Diferencovatelnost reálných funkcí

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

ROBUST - 12.-16.září 2016, Kurzovní

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Definition

Let $\mathrm{D} \subset \mathbb{R}$, $\mathrm{D} \neq \emptyset$, $f : \mathrm{D} \to \mathbb{R}$ and $x \in \operatorname{int}(\mathrm{D})$. We say, f is differentiable at x (cz. diferencovatelná v bodě x) if there is $f'(x) \in \mathbb{R}$ such that for all $y \in \mathrm{D}$ we have

$$f(y) = f(x) + f'(x)(y - x) + |y - x| R_1(y - x; f, x),$$

where $\lim_{h \to \mathbf{0}} R_1(h; f, x) = 0$.
Equivalently, f is differentiable at x iff $\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = f'(x) \in \mathbb{R}$.
If $\mathrm{S} \subset \operatorname{int}(\mathrm{D})$, then we say f is differentiable at S (cz. diferencovatelná v množině S), if it is differentiable at each point $x \in \mathrm{S}$.

Jeden argument

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Lemma

If $\mathrm{D} \subset \mathbb{R}$, $\mathrm{D} \neq \emptyset$ *and* $f : \mathrm{D} \to \mathbb{R}$ *is differentiable at* $x \in \mathrm{int}(\mathrm{D})$ *then* f *is continuous at* x.

Lemma

Let $a, b \in \mathbb{R}$, $a < b$, $f : [a, b] \to \mathbb{R}$ *be differentiable at* (a, b), *right-continuous at* a *and left-continuous at* b. *Then,*

$$\int_a^b f'(s)\,\mathrm{d}s = f(b) - f(a).$$

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Více argumentů - 0

Definition

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$, $f : \mathrm{D} \to \mathbb{R}$, $x \in \mathrm{int}(\mathrm{D})$ and $h \in \mathbb{R}^n$. We say,
f is differentiable at x in direction h (cz. diferencovatelná v bodě x ve směru h) if there is $f'(x; h) \in \mathbb{R}$ such that for all $t \in \mathbb{R}$, $x + th \in \mathrm{D}$ we have

$$f(x + th) = f(x) + f'(x; h)\, t + |t|\, R_1(t; f, x, h),$$

where $\lim_{s \to 0} R_1(s; f, x, h) = 0$.
Equivalently, f is differentiable at x in direction h iff $\lim_{t \to 0} \frac{f(x+th) - f(x)}{t} = f'(x; h) \in \mathbb{R}$.

Více argumentů - 1

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Definition

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$, $f : \mathrm{D} \to \mathbb{R}$, $x \in \mathrm{int}(\mathrm{D})$. For $i \in \{1, 2, \ldots, n\}$, we say f possesses a partial derivative at x w.r.t. x_i (cz. parciální derivace v bodě x vzhledem k x_i) if f is differentiable at x in direction $\mathrm{e}_{i:n}$ and we set

$$\frac{\partial f}{\partial x_i}(x) = f'(x; \mathrm{e}_{i:n}).$$

If f possesses a partial derivative at x w.r.t. x_i for all $i \in \{1, 2, \ldots, n\}$ we say f possesses a gradient at x (cz. gradient v bodě x) and we denote

$$\nabla f(x) = \left(\frac{\partial f}{\partial x_i}(x)\right)_{i=1}^{n}.$$

Více argumentů - 2

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Definition

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$, $f : \mathrm{D} \to \mathbb{R}$ and $x \in \operatorname{int}(\mathrm{D})$. We say, f is differentiable at x (or, possesses total differential at x, Fréchet differentiable at x) (cz. diferencovatelná v bodě x) if f possesses a gradient $\nabla f(x) \in \mathbb{R}^n$ and for all $y \in \mathrm{D}$ we have

$$f(y) = f(x) + \langle \nabla f(x), y - x \rangle + \|y - x\| R_1(y - x; f, x),$$

where $\lim_{h \to \mathbf{0}} R_1(h; f, x) = 0$.
If $\mathrm{S} \subset \operatorname{int}(\mathrm{D})$, then we say f is differentiable at S (cz. diferencovatelná v množině S), if it is differentiable at each point $x \in \mathrm{S}$.

Více argumentů - 3

Definition

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$, $f : \mathrm{D} \to \mathbb{R}$ and $x \in \operatorname{int}(\mathrm{D})$.
We say, f is continuously differentiable at x (cz. spojitě diferencovatelná v bodě x), if there is $\delta > 0$ such that $\mathcal{U}(x,\delta) \subset \mathrm{D}$, f is differentiable at $\mathcal{U}(x,\delta)$ and gradient ∇f is continuous at x.
We say, f is continuously differentiable at a neighborhood of x (cz. spojitě diferencovatelná v okolí bodu x), if there is $\delta > 0$ such that $\mathcal{U}(x,\delta) \subset \mathrm{D}$, f is differentiable at $\mathcal{U}(x,\delta)$ and gradient ∇f is continuous at $\mathcal{U}(x,\delta)$.

