I kolokwium zad 4 – rozwiązanie wzorcowe i zasady oceniania.

Niech $K \subset \mathbb{R}^3$ będzie zwartym zbiorem wypukłym zawierającym pewną kulę jednostkową $B(\mathbf{0}, r)$ dla $r > 0$. Niech $f : K \to \mathbb{R}$ będzie funkcją ciągłą na K, różniczkowalną w $\operatorname{int} K$, taką że

$$(1) \quad x\frac{\partial f}{\partial x}(x,y,z) + z\frac{\partial f}{\partial z}(x,y,z) > 0 \quad \text{dla } (x,y,z) \in \operatorname{int} K \setminus \{(x,y,z) \mid x = 0 \text{ i } z = 0\}$$

oraz

$$(2) \quad y\frac{\partial f}{\partial y}(x,y,z) > 0 \quad \text{dla } (x,y,z) \in \operatorname{int} K \setminus \{(x,y,z) \mid y = 0\}.$$

Wykaż, że f przyjmuje minimum w $\mathbf{0} = (0,0,0)$ a maksimum na brzegu K.

Zadady oceniania

Do zdobycia było 10 punktów. Oceniając zadania przyjąłem następujące kryteria:

- 7-10 punktów. Pełne rozwiązanie zadania. Obcinałem pojedyncze punkty za mniejsze lub większe usterki, np. brak wskazania gdzie korzysta się z wypukłości K, brak uzasadnienia dlaczego w danym punkcie można różniczkować, stwierdzenie jakiejś nieprawdziwej własności badanej funkcji.
- 4-7 punktów. Rozwiązanie zawierające istotne pomysły, ale bez należytego uzasadnienia, bądź też mające istotne luki w rozumowaniu.
- 2-3 punkty. Wykazanie, że f nie ma punktów krytycznych w $\operatorname{int} K \setminus \{\mathbf{0}\}$ i wywnioskowanie stąd, że przyjmuje swoje kresy na $\partial K \cup \{\mathbf{0}\}$. Dodatkowe punkty można było uzyskać za przydatne obserwacje wykorzystujące dodatniość odpowiednich pochodnych kierunkowych.
- 1-2 punkty. Drobne acz przydatne obserwacje dotyczące zadania.

Rozwiązanie wzorcowe (szkic)

Zauważmy, że nierówności (1) i (2) zachodzą w słabym sensie (tzn z „$\geq$” zamiast „$>$”) dla dowolnych $(x,y,z) \in \operatorname{int} K$. Dodając je stronami dostaniemy nierówność $x\frac{\partial f}{\partial x}(x,y,z) + y\frac{\partial f}{\partial y}(x,y,z) + z\frac{\partial f}{\partial z}(x,y,z) \geq 0$ prawdziwą w dowolnym punkcie $(x,y,z) \in \operatorname{int} K$. Co więcej nierówność ostra zachodzi o ile któraś z nierówności (1)-(2) jest ostra,[1] a zatem

$$(3) \qquad x\frac{\partial f}{\partial x}(x,y,z) + y\frac{\partial f}{\partial y}(x,y,z) + z\frac{\partial f}{\partial z}(x,y,z) > 0 \quad \text{dla } (x,y,z) \in \operatorname{int} K \setminus \{\mathbf{0}\}.$$

Rozważmy teraz dowolny punkt wewnętrzny $(x_0, y_0, z_0) \in \operatorname{int} K \setminus \{\mathbf{0}\}$. Na mocy założenia o wypukłości K i o tym, że $\mathbf{0} \in \operatorname{int} K$ wnioskujemy, że dla pewnego $\varepsilon > 0$ cały odcinek $t \mapsto (t \cdot x_0, t \cdot y_0, t \cdot z_0)$ dla $t \in [0, 1+\varepsilon]$ leży we wnętrzu K. Wobec tego obcięcie f do tego odcinka, czyli $g(t) := f(t \cdot x_0, t \cdot y_0, t \cdot z_0)$ jest funkcją różniczkowalną dla $t \in (0, 1+\varepsilon)$.[2]

[1] De facto nie potrzeba nierówności ostrej – rozumowanie przechodzi przy nierówności słabej.
[2] Takiego spostrzeżenia brakło w większości rozwiązań.

