Nierówności, wypukłość i ekstrema

Jarosław GÓRNICKI, Rzeszów

Analizując konkretny problem niejednokrotnie musimy szukać wielkości ekstremalnych, rozstrzygać o ich ilości, wreszcie wyznaczać je. Najczęściej robimy to tak: staramy się opisać dane zjawisko za pomocą funkcji rzeczywistych różniczkowalnych, a następnie, posługujac się dobrze znanymi metodami rachunku różniczkowego, udzielamy odpowiedzi na postawione pytania. Z nauki szkolnej można nawet wynieść wrażenie, że jest to jedyny sposób podejścia do tego typu zadań.

Oczywiście tak nie jest, niektóre problemy ekstremalne, dotyczące funkcji rzeczywistych jednej i wielu zmiennych rzeczywistych, można rozstrzygać innymi metodami. Wymaga to pewnej fantazji, pomysłowości – ale czy nie jest to urok matematyki?

I. Zastosowanie prostych nierówności

Przykład 1

Oto elementarne rozumowanie: jeżeli $x+y=s$ oraz $x=y=\frac{s}{2}$, to $x\cdot y=\frac{s}{4}^2$. Gdy $x\neq y$, to $x\neq\frac{s}{2}$ i $(x-\frac{s}{2})^2>0$. Stąd oczywiście $x\cdot s-x^2<\frac{s}{4}^2$ i $x\cdot y= =x(s-x)$, więc $x\cdot y<\frac{s}{4}^2$.

Widzimy zatem, że *iloczyn dwóch dodatnich liczb rzeczywistych o danej sumie ma wartość największą, gdy czynniki są równe.* Geometrycznie oznacza to, że:

Ze wszystkich prostokątów o danym obwodzie największe pole ma kwadrat!

Przykład 2

Udowodnij, że iloczyn trzech liczb rzeczywistych dodatnich o danej sumie ma wartość największą, gdy czynniki są równe.

Dowód. Dla dowolnych liczb rzeczywistych x, y prawdziwa jest nierówność $x^2+y^2\geq 2xy$, a w konsekwencji

$$(x+y)^2=x^2+2xy+y^2\geq 4xy,$$

$$(x+y)^2\geq 4xy. \tag{1}$$

Niech teraz x, y, z będą liczbami rzeczywistymi dodatnimi takimi, że $x+y+z=s$. Jeżeli $x\neq y$, to $(x-y)^2>0$ oraz z (1) mamy $(x+y)^2>4xy$. Niech $3d=x+y+z$, zatem $(z+d)^2\geq 4zd$, $((x+y)+(z+d))^2\geq 4(x+y)(z+d)$, a stąd $(x+y+z+d)^4>4^2\cdot 4xy\cdot 4zd$, czyli $(x+y+z+d)^4>4^4xyzd$. Ponieważ $x+y+z=3d$, więc $(4d)^4>4^4xyzd$, a zatem $d^3>xyz$. Innymi słowy

$$xyz<\left(\frac{x+y+z}{3}\right)^3=\frac{s^3}{27}.$$

Gdy $x=y=z=\frac{s}{3}$, to $xyz=\frac{s^3}{27}$, co kończy dowód.

W dalszej części wykażemy, że prawdziwy jest wynik ogólniejszy: *iloczyn $n>1$ liczb rzeczywistych dodatnich o danej sumie ma wartość największą, gdy czynniki są równe.*

Tymczasem wykorzystamy fakt podany w przykładzie 2.

Przykład 3

Ze wszystkich prostopadłościanów o danej powierzchni największą objętość ma sześcian.

Dowód. Niech x, y, z będą długościami boków prostopadłościanu, zaś $6p$ jego powierzchnią. Mamy więc $xy+yz+zx=3p$.

Jeżeli $x=y=z$, to $x^2=p$, a więc $xyz=\sqrt{p^3}$. Gdy $x\neq y$, to również $\frac{1}{x}\neq\frac{1}{y}$ i na podstawie rozważań z przykładu 2,

$$\frac{27}{xyz}<\left(\frac{1}{x}+\frac{1}{y}+\frac{1}{z}\right)^3.$$

Ponieważ
$$\frac{1}{x}+\frac{1}{y}+\frac{1}{z}=\frac{xy+yz+zx}{xyz}=\frac{3p}{xyz},$$
więc $\frac{27}{xyz} < \left(\frac{3p}{xyz}\right)^3$, skąd $(xyz)^2 < p^3$, czyli $xyz < \sqrt{p^3}$.