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Více argumentů - 4

Gradient is necessary for expansion (1).

Lemma

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$, $f : \mathrm{D} \to \mathbb{R}$ *and* $x \in \operatorname{int}(\mathrm{D})$. *Let* f *fulfills an expansion for all* $y \in \mathrm{D}$

$$f(y) = f(x) + \langle \xi, y - x \rangle + \|y - x\| \, R_1(y - x; f, x),$$

where $\xi \in \mathbb{R}^n$ *and* $\lim_{h \to \mathbf{0}} R_1(h; f, x) = 0$.
Then f *is differentiable at* x, $\xi = \nabla f(x)$ *and*
$f'(x; h) = \langle \nabla f(x), h \rangle$ *for all directions* $h \in \mathbb{R}^n$.

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Více argumentů - 4hint

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Proof.
Using (1) for a direction $h \in \mathbb{R}^n$ and $t \in \mathbb{R}$ small enough, we have

$$f(x+th) = f(x) + \langle \xi, th \rangle + \|th\| R_1(th; f, x),$$

where $\lim_{h \to \mathbf{0}} R_1(h; f, x) = 0$.
Consider derivative ratio and let $t \to 0$:

$$\lim_{t \to 0} \frac{f(x+th) - f(x)}{t} = \langle \xi, h \rangle + \|h\| \lim_{t \to 0} \frac{|t|}{t} R_1(th; f, x) = \langle \xi, h \rangle .$$

Setting $h = \mathrm{e}_{i:n}$, we receive $\xi_i = \frac{\partial f}{\partial x_i}(x)$.
We have verified ξ is the gradient of f at x, f is differentiable at x and directional derivatives possess announced form. □

Více argumentů - 5

Lemma

If $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$ *and* $f : \mathrm{D} \to \mathbb{R}$ *is differentiable at* $x \in \operatorname{int}(\mathrm{D})$ *then* f *is continuous at* x*.*

Proof.

Continuity of f at x follows immediately (1). □

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Více argumentů - 6

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Lemma

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$ *and* $f : \mathrm{D} \to \mathbb{R}$. *Consider* $x \in \mathrm{D}$ *and* $h \in \mathbb{R}^n$ *such that* $x + th \in \mathrm{D}$ *for all* $0 \leq t \leq 1$. *Define function* $\varphi : [0,1] \to \mathbb{R} : t \in [0,1] \to f(x+th)$.

(i) *If* $0 < t < 1$, $x + th \in \mathrm{int}(\mathrm{D})$ *and* f *is differentiable at* $x + th$ *then* φ *is differentiable at* t *and* $\varphi'(t) = \langle \nabla_x f(x+th), h \rangle$.

(ii) *If* $x + th \in \mathrm{int}(\mathrm{D})$ *and* f *is differentiable at* $x + th$ *for all* $0 < t < 1$, φ *is continuous at* 0 *from right and* φ *is continuous at* 1 *from left then*

$$f(x+h) - f(x) = \varphi(1) - \varphi(0) = \int_0^1 \langle \nabla_x f(x+th), h \rangle \, \mathrm{d}t.$$

Vektorové funkce - 0

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Start with a curve.

Definition

Let $\mathrm{D} \subset \mathbb{R}$, $\mathrm{D} \neq \emptyset$, $f : \mathrm{D} \to \mathbb{R}^m$ and $t \in \operatorname{int}(\mathrm{D})$. Express the function as a vector of functions $f = (f_1, f_2, \ldots, f_m)^\top$. We say,

- f is differentiable at t if f_j is differentiable at t for each $j \in \{1, 2, \ldots, m\}$. We denote the derivative by $f'(t) = (f_1'(t), f_2'(t), \ldots, f_m'(t))^\top$.
- If $\mathrm{S} \subset \operatorname{int}(\mathrm{D})$, f is differentiable at S if f_j is differentiable at S for each $j \in \{1, 2, \ldots, m\}$.

Vektorové funkce - 1

And now a general case. We start with a notion of multidimensional scalar product.

Definition

Let $n, m \in \mathbb{N}$, $x \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times m}$. We define denotation

$$\lessdot A, x \gtrdot \;=\; \left(\langle A_{\cdot,1}, x \rangle, \langle A_{\cdot,2}, x \rangle, \ldots, \langle A_{\cdot,m}, x \rangle\right)^\top .$$

Using matrix notation, we can write $\lessdot A, x \gtrdot = A^\top x$.

