\mu} \\ \hline \boldsymbol{\nu} \end{array}\right) = \left(\begin{array}{ccc} \lambda_1 & \lambda_2 & \lambda_3 \\ \hline \mu_1 & \mu_2 & \mu_3 \\ \hline \nu_1 & \nu_2 & \nu_3 \end{array}\right).$$

Then $\mathbf{b}_i^*(\mathbf{b}_j) = \delta_{ij}$ means

$$\left(\begin{array}{c} \mathbf{b}_1^* \\ \hline \vdots \\ \hline \mathbf{b}_n^* \end{array}\right) \left(\mathbf{b}_1 \,\middle|\, \cdots \,\middle|\, \mathbf{b}_n\right) = I \quad \text{i.e.} \quad \left(\begin{array}{c} \mathbf{b}_1^* \\ \hline \vdots \\ \hline \mathbf{b}_n^* \end{array}\right) = \left(\mathbf{b}_1 \,\middle|\, \cdots \,\middle|\, \mathbf{b}_n\right)^{-1}.$$

Example: if $\mathbf{b}_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}$ and $\mathbf{b}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, then

$$\left(\begin{array}{c} \mathbf{b}_1^* \\ \hline \mathbf{b}_2^* \end{array}\right) = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1/2 & -1/2 \\ 0 & 1 \end{pmatrix}.$$

So $\mathbf{b}_1^* = \begin{pmatrix} 1/2 & -1/2 \end{pmatrix}$ and $\mathbf{b}_2^* = \begin{pmatrix} 0 & 1 \end{pmatrix}$.

Multilinear functionals

Definition: $\phi : V \oplus W \to \mathbb{R}$ is *bilinear* if it is linear in each slot. Multilinear functions $\phi : \bigoplus_{i=1}^{k} V_i \to \mathbb{R}$ are defined similarly.

Examples: linear functional ($k = 1$), dot product ($k = 2$), determinant (as a function not on matrices but on k vectors in a k-dimensional vector space).

Definition of tensor product of dual spaces: for $\boldsymbol{\lambda} \in V^*$ and $\boldsymbol{\mu} \in W^*$,

$$(\boldsymbol{\lambda} \otimes \boldsymbol{\mu}) : V \oplus W \to \mathbb{R} := \boldsymbol{\lambda}(\mathbf{v})\boldsymbol{\mu}(\mathbf{w}).$$

This is a bilinear map; similarly, tensor products of more than two dual spaces consist of multilinear maps.

In coordinates:

$$(\boldsymbol{\lambda} \otimes \boldsymbol{\mu})(\mathbf{v}, \mathbf{w}) = \boldsymbol{\lambda} \otimes \boldsymbol{\mu} \left(\sum_{i=1}^{m} v_i \mathbf{e}_i, \sum_{j=1}^{n} w_j \mathbf{f}_j \right) = \sum_{i=1}^{m} \sum_{j=1}^{n} v_i w_j \boldsymbol{\lambda}(\mathbf{e}_i)\boldsymbol{\mu}(\mathbf{f}_j).$$

Now, the $\boldsymbol{\lambda}(\mathbf{e}_i)$'s and $\boldsymbol{\mu}(\mathbf{f}_j)$'s uniquely specify row vectors as before. So now, for tensor products of row vectors, just as we had with tensor products of column vectors, we have tensors as 2D arrays in coordinates. Guises (1), (2), and (3) have been unified! Next: guise (4), by route of first going back and shedding more light on guises (2) and (3).

Change of basis

We call elements of V, $V \otimes W$, etc. *contravariant tensors*; we call elements of V^*, $V^* \otimes W^*$, etc. *covariant tensors*. Why?

Notation: E, F are bases for V. Write $[\mathbf{v}]_E$, $[\boldsymbol{\lambda}]_E$ for column/row vectors in E coordinates. Similarly for $[\mathbf{v}]_F$ and $[\boldsymbol{\lambda}]_F$.

Example:

$$E = \{\mathbf{e}_1, \mathbf{e}_2\} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}, \qquad F = \{\mathbf{f}_1, \mathbf{f}_2\} = \left\{ \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\};$$

$$[\mathbf{v}]_E = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \qquad [\boldsymbol{\lambda}]_E = \begin{pmatrix} 3 & 4 \end{pmatrix}.$$

What are $[\mathbf{v}]_F$ and $[\boldsymbol{\lambda}]_F$?

$$\begin{pmatrix} v_{1,E} \\ v_{2,E} \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix} = v_{1,F} \begin{pmatrix} 2 \\ 0 \end{pmatrix} + v_{2,F} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} v_{1,F} \\ v_{2,F} \end{pmatrix};$$

$$\begin{pmatrix} v_{1,F} \\ v_{2,F} \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} 1/2 & -1/2 \\ 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 3 \\ 4 \end{pmatrix} = \begin{pmatrix} -1/2 \\ 4 \end{pmatrix}.$$

Explicit change of basis for column and row vectors, continued

Definition: The change-of-basis matrix Q from basis E to basis F, so that $[\mathbf{v}]_E = Q[\mathbf{v}]_F$, is

$$Q = \left(\mathbf{f}_1 \,\middle|\, \cdots \,\middle|\, \mathbf{f}_n \right)_E .$$

