The tensor description of Euclidean spaces

Our primary object is the invariant position vector $\mathbf{R}$—invariant because it is independent of the coordinate system. It is convenient, however, to introduce an arbitrary coordinate system $Z = \{Z^i\} = \{Z^1, Z^2, Z^3\}$, so that $\mathbf{R} = \mathbf{R}(Z)$. The $\{Z^i\}$ are tacitly assumed independent. In words, $\mathbf{R}(Z)$ represents a vector-valued function, parametrized by Z, whose value at every point in space agrees with the position vector $\mathbf{R}$. The coordinate system Z induces a natural basis, the *covariant basis* $\{\mathbf{Z}_i\}$, defined by

$$\mathbf{Z}_i := \frac{\partial \mathbf{R}(Z)}{\partial Z^i}.$$

In general, the covariant basis is a local basis, because the basis vectors depend on their position in space; i.e. $\mathbf{Z}_i = \mathbf{Z}_i(Z)$. We can always decompose an arbitrary vector $\mathbf{V}$ in the covariant basis with $\mathbf{V} = V^i \mathbf{Z}_i$, where the $\{V^i\}$ are called the *contravariant components* of $\mathbf{V}$. We further define the *covariant metric tensor* Z_{ij} as $Z_{ij} := \mathbf{Z}_i \cdot \mathbf{Z}_j$. Because the dot product is commutative, the covariant metric tensor is symmetric: $Z_{ij} = Z_{ji}$. The covariant metric tensor encodes all information about the dot product, since, for two arbitrary vectors $\mathbf{U}$ and $\mathbf{V}$, we have

$$\mathbf{U} \cdot \mathbf{V} = U^i \mathbf{Z}_i \cdot V^j \mathbf{Z}_j = U^i V^j Z_{ij}.$$

Let us further define the *contravariant basis* as the dual basis to the covariant basis, satisfying $\mathbf{Z}^i \cdot \mathbf{Z}_j := \delta^i_j$, as well as the *contravariant metric tensor* Z^{ij}, defined as $Z^{ij} := \mathbf{Z}^i \cdot \mathbf{Z}^j$. Expanding $\mathbf{Z}^i$ in the $\{\mathbf{Z}_k\}$ basis, we find

$$\mathbf{Z}^i = A^k \mathbf{Z}_k \quad \Longleftrightarrow \quad Z^{ij} = \mathbf{Z}^i \cdot \mathbf{Z}^j = A^k \mathbf{Z}_k \cdot \mathbf{Z}^j = A^k \delta^j_k = A^j \quad \Longleftrightarrow \quad \mathbf{Z}^i = Z^{ik} \mathbf{Z}_k,$$

and, in addition, the covariant and contravariant metric tensors are inverses of each other, as

$$\delta^i_j = \mathbf{Z}^i \cdot \mathbf{Z}_j = Z^{ik} \mathbf{Z}_k \cdot \mathbf{Z}_j = Z^{ik} Z_{kj}.$$

We have previously mentioned that the covariant (and contravariant) bases are generally coordinate-dependent. The dependence of these basis vectors on the coordinate system Z is given by the *Christoffel symbol* Γ^k_{ij}, defined via

$$\frac{\partial \mathbf{Z}_i}{\partial Z^j} =: \Gamma^k_{ij} \mathbf{Z}_k \quad \Longleftrightarrow \quad \Gamma^k_{ij} = \frac{\partial \mathbf{Z}_i}{\partial Z^j} \cdot \mathbf{Z}^k.$$

The dependence of the contravariant vectors can be established using the product rule, with

$$\Gamma^k_{ij} = \frac{\partial (\mathbf{Z}_i \cdot \mathbf{Z}^j)}{\partial Z^j} - \frac{\partial \mathbf{Z}^k}{\partial Z^j} \cdot \mathbf{Z}_i = \frac{\partial \delta^j_i}{\partial Z^j} - \frac{\partial \mathbf{Z}^k}{\partial Z^j} \cdot \mathbf{Z}_i = -\frac{\partial \mathbf{Z}^k}{\partial Z^j} \cdot \mathbf{Z}_i \quad \Longleftrightarrow \quad \frac{\partial \mathbf{Z}^k}{\partial Z^j} = -\Gamma^k_{ij} \mathbf{Z}^i.$$