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Vektorové funkce - 2

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Definition

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$, $n \geq 2$, $f : \mathrm{D} \to \mathbb{R}^m$ and $x \in \operatorname{int}(\mathrm{D})$. Express the function as a vector of functions $f = (f_1, f_2, \dots, f_m)^\top$. We say,

- f possesses a gradient at x if f_j possesses a gradient at x for each $j \in \{1, 2, \dots, m\}$. We denote $\nabla f(x) = (\nabla f_1(x), \nabla f_2(x), \dots, \nabla f_m(x))$.
- f is differentiable at x if f_j is differentiable at x for each $j \in \{1, 2, \dots, m\}$.
- If $\mathrm{S} \subset \operatorname{int}(\mathrm{D})$, f is differentiable at S if f_j is differentiable at S for each $j \in \{1, 2, \dots, m\}$.

Vektorové funkce - 3

Lemma

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$, $f : \mathrm{D} \to \mathbb{R}^m$ *and* $x \in \operatorname{int}(\mathrm{D})$. *Then,* f *is differentiable at* x *if and only if* f *possesses a gradient* $\nabla f(x) \in \mathbb{R}^{n \times m}$ *and for all* $y \in \mathrm{D}$ *we have*

$$f(y) = f(x) + \lessdot \nabla f(x), y - x \gtrdot + \|y - x\| R_1(y - x; f, x),$$

where $R_1(\cdot; f, x) : \mathbb{R}^n \to \mathbb{R}^m$ *and* $\lim_{h \to \mathbf{0}} R_1(h; f, x) = 0$.
The expression is more simple for $n = 1$. *Let* $\mathrm{D} \subset \mathbb{R}$, $\mathrm{D} \neq \emptyset$, $f : \mathrm{D} \to \mathbb{R}^m$ *and* $t \in \operatorname{int}(\mathrm{D})$. *Then,* f *is differentiable at* t *if and only if* f *possesses a derivative* $f'(t) \in \mathbb{R}^m$ *and for all* $s \in \mathrm{D}$ *we have*

$$f(s) = f(t) + (s - t)f'(t) + |s - t| R_1(s - t; f, t),$$

where $R_1(\cdot; f, x) : \mathbb{R} \to \mathbb{R}^m$ *and* $\lim_{h \to \mathbf{0}} R_1(h; f, x) = 0$.

Proof.

It is a straightforward rewriting of definition. □

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Řetízkové pravidlo - 0

Differentiability directly implies chain rule (cz. řetízkové pravidlo).

Lemma

Let $I \subset \mathbb{R}$, $\operatorname{int}(I) \neq \emptyset$, $\mathrm{D} \subset \mathbb{R}^n$, $\operatorname{int}(\mathrm{D}) \neq \emptyset$, $g : I \to \mathrm{D}$, $f : \mathrm{D} \to \mathbb{R}$ *and* $t \in \operatorname{int}(I)$ *such that* $g(t) \in \operatorname{int}(\mathrm{D})$. *If* f *is differentiable at* $g(t)$ *and* g *is differentiable at* t, *then* $f \circ g$ *is differentiable at* t *and*

$$(f \circ g)'(t) \;=\; \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(g(t))\, g_i'(t) = \langle \nabla f(g(t)), g'(t) \rangle .$$

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Druhá derivace - 0

Also, notion of the second derivative must be explained.

Definition

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$, $f : \mathrm{D} \to \mathbb{R}$ and $x \in \mathrm{int}\,(\mathrm{D})$. We say, f possesses the second partial derivatives at x (cz. má druhé parciální derivace v x), if f possesses a gradient on a neighborhood of x and all partial derivatives of gradient at x exists; i.e. $\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(x)$ exists for all indexes $i, j \in \{1, 2, \ldots, n\}$.

Then, we denote $\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(x)$ for all $i, j \in \{1, 2, \ldots, n\}$. Matrix of the second partial derivatives is denoted $\nabla^2 f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}(x)\right)_{i=1,j=1}^{n,n}$ and called Hessian matrix.

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Druhá derivace - 1

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Definition

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$, $f : \mathrm{D} \to \mathbb{R}$ and $x \in \mathrm{int}(\mathrm{D})$. We say, f is twice differentiable at x (or, Second Peano derivative) (cz. dvakrát diferencovatelná v x), if there is a gradient $\nabla f(x) \in \mathbb{R}^n$ and a symmetric matrix $H_f(x) \in \mathbb{R}^{n \times n}$ such that for all $y \in \mathrm{D}$ we have

$$\begin{aligned} f(y) &= f(x) + \langle \nabla f(x), y - x \rangle + \frac{1}{2} \langle y - x, H_f(x)(y - x) \rangle \\ &\quad + \|y - x\|^2 R_2(y - x; f, x), \end{aligned}$$

where $\lim_{h \to \mathbf{0}} R_2(h; f, x) = 0$.
If $\mathrm{S} \subset \mathrm{int}(\mathrm{D})$, then we say f is twice differentiable at S (cz. dvakrát diferencovatelná v množině S), if it is twice differentiable at each $x \in \mathrm{S}$.

Druhá derivace - 2

Matrix $H_f(x)$ can differ from Hessian matrix. The reasons are

- ∇f does not exist in any neighborhood of x,
- ∇f exists in a neighborhood of x and $\nabla^2 f(x)$ does not exist.
- ∇f exists in a neighborhood of x, $\nabla^2 f(x)$ exist, but, asymmetric.

