2.3 First variation

Here are some notes to compare the optimization process for functions in one dimension, functions in multiple (but finite) dimensions, and functionals. The idea is to try to make the functional case as analogous to the previous cases as possible.

Single variable calculus

1. Function $f(u)$: inputs are numbers, outputs are numbers
2. We say that number u is a critical point if $f'(u) = 0$.
3. If v is any test number, then

   $$\frac{d}{d\varepsilon}\left[f(u+\varepsilon v)\right]_{\varepsilon=0} = f'(u)\, v.$$

4. If quantity q is such that $q\, v = 0$ for all test numbers v, then q must be zero.

   Notice that we actually don't need to consider multiplication every single number in order to conclude that q is zero. It is enough to only use a single test number: $v = 1$.

5. Thus u is a critical point of f if

   $$\frac{d}{d\varepsilon}\left[f(u+\varepsilon v)\right]_{\varepsilon=0} = 0$$

   for all test numbers v.

Multivariable calculus

1. Function $f(\mathbf{u})$: inputs are vectors $\mathbf{u}$, outputs are numbers.
2. We say that vector $\mathbf{u}$ is a critical point if $\operatorname{grad} f(\mathbf{u}) = 0$.
3. If $\mathbf{v}$ is any test vector, then

   $$\frac{d}{d\varepsilon}\left[f(\mathbf{u}+\varepsilon \mathbf{v})\right]_{\varepsilon=0} = \operatorname{grad} f(\mathbf{u}) \cdot \mathbf{v}.$$

4. If quantity $\mathbf{q}$ is such that $\mathbf{q} \cdot \mathbf{v} = 0$ for all test vectors $\mathbf{v}$, then $\mathbf{q}$ must be zero.

   Notice that we don't need to consider the dot product with every single vector. It is enough to consider just a collection of vectors which is large enough to conclude that q is zero. For example, in $\mathbb{R}^2$ it is enough to consider just two test vectors: $v = \langle 1, 0 \rangle$ and $v = \langle 0, 1 \rangle$. (Why?)

5. Thus $\mathbf{u}$ is a critical point of f if

$$\frac{d}{d\varepsilon} \left[f(\mathbf{u} + \varepsilon \mathbf{v}) \right]_{\varepsilon=0} = 0$$

   for all test vectors $\mathbf{v}$.

Functional calculus (Preliminary version)

1. Functional $F(u)$: inputs are functions, outputs are numbers.

2. We say that function u is a critical point if $\operatorname{grad} F(u) = 0$.

   But at this stage we don't know what we mean by "the gradient of a functional"! Our resolution of this issue is to use the $\frac{d}{d\varepsilon}$ stuff to define what we mean by the gradient of the functional.

3. If v is any test function, then

$$\frac{d}{d\varepsilon} \left[F(u + \varepsilon v) \right]_{\varepsilon=0} = \operatorname{grad} F(u) \cdot v.$$

   Now there are two things that we don't have defined: We don't know what the gradient of a functional is, and we also don't know what it means to take the "dot product" of two functions.

   Actually, this formula gives us some hope: If we knew how to define the dot product, we could simply define the gradient of a functional to be whatever object appears in this formula.

4. If quantity q is such that $q \cdot v$ for all test functions v, then q must be zero.

Notice that we don't need to consider the "dot product" with every single function. It is enough to consider just a collection of functions which is large enough to conclude that q is zero.

This is our first clue about how to start making some definitions. Remember that if q is some quantity with

$$\int_\Omega q\,v\,dV = 0 \qquad \textit{for all } v \textit{ in } C_0^\infty$$

then it must be that $q = 0$.

This motivates us to define the "dot product," which we call an inner product, of two functions by "multiply and integrate."

5. Thus u is a critical function of F if

$$\frac{d}{d\varepsilon}\left[F(u+\varepsilon v)\right]_{\varepsilon=0} = 0$$

for all test functions v.

This actually doesn't need much commentary – we can simply use this as a definition of what it means for u to be a critical point of functional F.

Functional calculus (Refined version)

1. Functional $F(u)$:
   - Inputs: functions $u\colon \Omega \to \mathbb{R}^k$.
     Here Ω is some region in $\mathbb{R}^n$.
   - Outputs: numbers.
2. We want to treat our input functions like vectors:
   - We already know we can add/subtract them, as well as scale them.
   - We define the *inner product* of tow functions u and w to be

$$\langle u, v\rangle = \int_\Omega u\,v\,dV.$$

This plays the role of the dot product.

- Here is one important property of this inner product: If q is some quantity with $\langle q, v\rangle = 0$ for all test functions v, then $q = 0$.

3. We now define the gradient of a functional to be the object which satisfies

$$\frac{d}{d\varepsilon}\left[F(u+\varepsilon v)\right]_{\varepsilon=0} = \langle \operatorname{grad} F(u), v\rangle$$

for all test functions v.

This is called the *first variational formula* for functional F.

Just to be clear: we are defining the gradient of F at u is the object which fits in the box:

$$\frac{d}{d\varepsilon}\left[F(u+\varepsilon v)\right]_{\varepsilon=0} = \int_\Omega \boxed{}\, v dV$$

for all test functions v.

4. We say that u is a critical point of F if

$$\operatorname{grad} F(u) = 0.$$

This equation is called the *Euler-Lagrange* equation(s) of the functional.

5. Our definition of critical point means that u is a critical point of F if

$$\frac{d}{d\varepsilon}\left[F(u+\varepsilon v)\right]_{\varepsilon=0} = 0$$

for all test functions v. In practice, this is usually the best method for actually determining the Euler-Lagrange equation(s) of the functional.