II. Funkcje wypukłe i nierówność Jensena

Poznając matematykę spotykamy sie między innymi ze *zbiorami wypukłymi*, *funkcjami wypukłymi* [3]. Często nie doceniamy znaczenia tych pojęć. Herman Minkowski badając zbiory wypukłe wyraził swój zachwyt mówiąc: „interesuje mnie wszystko, co jest wypukłe".

Niech $I \subset \mathbf{R}$ będzie przedziałem. Funkcję $f : I \to \mathbf{R}$ nazywamy *wypukłą* w przedziale I, jeżeli dla dowolnych $x, y \in I$, $t \in [0,1]$,
$$(*) \qquad f(tx+(1-t)y) \le t \cdot f(x) + (1-t) \cdot f(y).$$
Gdy dla $x \ne y$, $t \in (0,1)$ jest nierówność ostra, wtedy funkcję f nazywamy *ściśle wypukłą*.

Geometrycznie oznacza to, że wykres funkcji f na dowolnym przedziale $[x,y]$, $x, y \in I$ leży pod lub na cięciwie o końcach $(x, f(x))$, $(y, f(y))$.

Siłę tego pojęcia ujawnia stwierdzenie: każda *funkcja rzeczywista zmiennej rzeczywistej wypukła w przedziale otwartym jest ciągła w tym przedziale.*

Wypukłość funkcji można też rozważać w innym znaczeniu: mówimy, że funkcja $f : \mathbf{R} \supset I \to \mathbf{R}$ jest *wypukła w sensie Jensena* w przedziale I, jeśli dla dowolnych $x, y \in I$,
$$f\left(\frac{x+y}{2}\right) \le \frac{1}{2}(f(x)+f(y)).$$
Gdy dla $x \ne y$ nierówność ta jest ostra, to mówimy o ścisłej wypukłości w *sensie Jensena*.

Związek pomiędzy tymi określeniami wypukłości jest następujący: *jeżeli funkcja f, określona w przedziale otwartym I, jest wypukła w sensie Jensena i jest ciągła, to spełnia ona warunek $(*)$ dla dowolnych $x, y \in I$, $t \in [0,1]$.*

Teraz naszą uwagę skierujemy na tzw. nierówność Jensena.

Twierdzenie 1 (nierówność Jensena, 1905). *Jeżeli $f : I \to \mathbf{R}$ jest funkcją wypukłą na przedziale $I \subset \mathbf{R}$, to dla różnych $x_1, x_2, \ldots, x_n \in I$, $n \ge 2$, oraz $0 < t_i < 1$ takich, że $\sum_{i=1}^{n} t_i = 1$ spełniona jest nierówność:*
$$f\left(\sum_{i=1}^{n} t_i x_i\right) \le \sum_{i=1}^{n} t_i \cdot f(x_i).$$
Gdy f jest ściśle wypukła, to nierówność jest ostra, zaś równość ma miejsce jedynie w przypadku, gdy $x_1 = x_2 = \ldots = x_n$.

Dowód. Dla $n = 2$ twierdzenie jest prawdziwe, gdyż nierówność Jensena pokrywa się z definicją funkcji wypukłej. Załóżmy prawdziwość twierdzenia dla $n-1$. Wtedy dla dowolnych $0 < t_i < 1$ takich, że $\sum_{i=1}^{n} t_i = 1$ mamy
$$a = \sum_{i=1}^{n-1} t_i > 0, \text{ a ponadto punkt } x = \frac{1}{a} \cdot \sum_{i=1}^{n-1} t_i x_i \text{ należy do } I.$$
Na mocy założenia indukcyjnego,
$$a \cdot f\left(\sum_{i=1}^{n-1} \frac{t_i}{a} x_i \cdot x\right) \le \sum_{i=1}^{n-1} t_i \cdot f(x_i).$$

Z wypukłości funkcji f, mamy

$$f\left(\sum_{i=1}^{n} t_i x_i\right) = f(ax + (1-a)x_n) \leq a \cdot f(x) + (1-a) \cdot f(x_n) \leq \sum_{i=1}^{n} t_i \cdot f(x_i),$$

co na mocy zasady indukcji matematycznej kończy dowód.