Let us note the difference from Hessian is not mentioned in [1].

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Druhá derivace - 3

Lemma

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$, $f : \mathrm{D} \to \mathbb{R}$ *and* $x \in \operatorname{int}(\mathrm{D})$. *If* f *is twice differentiable at* x *then matrix* $H_f(x)$ *is uniquely determined.*

Proof.

Since $H_f(x)$ is symmetric, its uniqueness follows an observation on quadratic forms from linear algebra. □

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Druhá derivace - 4

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Lemma

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$, $f : \mathrm{D} \to \mathbb{R}$ *and* $x \in \operatorname{int}(\mathrm{D})$. *If* f *is differentiable at a neighborhood of* x *and* ∇f *is differentiable at* x*, then,* $\nabla^2 f(x)$ *exists and* f *is twice differentiable at* x *with*

$$H_f(x) = \frac{1}{2}\nabla^2 f(x) + \frac{1}{2}\left(\nabla^2 f(x)\right)^{\top}.$$

If, moreover, Hessian matrix is symmetric, i.e. $\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$ *for all* $i, j \in \{1, 2, \ldots, n\}$, *then*

$$H_f(x) = \nabla^2 f(x).$$

Druhá derivace - 4hint

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Proof.
According to our assumptions, there is $\delta > 0$ such that $\mathcal{U}(x, \delta) \subset \mathrm{D}$ and for all $y \in \mathcal{U}(x, \delta)$, $h \in \mathbb{R}^n$, $\|h\| < \delta - \|y - x\|$ we have

$$f(y+h) - f(y) = \langle \nabla f(y), h \rangle + \|h\| R_1(h; f, y),$$
$$\nabla f(y) - \nabla f(x) = \lessdot \left(\nabla^2 f(x)\right)^\top, y - x \gtrdot + \|y - x\| R_1(y - x; \nabla f, x).$$

According to Lemma 10

$$f(x+h) - f(x) - \langle \nabla_x f(x), h \rangle = \int_0^1 \langle \nabla_x f(x+th) - \nabla_x f(x), h \rangle \, \mathrm{d}t.$$

Plugging in expansion of gradient, we are receiving the statement. □

Druhá derivace - 5

Lemma

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$, $f : \mathrm{D} \to \mathbb{R}$, $x \in \operatorname{int}(\mathrm{D})$ *and* $h \in \mathbb{R}^n$.

(i) *If* f *is twice differentiable at* x, *then*

$$\lim_{t \to 0} \frac{f(x+th) - f(x) - t\langle \nabla f(x), h \rangle}{t^2} = \frac{1}{2} \langle h, H_f(x)\, h \rangle .$$

(ii) *Let us denote* $\mathrm{D}_h = \{t \in \mathbb{R} : x + th \in \mathrm{D}\}$. *If* f *is differentiable at a neighborhood of* x *and* ∇f *is differentiable at* x, *then,* $\nabla^2 f(x)$ *exists and function* $\varphi : \mathrm{D}_h \to \mathbb{R} : t \in \mathrm{D}_h \to f(x+th)$ *possesses derivatives*

$$\begin{aligned} \varphi'(t) &= \langle \nabla f(x+th), h \rangle \quad \textit{for all } t \textit{ small enough}, \\ \varphi''(0) &= \langle h, \nabla^2 f(x)\, h \rangle . \end{aligned}$$

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Existence and continuity of gradient, resp. of Hessian, are sufficient conditions for differentiability in the sense of Definitions 6 and 17.

Lemma

Let $I \subset \mathbb{R}$, $\mathrm{int}\,(I) \neq \emptyset$, $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$, $g : I \to \mathrm{D}$, $f : \mathrm{D} \to \mathbb{R}$ *and* $t \in \mathrm{int}\,(I)$ *such that* $g(t) \in \mathrm{int}\,(\mathrm{D})$. *If gradient of* f *exists on a neighborhood of* $g(t)$ *and is continuous at* $g(t)$ *and* g *is differentiable at* t, *then* $f \circ g$ *is differentiable at* t *with*

$$(f \circ g)'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(g(t))\, g_i'(t) = \langle \nabla f(g(t)), g'(t) \rangle .$$

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Using Lemma 21, we derive differentiability of a function.

Lemma

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$, $f : \mathrm{D} \to \mathbb{R}$ *and* $x \in \mathrm{int}\,(\mathrm{D})$. *If gradient of* f *exists on a neighborhood of* x *and is continuous at* x*, then* f *is differentiable at* x *with*

$$f\,(x+h) = f\,(x) + \langle\, \nabla f\,(x)\,, h\,\rangle + \|h\|\, R_1\,(h; f, x)\,,$$
$$|R_1\,(h; f, x)\,| \leq \max\left\{\|\nabla f\,(x+uh) - \nabla f\,(x)\| \,:\, 0 \leq u \leq 1\right\}$$

if h *is sufficiently small.*

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Lemma

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$, $f : \mathrm{D} \to \mathbb{R}$ *and* $x \in \operatorname{int}(\mathrm{D})$. *Then,* f *is continuously differentiable at a neighborhood of* x *if and only if there is* $\delta > 0$ *such that* ∇f *exists at* $\mathcal{U}(x, \delta)$ *and is continuous at* $\mathcal{U}(x, \delta)$.

