An introduction to non-smooth analysis and geometry

Lecture 9: Upper gradients

1. The fundamental theorem of calculus revisited

Recall that if $f\colon \mathbb{R} \to \mathbb{R}$ is an absolutely continuous function, then

$$f(x) - f(y) = \int_x^y f' \, d\mathcal{L}^1.$$

Similarly, if $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is a $\mathcal{C}^1$-mapping, and $\gamma\colon [0,1] \to \mathbb{R}^n$ is a rectifiable path, then

$$f(\gamma(0)) - f(\gamma(1)) = \int_\gamma Df \, ds.$$

This is a very good exercise to do if it is not immediately obvious to you. Another good thing to think about: is this still true if f is only a Lipschitz function? (we'll come back to this point later)

Now, in keeping with the philosophy that a linear structure is not really necessary for analysis, we might ask ourselves which parts of the above statement make sense for a mapping defined on a metric space. The notion of a rectifiable curve makes perfect sense in any metric space (although there may not be any, like in the Von Koch snowflake curve). Integration with respect to arc length is fine, too, but the notion of derivative is problematic.

Let us consider a mapping $f\colon (X, d_X) \to (Y, d_Y)$ between metric spaces. Now, $f(x) - f(y)$ doesn't make sense, but its *absolute value* does, in the form of the metric. The information we have gotten rid of is the "direction" of the variation of f - what is left is only the magnitude. Moreover, since we don't have a notion of direction in the domain either, we can't talk about partial derivatives or directional derivatives. In the case of a $\mathcal{C}^1$-mapping $f\colon \mathbb{R}^n \to \mathbb{R}^m$, the "reduced information" version of the fundamental theorem is

$$|f(\gamma(0)) - f(\gamma(1))| \leq \int_\gamma ||Df|| \, ds.$$

where $||Df||$ is the operator norm of the derivative Df.

Definition 1.1. Let $f\colon (X, d_X) \to (Y, d_Y)$ be a mapping between metric spaces. A Borel function $g\colon (X, d_X) \to [0, \infty]$ is an *upper gradient* of f if for every rectifiable path $\gamma\colon [0,1] \to (X, d_X)$,

$$d_Y(f(\gamma(0)), f(\gamma(1))) \leq \int_\gamma g \, ds.$$

A few points worth noting

- The constant function $g = \infty$ is an upper gradient of every mapping.
- If g is an upper gradient of a mapping f, and $h\colon X \to [0, \infty]$ is another Borel function such that $g(x) \leq h(x)$ for each $x \in X$, then h is also an upper gradient of f. In particular, there is usually no uniqueness for upper gradients.
- If there are no rectifiable paths in X, then 0 is an upper gradient of every mapping defined on X. Hence, this concept is not so interesting in spaces with out a lot of curves.
- if $f\colon \mathbb{R}^n \to \mathbb{R}^m$ is a $\mathcal{C}^1$-mapping, then $||Df||$ is an upper gradient of f.
- if f is L-Lipschitz, then the constant function $g = L$ is an upper gradient of f.

The first result we would like to have is a reasonable existence theory for upper gradients for a large class of mappings.

Theorem 1.2. *Let $f\colon (X, d_X) \to \mathbb{R}$ be a locally Lipschitz mapping, i.e., for each compact set $K \subseteq X$ there is a constant L_K such that $f|_K$ is L_K-Lipschitz. Then* $\operatorname{lip}(f)\colon (X, d_X) \to [0, \infty)$ *is an upper gradient of f.*

Proof. In order to ensure that a path integral

$$\int_\gamma g \; ds$$

is well-defined, the mapping g must be a Borel function on X. To be thorough, we should check the following statement: since f is locally Lipschitz, $\operatorname{lip} f$ is a Borel function. We will not check this, although it is true.

Let $\gamma\colon [0, l] \to X$ be a rectifiable curve parameterized by arclength. Then γ is 1-Lipschitz, and since $\gamma([0, l])$ is a compact set, there is a constant L such that

$$f \circ \gamma\colon [0, l] \to \mathbb{R}$$

is an L-Lipschitz mapping. Thus, by the one-dimensional Rademacher theorem, $f \circ \gamma$ is differentiable $\mathcal{L}^1$-almost everywhere, and

$$|f(\gamma(u)) - f(\gamma(v))| \leq \int_{[u,v]} |(f \circ \gamma)'(t)| dt.$$

Now, let $t \in (0, l)$ be a point of differentiability for $f \circ \gamma$, and consider a sequence of scales $\{r_n\}_{n \in \mathbb{N}}$ decreasing to 0 such that

$$\operatorname{lip} f(\gamma(t)) = \lim_{n \to \infty} \sup_{y \in B_X(\gamma(t), r_n)} \frac{|f(y) - f(\gamma(t))|}{r_n}.$$

