

JENSENOVA NEROVNOSŤ AKO DÔLEŽITÝ MATEMATICKÝ NÁSTROJ A ÚČINNÝ PROSTRIEDOK PRE DOKAZOVANIE NEROVNOSTÍ

JENSEN'S INEQUALITY AS AN IMPORTANT MATHEMATICAL TOOL AND EFFECTIVE RESOURCE FOR PROOVING CLASSICAL INEQUALITIES

JOZEF FULIER

ABSTRAKT. *V príspevku sú z hľadiska prípravy budúcich učiteľov matematiky prezentované základné pojmy a súvislosti teórie konvexných funkcií: podmienky konvexnosti, nutné podmienky a postačujúce podmienky pre konvexnosť (resp. rýdzu konvexnosť funkcie). V článku sa zdôrazňuje, že potenciál vyučovania konvexných funkcií (najmä v súvislosti s Jensenovou nerovnosťou) nie je dostatočne využitý a tejto problematike by mala byť venovaná väčšia pozornosť. Sú uvedené dva varianty Jensenovej nerovnosti, vrátane jej fyzikálnej a geometrickej interpretácie. Využitím Jensenovej nerovnosti sú dokázané dôležité klasické nerovnosti (AG-nerovnosť, HG-nerovnosť, Cauchyho – Schwarzova nerovnosť, Hölderova a Minkowského nerovnosť).*

KLÚČOVÉ SLOVÁ: *konvexná (rýdzo konvexná) funkcia, konkávna (rýdzo konkávna) funkcia, konvexná množina, podmienky konvexnosti, Jensenova nerovnosť, klasické nerovnosti*

ABSTRACT. *The contribution in terms of training of future mathematics teachers presented the basic concepts and background theory of convex functions: convexity conditions, necessary conditions and sufficient conditions for convex (or strictly convex) functions. The article emphasizes that the potential for learning convex functions (particularly in relation to the Jensen's inequality) is not sufficiently exploited, and this issue should be given greater attention. There are two variants of the Jensen inequality, including its physical and geometrical interpretation. Using the Jensen inequality is proved important classical inequalities (AG-inequality, HG-inequality, Cauchy-Schwarz inequality, Holder and Minkowski inequality).*

KEY WORDS: *convex (strictly convex) function, concave (strictly concave) function, convex set convexity condition, Jensen's inequality, classical inequalities.*

CLASSIFICATION: D59, D89, I49

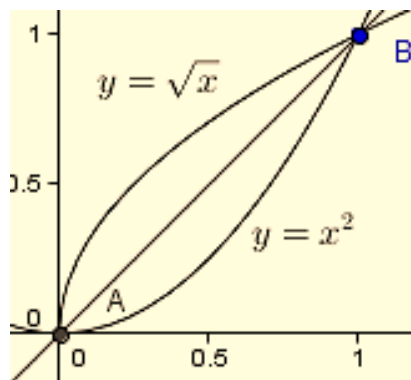
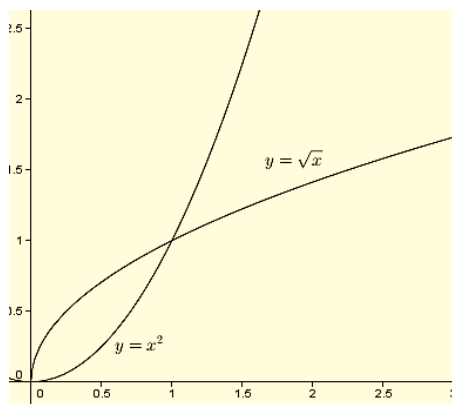
Úvod

V aplikáciách, pri riešení matematických problémov i v dôkazoch matematických viet sa veľmi často využívajú nerovnosti. V niektorých prípadoch je použitie nerovnosti unifikujúcim prvkom pri riešení celej triedy problémov (takýmto jednotiacim prvkom môže byť sústavné využívanie známej nerovnosti medzi aritmetickým a geometrickým priemerom, ktorú môžeme napríklad využiť pri riešení jednoduchých extrémnych úloh na elementárnej úrovni - bez využitia derivácií). Vedieť nájsť a následne dokazovať potrebné nerovnosti patrí k veľkému umeniu, pretože nerovnosti sú veľmi rôznorodé a ich dôkazy si spravidla vyžadujú rozličné techniky a prístupy, často prenikavú intuíciu a nemálo tvorivosti. Z uvedeného dôvodu je dobré mať prehľad o dôležitých nerovnostiach a o ich spôsoboch dokazovania. O tejto problematike bol vydaný už celý rad špecializovaných kníh, spomeňme napríklad monografiu (*Beckenbach & Bellman, 1961*), v ktorých je možné nájsť skutočne rôznorodé nerovnosti. Z odborného i pedagogického hľadiska je zaujímavé či aspoň tie najdôležitejšie nerovnosti nie je možné odvodiť či

dokázat' využitím čo „najmenšieho množstva“ elementárnych (alebo aspoň primerane jednoduchých) matematických nástrojov. Je vhodné si uvedomiť, že jedným z motorov rozvoja matematiky je proces zovšeobecnenia, ktorý spravidla spočíva v postupnom začleňovaní izolovaných poznatkov do všeobecnejšej formy, z ktorej tieto partikulárne prípady bezprostredne alebo sprostredkovane vyplývajú. Prirodzene asi ťažko môžeme očakávať, že sa nám podarí odvodiť akúsi *super-nerovnosť*, ktorá by zahrňovala všetky (či temer všetky) alebo aspoň významné skupiny dôležitých nerovností, pričom by bola navyše dostatočne jednoduchá, aby mohla byť vyučovaná (a azda i dokazovaná) už na strednej škole, prípadne v bakalárskom stupni vysokoškolského štúdia matematiky. V ďalšom pojednáme o jednej nerovnosti, ktorá si možno pomenovanie *super-nerovnosť* zaslúži. Ukážeme si prístup, ktorý v teórii nerovností, určite nepatrí k neznámym, ale rozhodne k málo využívaným. Ukážeme, že tento prístup si zasluhuje našu pozornosť, a to nielen z aspektu vyučovania základného kurzu matematickej analýzy na vysokej škole (napríklad v príprave učiteľov matematiky), ale svojou jednoduchosťou a kompaktnosťou aj z aspektu vyučovania matematiky na strednej škole. Ide o klasickú *Jensenovu nerovnosť* spolu s jej dôležitými aplikáciami.

Konvexné a konkávne funkcie, podmienky konvexnosti, konvexné množiny

Uvažujme na intervale $J = [0, \infty)$ dve funkcie $f: f(x) = x^2$, $g: g(x) = \sqrt{x}$. Tieto funkcie majú na intervale J veľa zhodných vlastností: obe sú spojité, rastúce, kladné a majú derivácie ľubovoľného rádu. Ich grafy sa však predsa len dosť odlišujú: funkcia f *zrýchľuje svoj rast*, čo sa prejavuje na jej grafe tým, že sa (s rastom argumentu) *zakrivuje nahor*, naopak funkcia g *spomaľuje svoj rast*, čo sa prejavuje na jej grafe tým, že sa *zakrivuje nadol*. Základnú odlišnosť týchto dvoch grafov môžeme ľahko identifikovať na grafoch reštrikcií týchto funkcií vzhľadom na interval $[0,1]$. Totiž grafy týchto dvoch funkcií majú dva priesečníky, body $A = (0,0)$, $B = (1,1)$. V intervale $[0,1]$ graf funkcie f (resp. graf funkcie g) leží (s výnimkou bodov A, B) celý pod (resp. celý nad) priamkou $\overline{AB}$ spájajúcej priesečníky A, B uvedených grafov. Tento fakt môžeme interpretovať i tak, že body A, B grafu *danej rastúcej funkcie* môžu byť spojené tromi základnými spôsobmi: buď úsečkou AB , alebo krivkou ležiacou celou (až na koncové body A, B) *pod úsečkou* AB , alebo krivkou ležiacou *nad touto úsečkou*¹.



Obr. 1 Grafy konvexnej funkcie $f(x) = x^2$ a grafy konkávnej funkcie $g(x) = \sqrt{x}$

¹ Analogické situácie nastanú i v prípade, ak dva body grafu spájame krivkami, ktoré sú grafmi *klesajúcich funkcií*.

O takýchto funkciách hovoríme, že majú rozličnú vypuklosť. Presnejšie, funkcia f sa nazýva *konvexnou* (obširnejšie: *rýdzo konvexnou*) *funkciou* a funkcia g sa nazýva *konkávnu* (rýdzo *konkávnu*) *funkciou* v intervale J .

