Lecture 8 (10.4.2019)

(partially translated and adapted from lecture notes by Martin Klazar)

Multivariable calculus

The following theorem gives a method how to compute multivariable Riemann integral by computing "ordinary" integrals.

Theorem 29 (Fubini)**.** *Let $X \subset \mathbb{R}^m$, $Y \subset \mathbb{R}^n$ and $Z = X \times Y \subset \mathbb{R}^{m+n}$ be $m-$, $n-$, and $m+n$-dimensional boxes, respectively. Let $f : Z \to \mathbb{R}$, $f \in \mathcal{R}(Z)$. Then integrals $\int_Z f$, $\int_X(\int_Y f)$ and $\int_Y(\int_X f)$ exist and are all equal.*

Integrals $\int_X(\int_Y f)$ and $\int_Y(\int_X f)$ have the following meaning. Define a function $F : X \to \mathbb{R}$ as $F(\mathbf{x}) = \int_Y f(\mathbf{x}, \mathbf{y})\ d\mathbf{y}$, whenever $\int_Y f(\mathbf{x}, \mathbf{y})\ d\mathbf{y}$ exists and by arbitrary value from the interval $[\underline{\int_Y} f(\mathbf{x}, \mathbf{y})\ d\mathbf{y}, \overline{\int_Y} f(\mathbf{x}, \mathbf{y})\ d\mathbf{y}]$ otherwise. We then interpret $\int_X(\int_Y f)$ as $\int_X F$. We define a function $G : Y \to \mathbb{R}$ and interpret $\int_Y(\int_X f)$ analogously as $\int_Y G$.

By repeated application of Fubini Theorem, one can derive the following.

Corollary 30. *Let $I = [a_1, b_1] \times \cdots \times [a_n, b_n]$ be a box and let $f : I \to \mathbb{R}$, $f \in \mathcal{R}(I)$. Then*

$$\int_I f = \int_{a_n}^{b_n} \left(\int_{a_{n-1}}^{b_{n-1}} \cdots \left(\cdots \int_{a_1}^{b_1} f(x_1, \ldots x_n)\ dx_1 \right) \cdots\ dx_{n-1} \right)\ dx_n.$$

Note that the order of variables can be chosen arbitrarily.

Directional derivative, partial derivative, total differential

Let $U \subset \mathbb{R}^m$ be a neighborhood of a point $\mathbf{a}$ and $f : U \to \mathbb{R}$ be a function. *Directional derivative* of f at a point $\mathbf{a}$ in direction $\mathbf{v} \in \mathbb{R}^m \setminus \{\mathbf{o}\}$ is defined as a limit

$$\mathrm{D}_{\mathbf{v}} f(\mathbf{a}) := \lim_{t \to 0} \frac{f(\mathbf{a} + t\mathbf{v}) - f(\mathbf{a})}{t}\ ,$$

if it exists. Imagine that U is an area in three dimensional Euclidean space, where f is a function of temperature in a given point and a particle moving through the area. Directional derivatives $\mathrm{D}_{\mathbf{v}} f(\mathbf{a})$ corresponds to immediate change of temperature of surroundings of a particle in a moment when it is at a point $\mathbf{a}$ and has velocity $\mathbf{v}$.

Partial derivative of a function f at a point $\mathbf{a}$ with respect to the i-th variable x_i is a directional derivative $\mathrm{D}_{\mathbf{e}_i} f(\mathbf{a})$, where $\mathbf{e}_i$ the i-th vector of canonical basis, i.e., $\mathbf{e}_i = (0, 0, \ldots, 0, 1, 0, 0, \ldots, 0)$ has i-th coordinate 1 and all other coordinates 0. We denote partial derivative by $\frac{\partial f}{\partial x_i}(\mathbf{a})$ (or, as a shortcut $\partial_i f(\mathbf{a})$. Thus, partial derivative equals to the following limit.

$$\frac{\partial f}{\partial x_i}(\mathbf{a}) = \lim_{h \to 0} \frac{f(a_1, \ldots, a_{i-1}, a_i + h, a_{i+1}, \ldots, a_m) - f(a_1, a_2, \ldots, a_m)}{h} .$$

The vector of values of all partial derivatives of a function f at a point $\mathbf{a}$ is called the *gradient* of f at $\mathbf{a}$ and is denoted $\nabla f(\mathbf{a})$.

$$\nabla f(\mathbf{a}) := \left(\tfrac{\partial f}{\partial x_1}(\mathbf{a}), \tfrac{\partial f}{\partial x_2}(\mathbf{a}), \ldots, \tfrac{\partial f}{\partial x_m}(\mathbf{a})\right) .$$

A function $f : U \to \mathbb{R}$, $U \subseteq \mathbb{R}^m$, is *differentable* at $\mathbf{a} \in U$ if there exists a linear mapping $L : \mathbb{R}^m \to \mathbb{R}$, such that

$$\lim_{\mathbf{h} \to \mathbf{o}} \frac{f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - L(\mathbf{h})}{\|\mathbf{h}\|} = 0 .$$

This mapping L is called *(total) differential (or total derivative)* of f at $\mathbf{a}$ and is denoted by $\mathrm{D}f(\mathbf{a})$.

