

Separation property of continuously differentiable functions*

SANJO ZLOBEC^{1,†}

¹ *Department of Mathematics and Statistics, McGill University, Burnside Hall, 805 Sherbrooke Street West, Montreal, Quebec, Canada H3A 2K6*

Received August 30, 2013; accepted November 26, 2013

Abstract. We show that every continuously differentiable function in several variables with a global Lipschitz derivative on a compact convex set with interior points has a separation property. It separates two classes of quadratic functions given in terms of either the function's convexifiers or its concavifiers. The separation is used to obtain new global characterizations of the derivative and zero derivative points.

AMS subject classifications: 26A06, 26A24, 26A36, 26B20, 90C30

Key words: function separation property, method of convexification, convexifier, concavifier, global properties of the gradient, zero-derivative point

1. Introduction

This paper is based on a decomposition theorem which says that every continuously differentiable function in several variables with a global Lipschitz derivative on a compact convex set can be represented on this set as the difference of a convex function and a convex quadratic function [8]. The decomposition is used in various areas of mathematics where it is sometimes referred to as the “method of convexification”. The basic idea of the method is to apply properties of convex functions to the leading convex function in the decomposition. This often yields new results for generally non-convex functions. For example, the well-known inequalities of Jensen and Karamata for convex functions were thus extended to non-convex functions, e.g., [4, 6]. A noise-determined equation in quantum mechanics can be established by means of Jensen's inequality and further improved by the method of convexification, as it is noted in [3]. We used the method of convexification to characterize zero-derivative points and the gradient in [7, 8]. Here we use the method to establish a separation property of continuously differentiable functions.

In Section 2, we show that every continuously differentiable function with a global Lipschitz derivative on a compact convex set with interior points separates classes of quadratic functions given in terms of free convexifiers and free concavifiers. The separation property is used in Section 3 to introduce new global characterizations of the gradient and zero derivative points and in Section 4 to determine regions with prescribed tolerances around a fixed point.

*Research partly supported by NSERC of Canada.

†Corresponding author. *Email address:* zlobec@math.mcgill.ca (S. Zlobec)

The term “convexification procedures” is used in a different context in numerical optimization where nonconvex programs with twice differentiable functions are transformed into ones which can be solved with the aid of primal-dual methods [1, 2].

2. The function separation theorem

We use the notions of “global Lipschitz property” of the gradient and the “uniform bound property” of functions. Following [7, 8, 9], the gradient of a continuously differentiable function f is said to have the global Lipschitz property on a set K if

$$\|\nabla^T f(x) - \nabla^T f(y)\| \leq L\|x - y\|$$

for some constant L and every x and y in K . Here $\|u\|$ denotes the Euclidean norm of an n -tuple u and u^T denotes its transpose. We note that L can only increase if the set K is enlarged by inclusion. The second notion is less familiar.

Definition 1 (Uniform bound property of functions). *Consider a compact convex set K in \mathbb{R}^n and its point x^* . A function $\varphi(x)$ in n variables is said to have the uniform bound property on $K \setminus \{x^*\}$ if φ is defined on $K \setminus \{x^*\}$ and if there exists a constant c such that $|\varphi(x)| \leq c$ for every x in $K \setminus \{x^*\}$.*

It is well known that a continuously differentiable function f with the global Lipschitz property of its gradient on a compact convex set K can be made convex by adding to it a suitable quadratic term. In particular, if $f(x)$ is a convex function then $C(x, \alpha) = f(x) + \frac{1}{2}\alpha\|x\|^2$ is convex for any “convexifier” $\alpha \geq 0$. However, in some cases $C(x, \alpha)$ can be convex also for $\alpha < 0$. In order to include this possibility, we extend the definition of convexifiers from [8] to “sign unrestricted” numbers which we call “free” convexifiers.

