

SENSITIVITY OF THE “INTERMEDIATE POINT” IN THE MEAN VALUE THEOREM: AN APPROACH VIA THE LEGENDRE-FENCHEL TRANSFORMATION

JEAN-BAPTISTE HIRIART-URRUTY¹

Abstract. We study the sensitivity, essentially the differentiability, of the so-called “intermediate point” c in the classical mean value theorem $\frac{f(b)-f(a)}{b-a} = f'(c)$; we provide the expression of its gradient $\nabla c(d, d)$, thus giving the asymptotic behavior of $c(a, b)$ when both a and b tend to the same point d . Under appropriate mild conditions on f , this result is “universal” in the sense that it does not depend on the point d or the function f . The key tool to get at this result turns out to be the LEGENDRE-FENCHEL transformation for convex functions.

1. INTRODUCTION

The basic mean value theorem (MVT for short), also called LAGRANGE’S MVT, is certainly one of the best known results in Calculus. Let us recall it. Given a differentiable function $f : I \rightarrow \mathbb{R}$ on some open interval I , for any $a < b$ in I , there exists $c \in (a, b)$ such that

$$\frac{f(b) - f(a)}{b - a} = f'(c). \tag{1}$$

We call (1) the MVT relation. The MVT theorem is mainly an *existence* result, it does not say more on this “intermediate point c ”..., especially if it is unique or not. In this note, we precisely consider the situation where such a c is unique for any $a < b$, and study the *sensitivity* of c as a function of both variables a and b (its continuity, and differentiability).

Since one does not know much about such c , c sometimes has the qualifier “mysterious”. Here are some preliminary comments about it:

- The intermediate point(s) c could be anywhere between a and b . We can say more on their location if we know more about the generating function f ; a typical example is when f is a polynomial function of degree n at most (see [10, Section 6.5] for that).
- When f' is strictly monotone on $(0, +\infty)$, $c(a, b) = (f')^{-1} \left[\frac{f(b)-f(a)}{b-a} \right]$ is a way of defining “means” of $a > 0$ and $b > 0$, but not all means are defined in such a way (*cf.* [8]). For example, with the help of the generating function $x > 0 \mapsto f(x) = \exp x$, one gets at $c(a, b) = \ln \left(\frac{e^b - e^a}{b - a} \right)$.

¹ Institut de Mathématiques, Université Paul Sabatier
118 route de Narbonne, Toulouse (France).
jbhu@math.univ-toulouse.fr

- One could expect for c some homogeneity properties like: for the same function f , the intermediate point $c(ta, tb)$ on the dilated interval (ta, tb) , with $t > 0$, is the dilated version of the intermediate point on (a, b) , that is $tc(a, b)$. This is not true, it is even exceptional. For example, $c(a, b)$ just above does not satisfy this property. According to a deep and nice result by S. GOLAB and S. LOJASIEWICZ [3], this happens only for three classes of functions (we restrict ourselves to strictly convex or strictly concave functions f on $(0, +\infty)$): With $\alpha \neq 0$,

$$\begin{aligned} f & : x \in (0, +\infty) \mapsto f(x) = \alpha x \ln x + \beta x + \gamma; \\ f & : x \in (0, +\infty) \mapsto f(x) = \alpha \ln x + \beta x + \gamma; \\ f & : x \in (0, +\infty) \mapsto f(x) = \alpha x^p + \beta x + \gamma \text{ (where } p \neq 0 \text{ and } p \neq 1). \end{aligned}$$

This is the occasion to propose various positively homogeneous means of the type $c(a, b)$:

- If $f(x) = x^2$, then $c(a, b) = \frac{a+b}{2}$ (arithmetic mean of a and b); if $f(x) = 1/x$, then $c(a, b) = \sqrt{ab}$ (geometric mean of a and b). But there is no function f for which the corresponding $c(a, b)$ would be the harmonic mean $\frac{2ab}{a+b}$ of a and b . Let us give a hint for that (from [8, pages 682 – 683]). According to what has been seen above, the (positively homogeneous) harmonic mean could only come from a function of the type $f : x > 0 \mapsto f(x) = x^p$ (where $p \neq 0$ and $p \neq 1$). But this is impossible. Indeed, if that was the case for some power p , one would have:

$$\begin{aligned} \frac{f(2) - f(1)}{2 - 1} & = p [H(1, 2)]^{p-1}, \\ \frac{f(3) - f(1)}{3 - 1} & = p [H(1, 3)]^{p-1}, \end{aligned}$$

where $H(a, b)$ stands for the harmonic mean of a and b . The first equation, that is $\left(\frac{2^p-1}{p}\right)^{\frac{1}{p-1}} = \frac{4}{3}$, leads to $p = -4.14457\dots$; the second one, that is $\left(\frac{3^p-1}{2p}\right)^{\frac{1}{p-1}} = \frac{3}{2}$, leads to $p = -4.36363\dots$.

- The logarithmic mean $c(a, b) = \frac{b-a}{\ln b - \ln a}$ comes from the generating function $\ln x$; the so-called identric mean $c(a, b) = \left(\frac{b^b}{a^a}\right)^{1/(b-a)} / e$ comes from the generating function $x \ln x$.
- We can play with the power p of the function x^p and get at some unusual (positively homogeneous) means of a and b . To take some examples,

$$\begin{aligned} c(a, b) & = \left(\frac{\sqrt{a}+\sqrt{b}}{2}\right)^2 \text{ if } f(x) = \sqrt{x}; \quad c(a, b) = \left(\frac{2a^2b^2}{a+b}\right)^{1/3} \text{ if } f(x) = \frac{1}{x^2}; \quad c(a, b) = \sqrt{\frac{a^2+ab+b^2}{3}} \text{ if } \\ f(x) & = x^3; \quad c(a, b) = \left(\frac{\sqrt{a}\sqrt{b}(\sqrt{a}+\sqrt{b})}{2}\right)^{2/3} \text{ if } f(x) = \frac{1}{\sqrt{x}}; \text{ etc.} \end{aligned}$$

For what we have to explore in this note, there is no loss of generality in assuming that f is defined on the whole of \mathbb{R} , that is to say $I = \mathbb{R}$.

Let us list here some qualitative properties related to this intermediate point c (see [7] for proofs):

- Such a point c is unique for any $a < b$, if and only if either f or $-f$ is strictly convex on \mathbb{R} . That militates for our choice thereafter: $f : \mathbb{R} \rightarrow \mathbb{R}$ is a differentiable strictly convex function.
- Such a point c can sometimes be expressed as a function of a and b , in a closed form; it depends on how easily the strictly increasing function f' can be inverted (see some examples above). There is however a simple basic situation in that respect: If $f(x) = \frac{1}{2}\alpha x^2 + \beta x + \gamma$, with $\alpha \neq 0$, then $c = \frac{a+b}{2}$ for all $a < b$. It is interesting to note that this property of c characterizes quadratic functions [7]. In the same vein, given $p \in (0, 1/2)$, it is impossible to have $c = \frac{pa+(1-p)b}{2}$ for all $a < b$ [7].

From now on, we define the “intermediate point function” $c : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ as follows:

$$\left\{ \begin{array}{l} \text{For } a \neq b, c(a, b) \text{ is the unique } c \text{ satisfying the MVT relation (1)} \\ c(d, d) = d \text{ for all } d \in \mathbb{R}. \end{array} \right. \quad (2)$$

The questions we address in this paper are:

- (The easiest one). Is c a continuous function of a and b ?
- (A more elaborate one). Is c a differentiable function of a and b ? If so, what is the gradient vector of c at any point (d, d) lying on the critical diagonal line of $\mathbb{R} \times \mathbb{R}$?

We answer both questions, delineating the appropriate (minimal) general assumptions on f for that.

2. BASIC TOOLS FROM ANALYSIS

2.1. The extended difference quotient

Given a differentiable strictly convex function $f : \mathbb{R} \rightarrow \mathbb{R}$, we define the extended difference quotient function $q : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ as follows:

$$\left\{ \begin{array}{l} \text{For } a \neq b, q(a, b) = \frac{f(b) - f(a)}{b - a}; \\ q(d, d) = f'(d) \text{ for all } d \in \mathbb{R}. \end{array} \right.$$

Property: q is a continuous function on $\mathbb{R} \times \mathbb{R}$.

