APPLICATIONS OF DIFFERENTIABILITY IN $\mathbb{R}^n$.

MATANIA BEN-ARTZI

April 2015

Functions here are defined on a subset $T \subseteq \mathbb{R}^n$ and take values in $\mathbb{R}^m$, where m can be smaller, equal or greater than n.

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 (as in the convergence definition below).

Usually, we use the Euclidean norm. The ball depends on the norm, but all the concepts are "norm independent".

CONVENTION: When we say that g is "differentiable at x_0" we mean that it is defined in some small (open) ball $B(x_0, \eta)$.

- DEFINITION: (**directional derivative of a function at a point**): Let $g(x) \in \mathbb{R}^m$ be defined in $B(x_0, \eta) \subseteq \mathbb{R}^n$ for some small $\eta > 0$. Let $u \in S^{n-1}$ (i.e., u is a UNIT VECTOR in $\mathbb{R}^n$). The **directional derivative of** g at x_0, in the $u-$direction, is defined by
$$D_u g(x_0) = \lim_{h \downarrow 0} \frac{g(x_0 + hu) - g(x_0)}{h}.$$
- `NOTATION`: Another common notation is $\frac{\partial g}{\partial u}(x_0)$.
- **CLAIM**: Suppose that g is differentiable at x_0. Then its directional derivative at x_0 exists (in any direction) and satisfies:
$$D_u g(x_0) = Dg(x_0)u.$$
- REMARK: Note that the right-hand side above is an application of a linear transformation (from $\mathbb{R}^n$ to $\mathbb{R}^m$) to a (unit) vector in $\mathbb{R}^n$, so the result is a vector in $\mathbb{R}^m$.
- REMARK: In terms of coordinates, the directional derivative is given by:
$$D_u g(x_0) = Jg(x_0)u.$$
Now $D_u g(x_0)$ and u are, respectively, the $m-$vector and $n-$ vector of coordinates.

****

THE CASE OF REAL FUNCTIONS

****

- REMARK: If $m = 1$ (i.e., g is a REAL function) then
$$D_u g(x_0) = \nabla g(x_0) \cdot u.$$
- **COROLLARY**: If g is a real function, differentiable at x_0, the direction of $\nabla g(x_0)$ is the direction of **STEEPEST INCREASE**.

The opposite direction, that of $-\nabla g(x_0)$, is that of **STEEPEST DECREASE**.

- REMARK : If $\nabla g(x_0) = 0$ then ALL DIRECTIONAL DERIVATIVES VANISH.
- MORE GENERALLY, suppose that
  (a) g is a real function differentiable at $x_0 \in \mathbb{R}^n$.
  (b) $(a, b) \ni t \hookrightarrow \gamma(t) \in \mathbb{R}^n$ is a differentiable curve.
  (c) $\gamma(c) = x_0, \quad$ for some $c \in (a, b)$.
- **CLAIM**: Under the above conditions, $g(\gamma(t))$ is differentiable at c and

$$\frac{d}{dt} g(\gamma(t))|_{t=c} = \nabla g(x_0) \cdot \gamma'(t)|_{t=c}.$$

- DEFINITION(**derivative along a curve**): The derivative $\frac{d}{dt} g(\gamma(t))$ is called the **derivative of g along γ.**
- EXAMPLE: Let $\gamma(t),\ t \in \mathbb{R}$, be a straight line in $\mathbb{R}^n$ given by

$$\gamma(t) = x_0 + tu, \quad u \in S^{n-1}.$$

  Then

$$\frac{d}{dt} g(\gamma(t))|_{t=0} = \nabla g(x_0) \cdot u,$$

  which is the **directional derivative of g at x_0, in the direction of u.**
- DEFINITION (**graph of a function**): The graph of g is the set

$$\{(x, g(x)),\, x \in D\} \subseteq \mathbb{R}^{n+1}.$$

- GEOMETRICALLY, if $\nabla g(x_0) = 0$ then $(x_0, g(x_0))$ is a “flat point” of the graph $y = g(x)$.
  On the other hand, if $\nabla g(x_0) \neq 0$, the the direction of the gradient is the direction of “steepest increase” on the graph (at the point $(x_0, g(x_0))$).