Definition 2 (Free convexifiers and concavifiers). *Consider a continuously differentiable function f in n variables defined on an open set of \mathbb{R}^n containing a compact convex set K . A number α such that $C(x, \alpha) = f(x) + \frac{1}{2}\alpha\|x\|^2$ is a convex function on K is called a free convexifier of f on K . A number β such that $\tilde{C}(x, \beta) = f(x) + \frac{1}{2}\beta\|x\|^2$ is a concave function on K is called a free concavifier of f on K .*

Example 1. *Consider $f(x) = x^4$ on $I = [1/\sqrt{6}, 1]$. Function $C(x, \alpha) = x^4 + \frac{1}{2}\alpha x^2$ is convex for $\alpha = -2$ and $\tilde{C}(x, \beta) = f(x) + \frac{1}{2}\beta\|x\|^2$ is concave for $\beta = -12$. Therefore $\alpha = -2$ and $\beta = -12$ are a free convexifier and a free concavifier, respectively, of $f(x)$ on I .*

Given a compact convex set K and its interior point x^* , let us use free convexifiers and concavifiers to show that f separates particular classes of quadratic functions. First we consider the tangent hyperplane $T(x^*, x) = f(x^*) + \nabla f(x^*) \cdot (x - x^*)$ at x^* and then we construct the “deviation function” $D(x^*, x) = f(x) - T(x^*, x)$. This function has a global separation property.

Theorem 1. *Let us consider a continuously differentiable function f in n variables defined on an open set of \mathbb{R}^n containing a compact convex set K . It is assumed that*

the gradient of f has the global Lipschitz property on K . If α and β are, respectively, a free convexifier and a free concavifier of f on K , then at every interior point x^* of K we have

$$-\frac{1}{2}\alpha\|x - x^*\|^2 \leq D(x^*, x) \leq -\frac{1}{2}\beta\|x - x^*\|^2, \quad (1)$$

for every x in K .

Proof. With a free convexifier α we know that $f(x) = C(x, \alpha) - \frac{1}{2}\alpha\|x\|^2$ where $C(x, \alpha) = f(x) + \frac{1}{2}\alpha\|x\|^2$ is a convex function on K . Therefore for an arbitrary x in K , $C(\lambda x + (1 - \lambda)x^*, \alpha) \leq \lambda C(x, \alpha) + (1 - \lambda)C(x^*, \alpha)$, for every $0 \leq \lambda \leq 1$. Hence, as in [7]

$$\begin{aligned} f(\lambda x + (1 - \lambda)x^*) &\leq \frac{1}{2}\alpha\|\lambda x + (1 - \lambda)x^*\|^2 + \lambda f(x) - \frac{1}{2}\alpha\lambda\|x\|^2 \\ &\quad + (1 - \lambda)f(x^*) + \frac{1}{2}\alpha(1 - \lambda)\|x^*\|^2. \end{aligned}$$

Using properties of the norm, and after division by $\lambda > 0$, this yields

$$\frac{f(x^* + \lambda(x - x^*)) - f(x^*)}{\lambda} \leq f(x) - f(x^*) + \frac{1}{2}\alpha(1 - \lambda)\|x - x^*\|^2.$$

On the left hand side we have a quotient of functions in the single variable λ of the type $0/0$. Using L'Hopital's rule, in the limit $\lambda \rightarrow 0$ this becomes

$$\nabla f(x^*) \cdot (x - x^*) \leq f(x) - f(x^*) + \frac{1}{2}\alpha\|x - x^*\|^2,$$

which is the left-hand side inequality in (1). But also $\tilde{C}(x, \beta) = f(x) + \frac{1}{2}\beta\|x\|^2$ is a concave function for an arbitrary free concavifier β . Using concavity of $\tilde{C}(\cdot, \beta)$, then switching to $f(x)$, rearrangement and using L'Hopital's rule we have

$$\nabla f(x^*) \cdot (x - x^*) \geq f(x) - f(x^*) + \frac{1}{2}\beta\|x - x^*\|^2,$$

the right-hand side inequality in (1). \square

The inequalities in Theorem 1 can be rearranged yielding the main result of the paper.

Theorem 2 (The function separation theorem). *Consider a continuously differentiable function f in n variables defined on an open set of \mathbb{R}^n containing a compact convex set K . It is assumed that the gradient of f has the global Lipschitz property on K . Let x^* be an interior point of K and let α and β , respectively, be a free convexifier and a free concavifier of f on K . Then*

$$-\frac{1}{2}\alpha\|x - x^*\|^2 + f(x^*) + \nabla f(x^*) \cdot (x - x^*) \leq f(x) \leq f(x^*) + \nabla f(x^*) \cdot (x - x^*) - \frac{1}{2}\beta\|x - x^*\|^2,$$

for every $x \in K$.