This is a standard result in differential calculus. Let us still provide a short proof for the convenience of the reader.

The only nontrivial point to check is that q is continuous at (d, d) .

Since f is convex and differentiable, it is C^1 . Therefore, f is (so-called) strictly differentiable at d , so that

$$\begin{aligned} f(a) - f(b) &= f'(d)(a - b) + (a - b)\varphi(a, b), \text{ with} \\ \lim_{\substack{a \rightarrow d \\ b \rightarrow d}} \varphi(a, b) &= 0. \end{aligned}$$

But, $q(a, b) - q(d, d) = \varphi(a, b)$. Thus $\lim_{\substack{a \rightarrow d \\ b \rightarrow d}} q(a, b) = q(d, d)$.

We can even go further on the properties of q , assuming more about the function f .

Property: Assume that f is C^2 . Then q is a continuously differentiable function on $\mathbb{R} \times \mathbb{R}$, with

$$\nabla q(d, d) = \frac{1}{2} \begin{pmatrix} f''(d) \\ f''(d) \end{pmatrix}. \quad (3)$$

This is again an exercise in differential calculus. Let us prove it quickly.

At a point (a, b) with $a \neq b$, we clearly have:

$$\begin{aligned} \frac{\partial q}{\partial a}(a, b) &= \frac{f(b) - f(a) - (b - a)f'(a)}{(b - a)^2}, \\ \frac{\partial q}{\partial b}(a, b) &= \frac{f(a) - f(b) - (a - b)f'(b)}{(b - a)^2}. \end{aligned}$$

Following second-order TAYLOR-LAGRANGE expansions of f at a and at b , there exist ξ_1 and ξ_2 in (a, b) such that $\frac{\partial q}{\partial a}(a, b) = \frac{1}{2}f''(\xi_1)$ and $\frac{\partial q}{\partial b}(a, b) = \frac{1}{2}f''(\xi_2)$. By the assumed continuity of f'' , we therefore get:

$$\lim_{\substack{a \rightarrow d \\ b \rightarrow d}} \frac{\partial q}{\partial a}(a, b) = \frac{1}{2}f''(d); \quad \lim_{\substack{a \rightarrow d \\ b \rightarrow d}} \frac{\partial q}{\partial b}(a, b) = \frac{1}{2}f''(d).$$

Thus, q is differentiable at (d, d) and (3) holds.

2.2. The LEGENDRE-FENCHEL transformation

Since the given function f is (strictly) convex, it is natural to evoke the LEGENDRE-FENCHEL transform for any property of f . Indeed, for our purposes its use turns out to be incredibly efficient.

The LEGENDRE-FENCHEL transform (or conjugate) f^* of f is defined as follows:

$$s \in \mathbb{R} \mapsto f^*(s) = \sup_{x \in \mathbb{R}} (sx - f(x)). \quad (4)$$

We use s to denote the variable because, in the geometrical interpretation of definition (4), s is a slope. We denote by I^* the interval on which the supremum in (4) is finite. When f' is invertible, which is our case (the strict convexity of f implies that f' is strictly increasing), another way of looking at f^* is:

$$f^*(s) = s \cdot (f')^{-1}(s) - f\left[(f')^{-1}(s)\right]. \quad (5)$$

This is the usual form of the so-called LEGENDRE transform of f , whether f is convex or not. In terms of derivatives, the relation between f and f^* is simple: the derivative function of f^* is the inverse of the derivative function of f :

$$f'(x) = s \Leftrightarrow (f^*)'(s) = x. \quad (6)$$

This property could even help obtain an antiderivative (or primitive function) of f , see an example below (Example 2.4). To get familiar with the transformation $f \rightsquigarrow f^*$, it is best to consider examples, which we do below.

