Matematika 1	**Izročki: 11. teden**

Optimizacija. 2. del.
Odvodi vektorskih funkcij. Vezani ekstremi v vektorski obliki.

Polona Oblak

Naj bo $F\colon \mathbb{R}^n \to \mathbb{R}^m$, $F(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_m(\mathbf{x}) \end{bmatrix}$ vektorska funkcija n spremenljivk $\mathbf{x} = (x_1, \ldots, x_n)$. Da bomo poudarili, da sta lahko n in m večja od 1, bomo označili vektor spremenljivk $\mathbf{x} = (x_1, \ldots, x_n)$ tudi kot $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$ in vektorsko funkcijo kot $F = \vec{F}$.

Spomnimo se, da je *odvod vektorske funkcije $\vec{F}$ po vektorju spremenljivk $\vec{x}$* je definiran kot

$$\frac{\partial \vec{F}}{\partial \vec{x}} = J_{\vec{F}}(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(\mathbf{x}) & \frac{\partial f_1}{\partial x_2}(\mathbf{x}) & \ldots & \frac{\partial f_1}{\partial x_n}(\mathbf{x}) \\ \frac{\partial f_2}{\partial x_1}(\mathbf{x}) & \frac{\partial f_2}{\partial x_2}(\mathbf{x}) & \ldots & \frac{\partial f_2}{\partial x_n}(\mathbf{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(\mathbf{x}) & \frac{\partial f_m}{\partial x_2}(\mathbf{x}) & \ldots & \frac{\partial f_m}{\partial x_n}(\mathbf{x}). \end{bmatrix} \tag{1}$$

Drugi odvod funkcije $f\colon \mathbb{R}^n \to \mathbb{R}$ (tu $m = 1$) pa kot

$$\frac{\partial^2 f}{\partial \vec{x}^2} = \frac{\partial}{\partial \vec{x}} \left(\frac{\partial f}{\partial \vec{x}} \right)^{\mathsf{T}}.$$

Funkcija $f\colon \mathbb{R}^n \to \mathbb{R}$ je *konveksna* na $\mathcal{D}$, če velja

$$f(t\,\mathbf{x} + (1-t)\,\mathbf{y}) \le t\,f(\mathbf{x}) + (1-t)\,f(\mathbf{y})$$

za vse $\mathbf{x}, \mathbf{y} \in \mathcal{D}$ in za vse $t \in [0,1]$. Funkcija f je *konkavna* na $\mathcal{D}$, če je funkcija $-f$ konveksna na $\mathcal{D}$.

Izrek 1. *Dvakrat zvezno odvedljiva funkcija $f\colon \mathcal{D} \subseteq \mathbb{R}^n \to \mathbb{R}$ je konveksna natanko tedaj, ko je $\frac{\partial^2 f}{\partial \vec{x}^2}$ pozitivno semidefinitna matrika na $\mathcal{D}$, in je konkavna natanko tedaj, ko je $\frac{\partial^2 f}{\partial \vec{x}^2}$ negativno semidefinitna na $\mathcal{D}$.*

Nekaj pravil:

(1) $\frac{\partial \vec{x}}{\partial \vec{x}} = I_n$

(2) Če je $A \in \mathbb{R}^{m\times n}$, potem $\frac{\partial A\vec{x}}{\partial \vec{x}} = A$.

(3) Če je $\vec{a} \in \mathbb{R}^n$, potem $\frac{\partial \vec{a}^\mathsf{T}\vec{x}}{\partial \vec{x}} = \vec{a}^\mathsf{T}$.
(4) Če je $A \in \mathbb{R}^{n\times n}$, potem $\frac{\partial(\vec{x}^\mathsf{T}A\vec{x})}{\partial \vec{x}} = \vec{x}^\mathsf{T}(A + A^\mathsf{T})$.
(5) Če je $A \in \mathbb{R}^{n\times n}$ simetrična matrika, potem velja $\frac{\partial(\vec{x}^\mathsf{T}A\vec{x})}{\partial \vec{x}} = 2\vec{x}^\mathsf{T}A$.
(6) $\frac{\partial||\vec{x}||^2}{\partial \vec{x}} = 2\vec{x}^\mathsf{T}$.
(7) Če $\vec{z} = \vec{z}(\vec{x})$ in $\vec{y} = \vec{y}(\vec{x})$, potem $\frac{\partial(\vec{y}^\mathsf{T}\vec{z})}{\partial \vec{x}} = \vec{z}^\mathsf{T}\frac{\partial \vec{y}}{\partial \vec{x}} + \vec{y}^\mathsf{T}\frac{\partial \vec{z}}{\partial \vec{x}}$.
(8) Če $G\colon \mathcal{D}_G \subseteq \mathbb{R}^m \to \mathbb{R}^n$ in $F\colon \mathcal{D}_F \subseteq \mathbb{R}^n \to \mathbb{R}^p$ in $H = F \circ G$, potem $\frac{\partial \vec{H}}{\partial \vec{x}} = \frac{\partial(\vec{F}\circ\vec{G})}{\partial \vec{x}} = \frac{\partial \vec{F}}{\partial \vec{G}}(\vec{G}(\vec{x})) \cdot \frac{\partial \vec{G}}{\partial \vec{x}}$.