The tensor property

We call a *variant* an object that can be defined by the same rule in all coordinate systems. The covariant basis vectors $\{\mathbf{Z}_i\}$ are variants that come from taking a partial derivative with respect to the $\{Z^i\}$ of the position vector $\mathbf{R}$. The covariant metric tensor Z_{ij} is a tensor, since it comes from taking the dot product $\mathbf{Z}_i \cdot \mathbf{Z}_j$. The Jacobian $J^i_{i'}$, however, is not a variant, because it expresses a relationship between *two* different coordinate systems and cannot be defined with just one. *Tensors* are special types of variants that transform in a certain manner. We call the variant T_i a *covariant tensor* if its values T_i and $T_{i'}$ in the coordinate systems Z and Z' respectively are related by

$$T_{i'} = T_i J^i_{i'},$$

and, similarly, the variant T^i is a *contravariant tensor* if its values T^i and $T^{i'}$ in the coordinate systems Z and Z' respectively are related by

$$T^{i'} = T^i J^{i'}_i.$$

Covariant tensors are so called because they transform in the same way as does the covariant basis $\{\mathbf{Z}_i\}$, and similarly for contravariant tensors; the bases transform as

$$\mathbf{Z}_{i'} = \frac{\partial \mathbf{R}}{\partial Z^{i'}} = \frac{\partial \mathbf{R}}{\partial Z^i} \frac{\partial Z^i}{\partial Z^{i'}} = \mathbf{Z}_i J^i_{i'} \quad \text{and} \quad \mathbf{Z}^{i'} \cdot \mathbf{Z}_{j'} = \delta^{i'}_{j'} = \delta^i_j J^{i'}_i J^j_{j'} = \mathbf{Z}^i \cdot \mathbf{Z}_j J^{i'}_i J^j_{j'} = (\mathbf{Z}^i J^{i'}_i) \cdot \mathbf{Z}_{j'} \quad \Longleftrightarrow \quad \mathbf{Z}^{i'} = \mathbf{Z}^i J^{i'}_i.$$

In addition, the contraction of a covariant and contravariant tensor is invariant, since

$$U' = S_{i'} T^{i'} = S_i J^i_{i'} T^j J^{i'}_j = S_i T^j \delta^i_j = S_i T^i = U.$$

Example. The components V^i of a generic vector $\mathbf{V} = V^i\mathbf{Z}_i$ are contravariant tensors, since

$$V^i\mathbf{Z}_i = \mathbf{V} = V^{i'}\mathbf{Z}_{i'} = V^{i'}J^i_{i'}\mathbf{Z}_i \quad\iff\quad V^i = V^{i'}J^i_{i'}.$$

The above definition of tensors readily generalizes to tensors of multiple indices.

Example. The covariant and contravariant metric tensors Z_{ij} and Z^{ij} are tensors, since

$$Z_{i'j'} = \mathbf{Z}_{i'}\cdot\mathbf{Z}_{j'} = \mathbf{Z}_i\cdot\mathbf{Z}_jJ^i_{i'}J^j_{j'} = Z_{ij}J^i_{i'}J^j_{j'} \quad\text{and}\quad Z^{i'j'} = \mathbf{Z}^{i'}\cdot\mathbf{Z}^{j'} = \mathbf{Z}^i\cdot\mathbf{Z}^jJ^{i'}_iJ^{j'}_j = Z^{ij}J^{i'}_iJ^{j'}_j.$$

Example. The Kronecker symbol δ^i_j is a tensor, since

$$\delta^{i'}_{j'} = \delta^i_jJ^{i'}_iJ^j_{j'}.$$

The tensor property is rather general. In particular, we can easily prove by definition that the sum of two tensors, product of two tensors, and contraction of two indices of a tensor all form tensors.

Tensor operators and covariant differentiation

The gradient and divergence are geometric operators—they depend on quantities intrinsic to space, and not on the parametrization of space. The general form of these operators are therefore independent of coordinates, so these operators must transform as invariants. As a corollary, results with these operators obtained in a special coordinate system must necessarily be common to all coordinate systems.