****

HYPERSURFACES and LEVEL SURFACES

****

- DEFINITION (**hypersurface**) Let $S \subseteq \mathbb{R}^{n+1}$. We say that S is an $n-$**dimensional hypersurface** if for every $y \in S$ there exists a $\delta > 0$ such that $S \cap B(y, \delta)$ is a graph of a function (with respect to a set of n out of the $n + 1$ coordinates).
  In the case $n = 2$ it is usually called a **surface** (two-dimensional geometric object).
- DEFINITION (**level surface**): Let f be a real function defined in $D \subseteq \mathbb{R}^{n+1}$, where D is open. Let $y_0 \in D$ and let $c = f(y_0)$. If the set $L_{y_0} = \{y \in D, \quad f(y) = c\}$ is a hypersurface, we call it the $c-$**level surface** of f (in D).
  REMARK: We can define the *local* $c-$level surface by restricting to $L^{\delta}_{y_0} = L_{y_0} \cap B(y_0, \delta)$ for some small $\delta > 0$.
- **RESTRICTION**: We assume always that the function expressing the hypersurface as a graph is **continuously differentiable**.
- EXAMPLE: For any real continuously differentiable function $g : T \hookrightarrow \mathbb{R}, \quad T \subseteq \mathbb{R}^n$ (T open set) the graph of g is the $0-$ level surface (in $T \times \mathbb{R} \subseteq \mathbb{R}^{n+1}$) of the function (of the $n + 1$ coordinates x, y) $f(x, y) = y - g(x)$.

****

GRADIENTS AND LEVEL SURFACES

****

FUNDAMENTAL QUESTION: Let $D \subseteq \mathbb{R}^{n+1}$ be an open domain (i.e., connected) and $f : D \to \mathbb{R}$ a real differentiable function.

Let $x_0 \in D$. *When is the set $L^{\delta}_{x_0} = \{x \in D \cap B(x_0, \delta),\ f(x) = f(x_0)\}$ a hypersurface (for sufficiently small $\delta > 0$)?*

- **LEMMA.** Suppose that $\frac{\partial f}{\partial x_1}(x_0) > 0$. Then there exists a ball $B(x_0, \delta)$ such that the set $\{x \in D \cap B(x_0, \delta),\ f(x) = f(x_0)\}$ is a hypersurface.
- NOTATION: For $x \in \mathbb{R}^{n+1}$ we write $x = (x_1, x'), \quad , x' = (x_2, ..., x_n, x_{n+1}) \in \mathbb{R}^n$.
- **PROOF.**
- Let $c = f(x_0)$.
- By continuity , there exists $\delta_1 > 0$ such that, $\frac{\partial f}{\partial x_1}(z) > 0, \quad z = (z_1, z'), \quad |z_1 - (x_0)_1| < \delta_1,\ |z' - (x_0)'| < \delta_1$.
- f is strictly increasing in x_1, so for some $\theta > 0$,

$$f((x_0)_1 + \delta_1, (x_0)') > c + \theta, \quad f((x_0)_1 - \delta_1, (x_0)') < c - \theta.$$

- By continuity there exists $0 < \delta_2 < \delta_1$ such that

$$f((x_0)_1 + \delta_1, x') > c + \theta, \quad f((x_0)_1 - \delta_1, x') < c - \theta, \quad |x' - (x_0)'| < \delta_2.$$

- CONCLUSION: Since $f(x_1, x')$ is continuous and strictly increasing in x_1, the **intermediate value theorem** implies that, for every $x' \in B((x_0)', \delta_2)$ there exists a UNIQUE $\widehat{x_1} = \widehat{x_1}(x') \in ((x_0)_1 - \delta_1, (x_0)_1 + \delta_1)$ such that

$$f(\widehat{x_1}, x') = c.$$

- PROPOSITION: The (real) function $\widehat{x_1}(x')$ is continuously differentiable on $B((x_0)', \delta_2)$.
- PROOF of the PROPOSITION.