Nierówność ta jest źródłem wielu innych ważnych nierówności. Na przykład funkcje wykładnicze $x \to a^x$ $(0 < a \neq 1)$ są ciągłe oraz ściśle wypukłe w sensie Jensena na zbiorze $\mathbb{R}$: gdy $x \neq y$, to $(a^{x/2} - a^{y/2})^2 > 0$, a więc $a^{(x+y)/2} < \frac{1}{2}(a^x + a^y)$. Wobec tego:

dla liczb $0 < p_i < 1$ *takich, że* $\sum_{i=1}^{n} p_i = 1$, $x_i > 0$, $i = 1, 2, \ldots, n$,

$$x_1^{p_1} \cdot x_2^{p_2} \cdot \ldots \cdot x_n^{p_n} \leq p_1 x_1 + p_2 x_2 + \ldots + p_n x_n,$$

przy czym równość ma miejsce tylko wtedy, gdy $x_1 = x_2 = \ldots = x_n$.

Dowód. Wystarczy tak wybrać liczby $t_i, i = 1, 2, \ldots, n$, że $x_i = 10^{t_i}$ i zastosować nierówność Jensena do funkcji wykładniczej.

Stąd, dla $p_1 = p_2 = \ldots = p_n = \frac{1}{n}$, otrzymujemy nierówność Augustyna Cauchy'ego z 1821 roku pomiędzy średnimi: geometryczną i arytmetyczną:

jeżeli $\mathbb{R} \ni x_i > 0$, $i = 1, 2, \ldots, n$, *to*

$$\frac{x_1 + x_2 + \ldots + x_n}{n} \geq \sqrt[n]{x_1 x_2 \ldots x_n}$$

i równość ma miejsce jedynie, gdy $x_1 = x_2 = \ldots = x_n$.

Korzystając z nierówności Cauchy'ego można w łatwy sposób rozwiązać zadanie z przykładu 2.

Przykład 4

Iloczyn $n > 1$ *liczb rzeczywistych dodatnich o danej sumie ma wartość największą, gdy czynniki są równe.*

Dowód. Przyjmijmy, że $x_1 + x_2 + \ldots + x_n = M$. Na podstawie nierówności Cauchy'ego:

$$\sqrt[n]{x_1 x_2 \ldots x_n} \leq \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{M}{n},$$

więc $x_1 x_2 \ldots x_n \leq \left(\frac{M}{n}\right)^n$ i równość ma miejsce jedynie w przypadku, gdy $x_1 = x_2 = \ldots = x_n = \frac{m}{n}$.

Geometrycznie oznacza to na przykład, że:

ze wszystkich prostopadłościanów o stałej sumie długości krawędzi największą objętość ma sześcian.

Dla Czytelnika znającego wzór Herona na obliczenie pola trójkąta wobec powyższego rozumowania oczywiste jest, że:

spośród wszystkich trójkątów o danym obwodzie największe pole ma trójkąt równoboczny.

Przykład 5

Jeżeli iloczyn $x_1 x_2 \cdot \ldots \cdot x_n$ *liczb dodatnich jest stały, to suma* $x_1 + x_2 + \ldots + x_n$ *ma najmniejszą wartość, gdy* $x_1 = x_2 = \ldots = x_n$.

Dowód. Przyjmijmy $c = x_1 x_2 \cdot \ldots \cdot x_n$. Stosując nierówność Cauchy'ego dostajemy $\frac{x_1 + x_2 + \ldots + x_n}{n} \geq \sqrt[n]{c}$ i równość ma miejsce tylko dla $x_1 = x_2 = \ldots = x_n$.

