QUANTITATIVE DIFFERENTIATION

JASON ZHAO

Contents

1. Lipschitz maps on $\mathbb{R}$

Let $I \subseteq \mathbb{R}$ denote a finite interval, and suppose $f : I \to \mathbb{R}$ is a Lipschitz continuous function with constant $||f||_{\mathrm{Lip}} \leq 1$. Classically, Rademacher's theorem makes the *qualitative* statement that the set of points where f is non-differentiable has measure zero. We would like establish a *quantitative* analogue, which states that for every $\varepsilon > 0$, there cannot be too many scales at which f fails to be ε-linear.

For simplicity, we will only consider discrete scales, dividing up the interval I into dyadic sub-intervals $I_{n,j} \subseteq I$ of size $|I_{n,j}| := 2^{-n}|I|$. Define the deviation of f from linearity on an interval I by

$$\mathsf{Deviation}(f, I) := \frac{1}{|I|} \inf_{\ell} ||f - \ell||_{L^\infty(I)},$$

where $\ell : I \to \mathbb{R}$ are taken over affine functions. It follows from compactness that the infimum is actually achieved. We say f is ε-linear on I if the deviation is less than ε.

Theorem 1 (Quantitative differentiation)**.** *Let $f : I \to \mathbb{R}$ be Lipschitz continuous with constant $||f||_{\mathrm{Lip}} \leq 1$. Then*

$$\sum_{\mathsf{Deviation}(f, I_{n,j}) \geq \varepsilon} |I_{n,j}| \lesssim \frac{|\log_2 \varepsilon|}{\varepsilon^2}|I|.$$

Remark. By density it suffices to prove the theorem for f continuously differentiable. Throughout we will work with quantities which are invariant under the rescaling

$$\mathsf{T}_{x_0,r} f(x) := \frac{1}{r}(f(x_0 + rx) - f(x_0)).$$

Note that f is differentiable at x_0 if and only if $\mathsf{T}_{x_0,r} f \to f'(x_0)x$ in the uniform norm.

1.1. Monotonicity formula.

For $g : I \to \mathbb{R}$ continuously differentiable, define the Dirichlet energy by

$$E(g, I) := \int_I |g'|^2 \, dx.$$

Observe that the minimisers subject to the Dirichlet boundary conditions are precisely affine functions. It follows from the fundamental theorem of calculus that for functions agreeing with f on the boundary, the minimum energy is given by

$$\min_{g_{|\partial I} = f_{|\partial I}} E(g, I) = \int_I |f'_I|^2 \, dx,$$

where f'_I denotes the average of f' on I,

$$f'_I = \frac{1}{|I|} \int_I f' \, dx.$$

Date: May 3, 2023.

Let $\ell_{f,I,n} : I \to \mathbb{R}$ denote the piecewise linear function that agrees with f at the endpoints of each dyadic interval $I_{n,j}$, then the DEFECT of f on I at scale 2^{-n} is defined, equivalently, by

$$\begin{aligned}\theta_n(f, I) &:= \sum_{j=1}^{2^n} \int_{I_{n,j}} (|f'|^2 - |f'_{I_{n,j}}|^2)\, dx \\ &:= E(f, I) - E(\ell_{f,I,n}, I)\end{aligned}$$

Lemma 2 (Monotonicity formula)**.** *Let $f : I \to \mathbb{R}$ be continuously differentiable. Then the defect is monotone non-increasing and in particular $\theta_n \downarrow 0$ as $n \to \infty$.*

Proof. It suffices to show that $E(\ell_{f,I,n}, I) \uparrow E(f, I)$. To show monotonicity, observe that for $n \leq m$, the piecewise linear functions $\ell_{f,I,n}$ and $\ell_{f,I,m}$ agree on the endpoints of the intervals $I_{n,j}$, thus the energy of the former is lower than the later on each interval. Thus

$$\begin{aligned}E(\ell_{f,I,n}, I) &= \sum_{j=1}^{2^n} E(\ell_{f,I,n}, I_{n,j}) \\ &\leq \sum_{j=1}^{2^n} E(\ell_{f,I,m}, I_{n,j}) = E(\ell_{f,I,m}, I)\end{aligned}$$