Let $\epsilon > 0$. For all sufficiently large n, choose $t_n = t + r_n(1 - \epsilon)$. Then, since γ is parameterized by arclength,

$$r_n(1 - \epsilon) = |t - t_n| = \operatorname{length}(\gamma|_{[t,t_n]}) \geq d_X(\gamma(t_n), \gamma(t)),$$

and so

$$\gamma(t_n) \in B_X(\gamma(t), r_n).$$

Thus

$$\begin{aligned} |(f \circ \gamma)'(t)| &= \lim_{n \to \infty} \frac{|f(\gamma(t_n) - f(\gamma(t))|}{|t - t_n|} \\ &= (1 - \epsilon)^{-1} \lim_{n \to \infty} \frac{|f(\gamma(t_n) - f(\gamma(t))|}{r_n} \\ &\leq (1 - \epsilon)^{-1} \lim_{n \to \infty} \sup_{y \in B_X(\gamma(t), r_n)} \frac{|f(y) - f(\gamma(t))|}{r_n} \\ &= (1 - \epsilon)^{-1} \operatorname{lip} f(\gamma(t)). \end{aligned}$$

Letting $\epsilon \to 0$ shows that $\operatorname{lip} f(\gamma(t))$ majorizes the norm of the derivative $|(f \circ \gamma)'(t)|$. Since almost every point $t \in (0, l)$ is a point of differentiability, this implies that

$$\int_{[0,l]} |(f \circ \gamma)'(t)| \; dt \leq \int_{[0,l]} \operatorname{lip} f(\gamma(t)) \; dt = \int_\gamma \operatorname{lip} f \; ds,$$

as desired. □

It is an interesting problem to try to weaken the locally Lipschitz condition on f above (it is not always possible). Moreover, the result is true if the target $\mathbb{R}$ is replaced by ANY metric space (this would be an interesting topic for a presentation).

2. Absolute continuity

What are the implications if a function has a non-trivial upper gradient? As one might guess based on how we motivated the defintion, we can sometimes recover something like the fundamental theorem of calculus for such functions.

Proposition 2.1. *Let $f\colon X \to \mathbb{R}$ be a function with an upper gradient $g\colon X \to [0, \infty]$. If $\gamma\colon [0, l] \to X$ is an arclength-parameterized path for which*

$$\int_\gamma g \; ds < \infty,$$

then $f \circ \gamma\colon [0, l] \to \mathbb{R}$ is an absolutely continuous function, and moreover, for almost every $t \in [0, l]$,

$$|(f \circ \gamma)'(t)| \leq g(\gamma(t)),$$

Proof. If $\{(a_i, b_i)\}_{i=1}^n$ is a collection of disjoint subintervals of $(0, l)$, then

$$\sum_{i=1}^n |f(\gamma(b_i) - f(\gamma(a_i))| \leq \sum_{i=1}^n \int_{(a_i,b_i)} g(\gamma(t)) \; dt = \int_{\bigcup_i (a_i,b_i)} g(\gamma(t)) \; dt.$$

Since

$$\int_\gamma g \; ds = \int_{[0,l]} g(\gamma(t)) \; dt < \infty,$$

the absolute continuity of the integral implies that for all $\epsilon > 0$, there is $\delta > 0$ such that if $E \subseteq [0, l]$ satisfies $\mathcal{L}^1(E) < \delta$, then

$$\int_E g(\gamma(t)) \; dt < \epsilon.$$

(now is a good time to prove this if you don't know the proof by heart). Hence, if

$$\sum_{i=1}^n |b_i - a_i| < \delta,$$

then

$$\sum_{i=1}^n |f(\gamma(b_i) - f(\gamma(a_i))| \leq \int_{\bigcup_i (a_i,b_i)} g(\gamma(t)) \; dt < \epsilon.$$

This shows that $f \circ \gamma$ is absolutely continuous, and thus differentiable almost everywhere on $[0, l]$

Now, by the Lebesgue differentiation theorem (applied only in $[0, l]$), for almost every $t \in [0, l]$,

$$\lim_{h \to 0} \frac{|f(\gamma(t)) - f(\gamma(t+h))|}{|h|} \leq \lim_{h \to 0} \frac{1}{|h|} \int_{[t,t+h]} g(\gamma(u)) \; du = g(\gamma(t)).$$

□

The moral of the story here is that if f has an upper gradient g that is "small", meaning integrable on many curves, then f is absolutely continuous on many curves. If there are enough such curves, then we might hope to be able to "paste" the different "derivatives along curves" together to form a "real derivative" of f.