Oto inne zastosowanie nierówności Jensena ([5]):

Przykład 6

Rozważmy dowolny n-kąt wpisany w koło o promieniu $r > 0$ i niech $x_1, x_2, \ldots, x_n$ będą miarami kątów środkowych opartych na kolejnych bokach wielokąta. Zatem $x_i \in [0, \pi]$ oraz $x_1 + \ldots + x_n = 2\pi$. Pole rozważanego wielokąta jest więc równe

$$\frac{1}{2} r^2 (\sin x_1 + \ldots + \sin x_n).$$

Ponieważ funkcja sinus jest *wklęsła* (we wzorze (*) nierówność jest przeciwna) w przedziale $[0,\pi]$, więc dla liczb $t_1 = t_2 = \ldots = t_n = \frac{1}{n}$ otrzymujemy:

$$\frac{1}{n}(\sin x_1 + \ldots + \sin x_n) \le \sin\left(\frac{1}{n}(x_1 + \ldots + x_n)\right) = \sin\frac{2\pi}{n}.$$

Wobec tego pole n-kąta wpisanego w koło o promieniu r jest nie większe niż $\frac{1}{2}r^2 n \cdot \sin\frac{2\pi}{n}$, przy czym maksimum jest osiągnięte gdy $x_1 = x_2 = \ldots = x_n$. Oznacza to, że:

w zbiorze wszystkich n-kątów ($n \ge 3$) wpisanych w dane koło o promieniu r największe pole ma n-kąt foremny.

III. Nierówność Bernoulliego

Chęć rozwiązania zadania ekstremalnego bez użycia rachunku różniczkowego może wymagać zastosowania specjalnie opracowanych nierówności. Oto jak można wykorzystać nierówności Bernoulliego w wersji ogólniejszej (nie tylko dla liczb naturalnych – [2]).

Twierdzenie 2 (nierówności Bernoulliego). *Jeżeli $x \ge -1$ i $0 < a < 1$, to wtedy*

(1) $$(1+x)^a \le 1 + ax,$$

natomiast jeżeli $a < 0$ lub $a > 1$, to

(2) $$(1+x)^a \ge 1 + ax.$$

Równość we wzorach (1) *i* (2) *ma miejsce tylko, gdy $x = 0$.*

Dowód. Załóżmy, że a z przedziału $(0,1)$ jest liczbą wymierną postaci $a = \frac{m}{n}$, gdzie m, n są liczbami naturalnymi i $1 \le m < n$. Wówczas

$$(1+x)^{\frac{m}{n}} = \sqrt[n]{(1+x)^m \cdot 1^{n-m}} = \sqrt[n]{\underbrace{(1+x)\ldots(1+x)}_{m} \cdot \underbrace{1\cdot\ldots\cdot 1}_{n-m}} \le$$

$$\le \frac{(1+x) + (1+x) + \ldots + (1+x) + 1 + \ldots + 1}{n} =$$

$$= \frac{m(1+x) + n - m}{n} = 1 + \frac{m}{n}x.$$

Znak równości otrzymujemy jedynie w przypadku, gdy wszystkie czynniki pod pierwiastkiem są identyczne, tj. $1 + x = 1$, $x = 0$. Jeżeli $x \ne 0$, to nierówność jest ostra. Przyjmijmy teraz, że a z przedziału $(0,1)$ jest liczbą niewymierną. Niech $r_1, r_2, \ldots$ będzie ciągiem liczb wymiernych z tego przedziału, których granicą jest liczba a. Wtedy

$$(1+x)^a = \lim_{n\to\infty}(1+x)^{r_n} \le \lim_{n\to\infty}(1 + r_n x) = 1 + ax.$$

Uzasadnienia wymaga teraz stwierdzenie, że nierówność ta jest ostra dla $x \ne 0$. Wybieramy liczbę wymierną r taką, że $a < r < 1$. Wówczas $(1+x)^{\frac{a}{r}} \le 1 + \frac{a}{r}x$, a stąd $(1+x)^a \le (1 + \frac{a}{r}x)^r$. Jeżeli $x \ne 0$, to $(1 + \frac{a}{r}x)^r < 1 + r\frac{a}{r}x = 1 + ax$, co kończy dowód pierwszej części.

Jeżeli $1 + ax < 0$, to nierówność (2) jest oczywista. Jeżeli $1 + ax \ge 0$, czyli $ax \ge -1$, to rozpatrzymy dwa przypadki.