On the other hand, it is clear that $\ell_{f,I,n} \to f$ pointwise as $n \to \infty$. In fact, Lipschitz continuity furnishes the uniform bound

$$||\ell_{f,I,n} - f||_{L^\infty} \leq 2^{-n}|I|.$$

Thus convergence of the energies follows from the reverse triangle inequality and Holder's inequality. □

1.2. **Quantitative rigidity.** We can view the deviation as an L^∞-integrability order zero regularity measure of the deviation from linearity, and the defect L^2-integrability order one regularity. Clearly the deviation vanishes if and only if the defect at unit scale vanishes. More quantitatively, trading integrability for regularity, we obtain

Lemma 3 (Defect rigidity). *Let $f : I \to \mathbb{R}$ be continuously differentiable with derivative $|f'| \leq 1$. For every $\varepsilon > 0$, if the scale-invariant defect satisfies the bound*

$$\frac{1}{|I|}\theta_0(f, I) \leq \varepsilon^2$$

then f is ε-linear on I,

$$\mathsf{Deviation}(f, I) \leq \varepsilon.$$

Proof. Let $\ell_{f,I,0} : I \to \mathbb{R}$ denote the affine function agreeing with f at the endpoints, which from the fundamental theorem of calculus we know has slope f'_I. Note also that the defect at unit scale can be rewritten as

$$\theta_0(f, I) = \int_I (|f'|^2 - |f'_I|^2)\, dx = \int_I |f' - f'_I|^2\, dx.$$

It follows then from the fundamental theorem of calculus and Cauchy-Schwartz that

$$\begin{aligned}\mathsf{Deviation}(f, I) &\leq \frac{1}{|I|}||f - \ell_{f,I,0}||_{L^\infty} \\ &\leq \frac{1}{|I|}\int_I |f' - f'_I|\, dx \leq \left(\frac{1}{|I|}\int_I |f' - f'_I|^2\, dx\right)^{1/2},\end{aligned}$$

completing the proof. □

Viewing $\ell_{f,I,N}$ as a linear approximation of f at scales 2^{-N} and from the monotonicity formula, we expect the approximation to improve relative to unit scale as $N \to \infty$. We refer to this improvement as the RELATIVE DEFECT of f on I, defined by

$$\mathsf{Defect}_N(f, I) := \theta_0(f, I) - \theta_N(f, I).$$

However, if the improvement is small between unit scale and scale 2^{-N}, then this suggests defect concentration below scales 2^{-N} and approximate linearity above scales 2^{-N}.

Lemma 4 (Relative defect rigidity). *Let $f : I \to \mathbb{R}$ be continuously differentiable with derivative $|f'| \leq 1$. For every $\varepsilon > 0$, if the scale-invariant relative defect satisfies the bound*

$$\frac{1}{|I|}\mathsf{Defect}_N(f, I) \leq \frac{\varepsilon^2}{4}$$

at scale $2^{-N} \leq \frac{\varepsilon}{2}$, then f is ε-linear on I,

$$\mathsf{Deviation}(f, I) \leq \varepsilon.$$

Proof. Since f and $\ell_{f,I,N}$ agree at the endpoints of I, they share a piecewise linear approximation at unit scale. Thus the relative defect of f at scale 2^{-N} is precisely the defect of $\ell_{f,I,N}$ at unit scale,

$$\mathsf{Defect}_N(f, I) = \theta_0(f, I) - \theta_N(f, I) = E(\ell_{f,I,N}, I) - E(\ell_{f,I,0}) = \theta_N(\ell_{f,I,N}, I).$$

It follows from rigidity of the defect then that

$$\mathsf{Deviation}(\ell_{f,I,N}, I) \leq \frac{\varepsilon}{2}.$$

Let ℓ witness the deviation from linearity of $\ell_{f,I,N}$ on I, then

$$\mathsf{Deviation}(f, I) \leq \frac{1}{|I|}||f - \ell||_{L^\infty} \leq \frac{1}{|I|}||f - \ell_{f,I,N}||_{L^\infty} + \mathsf{Deviation}(\ell_{f,I,N}, I) \leq \varepsilon,$$

where the second inequality follows from the triangle inequality and choice of ℓ, and the third inequality from Lipschitz continuity at scale $2^{-N} \leq \frac{\varepsilon}{2}$ and rigidity. □

