Tensors

Definition 1. ***Covariance, contravariance, invariance.*** Let ϕ be an arbitrary linear function of a vector space V and consider a change of basis $\{e_1, \cdots, e_n\} \to \{\tilde{e}_1, \cdots, \tilde{e}_n\}$ defined by $\tilde{e}_j = \sum_i e_i\,[\mathbf{T}]_{ij}$ where $\mathbf{T}$ is an invertible matrix. If ϕ obeys the same transformation law as the basis elements, it is a *covariant* quantity. If it obeys the opposite transformation law, it is a *contravariant* quatntity.

$$\phi(\tilde{e}_j) = \sum_i \phi(e_i)\,[\mathbf{T}]_{ij} \qquad (covariant) \qquad\qquad \phi(\tilde{e}_i) = \sum_j [\mathbf{T}^{-1}]_{ij}\,\phi(e_j) \qquad (contravariant) \tag{1}$$

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

Example 1. ***Vector coordinates are contravariant.*** Let $\{v_i\}$ be the coordinates of a vector v in the basis $\{e_i\}$, and let $\{\tilde{v}_j\}$ be its coordinates in $\{\tilde{e}_j = \sum_i e_i\,[\mathbf{T}]_{ij}\}$. Then the invariance of v implies that the coordinates are contravariant:

$$\sum_i e_i v_i = v = \sum_j \tilde{e}_j \tilde{v}_j = \sum_i e_i \sum_j [\mathbf{T}]_{ij}\tilde{v}_j \implies v_i = \sum_j [\mathbf{T}]_{ij}\tilde{v}_j \implies \tilde{v}_j = \sum_i [\mathbf{T}^{-1}]_{ji} v_i$$

Therefore, vector coordinates should be written with superscript indices $v_i \mapsto v^i$.

Definition 2. ***Linear functional.*** A *linear functional* $f : V \to \mathbb{C}$ is a scalar-valued function on V that satisfies $f(v+w) = f(v) + f(w)$ and $f(cv) = cf(v)$ for all $c \in \mathbb{C}$ and all $v, w \in V$. The *null functional* is given by $f_0(v) = 0\ \forall v \in V$.

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

$$(f+g)(v) \equiv f(v) + g(v) \qquad\qquad (cf)(v) \equiv cf(v) \tag{2}$$

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. ***The canonical dual basis.*** *If $\{e_1, \cdots, e_n\}$ is a basis for V then $\{e^1, \cdots, e^n\} \subset V^*$, defined by $e^i(e_j) = \delta^i_j$, is a basis for V^*. This "canonical dual basis" transforms contravariantly relative to $\{e_1, \cdots, e_n\}$.*

Proof: First, note that $e^j(v) = e^j(\sum_i e_i v^i) = \sum_i e^j(e_i)v^i = v^j$ for any $v \in V$. Therefore, $f(v) = f(\sum_i e_i v^i) = \sum_i f(e_i)v^i = \sum_i f(e_i)e^i(v)$ holds for all $f \in V^*$ and all $v \in V$, which implies that $f = \sum_i f(e_i)e^i$. This shows that $\mathrm{span}\{e^1, \cdots, e^n\} = V^*$. Second, assume there exist $c_j \in \mathbb{C}$ such that $\sum_j c_j e^j = f_0$, where f_0 is the null functional. Then $0 = f_0(e_i) = \sum_j c_j e^j(e_i) = c_i$ and all of the coefficients must be zero, showing that $\{e^1, \cdots, e^n\}$ is linearly independent. Since both sets have n elements, $\dim V = \dim V^*$. Since $v^j = e^j(v)$, the dual basis is contravariant.

Remark 1. If $\langle\cdot|\cdot\rangle$ is an inner product on V and $\mathbf{S}$ is the matrix of overlaps, $\langle e_i|e_j\rangle = [\mathbf{S}]_{ij}$, then the canonical dual basis is given by $e^i(\cdot) = \sum_j [\mathbf{S}^{-1}]_{ij}\,\langle e_j|\cdot\rangle$, which simplifies to $e^i(\cdot) = \langle e_i|\cdot\rangle$ when $\{e_1, \cdots, e_n\}$ is orthonormal.

Definition 4. ***Linear operator.*** A *linear operator* $\hat{T} : V \to V$ is a vector-valued function on V that satisfies $\hat{T}(v+w) = \hat{T}v + \hat{T}w$ and $\hat{T}(cv) = c\hat{T}v$. The *identity operator* is given by $\hat{1}v = v$ and the *null operator* is given by $\hat{0}v = v$.