Observations

If x^* in Theorem 2 is a zero-derivative point of f then for every $x \in K$

$$-\frac{1}{2}\alpha\|x - x^*\|^2 + f(x^*) \leq f(x) \leq -\frac{1}{2}\beta\|x - x^*\|^2 + f(x^*).$$

The theorem is a statement about the two variables $x \in \mathbb{R}^n$ and $x^* \in \mathbb{R}^n$. Its depiction may be given in \mathbb{R}^{2n+1} and, for a fixed x^* in \mathbb{R}^{n+1} .

Example 2. Consider $f(x) = \sin x$. We can specify a free convexifier $\alpha = 1$ and a free concavifier $\beta = -1$ of f on every compact interval around $x^* = 0$. Theorem 2 says that

$$x - \frac{1}{2}x^2 \leq \sin x \leq x + \frac{1}{2}x^2,$$

for every x . The separation property of $f(x)$ at this fixed x^* is depicted in Figure 1.

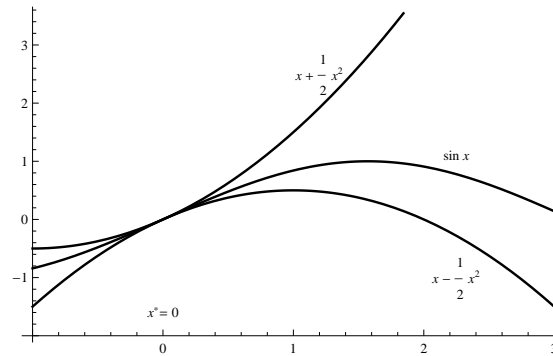


Figure 1: Function separation theorem

The separation property, at the fixed extreme point $x^* = \frac{\pi}{2}$, is depicted in Figure 2.

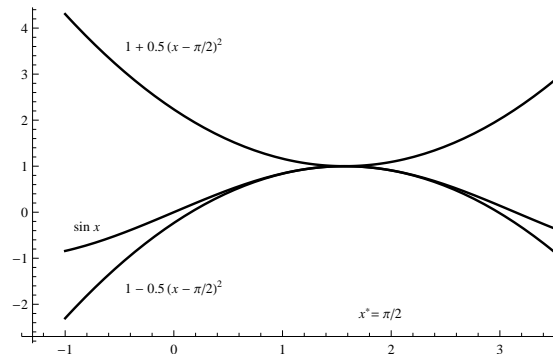


Figure 2: Function separation theorem at zero-derivative point

3. Global characterizations of the gradient

Using Theorem 2 one can globally characterize the gradient.

Theorem 3 (Global characterization of the gradient). *Consider a continuously differentiable function f in n variables defined on an open set of \mathbb{R}^n containing a compact convex set K with a nonempty interior. It is assumed that the gradient of f has the global Lipschitz property on K . Let x^* be an interior point of K and let α and β , respectively, be a free convexifier and a free concavifier of f on K . Denote by $G = G(x^*)$ an arbitrary row n -tuple. Then $G = \nabla f(x^*)$ if, and only if*

$$-\frac{1}{2}\alpha\|x - x^*\|^2 \leq f(x) - f(x^*) - G \cdot (x - x^*) \leq -\frac{1}{2}\beta\|x - x^*\|^2, \quad (2)$$

for every x in K .

Proof. In view of Theorem 2 we only have to show that the n -tuple $G = (G_i)$ in (2) is the gradient. The idea is to separate G from $x - x^*$ in $G \cdot (x - x^*)$. This can be done if we specify x to be $x = x^* + h$, where $h = (h_i)$ is a column n -tuple with zero components everywhere except in the i -th place where we set $h_i \neq 0$. Since $x \neq x^*$, we can divide the inequalities in (2) by $\|x - x^*\| = |h_i|$ to obtain

$$-\frac{1}{2}\alpha|h_i| \leq \frac{f(x^* + h) - f(x^*) - G_i h_i}{h_i} \leq -\frac{1}{2}\beta|h_i|.$$

After sending $h_i \rightarrow 0$, the above yields $G_i = \partial f(x^*)/\partial x_i$. Repeating this process for every $i = 1, \dots, n$ yields $G = \nabla f(x^*)$. \square

After specifying $G = 0$, Theorem 3 gives sharper characterizations of zero-derivative points than those given in [7, 8]. The above results can be simplified. We note that the convexifier α and the concavifier β in Theorem 3 are some fixed numbers. Regardless of their signs and magnitudes there exists a number $\Lambda \geq 0$ such that $\Lambda \geq \frac{1}{2}\alpha$ and $\Lambda \geq -\frac{1}{2}\beta$. For this Λ , using the absolute value function, the inequalities in (2) yield $|f(x) - f(x^*) - G \cdot (x - x^*)| \leq \Lambda\|x - x^*\|^2$. Theorem 3 thus recovers the following characterizations of the gradient from [9] where they were proved differently.