1.3. **Pigeonhole principle.** We can control the measure of the scales on which the scale-invariant relative defect is large by a pigeonhole principle argument, namely Markov's inequality. Furthermore by quantitative rigidity this also controls the scales on which the deviation is large: for $N \sim |\log_2 \varepsilon|$, we have

$$\sum_{\mathsf{Deviation}(f, I_{n,j}) > \varepsilon} |I_{n,j}| \leq \sum_{\frac{1}{|I_{n,j}|}\mathsf{Defect}_N(f, I_{n,j}) > \frac{\varepsilon^2}{16}} |I_{n,j}| \leq \frac{4}{\varepsilon^2}\sum_{I_{n,j}} \mathsf{Defect}_N(f, I_{n,j}).$$

To conclude the quantitative differentiation theorem, it remains to control the sum of the relative defect at fixed scale 2^{-N} over all dyadic intervals $I_{n,j}$. Rewriting the sum, we see that it forms the following telescoping series,

$$\begin{aligned}\sum_{I_{n,j}} \mathsf{Defect}_N(f, I_{n,j}) &= \sum_{n=0}^{\infty}\sum_{j=1}^{2^n} \theta_0(f, I_{n,j}) - \theta_N(f, I_{n,j}) \\ &= \sum_{n=0}^{\infty} \theta_n(f, I) - \theta_{n+N}(f, I) = \sum_{n=0}^{N-1} \theta_n(f, I).\end{aligned}$$

From the normalisation $|f'| \leq 1$, we conclude the bound

$$\sum_{I_{n,j}} \mathsf{Defect}_N(f, I_{n,j}) \leq N \cdot E(f, I) \lesssim |\log_2 \varepsilon| \cdot |I|.$$

1.4. **Rademacher's theorem.** Given the quantitative differentiation, Rademacher's theorem follows from a Vitali covering argument. We first check that differentiability is equivalent to convergence of the blow-ups to a linear function. Define $\mathsf{T}_{x_0,r}f : [-1, 1] \to \mathbb{R}$ the blow-up of f centered at $x_0 \in I$ at scale $r \ll 1$ by

$$\mathsf{T}_{x_0,r}f(x) := \frac{f(x_0 + rx) - f(x_0)}{r}.$$

Proposition 5. *A function $f : I \to \mathbb{R}$ is differentiable at $x_0 \in I$ if and only if*

$$||\mathsf{T}_{x_0,r}f - ax||_{L^\infty([-1,1])} \xrightarrow{r \to 0} 0$$

for some $a \in \mathbb{R}$.

Proof. It suffices to show the following statements are equivalent,

$$\left|\frac{f(x_0 + h) - f(x_0) - f'(x_0)h}{h}\right| < \varepsilon, \qquad \text{for all } |h| \leq \delta \tag{1}$$

$$\sup_{x \in [-1,1]}\left|\frac{f(x_0 + rx) - f(x_0)}{r} - f'(x_0)x\right| < \varepsilon, \qquad \text{for all } r < \delta. \tag{2}$$

Indeed, (2) implies (1) by choosing $x = \pm 1$ and $|h| = r$, while the converse implication follows from the change of variables $h = rx$ and $|x| \leq 1$. $\square$

Theorem 6 (Rademacher's theorem)**.** *Let $f : I \to \mathbb{R}$ be Lipschitz continuous, then f is differentiable a.e.*

Theorem 7 (Vitali covering theorem)**.** *Let $E \subseteq \mathbb{R}^d$ be a measurable set with finite Lebesgue measure, and let $\mathcal{V}$ be a Vitali covering for E. Then there exists an at most countable disjoint subcollection $\{U_j\}_j \subseteq \mathcal{V}$ such that*

$$\left| E \setminus \bigcup_j U_j \right| = 0.$$

Corollary 8. *Suppose $\mathcal{V}$ has finite measure and forms a Vitali cover of E, then $|E| = 0$.*