Example 2.1. Let $f(x) = \frac{1}{2}\alpha x^2 + \beta x + \gamma$, with $\alpha > 0$. Then

$$f^*(s) = \frac{s^2}{2\alpha} - \frac{\beta}{\alpha}s + \frac{\beta^2}{2\alpha} - \gamma.$$

Here $I^* = \mathbb{R}$ and f^* is of the same kind (*i.e.*, quadratic) as f .

Example 2.2. Let $f(x) = \exp x$. Then, for $s \geq 0$,

$$f^*(s) = s \ln s - s.$$

As usual, we understand that $0 \ln 0 = 0$. In this example, $I^* = [0, +\infty)$. This is certainly one of the simplest examples to illustrate the relationship (6) between the derivatives of f and f^* : $f'(x) = \exp x$ for all x , while $(f^*)'(s) = \ln s$ for all $s > 0$.

Example 2.3. Let $f(x) = \sqrt{1+x^2}$. Then, for $|s| \leq 1$,

$$f^*(s) = -\sqrt{1-x^2}.$$

Here $I^* = [-1, 1]$.

Example 2.4. Let $f(x) = \cosh x$. Then, a simple calculation of $f^*(s)$ defined in (4) leads to

$$\begin{aligned} f^*(s) &= s \cdot \sinh^{-1} s - \cosh(\sinh^{-1} s) \\ &= s \cdot \sinh^{-1} s - \sqrt{1+s^2}. \end{aligned} \quad (7)$$

Since $f' = \sinh$ and its inverse $(f')^{-1}$ is the derivative of f^* , the above calculation is an alternate way of determining an antiderivative (or primitive) function of \sinh^{-1} : it is the function given in (7).

All these examples illustrate formulas (5) and (6); we will even go further, in due time, with a relationship between second derivatives of f and f^* .

The properties of the LEGENDRE-FENCHEL transformation are well-known and used in convex analysis (see for example [11, section 26] or [5, chapter E]). We list below those we need for our purposes. Recall that we have assumed that $f : \mathbb{R} \rightarrow \mathbb{R}$ is a strictly convex differentiable function. We then have:

- (1) I^* is an interval with a nonempty interior, and f^* is differentiable on $\text{int}(I^*)$.
- (2) $\lim_{k \rightarrow \infty} |f'(x_k)| = +\infty$ whenever (x_k) is a sequence in $\text{int}(I^*)$ converging to an endpoint of $\text{int}(I^*)$.

Note that (2) holds vacuously when I^* is the whole of \mathbb{R} .

Said in different words: $\text{int}(I^*)$ is parameterized by the injective function f' , and for $s \in \text{int}(I^*)$, $f^*(s) = s \cdot x - f(x)$, where x is the unique solution of the equation $s = f'(x)$.

Examples 2.1–2.4 illustrate perfectly well the above-displayed properties of f^* .

We complement the relationship $(f^*)' = (f')^{-1}$ with one on the second-derivatives [2, 12, 13]:

- (3) Suppose f is C^2 with $f'' > 0$ on \mathbb{R} . Then so is f^* on the interior of I^* , with:

$$(f^*)''(s) = \frac{1}{f''(x)}, \text{ where } x \text{ is the unique solution of } s = f'(x). \tag{8}$$

Let us quickly illustrate the property above with the functions displayed in Example 2.3. We have

$$\begin{aligned} f'(x) &= \frac{x}{\sqrt{1+x^2}}, \quad f''(x) = \frac{1}{(1+x^2)^{\frac{3}{2}}}, \\ \text{while } (f^*)'(s) &= \frac{1}{\sqrt{1-s^2}}, \quad (f^*)''(s) = \frac{1}{(1-s^2)^{\frac{3}{2}}}. \end{aligned}$$

Indeed, $s = x/\sqrt{1+x^2}$, $x \in \mathbb{R}$, is a way of parameterizing the interior of $I^*(= (-1, 1))$, and, for all $s = x/\sqrt{1+x^2}$, we indeed have $(f^*)''(s) = 1/f''(x)$.

