Izpeljava Jensenove in Hölderjeve neenakosti ter neenakosti Minkowskega

1. Najosnovnejše o konveksnih funkcijah

Definicija. Naj bo X vektorski prostor in $D \subset X$ konveksna množica. Funkcija $\varphi\colon D \to \mathbb{R}$ je *konveksna*, če za poljubna $a, b \in D$ ter poljuben $\theta \in [0,1]$ velja:

$$\varphi\big((1-\theta)a + \theta b\big) \leq (1-\theta)\varphi(a) + \theta\varphi(b) \tag{1.1}$$

Funkcija φ je torej konveksna, če za poljubna a in b graf njene zožitve na daljico ab leži pod daljico s krajiščema $\big(a, \varphi(a)\big)$ in $\big(b, \varphi(b)\big)$. Slika!

Konveksne funkcije, definirane na intervalih na realni osi, lahko karakteriziramo kot funkcije, ki imajo naraščajoč diferenčni količnik.

Trditev 1.1. *Naj bo $I \subset \mathbb{R}$ interval in naj bo dana funkcija $\varphi\colon I \to \mathbb{R}$. Naslednje trditve so ekvivalentne.*

(1) *Funkcija φ je konveksna na intervalu I.*

(2) *Za poljubna števila $a < b$ in $c \in [a,b]$ iz intervala I velja:*

$$\varphi(c) \leq \frac{b-c}{b-a}\,\varphi(a) + \frac{c-a}{b-a}\,\varphi(b) \tag{1.2}$$

(3) *Za poljubna števila $a_1, a_2, b_1, b_2 \in I$, za katera je $a_1 \leq a_2$, $b_1 \leq b_2$, $a_1 < b_1$ in $a_2 < b_2$, velja (slika!):*

$$\frac{\varphi(b_1) - \varphi(a_1)}{b_1 - a_1} \leq \frac{\varphi(b_2) - \varphi(a_2)}{b_2 - a_2} \tag{1.3}$$

Dokaz. (1) ⇒ (2): V enačbo (1.1) vstavimo $\theta := (c-a)/(b-a)$ in dobimo (1.2).
(2) ⇒ (1): V enačbo (1.2) vstavimo $c := (1-\theta)a + \theta b$ in dobimo (1.1).
(2) ⇒ (3): Za $a < c \leq b$ iz (1.2) dobimo (slika!):

$$\frac{\varphi(c) - \varphi(a)}{c - a} \leq \frac{\varphi(b) - \varphi(a)}{b - a} \tag{1.4}$$

za $a \leq c < b$ pa prav tako iz (1.2) dobimo (slika!):

$$\frac{\varphi(b) - \varphi(a)}{b - a} \leq \frac{\varphi(b) - \varphi(c)}{b - c} \tag{1.5}$$

Iz (1.4) in (1.5) zdaj izpeljemo:

$$\frac{\varphi(b_1) - \varphi(a_1)}{b_1 - a_1} \leq \frac{\varphi(b_2) - \varphi(a_1)}{b_2 - a_1} \leq \frac{\varphi(b_2) - \varphi(a_2)}{b_2 - a_2}$$

kar je natančno (3).
(3) ⇒ (2): V (1.3) vstavimo npr. $a_1 = a$, $a_2 = b_1 = c$ in $b_2 = b$, pa dobimo (1.2). ■

Posledica 1.2. *Funkcija, ki je definirana na odprtem intervalu in tam odvedljiva, je konveksna natanko tedaj, ko je njen odvod naraščajoča funkcija.*

Posledica 1.3. *Funkcija φ, ki je definirana na odprtem intervalu in tam dvakrat odvedljiva, je konveksna natanko tedaj, ko je $\varphi'' \geq 0$.*

Trditev 1.4. *Naj bo $I \subset \mathbb{R}$ interval, $\varphi\colon I \to \mathbb{R}$ pa konveksna funkcija. Tedaj za vsak c iz notranjosti intervala I obstaja taka linearna funkcija $\ell\colon \mathbb{R} \to \mathbb{R}$, da velja $\ell(c) = \varphi(c)$ in $\ell(x) \leq \varphi(x)$ za vsak $x \in I$.* Slika!