Poznamenajme, že samotný pojem konvexná funkcia a aj príslušná pamäťová stopa v mysliach študentov sú viazané najmä na operatívnu činnosť pri zostrojovaní náčrtkov častí grafu funkcie na intervale, v ktorom je funkcia rýdzo konvexná: dva body grafu funkcie, ktorá je rýdzo konvexná na istom intervale, spojíme súvislou rýdzo konvexnou krivkou, ktorá sa veľmi silne podobá na časť paraboly otvorenej smerom nahor (alebo na časť paraboly otvorenej smerom nadol, pri rýdzo konkávnej funkcii). Zaujímavé je, že študenti uznávajú fakt, že z hľadiska konštrukcie dostatočne verného náčrtku grafu funkcie je určite dôležité či je funkcia v danom intervale rýdzo konvexná, alebo rýdzo konkávna. Avšak, ak použijú neutrálnu polohu (teda príslušné body jednoducho spoja úsečkou) sami to nepovažujú (a často ani ich učitelia) za tak vážnu chybu, čo by nejak znehodnocovalo celý náčrtok grafu funkcie. Je teda tento pojem úplne nedôležitý? Pokúsime sa dať odpoveď na túto otázku. Ukážeme, že konvexnosť funkcie je v mnohých aspektoch veľmi dôležitý pojem, ktorý si určite zaslúži našu pozornosť.

Poznamenajme, že pojmy *konvexnosť* a *konkávnosť* funkcie zaviedol dánsky matematik Johan Ludwig Jensen (1859-1925), ktorý sa o význame týchto pojmov vyjadril takto: *Zdá sa mi, že pojem konvexnej funkcie je rovnako dôležitý, ako pojem funkcie nezápornej alebo funkcie rastúcej. Ak sa v tomto nemýlim, tento pojem by mal nájsť svoje miesto v elementárnych textoch venovaných teórii reálnych funkcií.*

V základnom kurze matematickej analýzy vo vysokoškolskej príprave budúcich učiteľov (v minulosti dokonca aj v matematike stredoškolskej) sa definuje pojem konvexnosti a konkávnosti pre funkciu jednej premennej napríklad takto:

Definícia 1. Hovoríme, že funkcia f je *konvexná na intervale* $J \subset D(f)$, ak platí

$$\forall x_1, x, x_2 \in J: \quad x_1 < x < x_2 \Rightarrow f(x) \leq f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1} (x - x_1). \quad (1)$$

Výrok daný vzťahom (1) budeme nazývať **podmienkou konvexnosti** funkcie f v intervale J .

Geometrický význam podmienky konvexnosti (1): na každom intervale $[x_1, x_2] \subset J$ všetky body grafu reštrikcie funkcie f vzhľadom na interval $[x_1, x_2]$ *ležia pod alebo na* spojnici krajných bodov tohto grafu, t.j. bod $A = (x, f(x))$ *neleží nad úsečkou* spájajúcou body $A_1 = (x_1, f(x_1))$, $A_2 = (x_2, f(x_2))$.

Definíciu *funkcie konkávnej*, resp. *rýdzo konvexnej*, resp. *rýdzo konkávnej na intervale* J dostaneme, ak vo vzťahu (1) znak nerovnosti $\leq$ postupne nahradíme znakmi nerovností $\geq$, $<$, $>$. Príslušné výroky (so zmenou vyššie spomínaného znaku nerovnosti) nazývame *podmienkou konkávnosti*, resp. *podmienkou rýdzej konvexnosti*, resp. *podmienkou rýdzej konkávnosti*. Je zrejmé, že funkcia f je konkávna, resp. rýdzo konkávna na intervale J , práve vtedy, ak funkcia $(-f)$ je konvexná (resp. rýdzo konvexná) na intervale J . Preto stačí pozornosť venovať iba *konvexným*, resp. *rýdzo konvexným funkciám*.

Sformulujme niekoľko ďalších tvarov *podmienok konvexnosti*, ktoré sú ekvivalentné s podmienkou konvexnosti (1).

- a) Ľahko nahliadneme, že rovnicu priamky určenej bodmi A_1, A_2 , t.j. lineárnu funkciu $y = f(x_1) + \frac{f(x_2)-f(x_1)}{x_2-x_1}(x-x_1)$ môžeme jednoduchou úpravou pretransformovať na tzv. *Lagrangeov interpolačný polynóm* známy z numerickej matematiky, ktorý má tvar
- $$y = \frac{x_2-x}{x_2-x_1}f(x_1) + \frac{x-x_1}{x_2-x_1}f(x_2).$$

Preto podmienku konvexnosti (1) môžeme zapísať aj v tvare

$$\forall x_1, x, x_2 \in J: \quad x_1 < x < x_2 \Rightarrow f(x) \leq \frac{x_2-x}{x_2-x_1}f(x_1) + \frac{x-x_1}{x_2-x_1}f(x_2).. \quad (2)$$

- b) Ak v predchádzajúcej nerovnosti nahradíme $f(x) = f(x) \cdot 1 = f(x) \frac{(x_2-x)+(x-x_1)}{x_2-x_1}$ dostaneme ďalšiu ekvivalentnú podmienku konvexnosti

$$\forall x_1, x, x_2 \in J: \quad x_1 < x < x_2 \Rightarrow \frac{f(x)-f(x_1)}{x-x_1} \leq \frac{f(x_2)-f(x)}{x_2-x}. \quad (3)$$

Geometrický význam podmienky konvexnosti (3): smernica priamky $\overrightarrow{A_1A}$ nie je väčšia ako smernica priamky $\overrightarrow{AA_2}$.

- c) Bezprostredne je tiež možné overiť, že výrok daný vzťahom (1) môžeme využitím vlastností determinantu prepísať do nasledovného *symetrického tvaru*

$$\forall x_1, x, x_2 \in J: \quad x_1 < x < x_2 \Rightarrow$$

$$\begin{vmatrix} 1 & 1 & 1 \\ x_1 & x & x_2 \\ f(x_1) & f(x) & f(x_2) \end{vmatrix} = \begin{vmatrix} 1 & x_1 & f(x_1) \\ 1 & x & f(x) \\ 1 & x_2 & f(x_2) \end{vmatrix} \geq 0. \quad (4)$$

Geometrický význam podmienky konvexnosti (4):

absolútna hodnota determinantu vo vzťahu (4) určuje obsah trojuholníka s vrcholmi

$A_1 = (x_1, f(x_1))$, $A = (x, f(x))$, $A_2 = (x_2, f(x_2))$, pričom spomínaný determinant je kladný práve vtedy, ak vrcholy neležia na jednej priamke a cesta vrcholmi trojuholníka A_1, A, A_2 v postupnosti $A_1 \rightarrow A \rightarrow A_2 \rightarrow A_1$ je orientovaná proti chodu hodinových ručičiek.

- d) Veľmi užitočný tvar *podmienky konvexnosti funkcie* f dostaneme nasledovne.

Zvoľme ľubovoľne, ale pevne body $x_1, x, x_2 \in J$ také, že $x_1 < x < x_2$. Potom každý vnútorný bod $x \in [x_1, x_2]$ je možné jednoznačne zapísať v tvare $x = x_1 + \lambda_2(x_2 - x_1)$, kde $\lambda_2 = \frac{x-x_1}{x_2-x_1} \in (0, 1)$. Ak označíme $\lambda_1 := 1 - \lambda_2$, teda

$\lambda_1 = \frac{x_2-x}{x_2-x_1} \in (0, 1)$ potom bod $x \in (x_1, x_2)$ môžeme jednoznačne vyjadriť v tvare

$x = \lambda_1 x_1 + \lambda_2 x_2$, kde $\lambda_1, \lambda_2 \in (0, 1)$, $\lambda_1 + \lambda_2 = 1$.

To však znamená, že podmienku konvexnosti (2) môžeme prepísať do nasledovného tvaru

$$(\forall x_1, x_2 \in J, x_1 < x_2)(\forall \lambda_1, \lambda_2 \in (0, 1), \lambda_1 + \lambda_2 = 1):$$

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2)$$

Ak predpokladáme, že pre $x_1, x_2 \in J$ platí obrátená nerovnosť t.j. ak $x_2 < x_1$, potom vnútorný bod x intervalu $[x_2, x_1]$ možno tiež jednoznačne zapísať v tvare $x = \lambda_1 x_1 + \lambda_2 x_2$, kde $\lambda_1, \lambda_2 \in (0, 1)$, $\lambda_1 + \lambda_2 = 1$. Zopakovaním predchádzajúceho postupu dospejeme k tej istej nerovnosti

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2).$$

Tento fakt nám umožňuje zapísať podmienku konvexnosti pre ľubovoľné $x_1, x_2 \in J$, pričom stačí iba predpokladať, aby $x_1 \neq x_2$. Dokonca môžeme vynechať aj podmienku $x_1 \neq x_2$, pretože pre konvexné funkcie posledná nerovnosť zostane v platnosti aj v prípade $x_1 = x_2$.