Policzmy jej pochodną:

$$g'(t) = x_0\frac{\partial f}{\partial x}(t\cdot x_0, t\cdot y_0, t\cdot z_0) + y_0\frac{\partial f}{\partial y}(t\cdot x_0, t\cdot y_0, t\cdot z_0) + z_0\frac{\partial f}{\partial z}(t\cdot x_0, t\cdot y_0, t\cdot z_0) =$$
$$\frac{1}{t}\left[tx_0\frac{\partial f}{\partial x}(t\cdot x_0, t\cdot y_0, t\cdot z_0) + ty_0\frac{\partial f}{\partial y}(t\cdot x_0, t\cdot y_0, t\cdot z_0) + tz_0\frac{\partial f}{\partial z}(t\cdot x_0, t\cdot y_0, t\cdot z_0)\right]$$

Na mocy nierówności (3), powyższa pochodna jest ściśle dodatnia dla $t > 0$, a zatem funkcja g jest ściśle rosnąca. Wnioskujemy, że

$$f(\mathbf{0}) = g(0) < g(1) = f(x_0, y_0, z_0) < g(1+\varepsilon) = f((1+\varepsilon)(x_0, y_0, z_0))\ .$$

Punkt (x_0, y_0, z_0) był dowolnym punktem wewnętrznym K różnym od $\mathbf{0}$. Wnioskujemy stąd, po pierwsze, że ani $f(x_0, y_0, z_0)$, ani $f(\mathbf{0})$ nie może być kresem górnym f na K. Ponieważ na mocy tw. Wierstrassa (K zwarte, f ciągłe) kres górny jest gdzieś przyjmowany, to musi zachodzić $\sup_K f = \sup_{\partial K} f$. Z drugiej strony widzimy, że $f(\mathbf{0})$ jest nie większe niż dowolna wartość f w punktach $\operatorname{int} K$. Ponieważ każdy punkt brzegu K możemy aproksymować punktami wewnętrznymi (i wartości f w tym punkcie wartościami f w punktach wewnętrznych – ciągłość f), z twierdzenia o zachowaniu nierówności słabej w granicy, $f(\mathbf{0})$ jest nie większe niż wartości f w dowolnym punkcie K. $\square$

Inne rozwiązania/rozwiązania częściowe (szkice)

1. Cześć rozwiązań nie wykorzystywała obliczenia pełnej pochodnej $g(t)$, a jedynie tę pochodną w $t = 0$, czyli pochodną kierunkową f w kierunku $[x, y, z]$ w punkcie (x, y, z). Na mocy (3) wiemy, że ta pochodna jest dodatnia. Oznacza to, że dla ustalonego $(x_0, y_0, z_0) \in \operatorname{int} K \setminus \{\mathbf{0}\}$ istnieje $t_0 > 0$ (a priori zależne od tego punktu) takie, że dla dowolnego $0 < \varepsilon \le t_0$

   $$f(x_0, y_0, z_0) < f((1+\varepsilon)(x_0, y_0, z_0))\ .$$

   Stąd można wywnioskować,[3] że $f(x_0, y_0, z_0) < f((1+\varepsilon)(x_0, y_0, z_0))$ dla każdego $\varepsilon > 0$ takiego, że $(1+\varepsilon)(x_0, y_0, z_0) \in \operatorname{int} K$.

   To rozumowanie dowodzi, że f rośnie na wszystkich półprostych przechodzących przez $\mathbf{0}$ i dalej rozumujemy analogicznie jak poprzednio.