Dokaz. Iz neenačbe (1.3) sledi:

$$\alpha := \sup_{a<c} \frac{\varphi(a) - \varphi(c)}{a - c} \leq \inf_{b>c} \frac{\varphi(b) - \varphi(c)}{b - c} =: \beta \tag{1.6}$$

Izberimo poljubno število $\alpha \leq k \leq \beta$ in definirajmo:

$$\ell(x) := \varphi(c) + k(x - c)$$

Brez težav preverimo, da funkcija ℓ izpolnjuje zahtevana pogoja. ■

Opomba. Iz slednjega dokaza je razvidno, da ima vsaka konveksna funkcija realne spremenljivke v notranjih točkah definicijskega območja levi in desni odvod: to sta namreč koeficienta α in β iz (1.6).

Trditev 1.5. *Če je $I \subset \mathbb{R}$ odprt interval, je vsaka konveksna funkcija $\varphi\colon I \to \mathbb{R}$ zvezna.*

Dokaz. Naj bo $c \in I$. Ker je interval I odprt, obstajata taki števili $a, b \in I$, da je $a < c < b$. Naj bo $c < x < b$. Po trditvi 1.1 velja:

$$\frac{\varphi(c) - \varphi(a)}{c - a} \leq \frac{\varphi(x) - \varphi(c)}{x - c} \leq \frac{\varphi(b) - \varphi(c)}{b - c}$$

od koder sledi, da je φ v točki c zvezna z desne (slika!). Podobno dokažemo, da je φ zvezna tudi z leve. ■

2. Jensenova neenakost

Izrek 2.1 (Jensenova neenakost). *Naj bo $(X, \mathcal{F})$ merljiv prostor s pozitivno mero μ, za katero velja $\mu(X) = 1$. Dan naj bo še interval $I \subset \mathbb{R}$, merljiva funkcija $g\colon X \to I$ in konveksna funkcija $\varphi\colon I \to \mathbb{R}$. Tedaj velja:*

$$\varphi\left(\int g\, d\mu\right) \leq \int (\varphi \circ g)\, d\mu \tag{2.1}$$

če seveda integrala obstajata.

Opomba. Iz predpostavk izreka avtomatično sledi, da je tudi funkcija $\varphi \circ g$ merljiva, saj je funkcija φ vsaj v notranjosti intervala I zvezna in zato merljiva. Prav tako sledi, da je $\int g\, d\mu \in I$.

Dokaz izreka 2.1. Najprej opazimo, da v (2.1) velja enakost, če je funkcija φ linearna. Po trditvi 1.4 obstaja taka linearna funkcija ℓ, da velja $\ell(\int g\, d\mu) = \varphi(\int g\, d\mu)$ in $\ell(x) \leq \varphi(x)$ za vsak $x \in I$. Sledi:

$$\varphi\left(\int g\, d\mu\right) = \ell\left(\int g\, d\mu\right) = \int (\ell \circ g)\, d\mu \leq \int (\varphi \circ g)\, d\mu$$

in izrek je dokazan. ■

Jensenova neenakost je posplošitev številnih drugih neeenakosti. V klasični, manj splošni obliki je Jensenova neenakost formulirana na naslednji način.

Posledica 2.2. *Naj bo $I \subset \mathbb{R}$ interval, $\varphi\colon I \to \mathbb{R}$ pa konveksna funkcija. Tedaj za poljubne točke $x_1, \ldots x_n \in I$ in poljubna števila $\alpha_1, \ldots \alpha_n \geq 0$ z vsoto 1 velja:*

$$\varphi\left(\alpha_1 x_1 + \ldots \alpha_n x_n\right) \leq \alpha_1 \varphi(x_1) + \ldots + \alpha_n \varphi(x_n) \tag{2.2}$$

Kot zgled lahko vzamemo eksponentno funkcijo $\varphi(x) = e^x$, ki je očitno konveksna, saj ima pozitiven drugi odvod. Če vstavimo še $x_i = \ln y_i$, kjer so y_i pozitivna realna števila, dobimo neenakost med aritmetično in geometrično sredino:

$$y_1^{\alpha_1} \ldots y_n^{\alpha_n} \leq \alpha_1 y_1 + \ldots \alpha_n y_n \tag{2.3}$$

Zgornja enakost velja za poljubne $y_1, \ldots y_n \geq 0$.

