Generalizing the Mean Value Theorem

M. Motiwala

Abstract

We discuss possible generalizations of the single variable Mean Value Theorem. We prove a weaker "Mean Value Inequality" which holds in higher-dimensions and give necessary and sufficient conditions for strengthening this result to an "exact" analogue of the single variable statement.

§1 Introduction

Recall the single-variable Mean Value Theorem: for a continuous function $f : \mathbb{R} \to \mathbb{R}$ differentiable on (a, b), there exists some $c \in (a, b)$ such that

$$f(b) - f(a) = f'(c)(b - a)$$

Many of these mathematical objects generalize to vector spaces. For vectors p, q, we denote by $[p, q]$ the line segment joining them. That is,

$$[p, q] = \{p + t(q - p) : t \in [0, 1]\}$$

In this more general context, we deal with functions $f : X \to Y$ between normed vector spaces X, Y, intervals $[a, b]$ become line segments $[p, q]$, and the multiplication of real numbers $f'(c)$ and $b - a$ becomes the application of the linear map $f'(c)$ to the vector $q - p$. So, we might hope for a theorem along the lines of: for a continuous function $f : X \to Y$ differentiable on $[p, q]$, there exists some $c \in (p, q)$ such that

$$f(q) - f(p) = f'(c)(q - p)$$

We will show that such an "exact" analogue to the single variable theorem holds if and only if Y is homeomorphic to $\mathbb{R}$. We'll also prove the following "Mean Value Inequality" which, while a weaker statement, holds in general:

Theorem 1.1 (Mean Value Inequality)**.** *Let X, Y be normed vector spaces. Let $U \subseteq X$ be open and $p, q \in U$ such that the line-segment joining p, q is contained in U. If $f : U \to X$ is differentiable on $[p, q]$ with $f'\big|_{[p,q]}$ bounded by M, then*

$$\|f(q) - f(p)\| \leq M\|q - p\|$$

§2 When do we have an "exact" Mean Value Theorem?

We'll begin with a simple example which immediately demonstrates that we can not hope for an exact analogue to the Mean Value Theorem to hold in general.

2.1 A concrete counter example.

Consider $X = U = \mathbb{R}$, $Y = \mathbb{R}^2$, and take $f : U \to Y$ to be $f(t) = (\cos(t), \sin(t))$. Then, for $t \in \mathbb{R}$,

$$f'(t)(x) = \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix} x$$

Take $p = 0$ and $q = 2\pi$, so that $f(q) - f(p) = 0$. But, for any $c \in \mathbb{R}$,

$$f'(c)(q-p) = \begin{bmatrix} -\sin(t) \\ \cos(t) \end{bmatrix} 2\pi \neq 0$$

since $\sin(t)$ and $\cos(t)$ are never simultaneously zero. Thus, the "exact" MVT is certainly not true in general.

2.2 Generalizing our counter example.

Here, we generalize our counter example from §2.1. Specifically, we will show that for X, Y non-trivial normed vector spaces with $\dim Y \geq 2$, there exists some $f : X \to Y$ for which the exact MVT does not hold.

Let X, Y be non-trivial with $\dim Y \geq 2$. Fix β, γ to be bases of X and Y. Since X is non-trivial, we have some $b \in \beta$, and since Y has dimension at least 2, we have distinct $v_1, v_2 \in \gamma$. Define

$$\hat{b} : X \to \mathbb{R} \qquad \hat{b}(rb + w) = r \quad \text{where } w \in \operatorname{span} \beta(\backslash\{b\})$$

to be the coefficient extraction map for b; that is, $\hat{b}(x)$ extracts the coefficient of b when x is written as a linear combination of elements in β. Note that $\hat{b}$ is bounded and linear so it is differentiable with $\hat{b}'(c) = \hat{b}$.

Next, we define $g : \mathbb{R} \to Y$ as follows,

$$g(t) = \cos(t)v_1 + \sin(t)v_2$$

Claim: g is differentiable with $g'(t) = x \mapsto -x\sin(t)v_1 + x\cos(t)v_2$.