Tým dostaneme ďalšiu podmienku konvexnosti v tvare

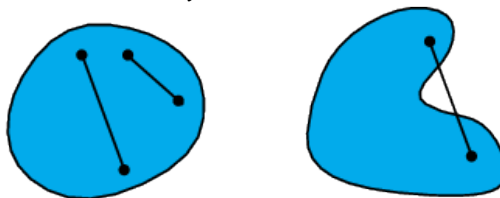
$$(\forall x_1, x_2 \in J)(\forall \lambda_1, \lambda_2 \in (0, 1), \lambda_1 + \lambda_2 = 1): \\ f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2). \quad (5)$$

Poznámka 1. Využitím podmienky konvexnosti (5) môžeme okamžite zaviesť definíciu konvexnej funkcie aj pre prípad funkcie n – premenných, t.j. funkcie $f: J \rightarrow \mathbb{R}$, kde $J \subset \mathbb{R}^n$, pričom vo všeobecnosti netreba žiadať, aby množina J bola intervalom v $\mathbb{R}^n$, ale stačí aby $J \subset \mathbb{R}^n$ bola konvexnou množinou². Toto súčasne napovedá, ktorú z podmienok (1) až (5) je výhodné vziať za definíciu konkávnej funkcie.

Je vhodné pripomenúť si, že s pojmom konvexnosti sa oboznamujú už žiaci základnej a strednej školy, keď hovoria o konvexných uhloch, konvexných mnohouholníkoch, konvexných telesách. Pojem konvexnej množiny patrí k základným pojmom geometrie i matematickej analýzy. Spravidla sa tento pojem definuje takto:

Definícia 2. Množinu $\mathcal{D} \subset \mathbb{R}^n$ nazývame konvexnou množinou, ak s ľubovoľnými bodmi $A, B \in \mathcal{D}$ patrí do množiny $\mathcal{D}$ celá úsečka spájajúca uvedené body. Algebricky zapísané: ak pre každé $A, B \in \mathcal{D}$ a každé $\lambda \in [0, 1]$ platí $\lambda A + (1 - \lambda)B \in \mathcal{D}$.

Ak každá z uvažovaných úsečiek $\overline{AB}$ leží celá (s možnou výnimkou krajných bodov A, B) vo vnútri $\mathcal{D}$, nazýva sa množina $\mathcal{D}$ rýdzo konvexnou.



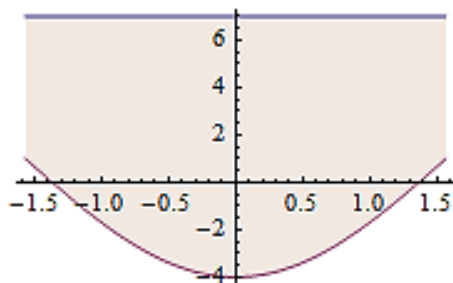
Obr. 2. Konvexná a nekonvexná množina v $\mathbb{R}^2$

Ako však súvisí pojem konvexnej funkcie s pojmom konvexná množina? Tento fakt, nebýva študentom celkom jasný. Jednou z možností, ako im môžeme ukázať súvislosť medzi pojmi konvexná funkcia - konvexná množina je, že definujeme tzv. nadgraf (epigraf) danej funkcie. Na chvíľu zvolíme trochu všeobecnejšie hľadisko a pojem nadgrafu zavedieme pre ľubovoľnú funkciu viacerých premenných.

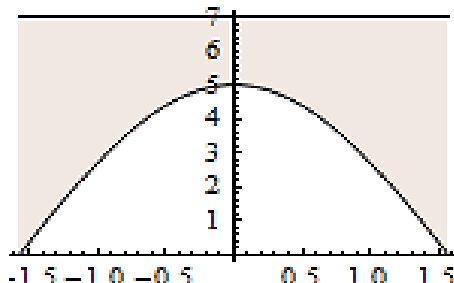
² Konvexná podmnožina $J \subset \mathbb{R}^1$ obsahujúca aspoň dva body je interval.

Definícia 3. Nech $f: D \rightarrow \mathbb{R}$, kde $D \subset \mathbb{R}^n$. Nadgrafom (epigrafom) funkcie f rozumieme množinu $G_f = \{(x, y) \in \mathbb{R}^{n+1}: x \in D, f(x) \leq y\}$.

Lahko nahliadneme, že k tomu, aby nadgraf funkcie f bol konvexnou množinou je nevyhnutné, aby aj definičný obor funkcie f bol konvexnou množinou. Preto nás iste neprekvapí, že tento predpoklad sa objavuje v predpokladoch každej vety, ktorá nám dáva nutnú a postačujúcu podmienku k tomu, aby funkcia f bola konvexná na množine D .



Obr. 3 Nadgraf rýdzo konvexnej funkcie
 $f: y = 1 - 5 \cos x, x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$



Nadgraf rýdzo konkávnej funkcie
 $g: y = 5 \cos x, x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$

Veta 1. Nech $D \subset \mathbb{R}^n$ je konvexná množina a nech $f: D \rightarrow \mathbb{R}$. Funkcia f je konvexná na množine D vtedy a len vtedy, ak jej nadgraf, t.j. množina G_f je konvexnou množinou.

Dôkaz. 1. $\Rightarrow$). Predpokladajme, že funkcia f je konvexná na množine $D \subset \mathbb{R}^n$, dokážeme že jej nadgraf G_f je konvexnou množinou. Nech $(x_1, y_1), (x_2, y_2) \in G_f$ a nech $\lambda \in [0, 1]$. Potom platí

$$\lambda(x_1, y_1) + (1 - \lambda)(x_2, y_2) = (\lambda x_1 + (1 - \lambda)x_2, \lambda y_1 + (1 - \lambda)y_2).$$

Z konvexnosti funkcie a definície nadgrafu dostaneme

$$\lambda y_1 + (1 - \lambda)y_2 \geq \lambda f(x_1) + (1 - \lambda)f(x_2) \geq f(\lambda x_1 + (1 - \lambda)x_2), \text{ teda}$$

$\lambda(x_1, y_1) + (1 - \lambda)(x_2, y_2) \in G_f$, čo znamená nadgraf G_f funkcie f je konvexná množina.

2. $\Leftarrow$). Nech $x_1, x_2 \in D$ a $\lambda \in [0, 1]$. Potom $(x_1, f(x_1)), (x_2, f(x_2)) \in G_f$ a teda $\lambda(x_1, f(x_1)) + (1 - \lambda)(x_2, f(x_2)) = (\lambda x_1 + (1 - \lambda)x_2, \lambda f(x_1) + (1 - \lambda)f(x_2)) \in G_f$. Toto však podľa definície nadgrafu znamená, že

$$\lambda f(x_1) + (1 - \lambda)f(x_2) \geq f(\lambda x_1 + (1 - \lambda)x_2),$$

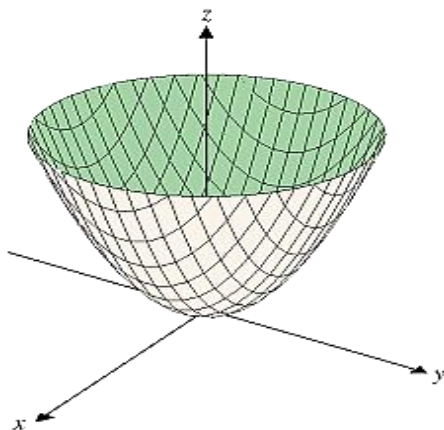
teda funkcia f je konvexná.

V prípade funkcie jednej, príp. dvoch premenných je možné doplniť uvedené úvahy geometrickou interpretáciou prostredníctvom (modelov) grafov konvexných resp. konkávnych funkcií na konvexnej množine D .