3. THE MAIN RESULTS

We begin with the following key observation on the function c defined in (2):

$$c = (f^*)' \circ q. \tag{9}$$

This is indeed true for $a \neq b$, since by definition

$$\begin{aligned} c(a, b) &= (f')^{-1} \left(\frac{f(b) - f(a)}{b - a} \right) \\ &= (f^*)' \left(\frac{f(b) - f(a)}{b - a} \right) = ((f^*)' \circ q)(a, b), \end{aligned}$$

but also at diagonal points (d, d) since

$$\begin{aligned} c(d, d) &= d = (f^*)'(f'(d)) \\ &= ((f^*)' \circ q)(d, d). \end{aligned}$$

It now remains to scroll the results, everything has been prepared for that.

Theorem 3.1 (continuity). *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be strictly convex and differentiable. For (a, b) in $\mathbb{R} \times \mathbb{R}$, let $c(a, b)$ be the intermediate point in the MVT, as defined in (2). Then c is a continuous function on $\mathbb{R} \times \mathbb{R}$.*

Proof. We have seen in Section 2.1 that the extended difference quotient q is a continuous function from $\mathbb{R} \times \mathbb{R}$ into $\text{int}(I^*)$. Moreover, $(f^*)' : \text{int}(I^*) \rightarrow \mathbb{R}$, the derivative function of a differentiable convex function, is continuous. Hence, c which is nothing but $(f^*)' \circ q$ (see (9)), is a continuous function. □

The result of Theorem 3.1 is not completely new, it appears in the following alternative form in the literature (in [3, 4, 9] for example): Assuming that f' is strictly monotone, $\theta(a, h)$ is defined as the unique parameter in $(0, 1)$ for which $(f(a + h) - f(a))/h = f'(a + \theta(a, h)h)$; then θ is a continuous function of a and h .

Theorem 3.2 (differentiability). *Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be twice continuously differentiable with $f'' > 0$. For (a, b) in $\mathbb{R} \times \mathbb{R}$, let $c(a, b)$ be the intermediate point in the MVT, as defined in (2). Then c is a continuously differentiable function on $\mathbb{R} \times \mathbb{R}$, with*

$$\nabla c(d, d) = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix} \text{ for all } (d, d) \text{ in } \mathbb{R} \times \mathbb{R}. \tag{10}$$

Proof. Again, $c = (f^*)' \circ q$, resulting from the composition of two continuously differentiable functions, is itself continuously differentiable.

As for $\nabla c(d, d)$, it suffices to apply the chain rule, with (3) and (8) as ingredients:

$$\begin{aligned} \nabla c(d, d) &= (f^*)''(f'(d)) \cdot \frac{1}{2} \begin{pmatrix} f''(d) \\ f''(d) \end{pmatrix} \\ &= \frac{1}{f''(d)} \cdot \frac{1}{2} \begin{pmatrix} f''(d) \\ f''(d) \end{pmatrix} = \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}. \end{aligned}$$

□

Remarks.

- By applying Theorem 3.2 partially, we recover the following result on the behavior of the intermediate point in the MVT (see [1] for example, but it was known long before): assume that f is twice continuously differentiable in a neighborhood of a fixed a , with $f''(a) \neq 0$, then the intermediate point c in the MVT property has the following asymptotic behavior

$$\lim_{b \rightarrow a} \frac{c - a}{b - a} = \frac{1}{2}. \tag{11}$$

Indeed, one can choose $f''(a) > 0$ without loss of generality, and extend f to the whole of \mathbb{R} so as to satisfy the assumptions in Theorem 3.2. Then, use the fact that $\frac{\partial c}{\partial b}(a, a) = \frac{1}{2}$.

- The “perfect” or asymptotic case is when f is quadratic, $f(x) = \frac{1}{2}\alpha x^2 + \beta x + \gamma$, with $\alpha > 0$ (see Example 2.1). In that case, $c(a, b) = \frac{a+b}{2}$ for all (a, b) in $\mathbb{R} \times \mathbb{R}$.

What Theorem 3.2 entails is that, under appropriate assumptions on f , say $f''(d) > 0$, $c(a, b) \sim \frac{a+b}{2}$ when a and b are close to a same point d .

