Tensors

Definition 1. ***Covariance, contravariance, and invariance.*** Let ϕ be an arbitrary-valued (vector, scalar, etc.) linear function of V and consider a change of basis $B \to \tilde{B}$ defined by $\tilde{e}_j = \sum_i e_i(\mathbf{T})_{ij}$ where $\mathbf{T}$ is invertible.[1] Let $S = \{\phi(e_1), \ldots, \phi(e_n)\}$ be the set of values of ϕ on B. This basis-dependent set can be characterised as follows.

> S is *covariant* with B if it obeys the same transformation law as B: $\qquad \phi(\tilde{e}_j) = \sum_i \phi(e_i)(\mathbf{T})_{ij}$
>
> S is *contravariant* to B if it obeys the inverse transformation law of B: $\qquad \phi(\tilde{e}_i) = \sum_j (\mathbf{T}^{-1})_{ij}\phi(e_j)$

The elements of a *covariant* set are typically denoted with subscript indices, $\phi_i = \phi(e_i)$, whereas the elements of a *contravariant* set are typically denoted with superscript indices, $\phi^i = \phi(e_i)$. Abstract quantities such as vectors, functionals, and operators are called *invariant* since they do not depend on the choice of basis.

Example 1. ***Vector coordinates are contravariant.*** A vector v is an *invariant* quantity and can be expressed as $v = \sum_i e_i v^i$ or as $v = \sum_i \tilde{e}_i \tilde{v}^i$, v^i and $\tilde{v}^i$ being coordinates with respect to B and $\tilde{B}$, respectively. Using the fact that the basis vectors themselves form a *covariant* set (by definition), we can show that the coordinates must form a *contravariant* set: noting that $\tilde{e}_i = \sum_k e_k(\mathbf{T})_{ki} \implies \sum_i \tilde{e}_i(\mathbf{T}^{-1})_{ij} = e_j$, the invariance of v implies $\sum_i \tilde{e}_i \tilde{v}^i = \sum_j e_j v^j = \sum_{ji} \tilde{e}_i(\mathbf{T}^{-1})_{ij} v^j \implies \tilde{v}^i = \sum_j (\mathbf{T}^{-1})_{ij} v^j$ where the second step follows from the fact that $\tilde{B}$ is linearly independent. That is, the *invariance* of $v = \sum_i e_i v^i$ requires that the *covariance* of the basis is cancelled by the *contravariance* of coordinates.

Definition 2. ***Linear functional.*** A *linear functional* $f : V \to \mathbb{C}$ is a scalar-valued function on V that satisfies linearity, i.e. $f(v + w) = f(v) + f(w)$ and $f(cv) = cf(v)$ for all $c \in \mathbb{C}$ and all $v, w \in V$.

Definition 3. ***Dual space V^*.*** The *dual space* V^* of a vector space V is the space of linear functionals on V, which itself forms a vector space with vector addition, $(f + g) \in V^*$, and scalar multiplication, $(cf) \in V^*$ defined by

$$(f + g)(v) \equiv f(v) + g(v) \qquad\qquad (cf)(v) \equiv cf(v)$$

for all $f, g \in V^*$, $v \in V$, and $c \in \mathbb{C}$. Its dimension is $\dim V^* = \dim V$, which follows from Prop 1.

Proposition 1. ***Basis for V^* (canonical dual basis).*** *If $B = \{e_1, \ldots, e_n\}$ is a basis for for V then $B^* = \{e^1, \ldots, e^n\}$, with elements $e^i \in V^*$ defined by $e^i(e_j) = \delta^i_j$, is a basis for V^*. This "canonical dual basis" transforms contravariantly relative to B.*

Proof: Let f be an arbitrary element of V^*, let v be an arbitrary element of V whose basis expansion is $v = \sum_i e_i v^i$, and let $c_1, \ldots, c_n$ denote scalar values, $c_i \in \mathbb{C}$. Also, note that the identity $f(v) = f\left(\sum_i e_i v^i\right) = \sum_i f(e_i)v^i$ holds for all $f \in V^*$ by linearity. Finally, note that the null vector f_0 in V^* is defined by $f_0(v) = 0$ for all $v \in V$. Therefore:

1. ***$v^i = e^i(v)$.*** $\qquad e^i(v) = \sum_j e^i(e_j)v^j = \sum_j \delta^i_j v^j = v^i$
2. ***B^* spans the dual space.*** $\qquad f(v) = \sum_i f(e_i)v^i = \sum_i f(e_i)e^i(v) \implies f = \sum_i f(e_i)e^i$
3. ***B^* is linearly independent.*** $\qquad \sum_i c_i e^i = f_0 \implies 0 = f_0(e_i) = \sum_j c_j e^j(e_i) = \sum_j c_j \delta^j_i = c_i$

Point 2 shows that any $f \in V^*$ can be expanded as a linear combination of B^*, so that $\mathrm{span}B^* = V^*$. Point 3 shows that B^* is linearly independent since $c_1 e^1 + \cdots + c_n e^n = f_0$ is only possible for $c_1 = \cdots = c_n = 0$. This shows that B^* is a basis for V, and also implies that $\dim V^* = \dim V$. Point 1 implies that B^* transforms like the coordinates under change of basis, i.e. B^* is contravariant to B (see Ex 1).

Remark 1. If $\langle \cdot, \cdot \rangle$ is an inner product on V, and $\mathbf{S}$ is the matrix of overlaps $\langle e_i, e_j \rangle = (\mathbf{S})_{ij}$ for the basis vectors, then elements of the dual basis can be explicitly written as $e^i = \sum_j (\mathbf{S}^{-1})_{ij} \langle e_j, \cdot \rangle$ so that $e^i(e_j) = \sum_k (\mathbf{S}^{-1})_{ik} \langle e_k, e_j \rangle = \delta^i_j$. This shows that $\{\langle e_i, \cdot \rangle\}$ provides an alternative basis for the space of linear functionals. If the basis is orthonormal, we find that $e^i = \langle e_i, \cdot \rangle$ and the inner product basis becomes identical to the canonical dual basis.

Definition 4. ***Linear operator.*** A *linear operator* $\hat{T} : V \to V$ is a vector-valued function on V that satisfies linearity, i.e. $\hat{T}(v + w) = \hat{T}(v) + \hat{T}(w)$ and $\hat{T}(cv) = c\hat{T}(v)$ for all $c \in \mathbb{C}$ and all $v, w \in V$. Note that it is common practice to drop the parentheses around the argument and write $\hat{T}(v)$ as simply $\hat{T}v$. The *identity operator* is given by $\hat{1}v = v$ for all $v \in V$ and the *null operator* is given by $\hat{0}v = 0$ for all $v \in V$.

---

[1] Using concepts introduced below, $\mathbf{T}$ can be identified as the coordinate matrix of a linear transformation $\hat{T}$ that maps e_i into $\tilde{e}_i$. In this context, the form in which we have expressed the change of basis arises naturally as $\tilde{e}_j = \hat{T}e_j = \sum_i e_i e^i(\hat{T}e_j) = \sum_i e_i(\mathbf{T})^i_j$ via resolution of the identity.

Proposition 2. ***Resolution of the identity.*** *If B is a basis for V then the identity operator on V can be expressed with respect to the B as $\hat{1} = \sum_i e_i e^i$ for $e_i \in B$ and $e^i \in B^*$.*

Proof: Let $v = \sum_i e_i v^i$ be the expansion of v with respect to B. Then, using point 1 under Prop 1, we find that $\hat{1}(v) = v = \sum_i e_i v^i = \sum_i e_i e^i(v)$ holds for all $v \in V$ which implies $\hat{1} = \sum_i e_i e^i$.

Remark 2. ***Coordinte matrix of a linear operator.*** By applying two resolutions of the identity, any linear operator $\hat{T} : V \to V$ can be decomposed in the basis as $\hat{T} = \sum_{ij} e_i e^i(\hat{T} e_j) e^j$. Defining a matrix $\mathbf{T}$ with elements $T^i_j = e^i(\hat{T} e_j)$, this decomposition is expressed as $\hat{T} = \sum_{ij} T^i_j e_i e^j$, which identifies $\mathbf{T}$ as the *coordinates* of $\hat{T}$ in the space of vector-dual products, $\text{span}\{e_i e^j \mid e_i \in B, e^j \in B^*\}$.