Przypuśćmy, że $a > 1$. Wtedy z dowiedzionej już części, mamy

$$(1+ax)^{\frac{1}{a}} \le 1 + \frac{1}{a}ax = 1 + x.$$

Równość ma miejsce jedynie dla $x = 0$. Podnosząc obie strony ostatniej nierówności do potęgi a, otrzymujemy

$$1 + ax \le (1+x)^a.$$

Przypuśćmy teraz, że $a < 0$. Wybieramy liczbę $n \in \mathbb{N}$ taką, by $-\frac{a}{n} < 1$. Wówczas z pierwszej części dowodu, mamy

$$(1+x)^{-\frac{a}{n}} \le 1 - \frac{a}{n}x,$$

$$(1+x)^{\frac{a}{n}} \ge \frac{1}{1 - \frac{a}{n}x} \ge 1 + \frac{a}{n}x.$$

Podnosząc obie strony ostatniej nierówności do n-tej potęgi uzyskujemy

$$(1+x)^a \geq \left(1+\frac{a}{n}x\right)^n \geq 1+n\frac{a}{n}x = 1+ax,$$

i równość ma miejsce jedynie dla $x=0$.

Twierdzenie 3. *Jeżeli $a>0$, $b>1$, to funkcja*

$$f(x) = x^b - a\cdot x \ dla \ x \geq 0$$

osiąga minimum w punkcie $x = (a/b)^{\frac{1}{b-1}}$.

Dowód. Dla $b=2$ wynika to bezpośrednio z równości

$$x^2 - a\cdot x = \left(x-\frac{a}{2}\right)^2 - \frac{a^2}{4}.$$

W przypadku ogólnym posłużymy się nierównością Bernoulliego w postaci: *jeżeli $x \geq -1$ i $b>1$, to*

$$(1+x)^b \geq 1+bx$$

(równość ma miejsce tylko, gdy $x=0$).

Ponieważ $b>1$, więc $(1+z)^b \geq 1+bz$ dla $z \geq -1$, przy czym równość ma miejsce jedynie dla $z=0$. Przyjmując $1+z=y$, mamy $y^b \geq 1+b(y-1)$, $y^b - by \geq 1-b$ dla $y \geq 0$, przy czym równość ma miejsce jedynie dla $y=1$. Mnożąc ostatnią nierówność przez c^b ($c>0$), otrzymujemy

$$(cy)^b - b\cdot c^{b-1}(cy) \geq (1-b)\cdot c^b.$$

Przyjmując $x=cy$, $b\cdot c^{b-1}=a$ (stąd $c=(a/b)^{\frac{1}{b-1}}$), mamy

$$x^b - ax \geq (1-b)\cdot (a/b)^{\frac{b}{b-1}}$$

i równość ma miejsce tylko, gdy $x = c = (a/b)^{\frac{1}{b-1}}$, co kończy dowód.

Uwaga. Funkcja $g(x) = ax - x^b = -f(x)$ osiąga wartość największą w punkcie $x = (a/b)^{\frac{1}{b-1}}$.

Przejdźmy wreszcie do zastosowań ostatniego twierdzenia.

Przykład 7
Nie korzystając z rachunku różniczkowego wyznaczyć największą wartość funkcji $f(x) = \sin x \cdot \sin 2x$, $x \in \mathbf{R}$.

Rozwiązanie. Ponieważ $\sin 2x = 2\cdot \sin x \cdot \cos x$, więc

$$\sin x \cdot \sin 2x = 2\cdot \cos x \cdot \sin^2 x = 2\cdot \cos x \cdot (1-\cos^2 x) = 2(z-z^3),$$

gdzie $z = \cos x$, a stąd $-1 \leq z \leq 1$. Funkcja $z - z^3 = z(1-z^2)$ przyjmuje wartości ujemne dla $-1 \leq z < 0$, jest równa 0 dla $z=0$ i przyjmuje wartości dodatnie dla $0 < z \leq 1$. Zgodnie z uwagą po twierdzeniu 3 funkcja $z-z^3$, $z \geq 0$, przyjmuje wartość największą w punkcie $z_0 = \frac{1}{\sqrt{3}}$. Zatem funkcja $f(x) = \sin x \cdot \sin 2x$ przyjmuje największą wartość w punktach będących rozwiązaniem równania $\cos x = \frac{1}{\sqrt{3}}$ i wartość ta wynosi $\frac{4}{3\sqrt{3}} \approx 0,77$.