  (a) *Continuity:* Let $\varepsilon > 0$ be given. Take above $0 < \delta_1 < \varepsilon$. Then there is a suitable $\delta_2 > 0$. If $y, z \in B((x_0)', \delta_2) \subseteq \mathbb{R}^n$ Then $\widehat{x_1}(y), \widehat{x_1}(z) \in ((x_0)_1 - \delta_1, (x_0)_1 + \delta_1)$ so that

$$|\widehat{x_1}(y) - \widehat{x_1}(z)| \le 2\varepsilon.$$

  (b) *Differentiability:*

  Take $y, z \in B((x_0)', \delta_2) \subseteq \mathbb{R}^n$ so that only their first coordinates are different, namely,

$$y = (x_2, x_3, ..., x_{n+1}), \quad z = (x_2 + h, x_3, ..., x_{n+1}).$$

  Then by definition

$$f(\widehat{x_1}(y), y) - f(\widehat{x_1}(z), z) = c - c = 0.$$

  On the other hand

$$\begin{aligned} f(\widehat{x_1}(z), z) - f(\widehat{x_1}(y), y) &= [f(\widehat{x_1}(z), z) - f(\widehat{x_1}(y), z)] \\ &\quad + [f(\widehat{x_1}(y), z) - f(\widehat{x_1}(y), y)]. \end{aligned}$$

  Denote

$$\Delta\widehat{x_1} = \widehat{x_1}(z) - \widehat{x_1}(y).$$

Then from the mean value theorem

$$f(\widehat{x_1}(z), z) - f(\widehat{x_1}(y), z) = \frac{\partial f}{\partial x_1}(\xi_1, z)\Delta\widehat{x_1},$$

where ξ_1 is between $\widehat{x_1}(z)$ and $\widehat{x_1}(y)$.

Similarly

$$f(\widehat{x_1}(y), z) - f(\widehat{x_1}(y), y) = \frac{\partial f}{\partial x_2}(\widehat{x_1}(y), \xi_2, x_3, ..., x_{n+1})h,$$

where $\xi_2 \in (x_2, x_2 + h)$ (assuming $h > 0$).

Hence

$$0 = \frac{\partial f}{\partial x_1}(\xi_1, z)\Delta\widehat{x_1} + \frac{\partial f}{\partial x_2}(\widehat{x_1}(y), \xi_2, x_3, ..., x_{n+1})h.$$

- We conclude that

$$\frac{\Delta\widehat{x_1}}{h} = -\frac{\frac{\partial f}{\partial x_2}(\widehat{x_1}(y), \xi_2, x_3, ..., x_{n+1})}{\frac{\partial f}{\partial x_1}(\xi_1, z)}.$$

By the continuity property of $\widehat{x_1}$ we have

$$\xi_1 \to \widehat{x_1}(y), \quad \xi_2 \to x_2, \quad \text{as } h \to 0.$$

- We therefore obtain in the limit (as $h \to 0$), since the partial derivatives are continuous,

$$\frac{\partial\widehat{x_1}}{\partial x_2}(y) = -\Big[\frac{\partial f}{\partial x_1}(\widehat{x_1}(y), y)\Big]^{-1}\frac{\partial f}{\partial x_2}(\widehat{x_1}(y), y), \quad y \in B((x_0)', \delta_2).$$

- Replacing x_2 by any x_j, $3 \le j \le n+1$, we see that all partial derivatives of $\widehat{x_1}(x')$ exist and are continuous for $x' \in B((x_0)', \delta_2)$.
- **THIS CONCLUDES THE PROOF OF THE PROPOSITION AND THEREFORE THE LEMMA IS PROVED.**

****

GEOMETRIC INTERPRETATION OF THE GRADIENT

****

- <u>DEFINITION</u> (**orthogonality to a hypersurface**): Let $S \subseteq \mathbb{R}^{n+1}$ be a hypersurface and $y_0 \in S$. We say that a vector $N \in \mathbb{R}^{n+1}$ is **orthogonal** to S at y_0 if the following condition is satisfied:

  For every curve $\{\gamma(t), -\varepsilon < t < \varepsilon\} \subseteq S$ such that $\gamma(0) = y_0$, the tangent $\gamma'(0)$ is orthogonal to N, namely, $N \cdot \gamma'(0) = 0$.
- **THEOREM** (**the gradient as normal to level surfaces**): Let f be a real continuously differentiable function in a neighborhood of $y_0 \in \mathbb{R}^{n+1}$, and assume that $\nabla f(y_0) \neq 0$.