Theorem 4 ([9], Simplified characterizations of the gradient). *Consider a continuously differentiable function $f(x)$ in n variables defined on an open set of \mathbb{R}^n containing a compact convex set K . It is assumed that the gradient of f has the global Lipschitz property on K . At an arbitrary interior point x^* of K , consider an arbitrary row n -tuple $G = G(x^*)$. Then the following four statements are equivalent, i.e., any of them implies the other three:*

- (i) $G = \nabla f(x^*)$;
- (ii) There is a constant $\Lambda \geq 0$ such that $|f(x) - f(x^*) - G \cdot (x - x^*)| \leq \Lambda\|x - x^*\|^2$ for every x in K ;
- (iii) $\frac{1}{\Lambda}|f(x) - f(x^*) - G \cdot (x - x^*)| \leq \|x - x^*\|^2$ for every $\Lambda > 0$ sufficiently large and for every x in K ;

(iv) The ratio function $\varphi(x) = \frac{|f(x) - f(x^*) - G \cdot (x - x^*)|}{\|x - x^*\|^2}$ is uniformly bounded on $K \setminus \{x^*\}$.

One can specify $\Lambda = \frac{1}{2}L$ in (ii) and (iii). This Λ is an upper bound of the ratio function $\varphi(x)$ in (iv).

The above theorem provides multiple equivalent answers to the following basic question: Given a continuously differentiable function f , a row n -tuple G , and an interior point x^* in a compact convex set, what are necessary and sufficient conditions that G be the gradient of f at x^* ? Properties (ii), (iii) and (iv) are called, respectively, the “quadratic envelope”, “normalized quadratic envelope”, and “uniform bound” property of the gradient. Let us depict the latter.

Example 3. Consider the product function $f(x) = x_1x_2$ around some given interior point $x^* = (x_1^*, x_2^*)$ in $K = \{-1 \leq x_1, x_2 \leq 1\}$. We wish to know whether the row $G = (x_2^*, x_1^*)$ (note the reverse order of the components!) is the gradient of f at x^* . This is true if and only if

$$\frac{|(x_1 - x_1^*)(x_2 - x_2^*)|}{(x_1 - x_1^*)^2 + (x_2 - x_2^*)^2},$$

representing $\varphi(x)$ in (iv), is uniformly bounded on $K \setminus \{x^*\}$. The latter is verified after denoting $u_1 = x_1 - x_1^*$ and $u_2 = x_2 - x_2^*$ and noting that $(|u_1| - |u_2|)^2 \geq 0$. The graph of the ratio function $\varphi(x)$ around $x^* = 0$ is depicted by Figure 3. It is uniformly bounded on $K \setminus \{x^*\}$ and we conclude that $G = \nabla f(x^*) = (0, 0)$.

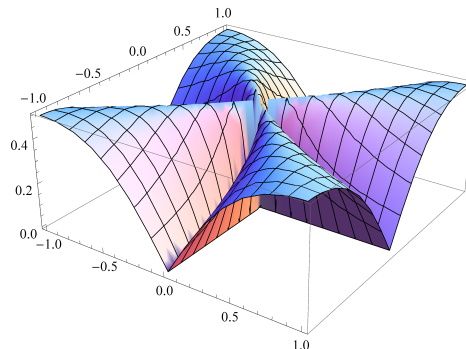


Figure 3: Uniform bound property of the gradient

We have verified the gradient without calculating partial derivatives!

4. Estimating bounds of functions

One of the applications of the Function separation theorem and Theorem 4(ii) is to determine regions around x^* where $f(x)$ approximates $f(x^*)$ within a prescribed tolerance.