Zainteresowanym takim podejściem polecam książkę [4].

IV. Zastosowania w analizie funkcjonalnej

Wszystkie wyżej przedstawione zadania można znacznie szybciej rozwiązać stosując rachunek różniczkowy. Z tego punktu widzenia dotychczasowa prezentacja może się wydać „przerostem formy nad treścią". Jak jednak postąpić w sytuacji, gdy nie możemy skorzystać z dobrodziejstw rachunku różniczkowego? Z taką sytuacją wielokrotnie spotykamy się na przykład w analizie funkcjonalnej, teorii optymalizacji. Poszukujemy tam często minimum lub maksimum jakiegoś funkcjonału.

W jednostajnie wypukłych przestrzeniach Banacha (jeśli ktoś wie, co to takiego) prawdziwy jest następujący rezultat ([1]):

Twierdzenie 4. *Niech C będzie niepustym, domkniętym, wypukłym i ograniczonym podzbiorem jednostajnie wypukłej przestrzeni Banacha E i niech $f : C \to [0, +\infty)$ będzie ciągła. Jeżeli f jest funkcją ściśle wypukłą, to f osiąga minimum dokładnie w jednym punkcie.*

Korzystając z tego faktu udowodnimy twierdzenie z 1965 roku o punktach stałych odwzorowań nieoddalających. Było ono znaczącym impulsem do badań nad metryczną teorią punktów stałych i geometrią przestrzeni Banacha.

Odwzorowanie $T : C \to C$ nazywamy *nieoddalającym* jeżeli

$$\|Tx - Ty\| \leq \|x - y\| \text{ dla dowolnych } x, y \in C.$$

Twierdzenie 5 (F. Brouwer, D. Ghde, W.A. Kirk). *Jeżeli C jest niepustym, domkniętym, wypukłym i ograniczonym podzbiorem jednostajnie wypukłej przestrzeni Banacha $(E, \|\cdot\|)$ i $T : C \to C$ jest odwzorowaniem nieoddalającym, to T ma punkt stały w zbiorze C.*

Dowód. Wybieramy punkt $x_0 \in C$ i tworzymy ciąg $x_n = T^n x_0$, $n = 1, 2, \dots$ Taki ciąg nie musi być zbieżny. Określamy więc funkcję $r(x) = \limsup_{n\to\infty} \|x_n - x\|$, $x \in C$. Łatwo sprawdzamy, że jest ona ciągła oraz ściśle wypukła. W takim razie na podstawie twierdzenia 4 w zbiorze C istnieje *asymptotyczne centrum ciągu* $\{x_n\}$, czyli $A(\{x_n\}) = \{x \in C : r(x) = \inf\{r(y) : y \in C\}\}$, i jest ono zbiorem jednoelementowym $A(\{x_n\}) = \{z\}$. Wtedy

$$r(Tz) = i_u\|x_n - Tz\| = i_u\|T^n x_0 - Tz\| \leq i_u\|T^{n-1}x_0 - z\| = i_u\|x_{n-1} - z\| = r(z),$$

czyli $Tz \in A(\{T^n x_0\})$. Stąd zaś wynika, że $Tz = z$.

Pominięte szczegóły oraz dalsze zastosowania techniki asymptotycznego centrum znajdzie Czytelnik w pracy [1].

Literatura

[1] Goebel K., Reich S., *Uniform convexity, hyperbolic geometry, and nonexpansive mappings*, M. Dekker, New York and Basel 1984.

[2] Korovkin P.P., *Inequalities*, Mir, Moskwa 1975.

[3] Kuczma M., *An introduction to the theory of functional equations and inequalities, Cauchy's equations and Jensen's inequality*, PWN, Warszawa–Kraków–Katowice 1985.

[4] Niven I., *Maxima and minima without calculus*, MAA, Washington DC., 1981.

[5] Płoski A., *Nierówności elementarne i ekstrema*, Delta 8/1978.

Wydawnictwa Uczelniane WSRP w Siedlcach

Wydanie I. Nakład 500 egz. Ark. wyd. 4,9. Ark. druk. 6,25. Format A-4. Papier kl. III. Oddano do druku: lipiec 1993 r. Druk ukończono: sierpień 1993 r.

Zamówienie nr 12/W/93 Cena: 16.000 zł