  Let $L^\delta_{y_0} = \{y \in B(y_0, \delta), \quad f(y) = f(y_0)\}$ be the level surface through y_0, for some sufficiently small $\delta > 0$. Then $\nabla f(y_0)$ is orthogonal to $L^\delta_{y_0}$ (at y_0).
- **PROOF**: Let $\{\gamma(t), -\varepsilon < t < \varepsilon\} \subseteq L_{y_0}$ be a curve such that $\gamma(0) = y_0$. Then $f(\gamma(t)) \equiv f(y_0)$ (why?) so that by the chain rule

$$0 = \frac{d}{dt}f(\gamma(t)) = \nabla f(\gamma(t)) \cdot \gamma'(t), \quad t \in (-\varepsilon, \varepsilon).$$

In particular, this is true at $t = 0$, so that $\nabla f(y_0)$ is orthogonal to $\gamma'(0)$.

- EXAMPLE(**tangent plane of a surface in** $\mathbb{R}^3$): Let $f(x,y,z)$ be a continuously differentiable function in a ball $B(P_0,\delta) \subseteq \mathbb{R}^3,\ P_0 = (x_0,y_0,z_0)$. If $\nabla f(P_0) \neq 0$, then the equation $f(x,y,z) = f(P_0)$ defines a surface in a small neighborhood of P_0, and the equation of the tangent plane to this surface, at P_0, is given by

$$\nabla f(P_0)\cdot(x-x_0,y-y_0,z-z_0) = \frac{\partial f}{\partial x}(P_0)(x-x_0)+\frac{\partial f}{\partial y}(P_0)(y-y_0)+\frac{\partial f}{\partial z}(P_0)(z-z_0) = 0.$$

****

EXTREMUM

****

- We assume that $T \subseteq \mathbb{R}^n$ is an open domain and g is a real continuous function defined in T.
- DEFINITION (**critical point**): If the real function $g(x)$ is differentiable at x_0, and $\nabla g(x_0) = 0$ we say that x_0 is a **critical point** of g.
- DEFINITION (**local extremum**): Let g be a real function defined in $T \subseteq \mathbb{R}^n$. We say that g has a **local maximum** (resp. **minimum**) at $x_0 \in T$ if there exists $\delta > 0$ such that

$$g(x_0) \geq g(y), \quad y \in T \cap B(x_0,\delta)$$

(resp. $g(x_0) \leq g(y), \quad y \in T \cap B(x_0,\delta)$).
- **THEOREM**: Let g be a real function defined in $B(x_0,\delta) \subseteq \mathbb{R}^n$ and differentiable at x_0. If g has a local extremum at x_0, then x_0 is a critical point of g.
- **PROOF.** In particular, for every $j = 1,2,...,n$, the function has an extremum (at x_0) along the x_j-direction (i.e., the one-dimensional function $g(x_0+te_j)$ has extremum at $t = 0$). It follows that

$$\frac{\partial g}{\partial x_j}(x_0) = 0, \quad j = 1,2,...,n.$$

- DEFINITION (**local extremum** with **constraints**): Let g be a real function defined in $T \subseteq \mathbb{R}^n$. Let ϕ be another real function defined in $T \subseteq \mathbb{R}^n$, such that $\phi(x_0) = 0$ at $x_0 \in T$. We say that g has a **local maximum** (resp. **minimum**) at x_0, **subject to the constraint** $\phi = 0$, if there exists $\delta > 0$ such that

$$g(x_0) \geq g(y), \quad y \in T \cap B(x_0,\delta) \cap \{\phi(y) = 0\}$$

(resp. $g(x_0) \leq g(y), \quad y \in T \cap B(x_0,\delta) \cap \{\phi(y) = 0\}$).
- **THEOREM**: Let g, ϕ be real functions defined in a ball $B(x_0,\delta) \subseteq \mathbb{R}^n$. Suppose that
  (1) g is differentiable at x_0.
  (2) ϕ is continuously differentiable in $B(x_0,\delta) \subseteq \mathbb{R}^n$, $\phi(x_0) = 0$ and $\nabla\phi(x_0) \neq 0$.