In the Cartesian system, the gradient ∇F of a scalar function F can be expressed as

$$\nabla F = \sum_i \frac{\partial F}{\partial Z^i}\mathbf{Z}_i.$$

This is not invariant, however, because both the partial derivatives and the basis vectors transform covariantly. The simplest change we can make for invariance is to let

$$\nabla F = \sum_i \frac{\partial F}{\partial Z^i}\mathbf{Z}^i,$$

where we use the contravariant rather than the covariant basis. Because the two bases are identical in Cartesian coordinates—as the definition of the contravariant basis shows—, this invariant expression reduces to the known result in Cartesian coordinates, and hence it must be the general form of the gradient in an arbitrary coordinate system. The generalization of the divergence must wait until we have introduced the covariant derivative.

Tensor analysis allows for straightforward manipulation of tensors but does not generalize well to other variants. While most of the interesting objects we consider are indeed tensors, there is one notable exception: the derivative. In particular, the partial derivative of a tensor does *not* produce a tensor, and it is thus of interest to find some suitable generalization of the derivative operator that will produce a tensor from a tensor. To this end, consider the partial derivative

$$\frac{\partial\mathbf{V}}{\partial Z^i} = \frac{\partial(V^j\mathbf{Z}_j)}{\partial Z^i} = \frac{\partial V^j}{\partial Z^i}\mathbf{Z}_j + V^j\Gamma^k_{ji}\mathbf{Z}_k = \left(\frac{\partial V^m}{\partial Z^i} + V^j\Gamma^m_{ji}\right)\mathbf{Z}_m.$$

The leftmost expression is a tensor, since $\mathbf{V}$ is invariant and the partial derivative transforms as a covariant tensor. Hence the rightmost expression is also a tensor, and, because the covariant basis is a tensor, so too is the quantity

$$\frac{\partial V^m}{\partial Z^i} + V^j\Gamma^m_{ji} =: \nabla_iV^m,$$

which defines the *covariant derivative* ∇_i for contravariant tensors. As desired, the covariant derivative is a generalization of the partial derivative that yields a tensor from a tensor. In Cartesian coordinates, in which the Christoffel symbols Γ^i_{jk} are identically zero, the covariant derivative reduces to the partial derivative. Physically, the covariant derivative measures the rate of change of both the tensor and the basis vectors, and this presents a more complete picture than does the rate of change of the partial derivative alone. If we instead work with the contravariant basis $\{\mathbf{Z}^i\}$, then a similar analysis shows

$$\nabla_iV_m := \frac{\partial V_m}{\partial Z^i} - V^j\Gamma^i_{jm}$$

for covariant tensors. For higher-order tensors, the generalization proceeds as before; we demonstrate the process for the tensor $\mathsf{T} = T^{ij}_k \mathbf{Z}_i \otimes \mathbf{Z}_j \otimes \mathbf{Z}^k$. We have

$$\begin{aligned}
\frac{\partial \mathsf{T}}{\partial Z^m} &= \frac{\partial T^{ij}_k}{\partial Z^m}\mathbf{Z}_i \otimes \mathbf{Z}_j \otimes \mathbf{Z}^k + T^{ij}_k \frac{\partial \mathbf{Z}_i}{\partial Z^m} \otimes \mathbf{Z}_j \otimes \mathbf{Z}^k + T^{ij}_k \mathbf{Z}_i \otimes \frac{\partial \mathbf{Z}_j}{\partial Z^m} \otimes \mathbf{Z}^k + T^{ij}_k \mathbf{Z}_i \otimes \mathbf{Z}_j \otimes \frac{\partial \mathbf{Z}^k}{\partial Z^m} \\
&= \frac{\partial T^{ij}_k}{\partial Z^m}\mathbf{Z}_i \otimes \mathbf{Z}_j \otimes \mathbf{Z}^k + T^{ij}_k \Gamma^n_{im} \mathbf{Z}_n \otimes \mathbf{Z}_j \otimes \mathbf{Z}^k + T^{ij}_k \Gamma^n_{jm} \mathbf{Z}_i \otimes \mathbf{Z}_n \otimes \mathbf{Z}^k - T^{ij}_k \Gamma^k_{mn} \mathbf{Z}_i \otimes \mathbf{Z}_j \otimes \mathbf{Z}^n \\
&= \left(\frac{\partial T^{ij}_k}{\partial Z^m} + T^{nj}_k \Gamma^i_{mn} + T^{in}_k \Gamma^j_{nm} - T^{ij}_n \Gamma^n_{km} \right) \mathbf{Z}_i \otimes \mathbf{Z}_j \otimes \mathbf{Z}^k.
\end{aligned}$$