Proof.

A consequence of Lemma 22. □

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Lemma

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$, $f : \mathrm{D} \to \mathbb{R}$ *and* $x \in \operatorname{int}(\mathrm{D})$. *If* ∇f, $\nabla^2 f$ *exist on a neighborhood of* x *and* $\nabla^2 f$ *is continuous at* x, *then Hessian* $\nabla^2 f(x)$ *is a symmetric matrix and* f *is twice differentiable at* x *with*

$$\begin{aligned} f(x+h) &= f(x) + \langle \nabla f(x), h \rangle + \frac{1}{2} \langle h, \nabla^2 f(x) h \rangle + \\ &\quad + \frac{1}{2} \|h\|^2 R_2(h; f, x), \end{aligned}$$

$$|R_2(h; f, x)| \leq \max\left\{ \left\| \nabla^2 f(x+uh) - \nabla^2 f(x) \right\| : 0 \leq u \leq 1 \right\}$$

if h *sufficiently small. Moreover,* $H_f(x) = \nabla^2 f(x)$.

Konvexní funkce - 0

Convexity of a function can be verified by means of functions of one variable.

Theorem

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$ *be a convex set and* $f : \mathrm{D} \to \mathbb{R}$. *Then, function* f *is convex if and only if functions* $\varphi_{x,s} : \mathrm{D}_{x,s} \to \mathbb{R}$ *are convex for all* $x \in \mathrm{D}$ *and all* $s \in \mathbb{R}^n$, *where* $\varphi_{x,s}(t) = f(x + ts)$ *and* $\mathrm{D}_{x,s} = \{t : x + ts \in \mathrm{D}, t \in \mathbb{R}\}$. *(Let us recall set* $\mathrm{D}_{x,s}$ *is always an interval.)*

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Konvexní funkce - 1

Lemma

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$ *be a convex open set and* $f : \mathrm{D} \to \mathbb{R}$.

- *If* f *is differentiable at* D *and* $x \in \mathrm{D}$, $s \in \mathbb{R}^n$, $t \in \mathrm{D}_{x,s}$, *we have*

$$\varphi'_{x,s}(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x + ts)\, s_i = \langle \nabla f(x + ts), s \rangle .$$

- *If* f *is twice differentiable at* D *and* $x \in \mathrm{D}$, $s \in \mathbb{R}^n$, $t \in \mathrm{D}_{x,s}$, *we have*

$$\lim_{u \to 0} \frac{\varphi_{x,s}(t+u) - \varphi_{x,s}(t) - u \langle \nabla f(x+ts), s \rangle}{u^2} = \\ = \frac{1}{2} \langle s, H_f(x+ts)\, s \rangle .$$

- *If* f *is differentiable at* D *and* ∇f *is differentiable at* D, *then,* $\nabla^2 f$ *exists on* D *and for* $x \in \mathrm{D}$, $s \in \mathbb{R}^n$, $t \in \mathrm{D}_{x,s}$, *we have*

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Konvexní funkce - 2

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Theorem

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$ *be a convex open set and* $f : \mathrm{D} \to \mathbb{R}$ *be differentiable at* D. *Then,*

f *is convex* $\Leftrightarrow$ $t \in \mathrm{D}_{x,s} \mapsto \langle \nabla f(x+ts), s \rangle$ *is nondecreasing on* $\mathrm{D}_{x,s}$ *for all* $x \in \mathrm{D}$, $s \in \mathbb{R}^n$.

Konvexní funkce - 3

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Theorem

Let $\mathrm{D} \subset \mathbb{R}^n$, $\mathrm{D} \neq \emptyset$ *be a convex open set and* $f : \mathrm{D} \to \mathbb{R}$. *If* f *is differentiable at* D *and* ∇f *is differentiable at* D, *then,* $\nabla^2 f$ *exists on* D, f *is twice differentiable at* D *with*

$$H_f(x) = \frac{1}{2}\nabla^2 f(x) + \frac{1}{2}\left(\nabla^2 f(x)\right)^\top$$

and

$$f \text{ is convex} \quad \Leftrightarrow \quad H_f(x) \text{ is positively semidefinite for all } x \in \mathrm{D}.$$

Konvexní funkce - 4

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Definition

Let $D \subset \mathbb{R}^n$, $D \neq \emptyset$ be a set and $f : D \to \mathbb{R}$ be a function. We say, f possesses at $x \in D$ subgradient $a \in \mathbb{R}^n$ (cz. subgradient), if we have

$$f(y) - f(x) \geq \langle a, y - x \rangle \text{ for all } y \in D.$$

Set of all subgradients at x will be called subdifferential of f at x (cz. subdiferenciál) and will be denoted by $\partial f(x)$.

