II. Calculus

II.3. Mean Value Theorem

Lecture based on

https://en.wikipedia.org/wiki/Mean_value_theorem

https://github.com/gwthomas/math4ml (Garrett Thomas, 2018)

Prof. Dr. Christoph Lippert

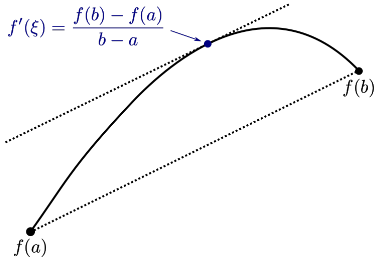
Digital Health & Machine Learning

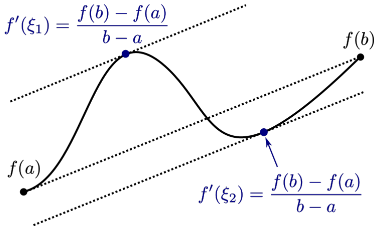
Theorem (Mean value theorem)

Let $f : [a, b] \to \mathbb{R}$ *be a continuous function on the closed interval* $[a, b]$*, and differentiable on the open interval* (a, b)*, where* $a \neq b$*.*

Then there exists some $c \in (a, b)$ *such that*

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





Mean value theorem in several variables

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function.

Fix points $\mathbf{x}, \mathbf{y}$, and define $g(t) = f\Big((1-t)\mathbf{x} + t\mathbf{y}\Big)$.

Since g is a differentiable function in one variable, the mean value theorem gives:

$$g(1) - g(0) = g'(c) \quad \text{for some } c \in [0,1].$$

But since $g(1) = f(\mathbf{y})$ and $g(0) = f(\mathbf{x})$, computing $g'(c)$ explicitly we have:

$$f(\mathbf{y}) - f(\mathbf{x}) = \nabla f\Big((1-c)\mathbf{x} + c\mathbf{y}\Big)^{\top}(\mathbf{y} - \mathbf{x})$$

By the **Cauchy–Schwarz inequality**, the equation gives the estimate:

$$\Big|f(\mathbf{y}) - f(\mathbf{x})\Big| \leq \Big|\nabla f\Big((1-c)\mathbf{x} + c\mathbf{y}\Big)\Big| \, \Big|\mathbf{y} - \mathbf{x}\Big|.$$

In particular, when the partial derivatives of f are bounded, f is **Lipschitz continuous**.

Mean value theorem for vector-valued functions

Let $\mathbf{f} : [a, b] \to \mathbb{R}^n$ be a continuous differentiable vector-valued function.

$$f_i(x + h) - f_i(x) = \nabla f_i(x + t_i h)^\top h$$

Generally there will not be a *single* t that fullfils this for all i.

However a certain type of generalization of the mean value theorem to vector-valued functions is obtained as follows:

Let f be a continuously differentiable real-valued function defined on an open interval I, and let $\mathbf{x}$ as well as $\mathbf{x}+h$ be points.

The mean value theorem in one variable tells us that there exists some t^* between 0 and 1 such that

$$f(x+h)-f(x)=f'(x+t^*h)\cdot h.$$

On the other hand, we have, by the fundamental theorem of calculus followed by a change of variables,

$$f(x+h)-f(x)=\int_x^{x+h} f'(u)\,du=\left(\int_0^1 f'(x+th)\,dt\right)\cdot h.$$

Thus, the value $f'(x+t^*h)$ at the particular point t^* has been replaced by the mean value $\int_0^1 f'(x+th)\,dt$.

Let $\mathbf{f}:\mathbb{R}^n\to'\mathbb{R}^m$ continuously differentiable, and $x,h\in\mathbb{R}^n$ be vectors.

If f is twice continuously differentiable, then

$$f(\mathbf{x}_0 + \mathbf{h}) \approx f(\mathbf{x}_0) + \nabla f(\mathbf{x}_0)^\top \mathbf{h} + \frac{1}{2}\mathbf{h}^\top \nabla^2 f(\mathbf{x}_0)\mathbf{h}$$

is a parabolic approximation to f that has the same Gradient and Hessian as f at $\mathbf{x}_0$.

Taylor's theorem has natural generalizations to functions of more than one variable.

Theorem (Taylor's theorem)

Suppose $f : \mathbb{R}^d \to \mathbb{R}$ is continuously differentiable, and let $\mathbf{h} \in \mathbb{R}^d$.
Then there exists $t \in (0, 1)$ such that

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \nabla f(\mathbf{x} + t\mathbf{h})^\top \mathbf{h}$$

Furthermore, if f is twice continuously differentiable, then

$$\nabla f(\mathbf{x} + \mathbf{h}) = \nabla f(\mathbf{x}) + \int_0^1 \nabla^2 f(\mathbf{x} + t\mathbf{h})\mathbf{h}\, \mathrm{d}t$$

and there exists $t \in (0, 1)$ such that

$$f(\mathbf{x} + \mathbf{h}) = f(\mathbf{x}) + \nabla f(\mathbf{x})^\top \mathbf{h} + \frac{1}{2}\mathbf{h}^\top \nabla^2 f(\mathbf{x} + t\mathbf{h})\mathbf{h}$$

This theorem is used in proofs about conditions for local minima of unconstrained optimization problems.