A GRADIENT THEOREM FOR LIPSCHITZ CONTINUOUS FUNCTIONS

MIGUEL LERMA

Abstract. We prove an extension of the Gradient Theorem to Lipschitz continuous functions.

1. Introduction

The Gradient Theorem (GT), that can be seen as a generalization of the Fundamental theorem of Calculus (FTC) to n dimensions, is typically formulated for functions $F : U \subseteq \mathbb{R}^n \to \mathbb{R}$ that are continuously differentiable on an open subset U of $\mathbb{R}^n$. If F is one such function, and $\gamma : [a, b] \to U$ is a smooth (continuously differentiable) path, then the theorem states that

$$\int_\gamma \nabla F(\mathbf{x}) \cdot d\mathbf{x} = F(\gamma(b)) - F(\gamma(a)) \,, \tag{1}$$

where $\mathbf{x} = (x_1, \dots, x_n)$, $\nabla F = \left(\frac{\partial F}{\partial x_1}, \dots, \frac{\partial F}{\partial x_n}\right)$, and $\cdot$ represents the dot product, so the integrand is $\nabla F(\mathbf{x}) \cdot d\mathbf{x} = \sum_{i=1}^n \frac{\partial f(\mathbf{x})}{\partial x_i} \, dx_i$. A proof can be found in [5].

In one dimension the GT becomes the FTC for continuously differentiable functions, i.e.:

$$\int_a^b F'(x) \, dx = F(b) - F(a) \,. \tag{2}$$

The FTC can be extended to absolutely continuous functions, as shown in [3] sec. 33.2, theorem 6. On the other hand, advances in some fields require the use of extensions of the GT that apply to classes of functions that are not necessarily everywhere differentiable (see e.g. [4], proposition 1.), so there are practical motivations to extend the GT in a similar manner if possible. However, although we can extend the definition of absolute continuity to $\mathbb{R}^n$ (in more than one way), this property poses some problems, e.g. the composition of two absolutely continuous functions may not be absolutely continuous. So we will restrict the extension of the GT to a smaller class of functions with better properties, namely Lipschitz continuous functions.

Definition 1. A function $F : S \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is *Lipschitz continuous* if there is a constant $K \geq 0$ such that $||f(\mathbf{x}) - f(\mathbf{y})|| \leq K||\mathbf{x} - \mathbf{y}||$ for every $\mathbf{x}, \mathbf{y} \in S$, where $|| \cdot ||$ represents Euclidean distance.

Date: March 17, 2021.

A good source on Lipschitz Analysis is [2].

In order to make sense of the integral of the derivative or gradient of a function, the given function at the least needs to be almost everywhere differentiable, i.e. differentiable everywhere except for a zero measure subset. The following theorem ensures that such is the case for Lipschitz continuous functions.

Rademacher's theorem. *If U is an open subset of $\mathbb{R}^n$ and $F : U \to \mathbb{R}^m$ is Lipschitz continuous, then f is differentiable almost everywhere in U.*

Proof. See e.g. [1], Theorem 3.1.6., or [2] Theorem 3.1. □

This may not be still enough for an extension of the GT to Lipschitz functions, because everywhere differentiability does not imply differentiable almost everywhere on a given path—e.g. the function $\max(x, y)$ is everywhere continuous, and almost everywhere differentiable on $\mathbb{R}^2$, but is not differentiable at any point of the line $x = y$. So (almost everywhere) differentiability on the path will need to be included as an additional premise.

2. Main Result

Now we will pose and prove the main result.

Theorem 1 (Gradient Theorem for Lipschitz Continuous Functions)**.** *Let U be an open subset of $\mathbb{R}^n$. If $F : U \to \mathbb{R}$ is Lipschitz continuous, and $\gamma : [a, b] \to U$ is a smooth path such that F is differentiable at $\gamma(t)$ for almost every $t \in [a, b]$, then*

$$\int_\gamma \nabla F(\mathbf{x}) \cdot d\mathbf{x} = F(\gamma(b)) - F(\gamma(a)) .$$

Proof. The path γ is continuously differentiable on a compact set (the interval $[a, b]$), hence it is Lipschitz continuous (because its derivative is continuous and so bounded on $[a, b]$). The composition of two Lipschitz continuous functions is Lipschitz continuous, hence $t \mapsto F(\gamma(t))$ is Lipschitz continuous, which implies absolutely continuous. By the Fundamental Theorem of Calculus for absolutely continuous functions we have

$$F(\gamma(b)) - F(\gamma(a)) = \int_a^b \frac{d}{dt} F(\gamma(t))\, dt .$$

By the multivariate chain rule we have

$$\frac{d}{dt} F(\gamma(t)) = \nabla F(\gamma(t)) \cdot \gamma'(t)$$

wherever F is differentiable (for almost every $t \in [a, b]$ by hypothesis). Hence

$$\int_a^b \frac{d}{dt} F(\gamma(t))\, dt = \int_a^b \nabla F(\gamma(t)) \cdot \gamma'(t)\, dt = \int_\gamma \nabla F(\mathbf{x}) \cdot d\mathbf{x} ,$$

and the result follows.

□

3. Conclusions

We have discussed difficulties of a possible extension of the Gradient Theorem to functions that are not everywhere differentiable. Then we have posed and proved one such extension to functions that are Lipschitz continuous.

References

[1] Federer, Herbert (1969). Geometric measure theory, Die Grundlehren der mathematischen Wissenschaften, 153, BerlinHeidelbergNew York: Springer-Verlag, pp. xiv+676, ISBN 978-3-540-60656-7, MR 0257325, Zbl 0176.00801.

[2] Heinonen, Juha (2004). Lectures on Lipschitz Analysis.
URL: `http://www.math.jyu.fi/research/reports/rep100.pdf`

[3] A. N. Kolmogorov, S. V. Fomin (1970). Introductory Real Analysis. Dover Books on Mathematics.

[4] Mukund Sundararajan, Ankur Taly, and Qiqi Yan (2017). Axiomatic attribution for deep networks. arXiv preprint `arXiv:1703.01365 [cs.LG]`

[5] Williamson, Richard and Trotter, Hale. (2004). Multivariable Mathematics, Fourth Edition,

Northwestern University, Evanston, USA
Email address: `mlerma@math.northwestern.edu`