General tensors

Three definitions of the term

Definition 1: A *tensor of order* (p,q) [hence of *rank* $p+q$] is a multilinear function

$$\underline{A}\colon \underbrace{\mathcal{V}^* \times \cdots \times \mathcal{V}^*}_{p \text{ times}} \times \underbrace{\mathcal{V} \times \cdots \times \mathcal{V}}_{q \text{ times}} \to \mathbf{R}.$$

(*Multilinear* means linear in each variable when the other $p+q-1$ variables are fixed.) One calls p the *contravariant rank* of $\underline{A}$ and q the *covariant rank*. The space of tensors of order (p,q) is

$$\mathcal{T}^p_q(\mathcal{V}) \equiv \underbrace{\mathcal{V} \otimes \cdots \otimes \mathcal{V}}_{p \text{ times}} \otimes \underbrace{\mathcal{V}^* \otimes \cdots \otimes \mathcal{V}^*}_{q \text{ times}}.$$

(The migration of the asterisk is deliberate!)

Definition 2: [The matrix of] a tensor of order (p,q) is a table of $(\dim \mathcal{V})^{p+q}$ numbers,

$$\{A^{j_1 \ldots j_p}_{k_1 \ldots k_q}\},$$

such that, under a change of basis which causes the coordinates of a ("contravariant") vector in $\mathcal{V}$ to transform as

$$v^j \mapsto S^j{}_l \, v^l \equiv {}^{\text{new}}v^j$$

[hence the coordinates of a covector in $\mathcal{V}^*$ to transform as $U_k \mapsto (S^{-1})^m{}_k \, U_m$], the components of the tensor transform according to the law

$$A^{j_1 \ldots j_p}_{k_1 \ldots k_q} \mapsto S^{j_1}{}_{l_1} \, S^{j_2}{}_{l_2} \cdots S^{j_p}{}_{l_p} \, (S^{-1})^{m_1}{}_{k_1} \cdots (S^{-1})^{m_q}{}_{k_q} \, A^{l_1 \ldots l_p}_{m_1 \ldots m_q}\,.$$

The indices $\{j_1, \ldots, j_p\}$ are called *contravariant indices*, and $\{k_1, \ldots, k_q\}$ are called *covariant indices*. [The mnemonic "Co—Low" can be used to remember which is which.]

Leibnitz would write

$${}^{\text{new}}A^{j_1 \ldots j_p}_{k_1 \ldots k_q} = \frac{\partial \xi^{j_1}}{\partial x^{l_1}} \cdots \frac{\partial \xi^{j_p}}{\partial x^{l_p}} \, \frac{\partial x^{m_1}}{\partial \xi^{k_1}} \cdots \frac{\partial x^{m_q}}{\partial \xi^{k_q}} \, A^{l_1 \ldots l_p}_{m_1 \ldots m_q}\,,$$

which is easier to remember. Here $A^{\cdots}_{\cdots}$ are the components in the x coordinate system, and ${}^{\text{new}}A^{\cdots}_{\cdots}$ are the components in the ξ coordinate system.

Definition 3: Start with the Cartesian product

$$\underbrace{\mathcal{V} \times \mathcal{V} \times \cdots \times \mathcal{V}}_{p} \times \underbrace{\mathcal{V}^* \times \cdots \times \mathcal{V}^*}_{q};$$

form the (huge, infinite-dimensional) vector space $\mathcal{W}$ of all formal linear combinations of elements of this set; consider the subspace $\mathcal{X} \subset \mathcal{W}$ spanned by elements of the form

$$(\alpha\vec{v}_1 + \vec{u}_1, \ldots) - \alpha(\vec{v}_1, \ldots) - (\vec{u}_1, \ldots) \qquad (\alpha \in \mathbf{R})$$

and similar elements associated with the other $p+q-1$ "slots"; then

$$\mathcal{T}^p_q(\mathcal{V}) \equiv \mathcal{W}/\mathcal{X},$$

and the equivalence class containing $(\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_p, \tilde{V}^1, \ldots, \tilde{V}^q)$ is called $\vec{v}_1 \otimes \vec{v}_2 \otimes \cdots \otimes \tilde{V}^q$.