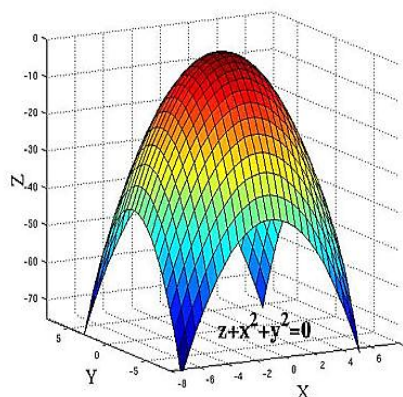
Konkrétne na Obr. 4a je graf rýdzo konvexnej funkcie $f: f(x, y) = x^2 + y^2$ definovanej na konvexnej množine $D = \mathbb{R}^2$ a na Obr. 4b je znázornený graf funkcie $g: g(x, y) = -x^2 - y^2$, ktorá je rýdzo konkávna na (konvexnej) množine $D = [-8, 5] \times [-8, 3]$.

Čo sa týka smeru zovšeobecňovania našich úvah o konvexnosti funkcie pre *funkcie viacerých premenných*, tento smer (napriek tomu, že je dôležitý) nebudeme ďalej rozvíjať, pretože pre našu, v podstate informatívnu úroveň výkladu, stačí využiť podmienky konvexnosti platné pre funkciu *jednej premennej*. Platí totiž nasledovná veta.

Veta 2. Funkcia $f: D \rightarrow \mathbb{R}$, kde $D \subset \mathbb{R}^n$, je konvexná množina, je *konvexnou funkciou* vtedy a len vtedy, ak zložená funkcia jednej premennej $g: g(t) = f(x + tv)$ je *konvexná* (vzhľadom na premennú t) pre všetky prípustné $x \in D, v \in \mathbb{R}^n$.



Obr. 4a: Graf rýdzo konvexnej funkcie dvoch premenných $f: f(x, y) = x^2 + y^2$



Obr. 4b: Graf rýdzo konkávnej funkcie dvoch premenných $g: g(x, y) = -x^2 - y^2$

Nutné podmienky, postačujúce podmienky pre konvexnosť funkcie

V základných kurzoch matematickej analýzy, v snahe dostať sa rýchle k aplikáciám, je venovaná hlavná pozornosť najmä postačujúcim podmienkam pre konvexnosť, resp. rýdzu konvexnosť funkcie, ktoré sú viazané najmä na existenciu derivácie prvého a druhého rádu danej funkcie. Menej pozornosti je venované *nutným podmienkam* pre konvexnosť funkcie. Vzhľadom na to, že *nadgraf konvexnej funkcie* na intervale J je konvexnou množinou, ktorá je nevyhnutne i *súvislou množinou*, dá sa očakávať, že každá konvexná funkcia musí byť nevyhnutne *spojitou funkciou* v každom vnútornom bode intervalu J . Vzhľadom na nerovnosť v podmienke konvexnosti danej vzťahom (3), *diferenčné podiely* vystupujúce v tejto nerovnosti vykazujú istý druh monotónnosti a sú ohraničené, čo znamená, že ich jednostranné limity, ktoré sú jednostrannými deriváciami funkcie f existujú v každom vnútornom bode intervalu J . Presnejšie, to môžeme sformulovať do nasledovnej vety takto:

Veta 4. Nech funkcia f je *konvexná* na intervale J . Potom v každom vnútornom bode x_0 intervalu J funkcia f :

- je *spojitá*,
- má konečné obe *jednostranné derivácie*, pričom $f'_+(x_0) \geq f'_-(x_0)$ a platí
 $(\forall x \in J, x < x_0): (f(x) \geq f(x_0) + f'_-(x_0)(x - x_0),$
 $(\forall x \in J, x > x_0): (f(x) \geq f(x_0) + f'_+(x_0)(x - x_0).$
- Množina bodov z intervalu J , v ktorých funkcia f nie je diferencovateľná, je *najviac spočítateľná*. Ak funkcia je diferencovateľná v každom bode intervalu

$(a, b) \subset J$, potom derivácia f' funkcie f *neklesajúca* na (a, b) , resp. *rastúca* na (a, b) , ak funkcia f je *rýdzo konvexná*.

Poznamenajme, že konvexná funkcia *nemusi byť spojitá v krajných bodoch* intervalu: stačí vziať napríklad funkciu $f: f(x) = 2 + \operatorname{sgn} x, x \in [-1, 0]$, ktorá je konvexnou funkciou, ale zjavne *nie je spojitá* v bode $x = 0$.

Ak k základným predpokladom kladených na danú funkciu f , pridáme predpoklad *existencie prvej derivácie*, či dokonca *druhej derivácie* v celom J , dostaneme veľmi jednoduché *postačujúce podmienky* pre konvexnosť, resp. rýdzu konvexnosť danej funkcie f .

Veta 5. Nech funkcia f je diferencovateľná na intervale (a, b) . Potom nasledujúce tvrdenia sú ekvivalentné:

- (a) funkcia f je rýdzo konvexná;
 - (b) funkcia f' je rastúca;
 - (c) pre ľubovoľné $x_0 \in (a, b)$ platí

$$(\forall x \in (a, b), x \neq x_0): (f(x) > f(x_0) + f'(x_0)(x - x_0)) \quad (6)$$
- (t.j. nech pre každý bod $x_0 \in (a, b)$ platí, body grafu funkcie f prislúchajúce bodom $x \neq x_0$ ležia nad dotyčnicou ku grafu funkcie f v bode x_0).

Veta 6. Nech funkcia f je dvakrát diferencovateľná na intervale (a, b) . Funkcia f je rýdzo konvexná (resp. rýdzo konkávna) na intervale (a, b) , vtedy a len vtedy, ak $f''(x) \geq 0$ (resp. $f''(x) \leq 0$) pre všetky $x \in (a, b)$ a keď neexistuje podinterval I intervalu (a, b) , v ktorom by $f'' \equiv 0$. Špeciálne, ak $f''(x) > 0$ pre všetky $x \in (a, b)$, tak funkcia f je rýdzo konvexná na intervale (a, b) .

Jensenova nerovnosť

Ukážme si ešte dve ekvivalentné *podmienky konvexnosti*, ktoré formálne zovšeobecňujú podmienku konvexnosti (5) (samozrejme, že nie obsahovo, pretože to budú opäť *podmienky konvexnosti*, ktoré sú *ekvivalentné* s podmienkami predchádzajúcimi). Tieto podmienky konvexnosti, vzhľadom na ich široké možnosti využitia pri odvodzovaní dôležitých nerovností, sformulujeme explicitne do matematických viet a obe budeme nazývať *Jensenovou nerovnosťou*.

Veta 7 (*Jensenova nerovnosť I*). Nech $n \in \mathbb{N}, n \geq 2$. Ak funkcia f je *konvexná* na intervale J , potom pre každé $x_1, x_2, \dots, x_n \in J$ a pre ľubovoľné kladné čísla $\lambda_1, \lambda_2, \dots, \lambda_n$ s vlastnosťou $\sum_{i=1}^n \lambda_i = 1$ platí

$$f(\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_n x_n) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2) + \dots + \lambda_n f(x_n) \quad (7)$$

Ak funkcia f je *rýdzo konvexná* na intervale J , potom za vyššie uvedených predpokladov platí nerovnosť (7), pričom rovnosť v uvedenom vzťahu nastane len v prípade, ak platí $x_1 = x_2 = \dots = x_n$.

Dôkaz. Tvrdenie ľahko dokážeme matematickou indukciou vzhľadom na n . Využitie indukčného predpokladu je zrejme z nasledovnej schémy:

Predpokladajme, že tvrdenie vety platí pre nejaké $n = k \geq 2$, potom pre $n = k + 1$ a pre $\lambda_{n+1} < 1$ (opačnom prípade, nemáme čo dokazovať) platí

$$f\left(\sum_{j=1}^{k+1} \lambda_j x_j\right) = f\left(\lambda_{k+1} x_{k+1} + (1 - \lambda_{k+1}) \sum_{j=1}^k \frac{\lambda_j}{1 - \lambda_{k+1}} x_j\right) \leq \lambda_{k+1} f(x_{k+1}) + (1 - \lambda_{k+1}) f\left(\sum_{j=1}^k \frac{\lambda_j}{1 - \lambda_{k+1}} x_j\right).$$

Využitím indukčného predpokladu a toho, že

$$\sum_{j=1}^k \frac{\lambda_j}{1 - \lambda_{k+1}} = \frac{1}{1 - \lambda_{k+1}} \sum_{j=1}^k \lambda_j = \frac{1 - \lambda_{k+1}}{1 - \lambda_{k+1}} = 1, \text{ dostaneme}$$

$$(1 - \lambda_{k+1}) f\left(\sum_{j=1}^k \frac{\lambda_j}{1 - \lambda_{k+1}} x_j\right) \leq (1 - \lambda_{k+1}) \left(\sum_{j=1}^k \frac{\lambda_j}{1 - \lambda_{k+1}} f(x_j)\right) = \sum_{j=1}^k \lambda_j f(x_j).$$

Veta 8 (*Jensenova nerovnosť II*). Nech $n \in \mathbb{N}, n \geq 2$. Ak funkcia f je konvexná na intervale J , potom pre každé $x_1, x_2, \dots, x_n \in J$ a pre ľubovoľné kladné čísla $\alpha_1, \alpha_2, \dots, \alpha_n$ platí vzťah

$$f\left(\frac{\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n}{\sum_{k=1}^n \alpha_k}\right) \leq \frac{\alpha_1}{\sum_{k=1}^n \alpha_k} f(x_1) + \frac{\alpha_2}{\sum_{k=1}^n \alpha_k} f(x_2) + \dots + \frac{\alpha_n}{\sum_{k=1}^n \alpha_k} f(x_n). \quad (8)$$

Ak funkcia f je rýdzo konvexná na intervale J , potom za vyššie uvedených predpokladov platí nerovnosť (8), pričom rovnosť v uvedenom vzťahu nastane len v prípade, ak platí $x_1 = x_2 = \dots = x_n$.