Remark 3. Note that resolution of the identity gives a natural motivation for the coordinate-space operations defined in linear algebra. For example, if $\hat{T}_1$ and $\hat{T}_2$ are both linear operators on V then $\hat{T}_1 \hat{T}_2 = \sum_{ijkl} (\mathbf{T}_1)^i_j (\mathbf{T}_2)^k_l e_i e^j(e_k) e^l = \sum_{ijl} (\mathbf{T}_1)^i_j (\mathbf{T}_2)^j_l e_i e^l = \sum_{il} (\mathbf{T}_1 \mathbf{T}_2) e_i e^l$ and we see that the coordinate matrix of $\hat{T}_1 \hat{T}_2$ is the matrix product of coordinate matrices for $\hat{T}_1$ and $\hat{T}_2$. Similarly, the action of $\hat{T}$ on a vector v is given by $\hat{T} v = \sum_{ij} T^i_j e_i e^j(v) = \sum_{ij} e_i T^i_j v^j$ so that the coordinates of $\hat{T} v$ are elements of $\mathbf{T}\mathbf{v}$, the matrix product of $\mathbf{T}$ with $\mathbf{v} = [v^i]$, a one-column matrix of vector coordinates.

Definition 5. ***Direct sum $V \oplus V'$.*** A *direct sum* of vector spaces V and V' is $\{v \oplus v' \mid v \in V, v' \in V'\}$, a new vector space with vector addition and scalar multiplication defined by

$$v_1 \oplus v_1' + v_2 \oplus v_2' = (v_1 + v_2) \oplus (v_1' + v_2') \qquad\qquad c(v \oplus v') = cv \oplus cv' \ .$$

Its dimension is $\dim(V \oplus V') = \dim V + \dim V'$. If $\{e_i\}$ is a basis for V and $\{e'_{i'}\}$ is a basis for V' then $\{e_i \oplus 0\} \cup \{0 \oplus e'_{i'}\}$ is a basis for $V \oplus V'$.

Definition 6. ***Tensor product $V \otimes V'$.*** A *tensor product* of vector spaces V and V' is $\{\sum v_i \otimes v'_{i'} \mid v_i \in V, v'_{i'} \in V'\}$, a new vector space with vector addition and scalar multiplication defined by

$$v_1 \otimes v' + v_2 \otimes v' = (v_1 + v_2) \otimes v' \qquad v \otimes v_1' + v \otimes v_2' = v \otimes (v_1' + v_2') \qquad c(v \otimes v') = cv \otimes v' = v \otimes cv' \ .$$

Its dimension is $\dim(V \otimes V') = \dim V \cdot \dim V'$. If $\{e_i\}$ is a basis for V and $\{e'_{i'}\}$ is a basis for V' then $\{e_i \otimes e'_{i'}\}$ is a basis for $V \otimes V'$.

Definition 7. ***Tensor.*** Most generally, a *tensor* is a vector in a tensor product space. An n^{th} *order tensor* is a vector in a product of n vector spaces, i.e. a member of $V_1 \otimes \cdots \otimes V_n$ When the product space has the form $V \otimes \cdots \otimes V \otimes V^* \otimes \cdots \otimes V^*$, we call its vectors *tensors* V. Given the importance of subcategory, the definition of *tensor* is sometimes restricted to mean only *tensors on* V. The space of *type-(m,n) tensors on* V is composed of m copies of V and n copies of V^*, sometimes denoted $T^m_n(V)$. A member t of $T^m_n(V)$ can be expanded in the basis as

$$t = \sum_{\substack{i_1 \cdots i_m \\ j_1 \cdots j_n}} t^{i_1 \cdots i_m}_{j_1 \cdots j_n} e_{i_1} \otimes \cdots \otimes e_{i_m} \otimes e^{j_1} \otimes \cdots \otimes e^{j_n}$$

with a coordinate array $\mathbf{t} = [t^{i_1 \cdots i_m}_{j_1 \cdots j_n}]$ indexed by m contravariant indices and n covariant indices. In this terminology, the space of type-$(0,0)$ tensors on V is given by $T^0_0(V) = \mathbb{C}$, the field of scalars. V itself can be identified as the space of type-$(1,0)$ tensors, $T^1_0(V) = V$, and its dual V^* is the space of type-$(0,1)$ tensors, $T^0_1(V) = V^*$. The space of type-$(1,1)$ tensors is isomorphic to the space of linear operators $\hat{T} : V \to V$ via the mapping $\sum_{ij} T^i_j e_i e^j \leftrightarrow \sum_{ij} T^i_j e_i \otimes e^j$. More generally, $T^n_n(V)$ is isomorphic to the space of multilinear mappings from $\underbrace{V \otimes \cdots \otimes V}_{n \text{ times}}$ to itself.

