4. Gradient as a Functional

The gradient of a scalar valued function is denoted by its list of partial derivatives. Should it be considered a vector or a functional? Note that the gradient is usually used to express the change in the value of a function along a direction vector. This is a typical role of a functional.

Time Derivative Example

Consider a surface and a parametrization $\vec{x}(\vec{u})$

Also consider a scalar valued function $f(\vec{u})$ and a path $\vec{c}(t)$ in the parameter space. The function and path have compatible expressions on the surface as $f(\vec{x})=f(\vec{x}(\vec{u}))$ and $\vec{x}(\vec{c}(t))$.

The path velocity vector $\frac{dc^k}{dt}$ at any point of the path in the parameter space has its transformed version in the tangent space $\frac{\partial x^i}{\partial u^k}\frac{dc^k}{dt}$.

Note that the time rate of change of the function along the path should not change in either context. We use the chain rule to calculate the time derivative:

$$\frac{d}{dt}f(\vec{x}(\vec{c}(t))) = \left(\frac{\partial f}{\partial x^i}\right)\left(\frac{\partial x^i}{\partial u^k}\frac{dc^k}{dt}\right)=\left(\frac{\partial f}{\partial x^i}\frac{\partial x^i}{\partial u^k}\right)\left(\frac{dc^k}{dt}\right) = \frac{\partial f}{\partial u^k}\frac{dc^k}{dt}$$

We can analyze the above middle equivalence in terms of functionals and vectors.

The left expression is the functional $\frac{\partial f}{\partial x^i}$ acting on the vector $\frac{\partial x^i}{\partial u^k}\frac{dc^k}{dt}$ in the tangent space.

The right expression is the functional $\frac{\partial f}{\partial x^i}\frac{\partial x^i}{\partial u^k}$ acting on the vector $\frac{dc^k}{dt}$ in the parameter space.

This example exhibits the role of transformation rules in conserving the result of a functional acting on a vector.

Later, we will see how to express the gradient as a vector.

Transformation of a Dot Product and g_{ij}
Consider a simple parameterization $\vec{x}(\vec{u})$ resulting from a change of variables (like polar coordinates). Also consider a dot product of two vectors expressed in Cartesian coordinates $x^i y^i$. The alternative basis vectors are constructed from derivatives of the parameterization and we get

$x^i y^i = \left(x'^m \frac{\partial x^i}{\partial u^m} \right) \left(y'^n \frac{\partial x^i}{\partial u^n} \right) = x'^m g_{mn} y'^n$ where $g_{mn} = \frac{\partial x^i}{\partial u^m} \frac{\partial x^i}{\partial u^n}$ is the array of all pairwise dot products of the basis in the tangent plane. In matrix terminology, this is also the transpose of the Jacobian matrix of $\vec{x}(\vec{u})$ multiplied by the original Jacobian matrix.

For polar coordinates it works out to $g_{ij} = \begin{bmatrix} 1 & 0 \\ 0 & r^2 \end{bmatrix}$

For the earth, it works out to $g_{ij} = \begin{bmatrix} R^2 \cos^2 \beta & 0 \\ 0 & R^2 \end{bmatrix}$ (recall β is latitude)

Arc Length
The arc length of a curve on a surface can be approximated by summing the length of vectors v^i tangent to the curve. To this end consider a surface parametrized by $\vec{x}(\vec{u})$ and a path $\vec{c}(t)$ that is carried to the surface as $\vec{x}(\vec{c}(t))$.

The arc length is approximated as $\sum \sqrt{\left(\frac{d}{dt} \vec{x}(\vec{c}(t)) \right) \cdot \left(\frac{d}{dt} \vec{x}(\vec{c}(t)) \right)} \; \Delta t$

The dot product that expresses vector length can be calculated in the parameter space using g_{ij} and the approximating sum converges to an integral.

$$\int \sqrt{v^i v^i} \, dt = \int \sqrt{\frac{d c^i}{dt} \frac{d c^j}{dt} g_{ij}} dt \quad .$$

Later we will discuss geodesics, that have the smallest path integral between its endpoints.

Note that the integrand expression specifying the arc length also can specify g_{ij} because of its symmetry properties.