Dôkaz. Tvrdenie vety 8 bezprostredne vyplýva z vety 7, v ktorej kladieme $\lambda_j = \frac{\alpha_j}{\sum_{k=1}^n \alpha_k}$, pričom zrejme $\sum_{j=1}^n \lambda_j = 1$.

Fyzikálny význam Jensenovej nerovnosti

Vlastnosť konvexných (resp. rýdzo konvexných) funkcií v tvare *Jensenovej nerovnosti II* má veľmi zaujímavú (a veľmi prirodzenú) **fyzikálnu interpretáciu** o polohe *ťažiska*³ (*hmotného streda*, *barycentra*) *sústavy hmotných bodov*, ktorá je pochopiteľná aj pre žiakov stredných škôl. Zopakujme si niektoré fakty.

V prvom rade, pod ťažiskom dvoch hmotných bodov A, B s hmotnosťami m_A, m_B budeme rozumieť taký bod T , že pre veľkosti $|AT|, |BT|$ úsečiek AT, BT platí $\frac{|AT|}{|BT|} = \frac{m_B}{m_A}$.

Toto je známa *podmienka rovnováhy na páke* (pozri Obr.5a). Karteziánske súradnice ťažiska T je možné vyjadriť pomocou súradníc bodov A, B veľmi jednoduchými vzťahmi:

$$x_T = \frac{m_A x_A + m_B x_B}{m_A + m_B}, y_T = \frac{m_A y_A + m_B y_B}{m_A + m_B}.$$

Túto definíciu ľahko zovšeobecníme aj na určenie *ťažiska (hmotného streda)* troch bodov. Určíme ho ako *hmotný stred* ťažiska prvých dvoch bodov a tretieho bodu. Vo všeobecnosti ťažisko n -bodov $A_1 = (x_1, y_1), A_2 = (x_2, y_2), \dots, A_n = (x_n, y_n)$ s hmotnosťami $m_1, m_2, \dots, m_n$ určíme ako ťažisko bodov $A_1, A_2, \dots, A_{n-1}$ a bodu A_n .

³ *Ťažisko (hmotný stred)* je pôsobisko tiažovej sily pôsobiace na teleso. V skutočnosti je medzi pojmami *ťažisko* a *hmotný stred* principiálny rozdiel. Ťažisko zavádzame ako pôsobisko výslednice tiažových síl pôsobiacich na jednotlivé časti telesa v tiažovom poli (môžeme tiež povedať, že je to bod, voči ktorému je výsledný moment pôsobiacich tiažových síl nulový). Pojem ťažiska teda stráca význam v beztiažovom stave. Hmotný stred je bod, ktorý je pevne určený tvarom telesa a rozložením hustoty. Nezávisí na prítomnosti vonkajšieho silového poľa. V homogénnom tiažovom poli (napr. v tesnej blízkosti zemského povrchu) oba pojmy splývajú a veľmi často sa používajú ako synonyma. V nehomogénnom tiažovom poli je však nutné oba pojmy rozlišovať. Ťažisko homogénneho geometrického pravidelného telesa leží v jeho geometrickom strede (v geometrickom ťažisku). Ťažisko telesa leží v priesečníku ťažníc pri postupnom zavesení telesa v najmenej dvoch rôznych bodoch.

Jednoduchým výpočtom sa dá ukázať, že súradnice takto určeného ťažiska T hmotných bodov $A_1, A_2, \dots, A_n$ nezávisí od poradia týchto bodov a platí:

$$x_T = \frac{m_1x_1 + m_2x_2 + \dots + m_nx_n}{m_1 + m_2 + \dots + m_n}, y_T = \frac{m_1y_1 + m_2y_2 + \dots + m_ny_n}{m_1 + m_2 + \dots + m_n}.$$

Lema. *Nech $\mathcal{U}$ je daný konvexný útvar, v ktorom leží sústava hmotných bodov $A_1, A_2, \dots, A_n$ s hmotnosťami $m_1, m_2, \dots, m_n$. Potom aj ťažisko (hmotný stred) T tejto sústavy bodov leží v útware $\mathcal{U}$.*

Dôkaz. Pre $n = 2$ tvrdenie lemy zrejme platí. Totiž ťažisko T dvoch bodov z konvexného útvaru leží na úsečke spájajúcej tieto body, a táto celá leží v konvexnom útware $\mathcal{U}$, teda aj $T \in \mathcal{U}$.

Predpokladajme, že tvrdenie platí pre $n = k$, dokážeme, že platí aj pre $n = k + 1$. To však platí, pretože podľa predpokladu ťažisko T_k bodov $A_1, A_2, \dots, A_k \in \mathcal{U}$ leží v útware $\mathcal{U}$. To však znamená, že ťažisko T bodov $A_1, A_2, \dots, A_{k+1}$ je ťažiskom dvoch bodov $T_k, A_{k+1} \in \mathcal{U}$, preto tiež leží v útware $\mathcal{U}$.

Prepíšme tento *zákon statiky* do matematickej reči. Pre jednoduchosť predpokladajme, že body $A_1, A_2, \dots, A_n$, ktorých x – ové súradnice sú po rade $x_1, x_2, \dots, x_n$, ležia na grafe konvexnej funkcie $y = f(x)$, čiže $A_k = (x_k, f(x_k))$, $k = 1, 2, \dots, n$. Ak uvážime, že body $A_1, A_2, \dots, A_n$ ležia v *nadgrafe* funkcie f , ktorý je podľa predpokladu *konvexnou množinou*, tak y – ová súradnica y_T ťažiska T útvaru $\mathcal{U}$ nie je menšia ako hodnota funkcie f v bode $x_T = \frac{m_1x_1 + m_2x_2 + \dots + m_nx_n}{m_1 + m_2 + \dots + m_n}$, t.j. $f(x_T) \leq y_T$ (pozri Obr. 5b). Teda platí

$$f\left(\frac{m_1x_1 + m_2x_2 + \dots + m_nx_n}{m_1 + m_2 + \dots + m_n}\right) \leq \frac{m_1f(x_1) + m_2f(x_2) + \dots + m_nf(x_n)}{m_1 + m_2 + \dots + m_n},$$

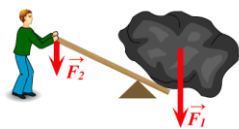
čo je *Jensenova nerovnosť* typu II.

Poznamenajme, že na úlohy typu: *nájsť ťažisko sústavy daných hmotných bodov*, môžeme pretransformovať aj veľa iných úloh z reálneho života. Konkrétne môže ísť i o takúto úlohu *optimalizačného typu*:

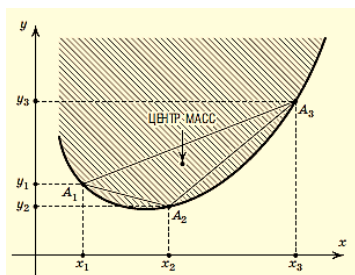
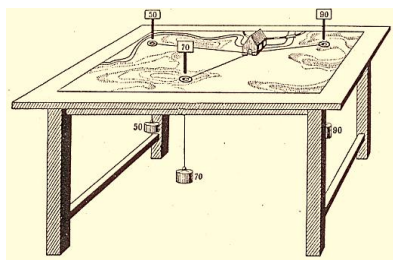
Úloha. *Pre n dedín sa má postaviť spoločná škola. V prvej dedine žije k_1 detí školského veku, v druhej dedine žije k_2 detí a n -tej dedine žije k_n detí školského veku. Akú geografickú polohu je potrebné zvoliť pre výstavbu novej školy, aby celková doba, ktorú potrebujú všetky deti k dochádzke do školy bola čo najmenšia?*

Túto úlohu môžeme riešiť *experimentálne* a jej riešenie je veľmi presvedčivé. Pre približné vyriešenie tejto úlohy totiž stačí položiť na stôl dostatočne podrobnú geografickú mapu oblasti, v ktorej sa nachádzajú všetky spomínané dediny, potom prevrátať v stole otvory⁴ v miestach, kde sa nachádzajú dediny, spustiť týmito otvormi n povrazov

⁴ Prirodzene dajú sa zvoliť i menej deštruktívne metódy.