Proof. Let $t \in \mathbb{R}$. Then,

$$\begin{aligned}
&\lim_{x \to t} \frac{\|g(x) - g(t) - (-(x-t)\sin(t)v_1 + (x-t)\cos(t)v_2)\|}{|x-t|} \\
&= \lim_{x \to t} \frac{\|v_1(\cos(x) - \cos(t)) + v_2(\sin(x) - \sin(t)) + (x-t)\sin(t)v_1 - (x-t)\sin(t)v_2\|}{|x-t|} \\
&\leq \lim_{x \to t} \|v_1\| \left| \frac{\cos(x) - \cos(t)}{x-t} + \sin(t) \right| + \|v_2\| \left| \frac{\sin(x) - \sin(t)}{x-t} - \cos(t) \right|
\end{aligned}$$

The expressions inside the absolute values are the difference quotients of cos and sin, so by continuity of the absolute value, we obtain

$$\begin{aligned}
&= \|v_1\|(\cos'(t) + \sin(t)) + \|v_2\|(\sin'(t) - \cos(t)) \\
&= 0
\end{aligned}$$

By the Squeeze theorem, the original limit is zero, and hence we have the result. □

Finally, we'll take $h : X \to Y$ to be $h := g \circ \hat{b}$. As a composition of differentiable functions, h is differentiable with

$$h'(c)(x) = g'(\hat{b}(c))(\hat{b}'(c)(x)) = g'(\hat{b}(c))(\hat{b}(x))$$

Taking $p = 0$ and $q = 2\pi b$, we have $h(q) - h(p) = 0$ but for any $c \in X$,

$$h'(c)(q-p) = g'(t)(2\pi) = -2\pi\sin(t)v_1 + 2\pi\cos(t)v_2 \qquad \text{where } t = \hat{b}(c)$$

For this to be zero, we must have (by linear independence) both $-2\pi\sin(t) = 0 = 2\pi\cos(t)$ which is impossible.

Thus, the pathology observed in §2.1 is in fact quite common; if Y is sufficiently large, we will always find differentiable functions for which no direct analogue to the Mean Value Theorem holds.

2.3 Necessary and sufficient conditions for an exact MVT.

In this section, we prove one of our main results: the natural generalization of the single-variable Mean Value Theorem holds for functions $f : U \to Y$ if and only if Y is homeomorphic to $\mathbb{R}$ (or one of X, Y is trivial).

Note that §2.2 shows one direction: if X, Y are both non-trivial and Y is not homeomorphic to $\mathbb{R}$, then $\dim Y \geq 2$ and so our construction from §2.2 disproves the exact MVT.

Now, we prove the other direction.

Theorem 2.1. *Let X, Y be normed vector spaces. Let $U \subseteq X$ be open and $p, q \in U$ be such that $[p, q] \subseteq U$. Let $f : U \to Y$ be differentiable on $[p, q]$. If either of X, Y is trivial, or if $Y \simeq \mathbb{R}$, then there exists some $c \in (p, q)$ such that*

$$f(q) - f(p) = f'(c)(q - p)$$

where $f'(c)(q - p)$ is the application of $f'(c)$ to the vector $q - p$.

Proof. Let us first dispense with the trivial cases. If Y is trivial, then the only function into Y is the zero map which trivially satisfies the claim. Similarly, if X is trivial, the only functions from X to Y are constant, so again trivially satisfy the claim.

Now, suppose $Y \simeq \mathbb{R}$. This must mean $\dim Y = 1$, so that in fact, we can assume Y is *linearly* homeomorphic to $\mathbb{R}$.

So, there exists a continuous linear bijection $\phi : Y \to \mathbb{R}$. Since ϕ is linear and bounded, it is differentiable with $\phi'(c) = \phi$.

Define $\alpha : [0, 1] \to U$ by

$$\alpha(t) = t(q - p) + p$$

Note that α is well-defined by our assumption that $[p, q] \subseteq U$. Furthermore, α is totally differentiable with $\alpha'(c) = t \mapsto t(q - p)$.

So, by the chain rule, the function

$$g := \phi \circ f \circ \alpha : [0, 1] \to \mathbb{R}$$

is totally differentiable. Recall that under the interpretation of g' as a function into $B(\mathbb{R}, \mathbb{R})$, we have

$$\left.\frac{\mathrm{d}g}{\mathrm{d}x}\right|_{x=c} = g'(c)(1)$$

By the single variable mean value theorem, we have some $\theta \in (0, 1)$ such that

$$g(1) - g(0) = g'(\theta)(1) \cdot (1 - 0)$$

By the chain rule, we have

$$\phi(f(\alpha(1))) - \phi(f(\alpha(0))) = \phi(f'(\alpha(\theta))(\alpha'(\theta)(1)))$$

Of course, $\alpha(1) = q, \alpha(0) = p$ and $\alpha'(\theta)(1) = q - p$, so we get

$$\phi(f(q)) - \phi(f(p)) = \phi(f'(\alpha(\theta))(q - p))$$

By linearity of ϕ, this turns into

$$\phi(f(q) - f(p)) = \phi(f'(\alpha(\theta))(q - p))$$

so by injectivity of ϕ, we have

$$f(q) - f(p) = f'(\alpha(\theta))(q - p)$$

and hence — taking $c = \alpha(\theta)$ — we have the result. □

§3 Proof of the Mean Value Inequality

Now, we prove **Theorem** 1.1. First, we establish the claim for functions $f : [a, b] \to Y$.