Konvexní funkce - 5

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Subgradient and subdifferencial are helpful tools for describing local minima of a convex function.

Lemma

Let $\mathcal{G} \subset \mathbb{R}^n$ be a nonempty open convex set, $f : \mathcal{G} \to \mathbb{R}$ be a convex function and $y \in \mathcal{G}$. Hence, the following is equivalent:

1. *f is differentiable at y and $\partial f(y) = \{\nabla f(y)\}$.*
2. *$\partial f(y)$ is an one-point set.*
3. *f possesses a gradient at y.*

Konvexní funkce - 6

Results on separation of convex bodies have consequences for convex function.

Theorem

Let $\mathrm{D} \subset \mathbb{R}^n$ *be a nonempty convex set and* $f : \mathrm{D} \to \mathbb{R}$ *be a convex function. Then,* $\partial f(x) \neq \emptyset$ *for each* $x \in \mathrm{rint}(\mathrm{D})$*.*

Equivalent description of a convex function using non-emptiness of subdifferentials is in power if function definition region is an open set.

Theorem

Let $\mathrm{D} \subset \mathbb{R}^n$ *be an open convex set and* $f : \mathrm{D} \to \mathbb{R}$*. Then,* f *is a convex function if and only if* $\partial f(x) \neq \emptyset$ *for each* $x \in \mathrm{D}$*.*

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Konvexní funkce - 7

For a continuous function, the characterization is also in power.

Theorem

Let $\mathrm{D} \subset \mathbb{R}^n$ *be a convex set and* $f : \mathrm{D} \to \mathbb{R}$ *be a continuous function. Then,* f *is a convex function if and only if* $\partial f(x) \neq \emptyset$ *for each* $x \in \operatorname{rint}(\mathrm{D})$*.*

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Kužele - 0

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Definition

Let $K \subset \mathbb{R}^n$ be a cone. We define polar of K (cz. polára K)

$$K^o = \{v \in \mathbb{R}^n : \forall\, x \in K \text{ we have } \langle v, x \rangle \le 0\}.$$

and bipolar of K (cz. bipolára K)

$$K^{oo} = K^{oo} = \{w \in \mathbb{R}^n : \forall\, v \in K^o \text{ we have } \langle w, v \rangle \le 0\}.$$

Kužele - 1

Basic properties of polar.

Lemma

If $K \subset \mathbb{R}^n$ is cone, then K^o is a closed convex cone and $K^{oo} = \mathrm{clo}\left(\mathrm{conv}\left(K\right)\right)$.

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Kužele - 2

Definition

Let $\mathrm{M} \subset \mathbb{R}^n$, $\widetilde{x} \in \operatorname{clo}(\mathrm{M})$. We define Tangent Cone to M at $\widetilde{x}$ (or, Cone of Tangents) (cz. tečný kužel k množině M v bodě $\widetilde{x}$) by

$$T_{\mathrm{M}}(\widetilde{x}) = \left\{ s \in \mathbb{R}^n : \begin{array}{c} \exists\ x_k \in \mathrm{M}, \lambda_k > 0\ k \in \mathbb{N} \\ \text{s.t. } x_k \to \widetilde{x},\ \lambda_k (x_k - \widetilde{x}) \to s. \end{array} \right\}.$$

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Kužele - 3

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Lemma

If $\mathsf{M} \subset \mathbb{R}^n$, $\widetilde{x} \in \operatorname{clo}(\mathsf{M})$, *then* $T_{\mathsf{M}}(\widetilde{x})$ *is a closed cone.*

Lemma

If $\mathsf{M} \subset \mathbb{R}^n$ *is a convex set and* $\widetilde{x} \in \operatorname{clo}(\mathsf{M})$, *then* $T_{\mathsf{M}}(\widetilde{x})$ *is a closed convex cone.*

Lemma

Let $\mathsf{M} \subset \mathbb{R}^n$, $x \in \operatorname{clo}(\mathsf{M})$ *and* $\mathsf{S} \subset \mathbb{R}^n$, $x \in \operatorname{int}(\mathsf{S})$. *Then,* $T_{\mathsf{M}\cap\mathsf{S}}(x) = T_{\operatorname{clo}(\mathsf{M})\cap\operatorname{clo}(\mathsf{S})}(x) = T_{\mathsf{M}}(x) = T_{\operatorname{clo}(\mathsf{M})}(x)$.