What about $[\boldsymbol{\lambda}]_F$? Regardless of basis used, $\boldsymbol{\lambda}(\mathbf{v})$ should be the same number:

$$[\boldsymbol{\lambda}]_E[\mathbf{f}_1]_E = \begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 6; \qquad [\boldsymbol{\lambda}]_E[\mathbf{f}_2]_E = \begin{pmatrix} 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 7.$$

So,

$$[\boldsymbol{\lambda}]_F[\mathbf{f}_1]_F = \begin{pmatrix} \lambda_1 & \lambda_2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 6; \qquad [\boldsymbol{\lambda}]_F[\mathbf{f}_2]_F = \begin{pmatrix} \lambda_1 & \lambda_2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 7.$$

So $[\boldsymbol{\lambda}]_F = \begin{pmatrix} \lambda_1 & \lambda_2 \end{pmatrix} = \begin{pmatrix} 6 & 7 \end{pmatrix}$. Stacking these equations together and generalizing for arbitrary n, we have

$$[\boldsymbol{\lambda}]_E \left(\mathbf{f}_1 \,\middle|\, \cdots \,\middle|\, \mathbf{f}_n \right)_E = [\boldsymbol{\lambda}]_F \left(\mathbf{f}_1 \,\middle|\, \cdots \,\middle|\, \mathbf{f}_n \right)_F = [\boldsymbol{\lambda}]_F I, = [\boldsymbol{\lambda}]_F,$$

i.e.

$$[\boldsymbol{\lambda}]_E \, Q = [\boldsymbol{\lambda}]_F .$$

In summary, column vectors and row vectors transform differently on change of coordinates:

$$[\mathbf{v}]_E \xleftarrow{Q} [\mathbf{v}]_F; \qquad [\boldsymbol{\lambda}]_E \xrightarrow{Q} [\boldsymbol{\lambda}]_F .$$

Tensors in physics

Instead of vector spaces V and V^* with real coefficients, use tangent bundles/cobundles TM and T^*M over a manifold M. The coefficients are $C^\infty(M)$ functions. (There are vector spaces attached to each point on the manifold, and the coefficients vary smoothly as you move from one point to another.)

The change-of-basis matrix in coordinates is $Q = \partial x_i / \partial y_j$, whose inverse is $Q^{-1} = \partial y_i / \partial x_j$. Then

$$[\lambda_i]_F = \sum_{j=1}^{n} \frac{\partial x^j}{\partial y^i} [\lambda_j]_E \qquad \text{covariant;}$$

$$[v^i]_F = \sum_{j=1}^{n} \frac{\partial y^i}{\partial x^j} [v^j]_E \qquad \text{contravariant.}$$

Conventions: subscripts outside derivative denominators, and superscripts inside derivative denominators, are called covariant indices. Superscripts outside derivative denominators, and subscripts inside derivative denominators, are called contravariant indices.

Physicists say that a covariant/contravariant tensor, respectively, is *anything* which transforms this way on change of coordinates.

Aside: Tensor algebras and mixed tensors

The tensor product of vector spaces is another vector space. You can add two tensors together, or multiply one tensor by a scalar. But if you can moreover define some way $\mathbf{u} * \mathbf{v}$ to multiply one tensor $\mathbf{u}$ by another tensor $\mathbf{v}$, you have an algebra. The product is typically of higher dimension than either of the operands: if $\mathbf{u} \in V \otimes V$ and $\mathbf{v} \in V$, then $\mathbf{u} * \mathbf{v} \in V \otimes V \otimes V$. This leads to a graded algebra which I would discuss more today, if there were time.

If $\mathbf{u} * \mathbf{v} = \pm \mathbf{v} * \mathbf{u}$ then we get symmetric and alternating tensor algebras, respectively: $\mathbf{uv}$ and $\mathbf{u} \wedge \mathbf{v}$.

One can tensor together $V \otimes V^*$, etc. Such tensors are of mixed valence, not purely covariant or contravariant. In fact, plain old linear transformations may be viewed as being of this form.

There's more . . .

Tensors in all four of their guises have been unified — they really are all the same thing. There is a fifth guise, which I (unfortunately) stopped thinking about shortly after (on the second attempt) I (high-)passed my geometry-topology qualifying exam: the geometrical guise.

Namely:

- The bivector $\mathbf{u} \wedge \mathbf{v}$ is the equivalence class of all vector pairs coplanar with $\mathbf{u}$ and $\mathbf{v}$, with the same signed area.
- The symmetric product $\mathbf{uv}$ is all pairs coplanar to $\mathbf{u}$ and $\mathbf{v}$ with the same inner product.
- The bifunctional $\mathbf{u}^* \wedge \mathbf{v}^*$ measures the area spanned by two other vectors, projected onto the plane spanned by $\mathbf{u}$ and $\mathbf{v}$.
- We can also form equivalence classes using things like

$$(\boldsymbol{\lambda} \wedge \boldsymbol{\mu})(\mathbf{u}, \mathbf{v}) = (\boldsymbol{\lambda} \wedge \boldsymbol{\mu})(\mathbf{u} \wedge \mathbf{v}) = \det(\boldsymbol{\lambda}|\boldsymbol{\mu}) \det(\mathbf{u}|\mathbf{v}).$$

These generalize to three and more dimensions. Although I don't know the history well enough, I suspect that it was precisely these geometric notions which led to the axioms we have today.

Thanks for attending!