Lemma 3.1. *Let Y be a normed vector space. Let $f : [a, b] \to Y$ be totally differentiable with f' bounded by M. Then,*

$$\|f(b) - f(a)\| \leq M(b - a)$$

Proof. We will show $\|f(b) - f(a)\| \leq (M + \varepsilon)(b - a)$ for all $\varepsilon > 0$.

Let $\varepsilon > 0$ and consider the set

$$C = \{s \in [a, b] : \|f(t) - f(a)\| \leq (M + \varepsilon)(t - a) \text{ for all } t \in [a, s]\}$$

Clearly, $a \in C$ and C is bounded, so $S := \sup(C)$ exists.

We would like $S \in C$, for which it suffices to show C is closed.

Claim: C is closed.

Proof. Let $(s_n)_{n=1}^{\infty}$ be a sequence in C and suppose $s_n \to s$. We aim to show $s \in C$. Now, if $t < s$, then there exists $t \leq s_n \leq s$ so that $t \in [a, s_n]$ and hence $\|f(t) - f(a)\| \leq (M + \varepsilon)(t - a)$. So, it remains to show that s itself satisfies the decisive inequality. We show that

$$\|f(s) - f(a)\| \leq (M + \varepsilon)(s - a) + \xi$$

for all $\xi > 0$.

Let $\xi > 0$. Since f is totally differentiable and hence continuous, $f(s_n) \to f(s)$. So, choose $n \in \mathbb{N}$ large enough so that $\|f(s) - f(s_n)\| < \xi/2$ and $|s_n - s| < \xi/2(M + \varepsilon)$. Then,

$$\begin{aligned}\|f(s) - f(a)\| &\leq \|f(s) - f(s_n)\| + \|f(s_n) - f(a)\| \\ &\leq \frac{\xi}{2} + (M+\varepsilon)(s_n - a) \\ &= \frac{\xi}{2} + (M+\varepsilon)(s-a) + (M+\varepsilon)(s_n - s) \\ &< \frac{\xi}{2} + (M+\varepsilon)(s-a) + (M+\varepsilon)\frac{\xi}{2(M+\varepsilon)} \\ &= (M+\varepsilon)(s-a) + \xi\end{aligned}$$

which shows $\|f(s) - f(a)\| \leq (M+\varepsilon)(s-a)$.

Thus, $\|f(t) - f(a)\| \leq (M+\varepsilon)(t-a)$ for all $t \in [a, s]$ and so $s \in C$.

So, C is closed. □

From this, it follows that $S = \sup(C) \in C$.

We will show that $S = b$. That $\sup(C) \leq b$ is clear – since $C \subseteq [a, b]$ and hence is bounded by b – so suppose for the sake of contradiction that $\sup(C) < b$.

By differentiability of f at $S \in [a, b]$, choose $0 < \delta < b - \sup(C)$ such that for all h with $|h| < \delta$, we have

$$\frac{\|f(S+h) - f(S) - f'(S)(h)\|}{|h|} < \varepsilon$$

Let $u = S + \delta/2$. We'll show that $u \in C$, contradicting maximality of S.

Let $t \in [a, u]$. If $t \leq S$, then $t \in [a, S]$ and so we already have $\|f(t) - f(a)\| \leq M(t-a)$. Otherwise, $t = S + d$ for some $0 < d < \delta/2$. So, we have

$$\begin{aligned}\|f(t) - f(S)\| - \|f'(S)(t-S)\| &= \|f(S+d) - f(S)\| - \|f'(S)(d)\| \\ &\leq \Big| \|f(S+d) - f(S)\| - \|f'(S)(d)\| \Big| \\ &\leq \|f(S+d) - f(S) - f'(S)(d)\| && \text{by reverse triangle inequality} \\ &< \varepsilon d && \text{since } d = |d| < \delta\end{aligned}$$