3. Hölderjeva neenakost

Naj bo $(X, \mathcal{F})$ merljiv prostor s pozitivno mero μ in naj bo $f\colon X \to \mathbb{F}$, kjer je $\mathbb{F} \in \{\mathbb{R}, \mathbb{C}\}$, merljiva funkcija. Za $1 \leq p < \infty$ označimo:

$$\|f\|_p := \left(\int f^p\, d\mu\right)^{1/p}$$

Označimo še:

$$\|f\|_\infty := \operatorname{ess\,sup}|f|$$

Nadalje označimo:

$$L^p(\mu) := \{f\colon X \to \mathbb{F} \text{ merljiva}\ ; \|f\|_p < \infty\}$$

Definicija. Naj bo $1 \leq p, q, \leq \infty$. Števili p in q imenujemo *konjugirana eksponenta*, če je bodisi eden od njiju enak 1, drugi pa ∞, bodisi $1/p + 1/q = 1$.

Izrek 3.1 (Hölderjeva neenakost). *Naj bosta p in q konjugirana eksponenta, dani pa naj bosta še funkciji $f \in L^p(\mu)$ in $g \in L^q(\mu)$. Tedaj je funkcija fg absolutno integrabilna glede na μ in velja:*

$$\left| \int fg \, d\mu \right| \leq \|f\|_p \|g\|_q \tag{3.1}$$

Preden gremo dokazat izrek, bomo izpeljali še eno neenakost. Naj bo $x, y \geq 0$ in naj bosta $1 < p, q < \infty$ konjugirana eksponenta. Če v neenakost (2.3) vstavimo $n = 2$, $y_1 = x^p$, $y_2 = y^q$, $\alpha_1 = 1/p$ in $\alpha_2 = 1/q$, dobimo *Youngovo neenakost*:

$$xy \leq \frac{x^p}{p} + \frac{y^q}{q} \tag{3.2}$$

Zgornjo neenakost pa lahko še malo izostrimo: vzemimo $k > 0$ in spremenljivko x pomnožimo, y pa delimo s k. Dobimo:

$$xy \leq \frac{k^p}{p} x^p + \frac{1}{qk^q} y^q \tag{3.3}$$

Zgornja ocena je maksimalna možna izostritev ocene (3.2), saj velja naslednja trditev.

Trditev 3.2. *Za poljubni števili $x, y \geq 0$ in konjugirana eksponenta $1 < p, q < \infty$ velja:*

$$xy = \inf_{k>0} \left[\frac{k^p}{p} x^p + \frac{1}{qk^q} y^q \right] \tag{3.4}$$

Dokaz. Dovolj je dokazati, da za poljubna $x, y \geq 0$ v oceni (3.3) za neki $k > 0$ ali vsaj v limiti velja enakost. Če je $x = 0$, velja enakost v limiti, ko gre k proti neskončno; če je $y = 0$, velja enakost v limiti, ko gre k proti nič. Za $x, y > 0$ pa enakost velja, če postavimo:

$$k := \frac{y^{1/p}}{x^{1/q}}$$

∎

Dokaz izreka 3.1. Dovolj je dokazati neenakost (3.1), in sicer le za primer, ko je $f, g \geq 0$. Za primer, ko je kateri od konjugiranih eksponentov neskončen, je neenakost očitna. Za $1 < p, q < \infty$ pa po (3.3) za vsak $x \in X$ in vsak $k > 0$ velja:

$$f(x)g(x) \leq \frac{k^p}{p} f(x)^p + \frac{1}{qk^q} g(x)^q$$

Ko neenakost integriramo po μ, dobimo:

$$\int fg \, d\mu \leq \frac{k^p}{p} \|f\|_p^p + \frac{1}{qk^q} \|g\|_q^q$$

Ko naredimo infimum po k in uporabimo trditev 3.2, dobimo natančno to, kar smo iskali. ∎

Hölderjeva neenakost nam da še drugačno karakterizacijo količin $\|f\|_p$, ki je tesno povezana z dualnostjo prostorov L^p in L^q.

