Gradient contravariance and deep down the dual hole

Tomáš Faikl

5th July 2020

In second year of my bachelor's degree in Mathematical Physics, covariance and contravariance of mathematical objects was introduced (in Theoretical Physics). Beside other, we derived how grad U (let U be a scalar field) transforms. But it was not clear to me until now what it means in full context of 1-forms and analysis.

I will try to give some mathematical background but mostly just to familiarize ourselves with the notation that I will be using and not everything will be commented in detail. For more complete overview of tensors, differential 1-forms and other topics, please use literature if you are not familiar with these concepts. But in my opinion, if you are not familiar with these you will not be very interested in this post. But who knows :-).

See for ex. this document about more abstract treatment of tensors (as multilinear maps).

1 Introduction

Let V^n be a vector space of dimension n and let $\boldsymbol{B} = (\boldsymbol{e_1}, \boldsymbol{e_2}, \dots, \boldsymbol{e_n})$ denote a basis of V^n and let $V_n = (V^n)^*$, $\boldsymbol{B^*} = (\boldsymbol{e^1}, \dots, \boldsymbol{e^n})$ denote its dual vector space and its basis respectively.

Also, we will be using upper indices to denote vectors $\boldsymbol{v} = v^i \boldsymbol{e_i}$ (sometimes also denoted $\overrightarrow{v}$) that transform contravariantly (opposite to) $\boldsymbol{B}$ and lower indices to denote (co)vectors $\boldsymbol{\alpha} = \alpha_i \boldsymbol{e^i}$ (sometimes also denoted $\underleftarrow{\alpha}$) that transform covariantly (same as) $\boldsymbol{B}$. And of course, basis $\boldsymbol{B}$ will transform with $\tilde{\boldsymbol{e_j}} = S^i_j \boldsymbol{e_i}$ into $\tilde{\boldsymbol{B}} = (\tilde{\boldsymbol{e_1}}, \tilde{\boldsymbol{e_2}}, \dots, \tilde{\boldsymbol{e_n}})$, $\mathbb{S} \in \mathrm{GL}(n, \mathbb{R})$.

More abstractly, a covector can be thought of as a map from vector space to reals $\boldsymbol{\alpha} : V^n \to \mathbb{R}$ while a vector is a map which takes a covector and returns a real number, $\boldsymbol{v} : V_n \to \mathbb{R}$. See the linked document above for more information.

1.1 1-forms

A basic observation of mathematical analysis is that a total derivative of some function $f : M \to N$ (M, N affine spaces) calculated in point $x_0 \in M$ is a linear map from $\vec{M}$ to $\vec{N}$, $f'(x_0) \in \mathrm{Hom}(\vec{M}, \vec{N})$.

A differential 1-form is such an object

$$\boldsymbol{\omega} = \omega_i \underleftarrow{\mathrm{d}x^i}, \qquad \omega_i : M \to N, \quad \underleftarrow{\mathrm{d}x^i} \in \boldsymbol{B^*}, \quad \boldsymbol{\omega} \in \Lambda^1$$

where $\boldsymbol{B^*}$ is a basis of $\vec{M}^*$ and pointwise

$$\underleftarrow{\boldsymbol{\alpha}(x_0)} = \omega_i(x_0)\underleftarrow{\mathrm{d}x^i}$$

More symbols later, for $f : \mathbb{R}^n \to \mathbb{R}$ we define

$$(\mathrm{d}f)(x_0) = f'(x_0) \in (V^n)^*$$

and also call $\boldsymbol{\omega}$ an **exact** differential if there exists a function f such that $\mathrm{d}f = \boldsymbol{\omega}$. Thus, $\mathrm{d}f$ is trivially an exact differential.

2 The problem

When we introduced basic terms, we can begin to work towards the initially confusing part.

Let M be an affine space, $\vec{M}^n = V^n$ be its focus, g an arbitrary nondegenerate bilinear form on V^n and let $U = U(x),\ x \in M$ denote a scalar field on M. Then

$$(\mathrm{d}U)(x) \in (V^n)^* \quad \forall x \in M$$

and vector

$$(\mathrm{grad}U)(x) \in V^n \quad \forall x \in M$$

is defined such that $\mathrm{d}U(x)\boldsymbol{v} = g(\boldsymbol{v},\ (\mathrm{grad}\ U)(x)),\ \boldsymbol{v} \in V^n$ or more cleanly

$$\mathrm{d}U = g(\ \cdot\ ,\ \mathrm{grad}\ U).$$

In other words, g induces a natural (canonical) izomorphism between V^n and its dual V_n (generalized version of Riesz's theorem for scalar product).