Obr. 5a. Rovnováha na páke


 Obr. 5b. Ťažisko (↓) bodov bodov A_1, A_2, A_3


Obr. 5c. Určenie ťažiska sústavy hmotných bodov

(zhotovených z čo najľahšieho materiálu, napríklad zo silónového vlákna), ktorých horné konce zviažeme uzlom a k spodným koncom priviažeme v príslušnom poradí závažia s hmotnosťami $k_1, k_2, \dots, k_n$ vo vhodných hmotnostných jednotkách (napríklad v dekagramoch, aby sme mohli zanedbať vplyv hmotností jednotlivých povrazov). Riešenie sa nám objaví priamo pred očami. Škola má byť postavená na tom geografickom mieste, kde sa po vzniknutej rovnováhe objaví uzol. Na Obr. 5c je znázornené riešenie uvedenej úlohy pre $n = 3$, $k_1 = 50$, $k_2 = 70$, $k_3 = 90$, ktoré je uvedené v knihe (Steinhaus, 1953).

Využitie Jensenovej nerovnosti

Jensenovu nerovnosť sformulovanú pre ľubovoľnú konvexnú (resp. konkávnú) funkciu možno považovať za základnú v klasickej teórii nerovností. Voľbou konkrétnych konvexných (resp. konkávných) funkcií z nej možno získať viaceré dôležité nerovnosti. Poznamenajme, že v roku 1906 dánsky matematik J. L. Jensen uverejnil v časopise *Acta Mathematica* článok, v ktorom sa zaoberal Cauchyho dôkazom tzv. AG – nerovnosti, ktorú slávny francúzsky matematik L. A. Cauchy (1789 -1857) dokazoval tzv. spätnou indukciou. Pripomeňme si, že AG – nerovnosť je dôležitý vzťah medzi aritmetickým a geometrickým priemerom so širokým využitím v rozličných oblastiach matematiky.

Príklad 1 (AG - nerovnosť). Dokážte, že platí: Pre ľubovoľné nezáporné čísla $x_1, x_2, \dots, x_n$ platí

$$\sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} \leq \frac{x_1 + x_2 + \dots + x_n}{n}.$$

Pritom rovnosť nastane vtedy a len vtedy, ak $x_1 = x_2 = \dots = x_n$.

Dôkaz. Jensen v spomínanom článku urobil pri dokazovaní A-G nerovnosti takúto úvahu. Ak aspoň jedno z čísel $x_1, x_2, \dots, x_n$ je nula, A-G nerovnosť je triviálne splnená. Môžeme preto predpokladať, že všetky tieto čísla sú kladné. Vzhľadom na to, že funkcia $f: f(x) = -\ln x$ je klesajúca, tak logaritmovaním A-G nerovnosti dostaneme nerovnosť s ňou ekvivalentnú (pozor znak nerovnosti sa zmení na obrátený)

$$-\ln \frac{x_1 + x_2 + \dots + x_n}{n} \leq \frac{-\ln x_1 - \ln x_2 - \dots - \ln x_n}{n}. \quad (9)$$

Túto nerovnosť je možné prepísať do tvaru

$$f\left(\frac{x_1 + x_2 + \dots + x_n}{n}\right) \leq \frac{f(x_1) + f(x_2) + \dots + f(x_n)}{n}.$$

Jensen analýzou Cauchyho dôkazu zistil že túto nerovnosť je možné odvodiť s využitím jediného predpokladu o funkcii f :

pre ľubovoľné dve čísla x_1, x_2 patriace intervalu J platí

$$f\left(\frac{x_1 + x_2}{2}\right) \leq \frac{f(x_1) + f(x_2)}{2}.^{5)}$$

Poznamenajme, že táto nerovnosť je historicky prvým známym špeciálnym tvarom *Jensenovej nerovnosti* (pre $n = 2$, $\lambda_1 = \lambda_2 = \frac{1}{2}$). Funkcie spĺňajúce túto nerovnosť dnes označujeme názvom *J-konvexné funkcie*. Táto podmienka je slabšou podmienkou v porovnaní s podmienkami konvexnosti (1) až (8). Dá sa však dokázať nasledovné tvrdenie:

Každá J-konvexná funkcia na intervale J , ktorá je spojitá na intervale J , je konvexnou funkciou (v zmysle definície 1) na tomto intervale.

Vzhľadom na to, že funkcia $f: f(x) = -\ln x$ je J-konvexná a aj spojitá na $(0, \infty)$, tak je konvexná v zmysle definície 1. Túto skutočnosť môžeme štandardne overiť, priamo pomocou znamienka druhej derivácie funkcie f . Skutočne, pre každé $x \in (0, \infty)$ platí $f''(x) = (-\ln x)'' = \frac{1}{x^2} > 0$, preto funkcia f je rýdzo konvexná na intervale $(0, \infty)$, čo znamená, že platí nerovnosť (9) a teda platí aj AG – nerovnosť, ktorá je s ňou ekvivalentná.

Poznámka 1 (GH-nerovnosť). Ak v AG - nerovnosti nahradíme $x_k > 0$ výrazom $\frac{1}{x_k}$ pre $k = 1, 2, \dots, n$ dostaneme nerovnosť

$$\sqrt[n]{x_1 \cdot x_2 \cdot \dots \cdot x_n} \geq \frac{n}{\frac{1}{x_1} + \frac{1}{x_2} + \dots + \frac{1}{x_n}},$$

ktorá vyjadruje nerovnosť medzi *geometrickým* a *harmonickým priemerom* kladných čísel $x_1, x_2, \dots, x_n$ a preto ju niekedy nazývame *GH-nerovnosťou*.

Príklad 2 (Nerovnosť medzi váženým mocninným a váženým geometrickým priemerom). Dokážte, že platí: Pre ľubovoľné nezáporné čísla $x_1, x_2, \dots, x_n$ a ľubovoľné kladné reálne číslo p , pre ľubovoľné kladné čísla $\lambda_1, \lambda_2, \dots, \lambda_n$ s vlastnosťou $\sum_{i=1}^n \lambda_i = 1$ platí

$$(\lambda_1 x_1^p + \lambda_2 x_2^p + \dots + \lambda_n x_n^p)^{\frac{1}{p}} \geq x_1^{\lambda_1} \cdot x_2^{\lambda_2} \dots x_n^{\lambda_n}. \quad (10)$$

Pritom rovnosť nastane vtedy a len vtedy, ak $x_1 = x_2 = \dots = x_n$.

Dôkaz. Ak niektoré z čísel $x_1, x_2, \dots, x_n$ sa rovná nule, nerovnosť zrejme platí. Môžeme preto predpokladať, že každé z uvedených čísel je kladné. Logaritmovaním vzťahu (10) a využitím monotónnosti logaritmickej funkcie dostaneme s ňou ekvivalentnú nerovnosť

$$\frac{1}{p} \ln(\lambda_1 x_1^p + \lambda_2 x_2^p + \dots + \lambda_n x_n^p) \geq \ln(x_1^{\lambda_1} \cdot x_2^{\lambda_2} \dots x_n^{\lambda_n}),$$

ktorú jednoduchými ekvivalentnými úpravami prevedieme na nerovnosť

$$\ln(\lambda_1 x_1^p + \lambda_2 x_2^p + \dots + \lambda_n x_n^p) \geq \lambda_1 \ln x_1^p + \lambda_2 \ln x_2^p + \dots + \lambda_n \ln x_n^p.$$

⁵ Táto podmienka sa objavila pri sledovaní Cauchyho dôkazu spätnej indukcie, ktorý spočíval v troch krokoch: najprv sa dokázala A-G nerovnosť pre $n = 2$, potom pre všetky čísla $n = 2^k$, $k \in \mathbb{N}$ a až v poslednom kroku pre zvyšné prirodzené čísla. Špeciálne sa dokazovalo, že ak uvedená nerovnosť platí pre nejaké $n \in \mathbb{N}$, potom platí aj pre ľubovoľné prirodzené m , $2 \leq m < n$.