Trditev 3.3. *Naj bosta $1 \leq p < \infty$ in $1 < q \leq \infty$ konjugirana eksponenta. Tedaj za vsako funkcijo $f \in L^p(\mu)$ velja enakost:*

$$\|f\|_p = \sup\left\{\left|\int fg\,d\mu\right|; \|g\|_q = 1\right\} \tag{3.5}$$

Opomba. Za $p = \infty$ in $q = 1$ enakost ne velja vedno. Kot protiprimer vzemimo za X poljubno neštevno množico, za σ-algebro $\mathcal{F}$ pa vzamemo vse podmnožice množice X, ki so bodisi končne ali števno neskončne bodisi so taki njihovi komplementi. Mera μ naj bo na končnih in števno neskončnih množicah enaka 0, sicer pa ∞. Če vzamemo $f(x) = 1$ za vsak $x \in X$, ta funkcija očitno pripada $L^\infty(\mu)$ in velja $\|f\|_\infty = 1$, enakost 3.5 pa ne velja, saj prostor $L^1(\mu)$ vsebuje le funkcije, ki so skoraj povsod enake 0.

Dokaz trditve 3.3. Dovolj se je omejiti na primer, ko funkcija f ni skoraj povsod enaka 0. V tem primeru je $\|f\|_p > 0$. Dovolj je dokazati, da obstaja taka funkcija g z $\|g\|_q = 1$, za katero v oceni (3.1) velja enakost. Ločimo dva primera.

$p = 1, q = \infty$. V tem primeru postavimo:

$$g(x) := \begin{cases} \frac{|f(x)|}{f(x)}; & f(x) \neq 0 \\ 0; & \text{sicer} \end{cases}$$

Očitno je $\|g\|_\infty = 1$ in velja $fg = |f|$, zato v (3.1) velja enakost.

$1 < p, q < \infty$. V tem primeru postavimo:

$$g(x) := \begin{cases} \frac{|f(x)|^p}{f(x)\|f\|_p^{p/q}}; & f(x) \neq 0 \\ 0; & \text{sicer} \end{cases}$$

Tedaj za vsak $x \in X$ velja:

$$|g(x)|^q = \frac{|f(x)|^{(p-1)q}}{\|f\|_p^p}$$

Ker sta p in q konjugirana, je $(p-1)q = p$, zato je $\|g\|_q = 1$. Prav tako za vsak $x \in X$ pa velja tudi:

$$f(x)g(x) = \frac{|f(x)|^p}{\|f\|_p^{p/q}}$$

zato je končno:

$$\int fg\,d\mu = \frac{\|f\|_p^p}{\|f\|_p^{p/q}} = \|f\|_p$$

Trditev je dokazana. ■

Vaja. Dokaži, da trditev 3.3 velja tudi za primer, ko je $p = \infty$, $q = 1$, mera μ pa je σ-končna.

4. Neenakost Minkowskega

Izrek 4.1 (Neenakost Minkowskega). *Naj bo $(X, \mathcal{F})$ merljiv prostor s pozitivno mero μ. Tedaj za vsak $1 \leq p \leq \infty$ in poljubni funkciji $f, g \in L^p(\mu)$ velja:*

$$\|f + g\|_p \leq \|f\|_p + \|g\|_p \tag{4.1}$$

Neenakost Minkowskega torej pove, da je predpis $f \mapsto \|f\|_p$ seminorma na prostoru $L^p(\mu)$. Seveda to ni vedno norma, saj je enaka 0 na prostoru $N(\mu)$ vseh funkcij, ki so skoraj povsod glede na μ enake 0. Pač pa naša seminorma inducira normo na kvocientnem prostoru $\mathcal{L}^p(\mu) = L^p(\mu)/N(\mu)$.

Dokaz izreka 4.1. Za $p = \infty$ je enakost očitna. Naj bo zdaj $1 \leq p < \infty$. Najprej bomo pokazali, da je sploh $f + g \in L^p(\mu)$. Iz neenakosti (2.2) za $\varphi(x) = |x|^p$ sledi:

$$\left|\frac{f+g}{2}\right|^p \leq \frac{|f|^p + |g|^p}{2}$$

torej je $\|f + g\|_p^p \leq 2^{p-1}(\|f\|_p^p + \|g\|_p^p) < \infty$. Zdaj, ko vemo, da so vse tri funkcije v $L^p(\mu)$, pa lahko uporabimo trditev 3.3. Po Hölderjevi neenakosti za vsako funkcijo h, za katero je $\|h\|_q = 1$, velja:

$$\left|\int (f+g)h \, d\mu\right| \leq \left|\int fh \, d\mu\right| + \left|\int gh \, d\mu\right| \leq \|f\|_p + \|g\|_p$$

Ko naredimo supremum po h, dobimo natančno to, kar smo iskali. ■