The $\otimes$ symbol denotes the tensor product, which, as expected, encodes the concept of multiplying tensors together. For V, W tensors in $\mathbb{R}^m$ and $\mathbb{R}^n$ respectively, the tensor product $\otimes$ is a bilinear operator that forms the tensor $\mathsf{V} \otimes \mathsf{W} \in \mathbb{R}^{m \times n}$ out of the tensors V and W. We have previously labeled tensors by their components and glossed over the relevant basis; this basis is provided by the tensor product.

Example. The covariant metric tensor $\mathsf{Z} = Z_{ij} \mathbf{Z}^i \otimes \mathbf{Z}^j$ is a second-order tensor with the natural basis $\{\mathbf{Z}^i \otimes \mathbf{Z}^j\}$. We can confirm straightforwardly that each component of the basis is indeed a tensor, since

$$\mathbf{Z}^{i'} \otimes \mathbf{Z}^{j'} = J^{i'}_i J^{j'}_j \mathbf{Z}^i \otimes \mathbf{Z}^j.$$

Returning to our derivation, we have also used the fact that the product rule of differentiation applies to tensor products. To see this, consider the directional derivative

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}\epsilon} \mathbf{Z}_i \otimes \mathbf{Z}_j \otimes \mathbf{Z}^k (Z + \epsilon f)\bigg|_{\epsilon=0} &= \frac{\mathrm{d}}{\mathrm{d}\epsilon} \left(\mathbf{Z}_i \otimes \mathbf{Z}_j \otimes \mathbf{Z}^k(Z) + \epsilon f^m \frac{\partial \mathbf{Z}_i}{\partial \mathbf{Z}^m} \otimes \mathbf{Z}_j \otimes \mathbf{Z}^k(Z) + \ldots + O(\epsilon^2) \right)\bigg|_{\epsilon=0} \\
&= f^m \frac{\partial \mathbf{Z}_i}{\partial \mathbf{Z}^m} \otimes \mathbf{Z}_j \otimes \mathbf{Z}^k(Z) + f^m \mathbf{Z}_i \otimes \frac{\partial \mathbf{Z}_j}{\partial \mathbf{Z}^m} \otimes \mathbf{Z}^k(Z) + f^m \mathbf{Z}_i \otimes \mathbf{Z}_j \otimes \frac{\partial \mathbf{Z}^k}{\partial \mathbf{Z}^m}(Z).
\end{aligned}$$

The first equality involves a multivariate Taylor expansion for each of $\mathbf{Z}_i$, $\mathbf{Z}_j$, and $\mathbf{Z}^k$. Recall that we are notating Z as $\{Z^i\}$ and, by analogy, f as $\{f^i\}$; the more common vector notation would have as arguments $\mathbf{Z}$ and $\mathbf{Z} + \epsilon \mathbf{f}$. If we evaluate this expression using the chain rule, however, we have

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon} \mathbf{Z}_i \otimes \mathbf{Z}_j \otimes \mathbf{Z}^k (Z + \epsilon f)\bigg|_{\epsilon=0} = \frac{\mathrm{d}(Z^m + \epsilon f^m)}{\mathrm{d}\epsilon} \frac{\partial}{\partial (Z^m + \epsilon f^m)} \mathbf{Z}_i \otimes \mathbf{Z}_j \otimes \mathbf{Z}^k (Z + \epsilon f)\bigg|_{\epsilon=0} = f^m \frac{\partial}{\partial Z^m} \mathbf{Z}_i \otimes \mathbf{Z}_j \otimes \mathbf{Z}^k(Z).$$

By comparing the righthand expression with that above and noting that the f^m can be varied independently, we see that equality requires that the product rule hold.

rf