HIGHER ORDER DERIVATIVES, TAYLOR'S THEOREM

MATANIA BEN-ARTZI

May 2015

Functions here are **real** , defined on an open domain $T \subseteq \mathbb{R}^n$ and **continuously differentiable** (i.e., in $C^1(T)$).

The (open) ball of radius r, centered at x, is denoted by $B(x, r)$. It will be clear from the context what is the dimension .

We use the Euclidean norm. The corresponding operator norm is denoted by $\| \cdot \|$.

- Let $f \in C^1(T)$ be a real function.
- <u>DEFINITION</u>: (**Second Order Derivatives**): If $\frac{\partial f}{\partial x_j}$, has a partial derivative with respect to x_k, for some $k \in \{1, 2, ..., n\}$, at $x^{(0)} \in T$, we denote this derivative by $\frac{\partial^2 f}{\partial x_k \partial x_j}(x^{(0)})$.
- We call it the "mixed" (k, j) second order derivative (at $x^{(0)}$).
- REMARK: For the mixed derivative to exist at x_0 it is necessary that the first derivative $\frac{\partial f}{\partial x_j}$ exist in some ball $B(x^{(0)}, \delta)$.
- `NOTATION`: The "pure" second order derivative is denoted by $\frac{\partial^2 f}{\partial x_j^2}(x^{(0)})$.
-

  THEOREM. *Let $f \in C^1(T)$ be a real function.*

  Suppose that the mixed derivatives $\frac{\partial^2 f}{\partial x_k \partial x_j}$ and $\frac{\partial^2 f}{\partial x_j \partial x_k}$, $\quad j, k \in \{1, 2, ..., n\}$, exist in some ball $B(x^{(0)}, \delta)$.

  If they are continuous at $x^{(0)}$ then

$$\frac{\partial^2 f}{\partial x_k \partial x_j}(x^{(0)}) = \frac{\partial^2 f}{\partial x_j \partial x_k}(x^{(0)}).$$

- **PROOF.** It is enough to prove for a function of two variables, $f(x, y)$, $(x, y) \in T \subseteq \mathbb{R}^2$.

  Fix $(x_0, y_0) \in T$ and let $\delta > 0$ be sufficiently small so that if $|h|, |k| < \delta$ then the rectangle Q whose vertices are the four points four points

$$(x_0, y_0),\ (x_0 + h, y_0),\ (x_0 + h, y_0 + k),\ (x_0, y_0 + k)$$

  satisfies $\overline{Q} \subseteq T$.

  Define the functions (of a single variable):

$$\phi(x) = f(x, y_0 + k) - f(x, y_0), \quad \psi(y) = f(x_0 + h, y) - f(x_0, y).$$

  Then

$$\phi(x_0 + h) - \phi(x_0) = \psi(y_0 + k) - \psi(y_0).$$

But by the *Mean Value Theorem*:

$$\phi(x_0+h)-\phi(x_0)=h\phi'(\xi)=h\Big[\frac{\partial f(x,y_0+k)}{\partial x}-\frac{\partial f(x,y_0)}{\partial x}\Big]\Big|_{x=\xi},$$

$$\psi(y_0+k)-\psi(y_0)=k\psi'(\eta)=k\Big[\frac{\partial f(x_0+h,y)}{\partial y}-\frac{\partial f(x_0,y)}{\partial y}\Big]\Big|_{y=\eta},$$

where ξ is a point between x_0, x_0+h and η is a point between y_0, y_0+k.

Using again the mean value theorem we get,

$$\phi(x_0+h)-\phi(x_0)=hk\frac{\partial^2 f(x,y)}{\partial y\partial x}\Big|_{x=\xi,y=\eta_1},$$

$$\psi(y_0+k)-\psi(y_0)=hk\frac{\partial^2 f(x,y)}{\partial x\partial y}\Big|_{x=\xi_1,y=\eta},$$

where ξ_1 is a point between x_0, x_0+h and η_1 is a point between y_0, y_0+k. Clearly

$$\phi(x_0+h)-\phi(x_0)=\psi(y_0+k)-\psi(y_0),$$

so that

$$\frac{\partial^2 f(x,y)}{\partial y\partial x}\Big|_{x=\xi,y=\eta_1}=\frac{\partial^2 f(x,y)}{\partial x\partial y}\Big|_{x=\xi_1,y=\eta}.$$