Kužele - 4

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Definition

Let $\mathsf{S} \subset \mathbb{R}^n$, $\widetilde{x} \in \operatorname{clo}(\mathsf{S})$. We say, that $s \in \mathbb{R}^n$ is a Regular Normal to S at $\widetilde{x}$ (or, Normal to S at $\widetilde{x}$ in the Regular Sense), (cz. regulární normála k množině S v $\widetilde{x}$) if

$$\forall x \in \mathsf{S} \text{ we have } \langle s, x - \widetilde{x} \rangle \leq \|x - \widetilde{x}\| \, R\left(x - \widetilde{x}; s, \widetilde{x}\right),$$

where $R\left(x - \widetilde{x}; s, \widetilde{x}\right) \to 0$ provided $x \to \widetilde{x}$ and $x \in \mathsf{S}$.
Regular Normal cone to S at $\widetilde{x}$ (or, Cone of Regular Normals to S at $\widetilde{x}$) (cz. regulární normálový kužel)
$\widehat{N}_{\mathsf{S}}\left(\widetilde{x}\right)$ is a set of all regular normals to S at $\widetilde{x}$.

Definition

Let $S \subset \mathbb{R}^n$, $\widetilde{x} \in \mathrm{clo}(S)$. We say, that $s \in \mathbb{R}^n$ is
a Normal to S at $\widetilde{x}$ (or, Normal to S at $\widetilde{x}$ in the General Sense; Normal Vector to S at $\widetilde{x}$), (cz. normála k množině S v $\widetilde{x}$) if there are sequences $x_k \in S$, $s_k \in \widehat{N}_S(x_k)$ for $k \in \mathbb{N}$ such that $x_k \to \widetilde{x}$, $s_k \to s$.

Normal cone to S at $\widetilde{x}$ (or, Cone of Normals to S at $\widetilde{x}$), (cz. Normálový kužel k množině S v bodě $\widetilde{x}$)
$N_S(\widetilde{x})$ is the set of all normals to S at $\widetilde{x}$.

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Kužele - 6

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Perceive defined objects are really cones and normal cone always contains regular normal cone.

Lemma

If $\mathsf{S} \subset \mathbb{R}^n$ *and* $\widetilde{x} \in \mathrm{clo}\,(\mathsf{S})$, *then* $\widehat{N}_{\mathsf{S}}\,(\widetilde{x})$, $N_{\mathsf{S}}\,(\widetilde{x})$ *are cones and* $\widehat{N}_{\mathsf{S}}\,(\widetilde{x}) \subset N_{\mathsf{S}}\,(\widetilde{x})$.

Lemma

Let $\mathsf{M} \subset \mathbb{R}^n$, $x \in \mathrm{clo}\,(\mathsf{M})$ *and* $\mathsf{S} \subset \mathbb{R}^n$, $x \in \mathrm{int}\,(\mathsf{S})$. *Then,*
$\widehat{N}_{\mathsf{M}\cap\mathsf{S}}\,(x) = \widehat{N}_{\mathrm{clo}(\mathsf{M})\cap\mathrm{clo}(\mathsf{S})}\,(x) = \widehat{N}_{\mathsf{M}}\,(x) = \widehat{N}_{\mathrm{clo}(\mathsf{M})}\,(x)$
and
$N_{\mathsf{M}\cap\mathsf{S}}\,(x) = N_{\mathrm{clo}(\mathsf{M})\cap\mathrm{clo}(\mathsf{S})}\,(x) = N_{\mathsf{M}}\,(x) = N_{\mathrm{clo}(\mathsf{M})}\,(x)$.

Kužele - 7

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Theorem

If $S \subset \mathbb{R}^n$ *and* $\widetilde{x} \in \mathrm{clo}(S)$, *then* $T_S(\widetilde{x})^o = \widehat{N}_S(\widetilde{x})$, $\widehat{N}_S(\widetilde{x})^o \supset T_S(\widetilde{x})$.

Kužele - 8

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Polar of a normal cone has also certain importance.

Definition

For $\mathsf{S} \subset \mathbb{R}^n$ and $\widetilde{x} \in \mathrm{clo}(\mathsf{S})$ we define Regular Tangent cone to S at $\widetilde{x}$ (or, Cone of Regular Tangent Vectors of S at $\widetilde{x}$) (cz. regulární tečný kužel k množině S v bodě $\widetilde{x}$) by

$$\widehat{T}_{\mathsf{S}}(\widetilde{x}) = \left\{ s \in \mathbb{R}^n : \begin{array}{l} \text{for each } x_k \in \mathsf{S},\ k \in \mathbb{N},\ x_k \to \widetilde{x}, \\ \text{for each } \lambda_k > 0,\ k \in \mathbb{N},\ \lambda_k \nearrow +\infty, \\ \text{there is } \xi_k \in \mathsf{S},\ k \in \mathbb{N}, \\ \text{such that } \xi_k \to \widetilde{x},\ \lambda_k (\xi_k - x_k) \to s. \end{array} \right\}.$$

Kužele - 9

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

At first, consider basic properties of a regular tangent cone.