Primer 1. *Naj bo $A \in \mathbb{R}^{m\times n}$, $\operatorname{rank} A = m$, $\vec{b} \in \mathbb{R}^m$. Poiščite tisto rešitev sistema $A\vec{x} = \vec{b}$, ki ima najmanjšo normo. (Rešitev)*

1. Nadaljnje branje

(1) David A. Harville: Matrix Algebra From a Statistician's Perspective, Springer, 1997, Poglavje 15.
(2) J.E. Gentle: Matrix Algebra, Theory, Computations, in Applications in Statistics, Springer, 2017, Poglavje 4.
(3) K.B. Petersen, M.S. Pedersen: The Matrix Cookbook.

(Pazite na to, da so v vseh treh zgoraj omenjenih referencah odvodi definirani kot transponiranke matrike (1).)

2. Domača naloga

ϟ (1) Denimo, da funkcija $f\colon \mathbb{R}^2 \to \mathbb{R}$ doseže maksimalno vrednost na krivulji $x^2+2y^2 = 3$ v točki $(1,1)$. Ali sta $(\mathrm{grad} f)(1,1)$ in $[1,2]^T$ linearno odvisna?

ϟ (2) Naj bosta $\vec{a}, \vec{b} \in \mathbb{R}^n$ in funkciji $f, g\colon \mathbb{R}^n \to \mathbb{R}$ definirani kot $f(\vec{x}) = \vec{x}^\mathsf{T}\vec{a}$ in $g(\vec{x}) = \vec{x}^\mathsf{T}\vec{b}$. Izračunajte $\frac{\partial(f(\vec{x})g(\vec{x}))}{\partial \vec{x}}$.

ϟ (3) Naj bo $A \in \mathbb{R}^{n\times n}$ matrika z lastnostjo $A^\mathsf{T} = -A$. Izračunajte $\frac{\partial}{\partial \vec{x}}\left(\vec{x}^\mathsf{T}A\vec{x} + ||5\vec{x}||^2\right)$.

ϟ (4) Za matriko $A \in \mathbb{R}^{n\times n}$ izračunajte $\frac{\partial||A\vec{x}||^2}{\partial \vec{x}}$.

ϟ (5) Naj bodo $B, C, D \in \mathbb{R}^{n\times n}$ in $\vec{b}, \vec{d} \in \mathbb{R}^n$. Za funkcijo

$$f(\vec{x}) = (B\vec{x} + \vec{b})^\mathsf{T}C(D\vec{x} + \vec{d})$$

izračunajte $\frac{\partial f}{\partial \vec{x}}$ in $\frac{\partial^2 f}{\partial \vec{x}^2}$.

↯ (6) Naj bo $A \in \mathbb{R}^{n\times n}$ simetrična pozitivno definitna matrika ter $\vec{b} \in \mathbb{R}^n$. Poiščite in klasificirajte lokalne ekstreme funkcije $f\colon \mathbb{R}^n \to \mathbb{R}$, definirane s predpisom

$$f(\vec{x}) = \frac{1}{2}\vec{x}^\mathsf{T} A\vec{x} - \vec{x}^\mathsf{T}\vec{b}.$$

(7) Pokažite, da so naslednje funkcije konveksne:
- ↯ (a) $f(x, y) = e^y - \log x$,
- ↯ (b) $g(\vec{x}) = \frac{1}{2}\vec{x}^\mathsf{T} A\vec{x}$, kjer je A pozitivno semidefinitna matrika,
- ↯ (c) $h(\vec{x}) = ||\vec{x}||^2$,
- ↯ (d) $k(\vec{x}) = \vec{a}^\mathsf{T}\vec{x}$, kjer je $\vec{a} \in \mathbb{R}^n$.

↯ (8) Za dane vektorje $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k \in \mathbb{R}^n$ poiščite takšne $\vec{x} \in \mathbb{R}^n$, da bo povprečna kvadratna razdalja med vektorjem $\vec{x}$ in vektorji $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_k$ najmanjša možna.

(9) Naj bosta $\vec{a} \in \mathbb{R}^n$ ter $b \in \mathbb{R}$.
- ↯ (a) Poiščite najmanjšo in največjo vrednost $\vec{a}^\mathsf{T}\vec{x}$ za vse vektorje $\vec{x} \in \mathbb{R}^n$ s predpisano dolžino $||\vec{x}|| = b$.
- ↯ (b) Geometrijsko razložite svojo rešitev iz (9a) v primeru, ko je $n = 3$.

↯↯(10) Za pozitivno semidefinitno matriko $A \in \mathbb{R}^{n\times n}$, simetrično matriko $B \in \mathbb{R}^{n\times n}$ ter neničelni $b \in \mathbb{R}$ poiščite najmanjšo in največjo vrednost kvadratne forme $\vec{x}^\mathsf{T} A\vec{x}$ pri pogoju $\vec{x}^\mathsf{T} B\vec{x} = b$.

↯(11) Naj bodo dani vektorji in matriki

$$P = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \vec{q} = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \vec{r} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix} \text{ in } \vec{b} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Poiščite najmanjšo vrednost funkcije $\vec{x}^\mathsf{T} P\vec{x} + \vec{q}^\mathsf{T}\vec{x} + \vec{r}$ pri pogoju, da $\vec{x}$ reši linearni sistem $A\vec{x} = \vec{b}$.

(Naloge, označene s ↯ preverjajo razumevanje osnovnih pojmov in so primeri nalog s teoretičnih izpitov. Naloga, označena s ↯↯ je nekolika težja.)