Note that $\mathrm{d}U = (\mathrm{d}U)_i \boldsymbol{e^i}$. Also we can call the bilinear form g a pseudometric tensor (a second order tensor) and define $g_{ij} = g(\boldsymbol{e_i},\ \boldsymbol{e_j})$. Through linearity, we can write

$$\mathrm{d}U_i = g(\ \cdot\ ,\ \mathrm{grad}\ U)_i = g_{ij}(\mathrm{grad}U)^j = g_{ij}\mathrm{grad}U^j,$$

where $(\mathrm{d}U)_i$ and $\mathrm{grad}U^j$ are components of $\mathrm{d}U$ and $\mathrm{grad}U$ with respect to bases $\boldsymbol{B^*}$ and $\boldsymbol{B}$ respectively. Continuing, we can write (as $\mathrm{d}U$ is an exact differential)

$$\mathrm{d}U = \frac{\partial f}{\partial x^i}\mathrm{d}x^i,$$

where $\frac{\partial}{\partial x^i}$ is called a covariant derivative. When we would think of an object (a total derivative f') whose i-th component is $\frac{\partial f}{\partial x^i}$, the object transforms covariantly with basis transformation (just expand x^i into $g^{ij}x_j$).

Also, when we talk about lowering/raising and index, we mean transformation of a vector v^i to a covector $\boldsymbol{v_i}$ defined using g as $v_i = g_{ij}v^j$ and similarly the other way around.

$$g_{ij}\mathrm{grad}U^j = \mathrm{d}U_i = \frac{\partial f}{\partial x^i} \tag{1}$$

And after multiplying both sides with g^{ki} we have

$$\mathrm{grad}U^k = g^{ki}\frac{\partial f}{\partial x^i}.$$

But we have one last quest at that is to assign better meaning to this expression. This is going to be obtained using

$$\frac{\partial}{\partial x^i} = \frac{\partial}{\partial(g^{ik}x_k)} = \frac{1}{g^{ik}}\frac{\partial}{\partial x_k}$$

and consequently multiplying both sides by g^{ik} resulting together with (1) in

$$\boxed{\mathrm{grad}U^k = \frac{\partial U}{\partial x_k}}.$$

This object grad$U$ then transforms as is expected from proper vector; that is – contravariantly. On the other hand, its corresponding 1-form dU fullfills

$$\boxed{(\mathrm{d}U)_i = \frac{\partial U}{\partial x^i}}.$$

3 Conclusion

The issue I previously faced was using a bad definition of what exactly gradient means in more general spaces than Euclidian spaces with standard scalar product $\langle\cdot,\cdot\rangle : V^n \times V^n \to \mathbb{R}$ with positive signature. In this space both g_{ij} and g^{ij} (its inverse) are equal to diag$(+1, +1, \dots, +1)$, hence no difference between covariant and contravariant vectors can be observed. Together with other mistakes, this lead me to wrong conclusion that gradient is not properly defined in more general spaces, which was really strange so I tried once more.

The existence and a particular bilinear nondegenerate form choosen for some space depends on context. Eg. for Euclidian $E(3)$, $(g_{ij}) = \mathrm{diag}(+1, +1, +1)$, for Minkowski space $E(1, 3)$ used in special relativity, metric pseuoudotensor $(g_{ij}) = \mathrm{diag}(+1, -1, -1, -1)$ is chosen (so that it conserves spacetime interval). And for example on phase space Γ (differentiable manifold) in Hamiltonian formalism, one chooses symplectic (antisymmetric) form ω defined on cotangent bundle T^*M, where M is the system's configuration space, as $\boldsymbol{\omega} = \mathrm{d}p^i \wedge \mathrm{d}q^i$ in coordinates and $\omega = \left(\begin{smallmatrix} 0 & \mathbb{I} \\ -\mathbb{I} & 0 \end{smallmatrix}\right)$ in matrix notation. All matrix relations are in polar/symplectic basis.