More generally, a mapping $f : U \to \mathbb{R}^n$is *differentiable* at $\mathbf{a} \in U$, if there exists a linear mapping $L : \mathbb{R}^m \to \mathbb{R}^n$ satisfying

$$\lim_{\mathbf{h} \to \mathbf{o}} \frac{\|f(\mathbf{a} + \mathbf{h}) - f(\mathbf{a}) - L(\mathbf{h})\|}{\|\mathbf{h}\|} = 0$$

(note that norm in the norm in the denominator is in $\mathbb{R}^m$ and the norm in the numerator in $\mathbb{R}^n$). Again, we call L *differential* and denote it by $\mathrm{D}f(\mathbf{a})$. An important difference between directional and partial derivatives, which are simply real numbers, and the differential is, that the differential is a more complex object — a linear mapping.

Differentiability is a stronger property than existence of directional and partial derivatives. (Moreover, existence of all partial/directional derivatives at a point does not even imply continuity!)

One can calculate partial derivative with respect to x_i using the same methods as computing derivatives of functions of single variable — by treating all the variables except x_i as constants.

Theorem 31 (Properties of differential). *Let $f = (f_1, f_2, \ldots, f_n) : U \to \mathbb{R}^n$ be a mapping and $U \subset \mathbb{R}^m$ a neighborhood of $\mathbf{a}$.*

1. *Differential of f at $\mathbf{a}$ is unique (if it exists).*
2. *A mapping f is differentiable at $\mathbf{a}$, if and only if each coordinate function f_i is differentiable at $\mathbf{a}$.*
3. *If f is differentiable at $\mathbf{a}$, then f is continuous at $\mathbf{a}$.*

Theorem 32 (Differential $\Rightarrow \partial$). *Let $U \subset \mathbb{R}^m$ be a neighborhood of a point $\mathbf{a}$ and $f : U \to \mathbb{R}$ a function differentiable at $\mathbf{a}$. Then f has all partial derivatives at $\mathbf{a}$ and their values determine the differential:*

$$\begin{aligned} \mathrm{D}f(\mathbf{a})(\mathbf{h}) &= \frac{\partial f}{\partial x_1}(\mathbf{a}) \cdot h_1 + \frac{\partial f}{\partial x_2}(\mathbf{a}) \cdot h_2 + \cdots + \frac{\partial f}{\partial x_m}(\mathbf{a}) \cdot h_m \\ &= \langle \nabla f(\mathbf{a}), \mathbf{h} \rangle \end{aligned}$$

(i.e., value of the differential at $\mathbf{h}$ is a scalar product of $\mathbf{h}$ and a gradient of f at $\mathbf{a}$). Moreover, f then also has all directional derivatives at $\mathbf{a}$ and $\mathrm{D}_{\mathbf{v}} f(\mathbf{a}) = \mathrm{D}f(\mathbf{a})(\mathbf{v})$.

Proof. (Will be added.) □

The differential of a mapping $f : U \to \mathbb{R}^n$, a mapping $L = \mathrm{D}f(\mathbf{a}) : \mathbb{R}^m \to \mathbb{R}^n$, can be described by an $n \times m$ matrix, where $L(\mathbf{h})$ is the result of multiplication of $\mathbf{h}$ by the matrix:

$$L(\mathbf{h}) = \begin{pmatrix} L(\mathbf{h})_1 \\ L(\mathbf{h})_2 \\ \vdots \\ L(\mathbf{h})_n \end{pmatrix} = \begin{pmatrix} l_{1,1} & l_{1,2} & \dots & l_{1,m} \\ l_{2,1} & l_{2,2} & \dots & l_{2,m} \\ \vdots & \vdots & \dots & \vdots \\ l_{n,1} & l_{n,2} & \dots & l_{n,m} \end{pmatrix} \begin{pmatrix} h_1 \\ h_2 \\ \vdots \\ h_m \end{pmatrix} .$$

where i-th row of this matrix is a gradient of the coordinate function f_i at a point $\mathbf{a}$:

$$l_{i,j} = \frac{\partial f_i}{\partial x_j}(\mathbf{a}) \ .$$

Corollary 33 (Jacobi matrix). *Differential of a mapping $f : U \to \mathbb{R}^n$ at a point $\mathbf{a}$, where $U \subset \mathbb{R}^m$ is a neighborhood of $\mathbf{a}$ and f has coordinate functions $f = (f_1, f_2, \dots, f_n)$, is determined by* Jacobi matrix *if the mapping f at a point $\mathbf{a}$:*

$$\left(\frac{\partial f_i}{\partial x_j}(\mathbf{a}) \right)_{i,j=1}^{n,m} = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{a}) & \frac{\partial f_1}{\partial x_2}(\mathbf{a}) & \dots & \frac{\partial f_1}{\partial x_m}(\mathbf{a}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{a}) & \frac{\partial f_2}{\partial x_2}(\mathbf{a}) & \dots & \frac{\partial f_2}{\partial x_m}(\mathbf{a}) \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial f_n}{\partial x_1}(\mathbf{a}) & \frac{\partial f_n}{\partial x_2}(\mathbf{a}) & \dots & \frac{\partial f_n}{\partial x_m}(\mathbf{a}) \end{pmatrix} .$$

If the Jacobi matrix is a square matrix, its determinant is called *jacobian.*

Theorem 34 ($\partial \Rightarrow$ differential). *Let $U \subset \mathbb{R}^m$ is a neighborhood of a point $\mathbf{a} \in \mathbb{R}^m$. If a function $f : U \to \mathbb{R}$ has all partial derivatives on U and they are continuous at $\mathbf{a}$, then f is differentiable at $\mathbf{a}$.*

Geometry of partial derivatives and differentials.

We now generalize the notion of tangent line to a graph of a function of one variable to a tangent (hyper-)plane to a graph of a function of several variables. For simplicity, we consider only tangent planes for functions of two variables, general tangent hyperplanes are defined in an analogous way (but are hard to imagine).

Let $(x_0, y_0) \in U \subset \mathbb{R}^2$, where U is an open set in a plane, and $f : U \to \mathbb{R}$ is a function. Its graph

$$G_f = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in U, z = f(x, y)\}$$

is a surface in three dimensional Euclidean space. On G_f, there exists a point (x_0, y_0, z_0), such that $z_0 = f(x_0, y_0)$. Assume that f is differentiable at (x_0, y_0). Then, there exists a unique linear functions of two variables $L(x, y)$ (i.e. $L(x, y) = \alpha + \beta x + \gamma y$), such that a graph of $L(x, y)$ contains (x_0, y_0, z_0), and it satisfies

$$\lim_{(x,y)\to(x_0,y_0)} \frac{f(x, y) - L(x, y)}{d((x, y), (x_0, y_0))} = 0.$$

Specifically, it is a function

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

It follows from the uniqueness of a diffenential, because $T(x, y) = z_0 + \mathrm{D}f(x_0, y_0)(x - x_0, y - y_0)$. Graph of $T(x, y)$

$$G_T = \{(x, y, z) \in \mathbb{R}^3 \mid (x, y) \in \mathbb{R}^2, z = T(x, y)\}$$

is called the *tangent plane to the graph of f at* (x_0, y_0, z_0).

Equation of the tangent plane $z = T(x, y)$ can be rewritten in the form

$$\begin{aligned} \frac{\partial f}{\partial x}(x_0, y_0) \cdot (x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0) \cdot (y - y_0) - (z - z_0) &= 0 \ , \\ \text{alternatively} \quad \langle \mathbf{n}, (x - x_0, y - y_0, z - z_0) \rangle &= 0 \ , \end{aligned}$$

where $\mathbf{n} \in \mathbb{R}^3$ je vektor

$$\mathbf{n} = \left(\frac{\partial f}{\partial x}(x_0, y_0), \frac{\partial f}{\partial y}(x_0, y_0), -1 \right) .$$

Denoting $\mathbf{x} = (x, y, z)$ and $\mathbf{x_0} = (x_0, y_0, z_0)$, we can express G_T as

$$G_T = \{\mathbf{x} \in \mathbb{R}^3 \mid \langle \mathbf{n}, \mathbf{x} - \mathbf{x_0} \rangle = 0\} \ .$$

That is, the tangent plane consists of all points whose direction from $\mathbf{x_0}$ is perpendicular to $\mathbf{n}$. Vector $\mathbf{n}$ is called a *normal vector* of the graph of f at $\mathbf{x_0}$.