Familiar examples of tensor spaces

$$\begin{aligned}
&\mathcal{T}^0_0 = \mathbf{R} \\
&\mathcal{T}^0_1 = \mathcal{V}^* \\
&\mathcal{T}^1_0 = \mathcal{V}^{**} = \mathcal{V} \\
&\mathcal{T}^1_1 = \mathcal{V} \otimes \mathcal{V}^* = \mathcal{L}(\mathcal{V};\mathcal{V}) \\
&\mathcal{T}^0_2 = \mathcal{V}^* \otimes \mathcal{V}^* = \text{ space of bilinear forms on } \mathcal{V} \quad (\mathcal{V} \times \mathcal{V} \to \mathbf{R}) \\
&\mathcal{T}^2_0 = \mathcal{V} \otimes \mathcal{V} = \text{ space of bilinear forms on } \mathcal{V}^*
\end{aligned}$$

Relations among the definitions (indicated briefly)

<u>(3) → (1)</u>: $\vec{v}_1 \otimes \cdots \otimes \tilde{V}^q$ defines a multilinear functional on $\mathcal{V}^* \times \cdots \times \mathcal{V}$ by

$$(\vec{v}_1 \otimes \cdots \otimes \tilde{V}^q)(\tilde{U}^1, \ldots, \tilde{U}^p, \vec{u}_1, \ldots, \vec{u}_q) \\ \equiv \tilde{U}^1(\vec{v}_1) \cdots \tilde{U}^p(\vec{v}_p)\, \tilde{V}^1(\vec{u}_1) \cdots \tilde{V}^q(\vec{u}_q).$$

<u>(1) → (2)</u>: The matrix of a multilinear functional $\underline{A}$ with respect to a basis $\{\vec{d}_j\}$ for $\mathcal{V}$ and the dual basis $\{\tilde{D}^j\}$ for $\mathcal{V}^*$ is

$$\left\{ \underline{A}\left(\tilde{D}^{j_1}, \ldots, \tilde{D}^{j_p}, \vec{d}_{k_1}, \ldots, \vec{d}_{k_q} \right) \right\}$$

(Each index j_α or k_α ranges from 1 to $\dim \mathcal{V}$.)

<u>(3) → (2)</u>: The matrix of an element $\vec{v}_1 \otimes \cdots \otimes \tilde{V}^q \in \mathcal{T}^p_q$ is

$$\left\{ v_1^{j_1}\, v_2^{j_2} \cdots v_p^{j_p}\, V^1_{k_1} \cdots V^q_{k_q} \right\}.$$

Warning: Elements of the form $\vec{v}_1 \otimes \cdots \otimes \tilde{V}^q$ *span* $\mathcal{T}^p_q$, but they do not constitute *all* of it. The most general element is a linear combination of such products; it can't be factored as a single product. Example: We've noted that $\mathcal{T}^1_1$ is isomorphic to $\mathcal{L}(\mathcal{V};\mathcal{V})$, but the elements of $\mathcal{T}^1_1$ of the form $\vec{v} \otimes \tilde{V}$ are operators of rank 1 (or 0), in the old sense of "rank". (The range of such an operator is $\{\alpha\vec{v}\}$.) Further example: An example of a nonfactorable element of $\mathcal{V} \otimes \mathcal{V}$ (provided that $\dim \mathcal{V} \geq 4$) is

$$\vec{d}_1 \otimes \vec{d}_2 + \vec{d}_3 \otimes \vec{d}_4$$

when the four vectors are linearly independent. There is no way to write this as $\vec{v}_1 \otimes \vec{v}_2$.

Even when a tensor factors (is *simple*), that representation is not unique; for example,

$$\vec{v}_1 \otimes (\alpha\vec{v}_2) = (\alpha\vec{v}_1) \otimes \vec{v}_2 \qquad \forall \alpha \in \mathbf{R}.$$

A basis for $\mathcal{T}^p_q$

is $\left\{\vec{d}_{j_1} \otimes \cdots \otimes \vec{d}_{j_p} \otimes \tilde{D}^{k_1} \otimes \cdots \otimes \tilde{D}^{k_q}\right\}$. (Note that "co" indices are now up and "contra" indices down. This makes the summation convention work out right when an element of the space is expanded as the sum of its matrix elements times the corresponding basis elements.)