Proposition 2. ***Resolution of the identity.*** *Given a basis $\{e_i\}$, the identity operator can be expressed as $\hat{1} = \sum_i e_i e^i$.*

Proof: $\hat{1}(v) = v = \sum_i e_i v^i = \sum_i e_i e^i(v)$ holds for all $v \in V$ (see Prop 1), so $\hat{1} = \sum_i e_i e^i$.

Remark 2. Using Prop 2, we can identify the matrix of a linear operator, $\hat{T}e_j = \sum_i e_i\,[\mathbf{T}]_{ij}$, with $[\mathbf{T}]_{ij} = e^i(\hat{T}e_j)$. Therefore, the row index of $\mathbf{T}$ should be written as a superscript index, $[\mathbf{T}]_{ij} \mapsto [\mathbf{T}]^i_j$. Using two resolutions of the identity, $\hat{T}$ can be expressed as $\hat{T} = \sum_{ij} [\mathbf{T}]^i_j e_i e^j$, which identifies $\mathbf{T}$ as its coordinates in $V \times V^* = \mathrm{span}\{e_i e^j\}$.

Remark 3. The formula for matrix multiplication also follows from Prop 2: the matrix of $\hat{T}\hat{T}'$ is $e^i(\hat{T}\hat{T}'e_j) = \sum_k e^i(\hat{T}e_k)e^k(\hat{T}'e_j) = \sum_k [\mathbf{T}]^i_k [\mathbf{T}']^k_j$.

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

$$v_1 \oplus v_1' + v_2 \oplus v_2' = (v_1 + v_2) \oplus (v_1' + v_2') \qquad\qquad c(v \oplus v') = cv \oplus cv' \,. \tag{3}$$

If $\{e_i\}$ and $\{e'_{i'}\}$ are bases for V and V' then $\{e_i \oplus 0'\} \cup \{0 \oplus e'_{i'}\}$ is a basis for $V \oplus V'$, which has dimension $\dim V + \dim V'$.

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

$$v_1 \otimes v' + v_2 \otimes v' = (v_1 + v_2) \otimes v' \qquad v \otimes v_1' + v \otimes v_2' = v \otimes (v_1' + v_2') \qquad c(v \otimes v') = (cv) \otimes v' = v \otimes (cv') \tag{4}$$

If $\{e_i\}$ and $\{e'_{i'}\}$ are bases for V and V' then $\{e_i \otimes e'_{i'}\}$ is a basis for $V \otimes V'$, which has dimension $\dim V \cdot \dim V'$.

Remark 4. Since the sum $S + S'$ of two mutually orthogonal subspaces $S, S \subseteq V$ is isomorphic to their direct sum, $S \oplus S'$ and $S + S'$ are used interchangeably in this context. The reverse is also true. The direct sum of two distinct vector spaces $V \oplus V'$ contains copies of V and V' as mutually orthogonal subspaces, namely $V \simeq V \oplus 0'$ and $V' \simeq 0 \oplus V'$.

Definition 7. ***Direct sums and tensor products of inner product spaces.*** If V and V' are inner product spaces, then $V \oplus V'$ and $V \otimes V'$ are also inner product spaces with respect to the following inner products.

$$\langle v \oplus v' | w \oplus w' \rangle_{V\oplus V'} \equiv \langle v|w\rangle_V + \langle v'|w'\rangle_{V'} \qquad\qquad \langle v \otimes v' | w \otimes w' \rangle_{V\otimes V'} \equiv \langle v|w\rangle_V \cdot \langle v'|w'\rangle_{V'} \tag{5}$$

Definition 8. ***Tensor.*** A *tensor* is a vector in a tensor product space. An n^{th} *order tensor* lives in a product of n vector spaces, $V_1 \otimes \cdots \otimes V_n$. A *type-$(m,n)$ tensor on V* lives in $V \underset{m \text{ times}}{\otimes \cdots \otimes} V \otimes V^* \underset{n \text{ times}}{\otimes \cdots \otimes} V^*$, denoted $T^m_n(V)$.

Example 2. $T^0_0(V)$ is the scalar field $\mathbb{C}$, $T^1_0(V)$ is the vector space itself, $T^0_1(V)$ is the dual space V^*, and $T^1_1(V)$ is (up to isomorphism) the space of linear operators $V \otimes V^* \simeq V \times V^*$.