Theorem 5. Consider a continuously differentiable function $f(x)$ in n variables defined on an open set of \mathbb{R}^n containing a compact convex set K . It is assumed that the gradient of f has the global Lipschitz property on K with a Lipschitz constant $L \neq 0$. Let x^* be an arbitrary interior point of K and specify some $\epsilon \geq 0$. Then for every x satisfying

$$\|x - x^*\| \leq \left(\frac{2\epsilon}{L}\right)^{\frac{1}{2}}$$

we have

$$|f(x) - f(x^*) - \nabla f(x^*) \cdot (x - x^*)| \leq \epsilon. \quad (3)$$

Example 4. Consider $f(x) = \sin x$ around $x^* = 0$. Take $\epsilon = 0.01$. Let us find an interval around x^* where $|\sin x - x| \leq 0.01$ using Theorem 5. One can specify $L = 1$. We know, after substituting $f(x^*) = 0$ and $\nabla f(x^*) = 1$ in (3), that $\sin x - x \leq 0.01$ whenever $0 < x \leq (0.02)^{\frac{1}{2}}$. A bigger choice of L decreases the estimate for a region around x^* where $f(x)$ approximates $f(x^*)$ within the same tolerance ϵ . Thus for the choice $L = 3$, the bound (3) says that $\sin x - x \leq 0.01$ whenever $0 < x \leq (0.06)^{\frac{1}{2}}$. For another example see, e.g., [5, Example 2.105].

Similarly one can determine regions where the “average value” of the function falls within a prescribed ϵ from the value of f at x^* . For functions of the single variable the average value of f on an interval $I = [x^*, x]$ from x^* to x , where $x > x^*$ is defined as

$$A(x, x^*) = \frac{\int_{x^*}^x f(t) dt}{x - x^*}.$$

Assuming that $f(t)$ has the Lipschitz property on I with a Lipschitz constant L^* we can apply Theorem 4(ii) to the function

$$y(x) = \int_{x^*}^x f(t) dt.$$

Since $y'(x) = f(x)$, $y(x^*) = 0$ and we can set $\Lambda = 1/L^*$. After division by $|x - x^*| \neq 0$, Theorem 4(ii) yields the following result.

Corollary 1. Consider a continuous function of the single variable $f(x)$ on a compact interval I with an interior point x^* . Assume that f satisfies the global Lipschitz property on I with a constant L^* . Then

$$|A(x, x^*) - f(x^*)| \leq \frac{1}{2} L^* |x - x^*|,$$

for every $x \neq x^*$ in I .

Example 5. For $f(x) = \sin x$, from $x^* = 0$ to any $x > 0$, we have $A(x, x^*) = (1 - \cos x)/x$. Let us specify $L^* = 1$ and take $\epsilon = 0.01$. Now we conclude that the average value of $\sin x$ on the interval $0 < x \leq 0.02$ is bounded by 0.01.

5. Conclusion

We have shown that every continuously differentiable function with a global Lipschitz derivative on a compact convex set with interior points globally separates two classes of quadratic functions stated in terms of the function's either free convexifiers or its free concavifiers. The separation yields global characterizations of the derivative and zero-derivative points in the interior of the set. We have also estimated regions around a fixed point where values of the function fall within a specified tolerance. This may be useful in the study of average values of functions over an interval.

Acknowledgement

The author is indebted to the referee for his careful reading of the manuscript and constructive comments.

References

- [1] D. P. BERTSEKAS, *Convexification procedures and decomposition methods for nonconvex optimization problems*, J. Optim. Theory Appl. **29**(1979), 169–197.
- [2] C. A. FLOUDAS, C. E. GOUNARIS, *An overview of advances in global optimization during 2003-2008*, in: *Lectures on global optimization*, (P.M. Pardalos and T.F. Coleman, Eds.), Fields Institute Communications 55, 2009, 105–154.
- [3] R. HEESE, M. FREYBERGER, *Entropic uncertainty relation for pointer-based simultaneous measurements of conjugate observables*, Physical Review A **87**(2013), issue 1, available at <http://link.aps.org/doi/10.1103/PhysRevA.87.012123>.
- [4] A. M. KHAN, *Majorization theorem for convexifiable functions*, Math. Commun. **18**(2013), 61–65.
- [5] L. NERALIĆ, B. ŠEGO, *Matematika*, Second edition, Element, Zagreb, 2013.
- [6] S. ZLOBEC, *Jensen's inequality for nonconvex functions*, Math. Commun. **9**(2004), 119–124.
- [7] S. ZLOBEC, *On the behaviour of functions around zero-derivative points*, Int. J. Optim. Theory, Methods, Appl. **1**(2009), 329–340.
- [8] S. ZLOBEC, *Characterizing zero-derivative points*, J. Global Optimization **46**(2010), 155–161, (Published online on 2nd July 2009).
- [9] S. ZLOBEC, *Equivalent formulations of the gradient*, J. Global Optimization **50**(2011), 549–553.