2. Podobne rozumowanie wykorzystuje fakt, że gradient jest kierunkiem najszybszego wzrostu funkcji. Warunek (3) oznacza mianowicie, że kąt między gradientem funkcji $\nabla f(x, y, z)$ a wektorem $[x, y, z]$ jest ostry, a zatem gradient wskazuje na zewnątrz kuli $B(\mathbf{0}, \|(x, y, z)\|_2)$. Zatem na zewnątrz tej kuli można znaleźć punkty o większych wartościach niż na jej brzegu (i analogicznie wewnątrz punkty o wartościach mniejszych).

[3] Dokładne uzasadnienie wymaga nieco pracy i odwołuje się do zwartości odcinka domkniętego.

3. Inna strategia polegała na osobnym rozpatrzeniu $f(x,y,z)$ jako funkcji zmiennych (x,z) przy ustalonym y z wykorzystaniem nieróności (1), oraz osobnym rozpatrzeniu $f(x,y,z)$ jako funkcji zmiennej y przy ustalonych x i z z wykorzystaniem nierówności (2). Przez rozumowanie analogiczne jak w rozwiązaniu wzorcowym powinniśmy wówczas otrzymać nierówności $f(0,0,0) < f(x_0,0,z_0)$ oraz $f(x_0,0,z_0) < f(x_0,y_0,z_0)$, które wspólnie dają nierówność

$$f(0,0,0) \leq f(x_0,y_0,z_0)$$

dla dowolnego punktu $\boldsymbol{p}_0 = (x_0,y_0,z_0) \in \operatorname{int} K \backslash \{\mathbf{0}\}$. Problem z tym rozumowaniem jest taki, że z faktu, że $\mathbf{0}$ i $\boldsymbol{p}_0$ należą do K (a nawet do wnętrza) nie wynika, że $(x_0,0,z_0) \in K$.[4] To rozumowanie można poprawić tworząc odpowiednio drobne „schodki" (poruszamy się na przemian w płaszczyźnie (x,z) i wzdłuż osi y, aby nie wypaść z $\operatorname{int} K$). Składając nierówności z poszczególnych „stopni" otrzymamy żądaną nierówność

$$f(0,0,0) \leq f(x_0,y_0,z_0)\ .$$

4. Do wykazania, że supremum f jest przyjmowane na ∂K wystarczy nam znajomość nierówności (2). Istotnie, załóżmy że punkt $\boldsymbol{p} = (x_0,y_0,z_0) \in \operatorname{int} K$ jest takim supremum. Bez straty ogólności możemy założyć, że $y_0 \geq 0$. We wnętrzu K znajdziemy inny punkt $\boldsymbol{p}' = (x_0,y_1,z_0)$ o tej własności, że $y_1 > y_0$. Wobec $f(\boldsymbol{p}) \geq f(\boldsymbol{p}')$ na mocy Tw. Lagrange'a mamy $\frac{\partial f}{\partial y}(x_0,y,z_0) \leq 0$ w pewnym punkcie $y \in (y_0,y_1)$. Ale skoro $y > y_0 \geq 0$ z (2) mamy nierówność przeciwną. Sprzeczność.

5. Stosunkowo łatwo jest wykazać, że kresy f muszą być przyjmowane na brzegu K lub w $\mathbf{0}$. Mianowicie jeśli $\boldsymbol{p} \in \operatorname{int} K$ jest kresem funkcji f to musi być też jej punktem krytycznym, a zatem $\nabla f(\boldsymbol{p}) = [0,0,0]$. Tymczasem nierówności (1) i (2) implikują, że $\nabla f(\boldsymbol{p}) \neq [0,0,0]$ o ile $p \neq \mathbf{0}$.[5] Zatem $\boldsymbol{p} = \mathbf{0}$ jest jedynym możliwym punktem krytycznym f.

Michał Jóźwikowski, 20 listopada 2018.

[4] Ale jest to prawdą w pewnym otoczeniu $\mathbf{0}$ – korzystamy tutaj z założenia, że $\mathbf{0}$ jest punktem wewnętrznym.

[5] Uwaga: stąd wcale nie wynika (jeszcze), że $\nabla f(\boldsymbol{p}) = [0,0,0]$.