Adding $f'(S)(t-S)$ and writing $d = t - S$, we get

$$\|f(t) - f(S)\| < \varepsilon(t-S) + \|f'(S)(t-S)\|$$

and finally, by boundedness of f', we get

$$\|f(t) - f(S)\| < (M+\varepsilon)(t-S)$$

Thus, we have

$$\begin{aligned}\|f(t) - f(a)\| &\leq \|f(t) - f(S)\| + \|f(S) - f(a)\| \\ &< (M+\varepsilon)(t-S) + (M+\varepsilon)(S-a) && \text{by the above, and since } S \in C \\ &= (M+\varepsilon)(t-a)\end{aligned}$$

So, for all $t \in [a, u]$, $\|f(t) - f(a)\| \le (M + \varepsilon)(t - a)$ so $u \in C$. But, $u > S$ where $S = \sup(C)$. This is a contradiction, so we conclude that in fact, $\sup(C) = b$. In particular, we have $b \in C$ and so

$$\|f(b) - f(a)\| \le (M + \varepsilon)(b - a)$$

Since the above is true for all ε, we conclude that

$$\|f(b) - f(a)\| \le M(b - a)$$

which was to be shown. □

With this lemma in hand, the proof of **Theorem** 1.1 comes quickly.

Proof of 1.1. Let $p, q \in U$ such that $[p, q] \subseteq U$. Let $f : U \to Y$ and suppose f is totally differentiable on $[p, q]$ with $f'\big|_{[p,q]}$ bounded by M.

Consider the function $\alpha : [0, 1] \to U$ defined by

$$\alpha(t) = t(q - p) + p$$

Note that α is well-defined by our assumption $[p, q] \subseteq U$. Moreover, α is totally differentiable with $\alpha'(c) = t \mapsto t(q - p)$.

By the chain rule, the function

$$g := f \circ \alpha : [0, 1] \to Y$$

is totally differentiable with $g'(c) = f'(\alpha(c)) \circ \alpha'(c)$. In particular, for all $c \in [0, 1]$ and for all $t \in \mathbb{R}$ with $|t| \le 1$, we have

$$\begin{aligned}
\|g'(c)(t)\| &= \|f'(\alpha(c))(\alpha'(c)(t))\| \\
&\le M\|\alpha'(c)(t)\| && \text{since } f' \text{ is bounded by } M \\
&= M\|t(q - p)\| \\
&= M|t|\|p - q\| \\
&\le M\|p - q\| && \text{since } |t| \le 1
\end{aligned}$$

Thus, $\|g'(c)\|_{\text{op}} \le M\|p - q\|$ for all $c \in [0, 1]$, so g' is bounded by $M\|p - q\|$. By **Lemma** 3.1, we have

$$\|g(1) - g(0)\| \le M\|p - q\|(1 - 0)$$

Of course, $g(1) = f(q)$ and $g(0) = f(p)$, so we get

$$\|f(q) - f(p)\| \le M\|p - q\|$$

So, if $[p, q] \subseteq U$ with $f'\big|_{[p,q]}$ bounded by M, we have $\|f(p) - f(q)\| \le M\|p - q\|$, as needed. □

§4 Corollaries of the Mean Value Inequality

Here, we present some quick corollaries of the Mean Value Inequality.

Corollary 4.1. *Let X, Y be normed vector spaces. Let $U \subseteq X$ be open and convex. If $f : U \to Y$ is totally differentiable with f' bounded by M, then f is uniformly continuous.*

Proof. Let $\varepsilon > 0$ be given and choose $\delta = \varepsilon/M$. Let $p, q \in U$ such that $\|p - q\| < \delta$. Since U is convex, $[p, q] \subseteq U$. Furthermore, f' is globally bounded by M, so in particular $f'\big|_{[p,q]}$ is bounded by M. Thus, by the Mean Value Inequality, we have

$$\|f(q) - f(p)\| \leq M\|q - p\| < M\frac{\varepsilon}{M} = \varepsilon$$

Thus, f is uniformly continuous. □

Corollary 4.2. *Let X, Y be normed vector spaces. Let $U \subseteq X$ be open and convex. If $f : U \to Y$ is totally differentiable with $f' \equiv 0$, then f is constant.*

Proof. We show $f(p) = f(q)$ for all $p, q \in U$. Let $p, q \in U$. By convexity, $[p, q] \subseteq U$. Moreover, $f' \equiv 0$, so $f'\big|_{[p,q]}$ is bounded by $M = 0$, and hence by the Mean Value Inequality, we have

$$\|f(q) - f(p)\| \leq 0\|q - p\| = 0$$

By positive-definiteness, this implies $f(q) - f(p) = 0$ and hence $f(p) = f(q)$.

Thus, we have the result. □