Posledná nerovnosť je ekvivalentná s *Jensenovou nerovnosťou* pre funkciu $f: f(x) = -\ln x^p$, ktorá je rýdzo konvexná na intervale $(0, \infty)$, pretože v tomto intervale pre funkciu f platí $f''(x) = \frac{p}{x^2} > 0$. Toto už dokazuje platnosť nerovnosti (10).

Poznámka 2. Ak v dokázanej nerovnosti položíme $\lambda_1 = \lambda_2 = \dots = \lambda_n = \frac{1}{n}$ a $p = 1$ dostaneme AG-nerovnosť a pre $p = 2$ dostaneme známu nerovnosť medzi kvadratickým priemerom a geometrickým priemerom nezáporných čísel $x_1, x_2, \dots, x_n$.

Príklad 3 (Nerovnosti medzi mocninnými priermi). Dokážte, že platí: Pre ľubovoľné kladné čísla $x_1, x_2, \dots, x_n$ a ľubovoľné $p > q > 0$ platí

$$\sqrt[p]{\frac{x_1^p + x_2^p + \dots + x_n^p}{n}} \geq \sqrt[q]{\frac{x_1^q + x_2^q + \dots + x_n^q}{n}}.$$

Dôkaz. Vzhľadom na to, že $\frac{p}{q} > 1$, funkcia $f: f(u) = u^{\frac{p}{q}}$ je rýdzo konvexná na intervale $(0, \infty)$. Aplikáciou Jensenovej nerovnosti pre ľubovoľné kladné čísla $u_1, u_2, \dots, u_n$, a voľbou $\lambda_1 = \lambda_2 = \dots = \lambda_n = \frac{1}{n}$ dostaneme

$$\left(\frac{u_1 + u_2 + \dots + u_n}{n}\right)^{\frac{p}{q}} \leq \frac{u_1^{\frac{p}{q}} + u_2^{\frac{p}{q}} + \dots + u_n^{\frac{p}{q}}}{n}.$$

Umocnením poslednej nerovnosti s exponentom $\frac{1}{p}$ a dosadením za $u_k = x_k^q$ pre všetky $k = 1, 2, \dots, n$ dostaneme požadovanú nerovnosť.

Príklad 4 (Cauchyho – Schwarzova nerovnosť). Dokážte, že platí: Pre ľubovoľné kladné čísla $a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n$ platí

$$(a_1^2 + a_2^2 + \dots + a_n^2)(b_1^2 + b_2^2 + \dots + b_n^2) \geq (a_1 b_1 + a_2 b_2 + \dots + a_n b_n)^2.$$

Dôkaz. Funkcia $f: f(x) = x^2$ je rýdzo konvexná na $(-\infty, \infty)$. Napíšme pre túto funkciu *Jensenovu nerovnosť II*

$$\left(\frac{\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n}{\sum_{k=1}^n \alpha_k}\right)^2 \leq \frac{\alpha_1 (x_1)^2}{\sum_{k=1}^n \alpha_k} + \frac{\alpha_2 (x_2)^2}{\sum_{k=1}^n \alpha_k} + \dots + \frac{\alpha_n (x_n)^2}{\sum_{k=1}^n \alpha_k}, \quad (\alpha_k > 0).$$

Odtiaľ máme

$$(\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n)^2 \leq (\alpha_1 x_1^2 + \alpha_2 x_2^2 + \dots + \alpha_n x_n^2) \sum_{k=1}^n \alpha_k.$$

Položením $\alpha_k = b_k^2, x_k = \frac{a_k}{b_k}$ pre $k = 1, 2, \dots, n$ dostaneme požadovanú nerovnosť.

Príklad 5 (Hölderova nerovnosť). Dokážte, že platí: Nech p, q sú ľubovoľné kladné čísla s vlastnosťou $\frac{1}{p} + \frac{1}{q} = 1$. Potom pre ľubovoľné kladné čísla $a_1, a_2, \dots, a_n, b_1, b_2, \dots, b_n$ platí nerovnosť

$$\sum_{k=1}^n a_k b_k \leq \left(\sum_{k=1}^n a_k^p\right)^{\frac{1}{p}} \left(\sum_{k=1}^n b_k^q\right)^{\frac{1}{q}}.$$

Dôkaz. Ľahko nahliadneme, že $p > 1$, preto funkcia $f: f(x) = x^p$ je rýdzo konvexná na $(0, \infty)$. Podobne ako v predchádzajúcom príklade napíšme pre túto funkciu *Jensenovu nerovnosť II*

$$\left(\frac{\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n}{\sum_{k=1}^n \alpha_k} \right)^p \leq \frac{\alpha_1 (x_1)^p}{\sum_{k=1}^n \alpha_k} + \frac{\alpha_2 (x_2)^p}{\sum_{k=1}^n \alpha_k} + \dots + \frac{\alpha_n (x_n)^p}{\sum_{k=1}^n \alpha_k}, \quad (\alpha_k > 0), \text{ odkiaľ dostaneme}$$

$$\alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_n x_n \leq \left(\sum_{k=1}^n \alpha_k \right)^{\frac{p-1}{p}} \left(\sum_{k=1}^n \alpha_k x_k^p \right)^{\frac{1}{p}}.$$

Ak uvážime, že $\frac{p-1}{p} = q$, položením $\alpha_k = b_k^q$, $x_k = a_k b_k^{1-q}$ pre $k = 1, 2, \dots, n$, dostaneme požadovanú nerovnosť.

Príklad 6 (Minkowského⁶ nerovnosť). Dokážte, že platí: *Pre ľubovoľné kladné čísla $a_1, a_2, \dots, a_n$, $b_1, b_2, \dots, b_n$ platí nerovnosť*

$$\sqrt[n]{a_1 \cdot a_2 \cdot \dots \cdot a_n} + \sqrt[n]{b_1 \cdot b_2 \cdot \dots \cdot b_n} \leq \sqrt[n]{(a_1 + b_1) \cdot (a_2 + b_2) \cdot \dots \cdot (a_n + b_n)}.$$

Dôkaz. Pri dôkaze tejto nerovnosti využijeme funkciu $f: f(x) = \ln(1 + e^x)$, $x \in (-\infty, \infty)$. Ľahko overíme, že táto funkcia je rýdzo konvexná na intervale $(-\infty, \infty)$, pretože $f''(x) = \frac{e^x}{(1+e^x)^2} > 0$, pre všetky $x \in (-\infty, \infty)$. Aplikovaním *Jensenovej nerovnosti II*, pričom $\alpha_1 = \alpha_2 = \dots = \alpha_n = \frac{1}{n}$ dostaneme pre ľubovoľné $x_1, x_2, \dots, x_n$ nerovnosť

$$\ln \left(1 + e^{\sum_{k=1}^n \frac{1}{n} x_k} \right) \leq \sum_{k=1}^n \frac{1}{n} \ln(1 + e^{x_k}).$$

Využitím základných vlastností logaritmickej funkcie a následným odlogaritmovaním získanej nerovnice dostaneme

$$1 + e^{\sum_{k=1}^n \frac{1}{n} x_k} \leq (1 + e^{x_1})^{\frac{1}{n}} \cdot (1 + e^{x_2})^{\frac{1}{n}} \cdot \dots \cdot (1 + e^{x_n})^{\frac{1}{n}}.$$

Zvolením $x_k = \ln \frac{b_k}{a_k}$ pre $k = 1, 2, \dots, n$ po jednoduchей úprave dostaneme nerovnosť

$$1 + e^{\ln \left(\frac{b_1}{a_1} \right)^{\frac{1}{n}}} \cdot e^{\ln \left(\frac{b_2}{a_2} \right)^{\frac{1}{n}}} \cdot \dots \cdot e^{\ln \left(\frac{b_n}{a_n} \right)^{\frac{1}{n}}} \leq \sqrt[n]{\left(1 + e^{\frac{b_1}{a_1}} \right) \left(1 + e^{\frac{b_2}{a_2}} \right) \cdot \dots \cdot \left(1 + e^{\frac{b_n}{a_n}} \right)},$$

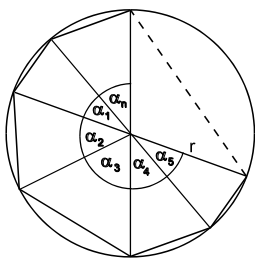
z ktorej po jednoduchей úprave a po vynásobení získanej nerovnosti vzt'ahom $\sqrt[n]{a_1 \cdot a_2 \cdot \dots \cdot a_n}$ dostaneme požadovanú nerovnosť.