Remark 5. A member t of $T^m_n(V)$ can be expanded in the basis $\{e_1, \cdots, e_n\}$ as follows.

$$t = \sum_{\substack{i_1\cdots i_m \\ j_1 \cdots j_n}} t^{i_1\cdots i_m}_{j_1\cdots j_n} e_{i_1} \otimes \cdots \otimes e_{i_m} \otimes e^{j_1} \otimes \cdots \otimes e^{j_n} \tag{6}$$

Its coordinate array $\mathbf{t} = [t^{i_1\cdots i_m}_{j_1\cdots j_n}]$ is indexed by m contravariant indices and n covariant indices.

Remark 6. Just as people often use the term "vector" to refer to the basis-dependent coordinate array v^i rather than v itself, more often than not the word "tensor" refers to the coordinate array $t^{i_1\cdots i_m}_{j_1\cdots j_n}$ rather than the abstract quantity t.

Definition 9. ***Tensor product $\mathbf{t} \otimes \mathbf{t'}$.*** The *tensor product* of $t \in T^m_n(V)$ and $t' \in T^{m'}_{n'}(V)$ is $t \otimes t' \in T^{m+m'}_{n+n'}(V)$ with coordinate array $[\mathbf{t} \otimes \mathbf{t}']^{i_1\cdots i_{m+m'}}_{j_1\cdots j_{n+n'}} = t^{i_1\cdots i_m}_{j_1\cdots j_n} t^{i_{m+1}\cdots i_{m+m'}}_{j_{n+1}\cdots j_{n+n'}}$. This generalizes the *matrix Kronecker product*.

Definition 10. ***Tensor contraction.*** A *(p,q) tensor contraction* is a map $\mathrm{tr}_{p,q} : T^m_n(V) \to T^{m-1}_{n-1}(V)$ given by tracing the p^{th} element in V with the q^{th} element in V^*:

$$\mathrm{tr}_{p,q}\left(e_{i_1} \otimes \cdots \otimes e_{i_p} \otimes \cdots \otimes e^{j_q} \otimes \cdots \otimes e^{j_n}\right) = \delta^{j_q}_{i_p}\, e_{i_1} \otimes \cdots \otimes \cancel{e_{i_p}} \otimes \cdots \otimes \cancel{e^{j_q}} \otimes \cdots \otimes e^{j_n} \tag{7}$$

The coordinate array of $\mathrm{tr}_{p,q}(t)$ is $[\mathrm{tr}_{p,q}(\mathbf{t})]^{i_1\cdots \cancel{i_p} \cdots i_m}_{j_1 \cdots \cancel{j_q} \cdots j_m} = \sum_{i_p, j_q} \delta^{j_q}_{i_p}\, t^{i_1\cdots i_m}_{j_1\cdots j_m}$, generalizing the *matrix trace*, $\mathrm{tr}(M) = \sum_i M^i_i$.

Definition 11. ***Cross-contraction.*** A *cross-contraction* traces a lower index on one tensor with an upper index on another, $\mathrm{tr}_{p,q'} : T^m_n(V) \otimes T^{m'}_{n'}(V) \to T^m_{n-1}(V) \otimes T^{m'-1}_{n'}$, generalizing the *matrix product* $\mathrm{tr}_{1,1'}(M \otimes M') = \sum_k M^i_k M^k_j$.

Notation 1. ***Einstein summation convention.*** In the *Einstein summation convention* any index which appears twice in a product, once as a covariant index and once as a contravariant one, is implicitly summed over: $\sum_i a_i b^i \mapsto a_i b^i$. For most purposes, this rule suffices to dispense with summation symbols altogether. The basis expansion of a vector now takes the form $v = e_i v^i$, that of a tensor takes the form $t = t^{i_1\cdots i_m}_{j_1\cdots j_n} e_{i_1} \otimes \cdots \otimes e_{i_m} \otimes e^{j_1} \otimes \cdots \otimes e^{j_n}$, and resolution of the identity can be written as $\hat{1} = e_i e^i$. The choice of symbol for a *contracted index* is is arbitrary, wherease 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.

Example 3. In Einstein notation, the matrix expression $\mathbf{C} = \mathbf{AB}$ is written as $C^i_j = A^i_k B^k_j$.

Example 4. The expression $a^{kl}_{ij} = \frac{1}{2} b^{vx}_{ij} c^{kl}_{vx} + \frac{1}{6} d^{xyz}_{ijv} e^{klv}_{xyz}$ is a balanced equation with free indices ${}^{kl}_{ij}$. Each term is an element of $T^2_2(V)$, so the addition (+) and assignment (=) operations are well-defined.