  Then:

  If g has a local extremum at x_0, subject to the constraint $\phi = 0$, then there exists a constant $\mu \in \mathbb{R}$ such that $\nabla g(x_0) = \mu\nabla\phi(x_0)$.
- **PROOF.**

  (a) By the conditions on ϕ we know that the level surface $\phi(x) = 0$ exists in a neighborhood of x_0, and without loss of generality we may assume that it is given as

$$x_n = \psi(x'),\ x' = (x_1, ..., x_{n-1}),$$

where $(x_0)_n = \psi(x'_0)$.

(b) Recall that

$$\frac{\partial \psi}{\partial x_j}(x'_0) = -\frac{\frac{\partial \phi}{\partial x_j}(x_0)}{\frac{\partial \phi}{\partial x_n}(x_0)}, \quad j = 1, 2, ..., n-1.$$

(c) Thus, in a neighborhood of x_0, the function g, subject to the constraint, can be expressed as a function of the $n-1$ coordinates x',

$$h(x') = g(x', \psi(x')), \quad x' \text{ in a neighborhood of } x'_0.$$

(d) The function h has an extremum at x'_0, so its gradient vanishes there. By the chain rule,

$$0 = \frac{\partial h}{\partial x_j}(x'_0) = \frac{\partial g}{\partial x_j}(x_0) + \frac{\partial g}{\partial x_n}(x_0)\frac{\partial \psi}{\partial x_j}(x'_0),\ j = 1, 2, ..., n-1,$$

so that

$$\frac{\partial g}{\partial x_j}(x_0) = \frac{\partial g}{\partial x_n}(x_0)\frac{\frac{\partial \phi}{\partial x_j}(x_0)}{\frac{\partial \phi}{\partial x_n}(x_0)} = \mu\frac{\partial \phi}{\partial x_j}(x_0), \quad j = 1, 2, ..., n-1,$$

where

$$\mu = \frac{\frac{\partial g}{\partial x_n}(x_0)}{\frac{\partial \phi}{\partial x_n}(x_0)}.$$

(e) Of course also

$$\frac{\partial g}{\partial x_n}(x_0) = \mu\frac{\partial \phi}{\partial x_n}(x_0).$$

■

- DEFINITION (**Lagrange multiplier**): The constant μ in the theorem is called the **Lagrange multiplier** of the constrained extremal problem.
  ****
- **THE GRADIENT CAN BE USED TO ESTIMATE FUNCTIONAL DIFFERENCES.**
  ****
- **THEOREM**: Let g be a real differentiable function in an open set $T \subseteq \mathbb{R}^n$. Let $a, b \in T$ be such that the line segment $l(a, b)$ connecting them is contained in T. Then

$$|g(b) - g(a)| \le \sup_{x \in l(a,b)} |\nabla g(x)| \cdot |b - a|.$$

- **PROOF.** Consider the function $f(t) = g(a + t(b - a))$, defined on the real interval $t \in [0, 1]$.

  By the chain rule it is differentiable with derivative given by

$$f'(t) = \nabla g(a + t(b - a)) \cdot (b - a).$$

Now apply the mean value theorem and the Cauchy-Schwarz inequality.

■

- HOWEVER, IT IS NOT NECESSARY FOR g TO HAVE A CRITICAL POINT ON l EVEN IF $g(b) = g(a)$ (different from Rolle's Theorem in one dimension).
- EXAMPLE: Take $n = 2$ and consider $T = \{(x, y), \quad 1 < x^2 + y^2 < 17\}$ and $g(x, y) = x^2 + y^2$.
- **THEOREM**: Let $T \subseteq \mathbb{R}^n$ be open and connected and let g be a real differentiable function in T. Suppose that $\nabla g \equiv 0$ in T. Then $g \equiv \textit{constant}$ in T.

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