Na záver tejto časti uved'me ešte riešenie jednej peknej extrémálnej úlohy.

Príklad 7. Dokážte, že zo všetkých konvexných n -uholníkov ($n \in \mathbb{N}, n \geq 3$) vpísaných do danej kružnice má pravidelný n -uholník najväčší obsah.

Dôkaz. Nech $n \in \mathbb{N}, n \geq 3$ ľubovoľne, ale pevne zvolené. Nech do kružnice so stredom S a polomerom r je vpísaný konvexný n -uholník. Vzhľadom na to, že hľadáme n -uholník s maximálnym obsahom, bez straty všeobecnosti môžeme predpokladať, že stred kružnice

⁶ Hermann Amandus Schwarz (1843-1921), Otto Ludwig Hölder (1859-1937) a Hermann Minkowski (1864 – 1909) boli významní nemeckí matematici



S je vnútorným bodom daného n -uholníka. V opačnom prípade je totiž možné zostrojiť vpísaný n -uholník s väčším obsahom, pričom stred kružnice S je už jeho vnútorným bodom. Z Obr. 6 ľahko nahliadneme, že každý takýto n -uholník sa skladá z n rovnoramenných trojuholníkov, ktoré majú v strede S danej kružnice spoločný vrchol. Označme jednotlivé stredové uhly (sú uhlami oproti základniám v jednotlivých

Obr. 6. Konvexný n -uholník

rovnoramenných trojuholníkov), postupne symbolmi: $\alpha_1, \alpha_2, \dots, \alpha_n$.

Zrejme $\alpha_1 + \alpha_2 + \dots + \alpha_n = 2\pi$, pričom $\alpha_k \in (0, \pi)$, pre $k = 1, 2, \dots, n$. Dôkaz bude ukončený, ak dokážeme, že n -uholník s maximálnym obsahom je taký, pre ktorý platí

$$\alpha_1 = \alpha_2 = \dots = \alpha_n.$$

Možno overiť, že obsah P_k rovnoramenného trojuholníka, ktorý prislúcha stredovému uhlu α_k sa rovná hodnote $\frac{1}{2}r^2 \sin \alpha_k$, preto pre obsah P celého vpísaného n -uholníka platí

$P = \sum_{k=1}^n P_k = \frac{1}{2}r^2 \sum_{k=1}^n \sin \alpha_k$. Ľahko overíme, že funkcia $f: f(x) = -\frac{1}{2}r^2 \sin x$ je v intervale $J = (0, \pi)$ rýdzo konvexná, preto pre túto funkciu platí Jensenova nerovnosť

$$f\left(\frac{\alpha_1 + \alpha_2 + \dots + \alpha_n}{n}\right) \leq \frac{f(\alpha_1) + f(\alpha_2) + \dots + f(\alpha_n)}{n}, \alpha_k \in (0, \pi), \text{ pre } k = 1, 2, \dots, n,$$

pričom rovnosť nastane vtedy a len vtedy, ak $\alpha_1 = \alpha_2 = \dots = \alpha_n$.

To však znamená, že pre obsah P ľubovoľného vpísaného n -uholníka platí nerovnosť

$P = \frac{1}{2}r^2(\sin \alpha_1 + \sin \alpha_2 + \dots + \sin \alpha_n) \leq \frac{1}{2}r^2 n \sin\left(\frac{\alpha_1 + \alpha_2 + \dots + \alpha_n}{n}\right) = \frac{1}{2}r^2 n \sin \frac{2\pi}{n}$, preto vždy platí $P \leq \frac{1}{2}r^2 n \sin \frac{2\pi}{n}$, pričom rovnosť nastane vtedy a len vtedy, ak

$\alpha_1 = \alpha_2 = \dots = \alpha_n = \frac{2\pi}{n}$. To je naozaj možné a nastane práve vtedy, ak príslušný vpísaný n -uholník je pravidelný.

Záver

Ukázali sme, že Jensenova nerovnosť je skutočne užitočným, účinným a najmä vysoko efektívnym nástrojom pre dokazovanie klasických nerovností, na ktorých sú založené dôkazy dôležitých viet, resp. s ich využitím je možné skracovať riešenia rôznych, najmä extrémálnych úloh (a to dokonca bez použitia derivácií). O význame konvexnosti funkcie svedčí fakt, že pri riešení extrémálnych problémov s často s úspechom využívané tvrdenie, v ktorej predpoklad konvexnosti funkcie (vo všeobecnosti funkcie n – premenných) je veľmi dôležitý. Svedčí o tom napríklad nasledovné tvrdenie, ktoré zovšeobecňuje známe tvrdenie z jednorozmerného prípadu, na prípad ľubovoľnej dimenzie n (dôkaz je možné nájsť napríklad v práci (Došlý, 2005) či (Borwein & Lewis, 2000)): Nech $f: D \rightarrow \mathbb{R}$ je konvexná funkcia na konvexnej množine $D \subset \mathbb{R}^n$. Potom ľubovoľné lokálne minimum funkcie f je aj globálnym minimom tejto funkcie. Množina bodov, v ktorých funkcia f nadobúda svoje minimum je konvexnou množinou. V prípade, že funkcia f je na množine rýdzo konvexná, tak táto množina nanajvýš jednoprvková. Ak funkcia f je navyše diferencovateľná na množine D a $x_0 \in D$ je jej stacionárnym bodom, potom bod x_0 je bodom globálneho minima.

Azda i toto dokumentuje, že konvexnosť funkcie je naozaj dôležitou vlastnosťou funkcie a určite stojí za to, aby sa ňou študenti bližšie zaoberali už v úvodných partiách matematickej analýzy. Bližšie oboznámenie sa s vlastnosťami konvexných funkcií, spolu s

matematickej analýzy. Bližšie oboznámenie sa s vlastnosťami konvexných funkcií, spolu s využitím *Jensenovej nerovnosti* pri odvodzovaní klasických nerovností, sa môže stať vhodnou propedeutikou pre štúdium konvexnej analýzy.

Literatúra

- [1] Beckenbach, E.F.- Bellman, R. (1961) Introduction to Inequalities. Springer-Verlag, Berlin, 1961.
- [2] Borwein J. M., Lewis A. S. (2000) Convex Analysis and Nonlinear Optimization. Springer-Verlag. New Yourk-Berlin-Heidelberg. 2000.
- [3] Došlý, O. (2005) Základy konvexnej analýzy a optimalizace v $\mathbb{R}^n$. Skriptum MU Brno, Brno 2005
- [4] Engel, A. (1998) Problem-Solving Strategies. Springer-Verlag, New York 1998.
- [5] Fulier, J. – Vrábek, P. (1997) Diferenciálny počet. UKF Nitra, Nitra 1997
- [6] Gunčaga, J. (2001) Jensenova nerovnosť, In: Disputationes Scientificalae, č. 2, 2001, KU, Ružomberok, s. 91 - 99.
- [7] Ižboldin, O.- Kurljandčik, L. (2000) Neravenstvo Jensena. Kvant, 2000/No 4, s. 7-10
- [8] Jensen J. L. W. V. (1906) Sur les fonctions convexes et les inégalités entre les valeurs moyennes. In: Acta Math. 30 (1906), strana 175 - 193
- [9] Kufner, A. (1975) Nerovnosti a odhady, ŠMM 39. Praha, Mladá fronta 1975
- [10] Larson, C. L. (1983) Problem Solving Through Problems, Springer Verlag GmbH & GKG New York Berlin Heidelberg Tokyo, 1983
- [11] Přinosil, M. (2008) Užití Jensenovy nerovnosti. Matematika - fyzika - informatika. č. 6/17, s. 321-330. Praha: Prometheus, 2008.
- [12] Solovjev, J.P. (2005) Neravenstva. Izd. Moskovskogo nepreryvnogo matematičeskogo obrazovania, Moskva 2005
- [13] Steel, J. M. (2004) The Cauchy-Schwarz Master Class - An Introduction to the Art of Mathematical Inequalities, Cambridge University Press, New York 2004.
- [14] Steinhaus, G. (1953) Matematický kaleidoskop. Přírodovědecké vydavatelství, Praha 1953

Článok prijatý dňa 1. júla 2012.

Adresa autora

*Prof. RNDr. Jozef Fulier, CSc.
Katedra matematiky
Fakulta prírodných vied
Univerzita Konštantína Filozofa
Tr. A. Hlinku 1
SK – 949 74 Nitra
e-mail: jfulier@ukf.sk*