Remark 4. Note that, working in the coordinate space over an implied basis, it is conventional to refer to the coordinate array $\mathbf{t} = [t^{i_1 \cdots i_m}_{j_1 \cdots j_n}]$ of a type-(m,n) tensor $t \in T^m_n(V)$ as "the tensor", similar to the practice of referring to the coordinate array $\mathbf{v} = [v^i]$ of a vector $v \in V$ as "the vector". When using this language, keep in mind that t, v and $\mathbf{t}, \mathbf{v}$ are, mathematically, very different kinds of objects: t and v are *invariants*, whereas $\mathbf{v}$ is *contravariant* and $\mathbf{t}$ is "contravariant of order m and covariant of order n." In this context the intrinstic, basis-independent quantities v and t that these coordinates represent are sometimes referred to as "the abstract tensor" and "the abstract vector", respectively.

Definition 8. ***Tensor product $t \otimes t'$.*** The *tensor product* of $t \in T_n^m(V)$ and $t' \in T_{n'}^{m'}(V)$ is $t \otimes t' \in T_{n+n'}^{m+m'}(V)$ given by

$$t \otimes t' = \sum_{\substack{i_1 \cdots i_{m+m'} \\ j_1 \cdots j_{n+n'}}} (\mathbf{t})_{j_1 \cdots j_n}^{i_1 \cdots i_m} (\mathbf{t}')_{j_{n+1} \cdots j_{n+n'}}^{i_{m+1} \cdots i_{m+m'}} e_{i_1} \otimes \cdots \otimes e_{i_{n+n'}} \otimes e^{j_1} \otimes \cdots \otimes e^{j_{m+m'}}$$

where $\mathbf{t}$ and $\mathbf{t}'$ are the coordinate arrays of t and t'. In coordinate representation, the tensor product becomes

$$(\mathbf{t} \otimes \mathbf{t}')_{j_1 \cdots j_{n+n'}}^{i_1 \cdots i_{m+m'}} = (\mathbf{t})_{j_1 \cdots j_n}^{i_1 \cdots i_m} (\mathbf{t}')_{j_{n+1} \cdots j_{n+n'}}^{i_{m+1} \cdots i_{m+m'}}$$

which is equivalent to a *matrix Kronecker product* for $t, t' \in T_1^1(V)$.

Definition 9. ***Tensor contraction.*** A *tensor contraction* on $T_1^1(V)$ is a mapping into $T_0^0(V) = \mathbb{C}$ which acts on basis elements as $e_i \otimes e^j \mapsto e^j(e_i) = \delta_i^j$. Applied to $t \in T_1^1(V)$ with coordinates $\mathbf{t} = [t_j^i]$, this mapping takes the form

$$t = \sum_{ij} t_j^i e_i \otimes e^j \mapsto \sum_{ij} t_j^i e^j(e_i) = \sum_i t_i^i$$

which shows that this is equivalent to a *matrix trace* in coordinate representation: $\operatorname{tr}(\mathbf{t}) = \sum_i t_i^i$. Since it sums a covariant index with a contravariant index, it can be shown that this mapping gives the same result in any basis.[2] Generalized to elements of $T_n^m(V)$, a *tensor contraction* is a mapping into $T_{n-1}^{m-1}(V)$ that acts on the basis as

$$[\cdots e_{i_{p-1}} \otimes e_{i_p} \otimes e_{i_{p+1}} \cdots e^{j_{q-1}} \otimes e^{j_q} \otimes e^{j_{q+1}} \cdots] \mapsto e^{j_q}(e_{i_p})[\cdots e_{i_{p-1}} \otimes e_{i_{p+1}} \cdots e^{j_{q-1}} \otimes e^{j_{q+1}} \cdots]$$

i.e. it represents a trace along one specific vector-dual pair. In coordinate representation, this mapping is simply

$$t_{j_1 \cdots j_{q-1} j_q j_{q+1} \cdots j_m}^{i_1 \cdots i_{p-1} i_p i_{p+1} \cdots i_n} \mapsto \sum_k t_{j_1 \cdots j_{q-1} k j_{q+1} \cdots j_m}^{i_1 \cdots i_{p-1} k i_{p+1} \cdots i_n} .$$