Since each index independently ranges up to $\dim \mathcal{V}$, it follows that

$$\dim\,\left(\mathcal{T}^p_q\right) = (\dim\,\mathcal{V})^{p+q}.$$

For contrast, note that

$$\dim\,[\underbrace{\mathcal{V} \oplus \cdots \oplus \mathcal{V}}_{p} \oplus \underbrace{\mathcal{V}^* \oplus \cdots \oplus \mathcal{V}^*}_{q}] = (p+q)\dim\,\mathcal{V}.$$

Recall that "$\oplus$" is effectively the same thing as "$\times$" (the Cartesian product) with addition and scalar multiplication defined.

$$(1,2,3) \oplus (0,4) \simeq (1,2,3,0,4); \qquad (1,2,3) \otimes (0,4) \simeq \begin{pmatrix} 0 & 0 & 0 \\ 4 & 8 & 12 \end{pmatrix}$$

Remarks

(A) All this can be generalized to tensor products of *arbitrary* vector spaces:

$\mathcal{V}_1 \otimes \mathcal{V}_2 \otimes \cdots \otimes \mathcal{V}_r$ consists of multilinear functions on $\mathcal{V}_1^* \times \mathcal{V}_2^* \times \cdots \times \mathcal{V}_r^*$.

(The analogues of definitions 2 and 3 are left to the reader.)

As an application of such a construction, consider a system of partial differential equations ($\equiv$ a PDE for a vector-valued unknown) with independent variable $\vec{x} \equiv \{x^\mu\} \in \mathbf{R}^3$ and dependent variable $\vec{\phi} \equiv \{\phi^j\} \in \mathcal{V}$, where $\mathcal{V}$ might be $\mathbf{C}^q$, say. [An example of the application is the spinor-valued wave functions of particles in quantum physics. In general, we have this situation whenever the value of the unknown function at a point is neither a scalar, nor a vector whose possible directions can be identified with the directions of physical space $\mathbf{R}^3$ (or whatever the domain space is).] The PDE could be of the form

$$\sum_{\mu=1}^{3} \frac{\partial^2 \vec{\phi}}{\partial x^{\mu 2}} + \sum_{\mu=1}^{3} \underline{M}^\mu \frac{\partial \vec{\phi}}{\partial x^\mu} = \vec{0}.$$

Here

$$\left\{ \frac{\partial \vec{\phi}}{\partial x^\mu} \right\} \equiv \{\nabla \phi^j\} \in \mathcal{V} \otimes (\mathbf{R}^3)^*,$$

$$\{\underline{M}^\mu\} \equiv \left\{ M^{\mu j}{}_k \right\} \in \mathcal{V} \otimes \mathcal{V}^* \otimes \mathbf{R}^3,$$

etc.

(B) Generalization of the isomorphism

$$\text{(operators)} \ \leftrightarrow \ \text{(bilinear functions)} :$$

Any element of $\mathcal{V}_1 \otimes \cdots \otimes \mathcal{V}_r$ can be thought of as a multilinear operator on $\mathcal{V}_j^* \times \mathcal{V}_{j+1}^* \times \cdots \times \mathcal{V}_r^*$ into $\mathcal{V}_1 \otimes \cdots \otimes \mathcal{V}_{j-1}$, for any j between 2 and $r-1$. (In fact, since $\mathcal{V} \otimes \mathcal{U}$ is isomorphic to $\mathcal{U} \otimes \mathcal{V}$, the same remark applies to any breakup of the list of factors into two parts.)

(C) A multilinear function on $\mathcal{V}_1 \times \cdots \times \mathcal{V}_r$ is equivalent to a *linear* functional on $\mathcal{V}_1 \otimes \cdots \otimes \mathcal{V}_r$ (and totally different from a linear functional on $\mathcal{V}_1 \oplus \cdots \oplus \mathcal{V}_r$). (See homework.)

(D) For $\underline{A} \in \mathcal{T}_q^p$ and $\underline{B} \in \mathcal{T}_s^r$, we can define $\underline{A} \otimes \underline{B} \in \mathcal{T}_{q+s}^{p+q}$ by

$$(\underline{A} \otimes \underline{B})(\vec{v}_1, \ldots, \tilde{V}^s) \equiv \underline{A}(\vec{v}_1, \ldots)\, \underline{B}(\vec{v}_{p+1}, \ldots)$$

(cf. Definition 1 of tensors); or by multiplying coefficients:

$$(\underline{A} \otimes \underline{B})^{j_1 \ldots j_{p+r}}_{k_1 \ldots k_{q+s}} \equiv A^{j_1 \ldots j_p}_{k_1 \ldots k_q}\, B^{j_{p+1} \ldots j_r}_{k_{q+1} \ldots k_s}$$

(cf. Definition 2 of tensors).