Letting h, $k\to 0$ and using the assumed continuity we finally obtain

$$\frac{\partial^2 f(x,y)}{\partial y\partial x}\Big|_{x=x_0,y=y_0}=\frac{\partial^2 f(x,y)}{\partial x\partial y}\Big|_{x=x_0,y=y_0}.$$

■

- REMARK: It is sufficient to assume that only *one* of the mixed derivatives is continuous at x_0. See R. Courant and F. John, Introduction to Calculus and Analysis, vol. 2, Ch. 1.4.d, p.36.
- Going to higher order, it should be clear what we mean by the "$l-th$ order mixed derivative" $\frac{\partial^l f}{\partial x_{i_1}...\partial x_{i_l}}$, $(i_1,...,i_l)\in\{1,2,...,n\}$.
- <u>DEFINITION</u>: The space $C^l(T)$ is the space of functions whose derivatives of all orders $\le l$ exist and are continuous in T.
- **Equality of Mixed Derivatives** : If $f\in C^l(T)$ then the value of $\frac{\partial^l f}{\partial x_{i_1}...\partial x_{i_l}}$ remains invariant under any permutation of the indices $i_1,i_2,...,i_l$.

****

TAYLOR'S THEOREM IN DIMENSION ≥ 2

****

- **TAYLOR'S THEOREM**: Let $f\in C^{l+1}(B(x,\delta))$ be a real function, for some $l\ge 1$. Then, for any $y\in B(x,\delta))$,

$$f(y)=f(x)+\sum_{i=1}^n\frac{\partial f(x)}{\partial x_i}(y_i-x_i)+\frac{1}{2!}\sum_{i,j=1}^n\frac{\partial^2 f(x)}{\partial x_i\partial x_j}(y_i-x_i)(y_j-x_j)+...+$$

$$\frac{1}{l!}\sum_{i_1,...,i_l=1}^n\frac{\partial^l f(x)}{\partial x_{i_1}...\partial x_{i_l}}(y_{i_1}-x_{i_1})(y_{i_2}-x_{i_2})...(y_{i_l}-x_{i_l})$$

$$+\frac{1}{(l+1)!}\sum_{i_1,...,i_l,i_{l+1}=1}^n\frac{\partial^{l+1} f(\xi)}{\partial x_{i_1}...\partial x_{i_{l+1}}}(y_{i_1}-x_{i_1})(y_{i_2}-x_{i_2})...(y_{i_l}-x_{i_l})(y_{i_{l+1}}-x_{i_{l+1}}),$$

where $\xi\in B(x,\delta)$ (actually, on the line segment connecting x to y).

- **PROOF.** Define the function of one variable

$$g(t) = f(x + t(y - x)), \quad 0 \le t \le 1,$$

and apply Taylor's theorem for a single variable

$$g(1) = g(0) + g'(0) + \frac{1}{2}g''(0) + ... + \frac{1}{l!}g^{(l)}(0) + \frac{1}{(l+1)!}g^{(l+1)}(\tau), \quad 0 \le \tau \le 1.$$

Now express the derivatives of g in terms of the partial derivatives of f, using the chain rule:

$$g^{(k)}(0) = \sum_{i_1,...,i_k=1}^{n} \frac{\partial^k f(x)}{\partial x_{i_1} ... \partial x_{i_k}} (y_{i_1} - x_{i_1})(y_{i_2} - x_{i_2})...(y_{i_k} - x_{i_k}).$$

****

The **HESSIAN**

****

- DEFINITION (**HESSIAN**): The symmetric matrix

$$H(x) = \Big(\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\Big)_{1 \le i,j \le n}$$

is called the **Hessian** matrix (at x) of f.

- REMARK: H is actually the Jacobian matrix of the map $D(Df)(x) \in Hom(\mathbb{R}^n, \mathbb{R}^n)$.

****

- **SUFFICIENT CONDITIONS FOR LOCAL EXTREMUM OF A C^2 FUNCTION.**

****

- **THEOREM.** *Let $f \in C^2(T)$ be a real function. Suppose that $x^{(0)} \in T$ is a critical point of f. Then:*

(a) If $H(x^{(0)})$ is positive definite–the point $x^{(0)}$ is a local minimum of f.

(b) If $H(x^{(0)})$ is negative definite–the point $x^{(0)}$ is a local maximum of f.

Institute of Mathematics, Hebrew University, Jerusalem 91904, Israel
E-mail address: mbartzi@math.huji.ac.il