Theorem

If $\mathsf{S} \subset \mathbb{R}^n$ *and* $\widetilde{x} \in \mathrm{clo}(\mathsf{S})$, *then* $\widehat{T}_{\mathsf{S}}(\widetilde{x})$ *is a closed convex cone.*

Lemma

Let $\mathsf{M} \subset \mathbb{R}^n$, $x \in \mathrm{clo}(\mathsf{M})$ *and* $\mathsf{S} \subset \mathbb{R}^n$, $x \in \mathrm{int}(\mathsf{S})$. *Then,*
$\widehat{T}_{\mathsf{M}\cap\mathsf{S}}(x) = \widehat{T}_{\mathrm{clo}(\mathsf{M})\cap\mathrm{clo}(\mathsf{S})}(x) = \widehat{T}_{\mathsf{M}}(x) = \widehat{T}_{\mathrm{clo}(\mathsf{M})}(x)$.

Theorem

If $\mathsf{S} \subset \mathbb{R}^n$ *and* $\widetilde{x} \in \mathrm{clo}(\mathsf{S})$, *then* $\widehat{T}_{\mathsf{S}}(\widetilde{x}) \subset T_{\mathsf{S}}(\widetilde{x})$.

Kužele - lokální uzavřenost

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Definition

Let $S \subset \mathbb{R}^n$, $\widetilde{x} \in \mathrm{clo}(S)$. We say, that set S is locally closed at $\widetilde{x}$ (cz. lokálně uzavřená v $\widetilde{x}$), if there is $\delta > 0$ such that $\mathcal{V}(\widetilde{x}, \delta) \cap S$ is a closed set.

Kužele - 10

Theorem

Let $\mathsf{S} \subset \mathbb{R}^n$ *and* $\widetilde{x} \in \operatorname{clo}(\mathsf{S})$. *If* S *is locally closed at* $\widetilde{x}$, *then*

$$\widehat{T}_{\mathsf{S}}(\widetilde{x}) = \left\{ s \in \mathbb{R}^n : \begin{array}{l} \text{For each } x_k \in \mathsf{S},\ k \in \mathbb{N},\ x_k \to \widetilde{x}, \\ \text{there are } s_k \in T_{\mathsf{S}}(x_k),\ k \in \mathbb{N} \\ \text{such that } s_k \to s. \end{array} \right\}$$

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Kužele - 11

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Theorem

Let $\mathsf{S} \subset \mathbb{R}^n$ *a* $\widetilde{x} \in \operatorname{clo}(\mathsf{S})$. *If* S *is locally closed at* $\widetilde{x}$, *then* $\widehat{T}_{\mathsf{S}}(\widetilde{x}) = N_{\mathsf{S}}(\widetilde{x})^{o}$, $\widehat{T}_{\mathsf{S}}(\widetilde{x})^{o} \supset N_{\mathsf{S}}(\widetilde{x})$.

Kužele - 12

Definition

Let $S \subset \mathbb{R}^n$, $\widetilde{x} \in S$. We say, S is regular at $\widetilde{x}$ in the Sense of Clarke, (cz. regulární ve smyslu Clarka), if S is locally closed at $\widetilde{x}$ and $N_S(\widetilde{x}) = \widehat{N}_S(\widetilde{x})$.

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Kužele - 13

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz

Lemma

Let $\mathsf{S} \subset \mathbb{R}^n$ *be convex,* $\widetilde{x} \in \mathsf{S}$*. Then,*

$$\begin{aligned} T_{\mathsf{S}}(\widetilde{x}) &= \operatorname{clo}\left(\{s \in \mathbb{R}^n : \exists \lambda > 0 \text{ such that } \widetilde{x} + \lambda s \in \mathsf{S}\}\right), \\ \operatorname{int}\left(T_{\mathsf{S}}(\widetilde{x})\right) &= \{s \in \mathbb{R}^n : \exists \lambda > 0 \text{ such that } \widetilde{x} + \lambda s \in \operatorname{int}(\mathsf{S})\}, \\ N_{\mathsf{S}}(\widetilde{x}) &= \widehat{N}_{\mathsf{S}}(\widetilde{x}) = \{s \in \mathbb{R}^n : \forall x \in \mathsf{S} \text{ we have } \langle s, x - \widetilde{x} \rangle \le 0\}. \end{aligned}$$

Therefore, convex set S *is regular at* $\widetilde{x}$ *in sense of Clarke if and only if* S *is locally closed at* $\widetilde{x}$*.*

Reference

- Bazara, M.S.; Sherali, H.D.; Shetty, C.M.: *Nonlinear Programming. Theory and Algorithms.* 2nd edition, Wiley, New York, 1993.
- Rockafellar, T.: *Convex Analysis.*, Springer-Verlag, Berlin, 1975.
- Rockafellar, T.; Wets, R. J.-B.: *Variational Analysis.* Springer-Verlag, Berlin, 1998.

Petr Lachout
KPMS MFF UK, Praha
petr.lachout@mff.cuni.cz