If arc length $= \int_{t_1}^{t_2} \sqrt{a\dot{x}^2 + 2b\,\dot{x}\,\dot{y} + d\,\dot{y}^2}\,dt$ then $g_{ij} = \begin{bmatrix} a & b \\ b & d \end{bmatrix}$

Surface Area (or Volume)

Consider a 2-d surface parametrized by $\vec{x}(\vec{u})$. A region on the surface has a corresponding region in the parameter space that can be divided into rectangles spanned by small vectors along the unit vectors $\Delta u^i \hat{u}_i$.

The spanning vectors of corresponding parallelograms in the embedding space are given by $\Delta u^k \frac{\partial x^i}{\partial u^k}$ in the tangent space to the surface. The summed areas of these parallelograms approximate the area. The parallelograms in the embedding space do not bend with the surface. They are in the tangent planes, pinned to the surface at the tangent plane origin.

We don't calculate the area using the vector cross product:

$A = \left| \left(\Delta u^1 \frac{\partial x^k}{\partial u^1} \right) \times \left(\Delta u^2 \frac{\partial x^k}{\partial u^2} \right) \right|$ because it doesn't generalize to higher dimensions.

Previously, we saw that the area^2 of a parallelogram embedded in a space of higher dimension is given by the determinant of the matrix of pairwise dot products of the vectors. In this context, the area squared of each parallelogram is given by the determinant of the matrix whose mn^{th} element is given by the dot product $\left(\Delta u^m \frac{\partial x^k}{\partial u^m} \right) \cdot \left(\Delta u^n \frac{\partial x^k}{\partial u^n} \right)$.

So,we use:

$$\text{Area}^2 = \left| \Delta u^m \frac{\partial x^k}{\partial u^m} \Delta u^n \frac{\partial x^k}{\partial u^n} \right| = \left| \frac{\partial x^k}{\partial u^m} \frac{\partial x^k}{\partial u^n} \right| (\Delta u^1)^2 (\Delta u^2)^2 = |g_{mn}| (\Delta u^1)^2 (\Delta u^2)^2$$

The above operation where the Δu^i's are factored out because they are constant multiples of a row or column.

So that the total area is calculated as:

$$\sum \sqrt{|g_{mn}|} \Delta u^1 \Delta u^2 \rightarrow \int \sqrt{|g_{mn}|} du^1 du^2$$

For polar coordinates $|g_{mn}|=r^2$ and area is calculated with the familiar $\int r\,dr\,d\theta$

Changing Parameterizations

Transition rules for vectors and functionals between parameter spaces in overlapping regions are the same as transition rules between a parameter space and a surface.

Consider two parameterizations of a surface in an embedding space with variables $\vec{z}$. Let the two parameterizations over some overlapping region be given by $\vec{z}(\vec{x})$ and $\vec{z}(\vec{y})$ where the parameter spaces have variables $\vec{x}$ and $\vec{y}$. Then the parameter spaces can parametrize each other by following a point between the surface and the parametrization spaces. We can find a transition function like $\vec{y}(\vec{x})$. This is like identifying the same point on two different geographical map projections.

Consider a vector v^i or a covector a_i in the parameter space with $\vec{y}$ variables, and corresponding vector v'^k or convector a'_k in the space with $\vec{x}$ variables. They must satisfy the dual transformation rules:

$v^i = v'^k \frac{\partial y^i}{\partial x^k}$ and $a'_k = a_i \frac{\partial y^i}{\partial x^k}$ where the derivative terms change from point to point.

Note that, these transformation rules do not need to refer to the embedded surface as long as the transition functions like $\vec{y}(\vec{x})$ are available.

The transformation rules could be written in a form where the derivatives of the inverse transition function is used for transforming functionals:

$$v^i = v'^k \frac{\partial y^i}{\partial x^k} \quad \text{and} \quad a_i = a'_k \frac{\partial x^k}{\partial y^i}$$

We don't see this form as often because the derivatives in one direction often have simpler expressions than the other. There are reasons to use the other form when studying induced functions.

If required, we can find $\frac{\partial x^k}{\partial y^i}$ by inverting the matrix $\frac{\partial y^i}{\partial x^k}$.