(E) If $\mathcal{V}$ is an inner-product space, we can "raise and lower indices" by $\underline{G}$:

$$A^k{}_j = g^{kl} A_{lj}\,, \qquad A_{lj} = g_{lk} A^k{}_j\,.$$

If we stick to ON bases, there is no need to distinguish contravariant from covariant indices. (See Sec. 35 and Simmonds). Note that

$$g_k{}^m \equiv g_{kl}\, g^{lm} = \delta_k{}^m,$$

and attaching a second $g^{\cdot\cdot}$ raises the other index, completing the conversion of $g_{\cdot\cdot}$ into $g^{\cdot\cdot}$. This vindicates the notational convention $\underline{G}^{-1} \leftrightarrow \{g^{\cdot\cdot}\}$.

Given the matrix representation of a (p,q)-tensor, we can obtain a $(p-1,q-1)$-tensor by setting any "up–down" pair of indices equal and summing: for example,

$$B_{jk} \equiv A^{\;\;\;l}_{jkl} \qquad \left(\sum_{l=1}^{N} \text{ understood}\right).$$

This operation is called *contracting* indices. B_{jk} has the right transformation property because

$$S^l{}_i (S^{-1})^m{}_l = \delta^m_i .$$

In other words, the contracted index pair is invariant under basis change for the same reason that the scalar quantity $\tilde{U}(\vec{v}) = U_j v^j$ is invariant.

Alternatively, consider a basis element

$$\tilde{V}^1 \otimes \vec{v}_2 \otimes \tilde{V}^3 \otimes \tilde{V}^4 .$$

The corresponding contracted ("B") tensor is

$$\tilde{V}^4(\vec{v}_2)\, \tilde{V}^1 \otimes \tilde{V}^3 ,$$

which is obviously intrinsically defined. Since each of the basis elements is arbitrary, this defines the contraction operation on all tensors in $\mathcal{T}^1_3$.

SPECIAL CASE: $\underline{A} \in \mathcal{T}^1_1 \;\Rightarrow\; B = A^j{}_j = \operatorname{tr} \underline{A} \in \mathbf{R}$.

Note in contrast that $\sum A^{jj}$ is *not* invariant under (nonorthogonal) basis changes; thus it has no intrinsic meaning. The same is true for any sum over indices which are both covariant or both contravariant.

ANOTHER EXAMPLE: In Riemannian geometry and general relativity one studies something called the *Riemann curvature tensor*, $R^j{}_{klm}$. From this are defined the *Ricci tensor*, $R_{km} \equiv R^j{}_{kjm}$, and the *curvature scalar*, $R \equiv R^k{}_k \equiv g^{kl} R_{lk}$. One can also construct higher-order scalar objects such as

$$R^{jk}{}_{lm}\, R^{lm}{}_{pq}\, R^{pq}{}_{jk} .$$

Totally antisymmetric tensors *(Chap. 8)*

We have run out of time, but I want to indicate briefly the relationships among antisymmetric tensors, determinants, volume, and integration. ("Totally antisymmetric" means antisymmetric under interchange of any pair of indices (or argument vectors). The indices or arguments must therefore be either all covariant or all contravariant.)

Certain integrals over hypersurfaces of dimension p are naturally described in terms of rank-p antisymmetric tensors, *without reference to a metric* (inner product). Such integrals generalize the notion of "flux through a surface" in classical vector analysis.

The classical formulation of the magnetic flux through a surface is

$$\Phi_S = \iint_S (\vec{B}\cdot\hat{n})\,dS,$$

where dS is the element of surface area. In terms of two coordinates, u and v, parametrizing S, the unit normal vector is

$$\hat{n} = \frac{\frac{\partial\vec{x}}{\partial u}\times\frac{\partial\vec{x}}{\partial v}}{\left\|\frac{\partial\vec{x}}{\partial u}\times\frac{\partial\vec{x}}{\partial v}\right\|}\,;$$

but the denominator just cancels a factor in the definition of dS, so that the integral in fact doesn't depend on the inner product used to define length and orthogonality. The upshot is that

$$\Phi_S = \iint_S \left[B_x\,dy\,dz + B_y\,dz\,dx + B_z\,dx\,dy\right],$$

which can be written as

$$\iint_S \frac{1}{2}\sum_{j,k}\Omega_{jk}\,dx^j\,dx^k$$

whre $\underline{\Omega}$ is the antisymmetric tensor associated to $\vec{B}$ via the cross product.

