University of Toronto – MAT237Y1 – LEC5201

Multivariable calculus!

Jean-Baptiste Campesato

January 1st, 2020

When the dimension of the codomain is greater than 1, there is no generalization of the MVT. Nevertheless, we have the following useful MVT-like inequality.

Theorem 1. *Let $U \subset \mathbb{R}^n$ be an open subset and $f : U \to \mathbb{R}^p$ be differentiable.*
Let $a, b \in U$. Assume that $\forall t \in [0,1]$, $(1-t)a + tb \in U$. Then

$$\|f(b) - f(a)\| \leq \left(\sup_{t\in(0,1)} \|Df\left((1-t)a + tb\right)\| \right) \|b - a\|$$

You'll find below two proofs of this theorem.

Proof 1 (using the MVT).
If $f(a) = f(b)$ then there is nothing to do, so we may assume that $f(b) \neq f(a)$.
Define $\psi : U \to \mathbb{R}$ by $\psi(x) = f(x) \cdot \left(\frac{f(b)-f(a)}{\|f(b)-f(a)\|} \right)$ where $\cdot$ denotes the dot product.
We may apply the MVT to ψ (since its codomain is $\mathbb{R}$), and we obtain that there exists $\theta \in (0,1)$ such that

$$\begin{aligned} \psi(b) - \psi(a) &= \nabla\psi\left((1-\theta)a + \theta b\right) \cdot (b-a) \\ &= d_{(1-\theta)a+\theta b}\psi(b-a) \\ &= \left(Df\big((1-\theta)a+\theta b\big)(b-a)\right) \cdot \left(\frac{f(b)-f(a)}{\|f(b)-f(a)\|} \right) \end{aligned}$$

(For the last equality, I used the fact that since $x \mapsto x \cdot v$ is linear then it is differentiable and that its differential is itself, together with the chain rule).

Then, after replacing ψ by its definition and simplifying, we obtain

$$\begin{aligned} \|f(b)-f(a)\| &\leq \left| \left(Df\big((1-\theta)a+\theta b\big)(b-a)\right) \cdot \left(\frac{f(b)-f(a)}{\|f(b)-f(a)\|} \right) \right| \\ &\leq \left\|Df\big((1-\theta)a+\theta b\big)(b-a)\right\| \left\| \frac{f(b)-f(a)}{\|f(b)-f(a)\|} \right\| \quad \text{by the Cauchy–Schwarz inequality} \\ &= \left\|Df\big((1-\theta)a+\theta b\big)(b-a)\right\| \\ &\leq \left\|Df\big((1-\theta)a+\theta b\big)\right\| \|b-a\| \quad \text{by sub-multiplicativity of the Frobenius norm} \\ &\leq \left(\sup_{t\in(0,1)} \|Df\left((1-t)a+tb\right)\| \right) \|b-a\| \end{aligned}$$

■

Proof 2 (using the FTC).

Claim 1. $f(b) - f(a) = \displaystyle\int_0^1 Df\left((1-t)a+tb\right)(b-a)\mathrm{d}t$
where the integral is computed componentwise
i.e. we apply the integral to the components of the vector in $M_{p,1}(\mathbb{R})$ obtained by multiplying the matrix $Df\left((1-t)a+tb\right) \in M_{p,n}(\mathbb{R})$ with the vector $(b-a) \in M_{n,1}(\mathbb{R})$.

Indeed, the k-th component of the RHS is

$$\begin{aligned}\int_0^1 \sum_{i=1}^n \frac{\partial f_k}{\partial x_i}((1-t)a+tb)(b_i-a_i)\mathrm{d}t &= \int_0^1 \varphi'(t)\mathrm{d}t \\ &= \varphi(1)-\varphi(0) \text{ by the Fundamental Theorem of Calculus} \\ &= f_k(b)-f_k(a)\end{aligned}$$

where $\varphi(t) = f_k\left((1-t)a+tb\right)$.
Which proves the claim.

Claim 2. $\left\|\left(\int_0^1 g_1(t)\mathrm{d}t, \ldots, \int_0^1 g_p(t)\mathrm{d}t\right)\right\| \leq \int_0^1 \left\|\left(g_1(t), \ldots, g_p(t)\right)\right\| \mathrm{d}t$

Indeed,

$$\begin{aligned}\left\|\left(\int_0^1 g_1(t)\mathrm{d}t, \ldots, \int_0^1 g_p(t)\mathrm{d}t\right)\right\|^2 &= \sum_{i=1}^n \left(\int_0^1 g_i(t)\mathrm{d}t\right)^2 \\ &= \sum_{i=1}^n \int_0^1 g_i(t)\mathrm{d}t \int_0^1 g_i(s)\mathrm{d}s \quad \text{since } t \text{ is a bound variable} \\ &= \int_0^1 \int_0^1 \sum_{i=1}^s g_i(t)g_i(s)\mathrm{d}t\mathrm{d}s \\ &= \int_0^1 \int_0^1 \left(g_1(t), \ldots, g_p(t)\right) \cdot \left(g_1(s), \ldots, g_p(s)\right) \mathrm{d}t\mathrm{d}s \\ &\leq \int_0^1 \int_0^1 \left|\left(g_1(t), \ldots, g_p(t)\right) \cdot \left(g_1(s), \ldots, g_p(s)\right)\right| \mathrm{d}t\mathrm{d}s \\ &\leq \int_0^1 \int_0^1 \left\|\left(g_1(t), \ldots, g_p(t)\right)\right\| \left\|\left(g_1(s), \ldots, g_p(s)\right)\right\| \mathrm{d}t\mathrm{d}s \text{ by Cauchy–Schwarz} \\ &= \int_0^1 \left\|\left(g_1(t), \ldots, g_p(t)\right)\right\| \mathrm{d}t \int_0^1 \left\|\left(g_1(s), \ldots, g_p(s)\right)\right\| \mathrm{d}s \\ &= \left(\int_0^1 \left\|\left(g_1(t), \ldots, g_p(t)\right)\right\| \mathrm{d}t\right)^2\end{aligned}$$

And the claim follows.

We go back to the proof of the theorem:

$$\begin{aligned}\|f(b)-f(a)\| &= \left\|\int_0^1 Df\left((1-t)a+tb\right)(b-a)\mathrm{d}t\right\| \quad \text{by Claim 1} \\ &\leq \int_0^1 \left\|Df\left((1-t)a+tb\right)(b-a)\right\| \mathrm{d}t \quad \text{by Claim 2} \\ &\leq \int_0^1 \left\|Df\left((1-t)a+tb\right)\right\| \|b-a\| \mathrm{d}t \quad \text{by sub-multiplicativity of the Frobenius norm} \\ &\leq \int_0^1 \left(\sup_{s\in(0,1)} \left\|Df\left((1-s)a+sb\right)\right\|\right) \|b-a\|\mathrm{d}t \\ &= \left(\sup_{s\in(0,1)} \left\|Df\left((1-s)a+sb\right)\right\|\right) \|b-a\|\end{aligned}$$

■