Theorem 3.2 also explains, *a posteriori*, what we have evoked at the beginning: given $p \in (0, 1/2)$, it is impossible to have $c = \frac{pa+(1-p)b}{2}$ for all $a < b$, except for affine functions f (for which $f'' = 0$).

- One cannot weaken the assumptions in Theorem 3.2 by just requiring f to be twice continuously differentiable and strictly convex. For example, when considering the function $x \in \mathbb{R} \mapsto f(x) = x^4$, its second derivative is $f''(x) = 12x^2$, which vanishes only at 0. Then, for the intermediate point, whose explicit expression here is $c(a, b) = \left(\frac{a^3+a^2b+ab^2+b^3}{4}\right)^{1/3}$, we have

$$\frac{\partial c}{\partial a}(0, 0) = \frac{\partial c}{\partial b}(0, 0) = \frac{1}{4^{1/3}}.$$

Final comments.

- As any differentiability result, Theorem 3.2 measures the “sensitivity” of $c(a, b)$ with respect to changes to the point (d, d) . The fact that $\frac{\partial c}{\partial a}(d, d) = \frac{\partial c}{\partial b}(d, d)$ just expresses the symmetry of this sensitivity of c with respect to the two variables a and b ; this is indeed a reasonable and expected result.

- Note that the value of $\nabla c(d, d)$ does not change with d , and is independent of the function f (satisfying the assumptions in Theorem 3.2)! It is a kind of “universal result”, as it sometimes happens in mathematics.

As a by-product of this result, we note that all the means of the form $c(a, b)$ (see the beginning of the paper) have the same asymptotic behavior, $c(a, b) \sim \frac{a+b}{2}$, when a and b are close to a same point d .

We would like to thank the two referees for their careful reading of the submitted manuscript and the subsequent remarks.

REFERENCES

- [1] A.G. AZPEITIA, *On the Lagrange remainder of the Taylor formula*, American Math. Monthly Vol. 89, n°5 (1982), 311 – 312.
- [2] J.-P. CROUZEIX, *A relationship between the second derivatives of a convex function and of its conjugate*. Math. Programming 13 (1977), 364 – 365.
- [3] S. GOLAB and S. LOJASIEWICZ, *Un théorème sur la valeur moyenne θ dans la formule des accroissements finis*. Annales Polonici Mathematici 3 (1956), 118 – 125.
- [4] J.V. CONÇALVES, *Sur l'inconnue θ du théorème des accroissements finis*. Portugaliae Mathematica 2 (1941), 121 – 128.
- [5] J.-B. HIRIART-URRUTY and C. LEMARÉCHAL. *Fundamentals of convex analysis*. Grundlehren Text Editions Springer (2001).
- [6] J.-B. HIRIART-URRUTY, *Mathematical Tapas. Vol. 1 (for Undergraduates)*. Springer Undergraduate Mathematics Series (2016).
- [7] J.-B. HIRIART-URRUTY, *Mean value theorems and convexity: an example of cross-fertilization of two mathematical items*. Elemente Der Mathematik, Vol. 75, N°2 (2020), 69 – 79.
- [8] M.E. MAYS, *Functions which parametrize means*. The American Mathematical Monthly 90, n°10 (1983), 677 – 683.
- [9] B. MORDUKHOVICH, *Qualitative relation of functions f and θ in Lagrange theorem*. Vestnik of Belorussian State University of V.I. Lenin, Series 1, n°1 (1970), 91 – 92 (in Russian).
- [10] Q.I. RAHMAN and G. SCHMEISSER, *Analytic Theory of Polynomials*. Clarendon Press, Oxford (2002).
- [11] R.T. ROCKAFELLAR, *Convex Analysis*. Vol. 28 of Princeton Math. Series, Princeton University Press (1970).
- [12] R.T. ROCKAFELLAR, *Higher derivatives of conjugate convex functions*. The International Journal of Applied Analysis, vol. 1, n°1, (1977), 41 – 42.
- [13] A. SEEGER, *Second derivatives of a convex function and of its Legendre-Fenchel transformate*. SIAM J. Optimization Vol. 2, n°3 (1992), 405 – 424.