To see how antisymmetric tensors relate to determinants, consider the elementary antisymmetrized monomial ($\{\hat{E}^j\}\equiv$ dual basis to the natural basis of $\mathbf{R}^N$)

$$\hat{E}^1\wedge\hat{E}^2 \equiv \hat{E}^1\otimes\hat{E}^2 - \hat{E}^2\otimes\hat{E}^1,$$
$$[\hat{E}^1\wedge\hat{E}^2](\vec{v},\vec{u}) \equiv v^1u^2 - v^2u^1.$$

If $\dim\mathcal{V}=2$, this is the determinant of the matrix whose columns are $\vec v$ and $\vec u$; furthermore, this number is, up to sign, the area of the parallelogram spanned by those two vectors. In $\mathbf{R}^3$, the area of the parallelogram determined by two vectors is $\|\vec v\times\vec u\|$, which is still the norm of an antisymmetric combination of the vectors. For three vectors in $\mathbf{R}^3$, the volume of the parallelepiped they span is

$$|\vec v\cdot(\vec u\times\vec w)| = \pm\begin{vmatrix} v^1 & v^2 & v^3 \\ u^1 & u^2 & u^3 \\ w^1 & w^2 & w^3 \end{vmatrix} \equiv [\hat E^1\wedge\hat E^2\wedge\hat E^3](\vec v,\vec u,\vec w).$$

This pattern continues to higher dimensions and provides the key to defining integration on hypersurfaces (p-dimensional curved sets in N-dimensional space) or manifolds (abstract p-dimensional spaces). The "wedge" operation, $\wedge$, incidentally, is defined for any list of covectors as arguments (and also for any list of vectors), although here we've used it only on the elements of an ON basis for $\mathcal{V}^*$. The definition of the wedge often differs by a factor of $p!$ from the one I've given here.

Remark: A general tensor of rank greater than 2 is *not* a sum of a totally symmetric and a totally antisymmetric part. More complicated intermediate symmetry types exist, symbolized by *Young diagrams*.

Duality

The connection between vectors and antisymmetric 2-tensors in $\mathbf{R}^3$ extends in $\mathbf{R}^N$ to an isomorphism

$$(\text{antisymmetric } p\text{-tensors}) \;\leftrightarrow\; (\text{antisymmetric } (N-p)\text{-tensors}).$$

The cross product corresponds to the case $N=3$, $p=2$, $N-p=1$. (An object with only one index counts as antisymmetric by default.) The classical notation for this duality isomorphism is

$$B^{j_1\ldots j_{N-p}} = \epsilon^{j_1\ldots j_{N-p}k_1\ldots k_p}\,A_{k_1\ldots k_p}\,,$$

where the tensor components are those with respect to an ON basis, and the object ϵ is defined by

$$\epsilon^{12\ldots N}\equiv 1,$$
$$\epsilon^{j_1\ldots j_N}\equiv 0 \text{ if any two indices are equal,}$$
$$\epsilon \text{ is antisymmetric under permutations.}$$

Nowadays it is more fashionable to give the definition by describing the action of the isomorphism on the ON basis elements rather than on the components of an arbitrary tensor:

$$*\left(\hat e_{k_1}\wedge\cdots\wedge\hat e_{k_p}\right) = \pm\hat e_{j_1}\wedge\cdots\wedge\hat e_{j_{N-p}}\,.$$

This mapping, called the *dual* or the *Hodge star*, requires

1) an inner product (metric), to define "ON basis";

2) an orientation, to fix the signs.

References on tensors, exterior algebra, and applications to differential geometry

B. Schutz, *Geometrical Methods of Mathematical Physics*, Cambridge U. P., 1980, Chaps. 2 and 4.

R. L. Bishop and S. I. Goldberg, *Tensor Analysis on Manifolds*, Dover, 1980, Chaps. 2 and 4.

V. I. Arnold, *Mathematical Methods of Classical Mechanics*, Springer, 1978, Chap. 7.

W. L. Burke, *Applied Differential Geometry*, Cambridge U. P., 1985, relevant chapters.

C. T. J. Dodson and T. Poston, *Tensor Geometry*, Pitman, 1977, relevant chapters.

I especially recommend the first half of Chapter 4 of Schutz.