A commonly encountered form of tensor contraction is a contraction *between* tensors t and t' in a tensor product $t \otimes t'$. For example, if t and t' are tensors in $T_1^1(V)$ then the following contraction of $t \otimes t' \in T_2^2(V)$

$$\sum_{\substack{i_1 i_2 \\ j_1 j_2}} t_{j_1}^{i_1} t_{j_2}^{i_2} e_{i_1} \otimes e_{i_2} \otimes e^{j_1} \otimes e^{j_2} \mapsto \sum_{i_1 j_2} \sum_k t_k^{i_1} t_{j_2}^k e_{i_1} \otimes e^{j_2}$$

is simply the matrix product of t and t'.

Notation 1. ***Einstein summation convention.*** The *Einstein summation convention* can be used to simplify algebraic manipulations by avoiding the use of summation symbols $\sum_{i_1 i_2 \cdots}$. Since the summations which appear in tensor manipulations generally have the form $\sum_i a_i b^i$ where $\{a_i\}$ is a covariant set, $\{b^i\}$ is a contravariant set, and the sum $\sum_i$ runs over basis vector indices, this convention takes any index which appears twice in a product, once as a covariant index and once as a contravariant one, to be implicitly summed over: $\sum_i a_i b^i \to a_i b^i$. A vector $v \in V$ can then be represented as

$$v = e_i v^i$$

where e_i are basis vectors and v^i are coordinates. More generally, an tensor $t \in T_n^m(V)$ is represented as follows.

$$t = t_{j_1 \cdots j_n}^{i_1 \cdots i_m} e_{i_1} \otimes \cdots \otimes e_{i_m} \otimes e^{j_1} \otimes \cdots \otimes e^{j_n}$$

The coordinates of $t'' \in T_1^1(V)$, the matrix product of t and $t' \in T_1^1(V)$, are

$$(\mathbf{t}'')_j^i = (\mathbf{t})_j^k (\mathbf{t}')_k^i$$

and the trace of $t \in T_1^1(V)$ is $\operatorname{tr}(\mathbf{t}) = t_k^k$. Note that the choice of symbol for a *contracted* (implicitly summed) *index* is arbitrary, whereas each *free* (uncontracted) *index* symbol must appear once in every term on the right- and left-hand sides of an equation, always with the same co- or contra-variance. Otherwise the equation is undefined. For example,

$$a_{ij}^{kl} = \frac{1}{2} b_{ij}^{vx} c_{vx}^{kl} + \frac{1}{6} d_{ijv}^{xyz} e_{xyz}^{klv}$$

is a balanced equation with free indices ${}_{ij}^{kl}$. Each term is an element of $T_2^2(V)$, so the addition (+) and assignment (=) operations are well-defined.

---

[2] A change of basis $B \to \tilde{B}$ given by $\tilde{e}_i = \sum_j e_j (\mathbf{T})_i^j$ transforms the coordinates of $\mathbf{t}$ into $\tilde{\mathbf{t}}$ given by $\tilde{t}_j^i = \sum_{kl} (\mathbf{T}^{-1})_k^i t_l^k (\mathbf{T})_j^l$ according to the co- and contra-variant transformation laws (see Def 1). Therefore, $\operatorname{tr}(\tilde{\mathbf{t}}) = \sum_i \tilde{t}_i^i = \sum_{ikl} (\mathbf{T}^{-1})_k^i t_l^k (\mathbf{T})_i^l = \sum_{kl} t_l^k (\mathbf{T}\mathbf{T}^{-1})_k^l = \sum_k t_k^k = \operatorname{tr}(\mathbf{t})$ which shows that the trace is basis-invariant.