Kolokwium z Optymalizacji II

(*Ściśle tajne* przed godz. 14 : 00 18 grudnia 2015.)

Proszę uważnie przeczytać treść zadań. Na ocenę bardzo duży wpływ będzie miała czytelność rozwiązań i poprawność uzasadnienia każdej odpowiedzi.

1. Iloczynem tensorowym funkcji $f\colon X \to \mathbb{R}$ i $g\colon Y \to \mathbb{R}$ jest funkcja $h = f \otimes g\colon X \times Y \to \mathbb{R}$ określona wzorem $h(\mathbf{x}, \mathbf{y}) \stackrel{\text{def}}{=} f(\mathbf{x})g(\mathbf{y})$.

   Zbadaj, które z podanych niżej warunków są dostateczne, aby iloczyn tensorowy dodatnich funkcji f i g klasy C^1 był funkcją pseudowypukłą:

a) funkcje f i g są pseudowypukłe,

b) funkcje $\sqrt{f}$ i $\sqrt{g}$ są pseudowypukłe,

c) funkcje $\log_a f$ i $\log_a g$ (dla $0 < a \neq 1$) są pseudowypukłe.

2. Znajdź globalne minimum i maksimum funkcji

$$f(x_1, x_2, y_1, y_2) \stackrel{\text{def}}{=} (|x_1 - 2x_2| + 2x_1^2 - x_1 + x_2^2)(y_1 + y_2)^3$$

   w zbiorze

$$W = \{\, (x_1, x_2, y_1, y_2)\colon \|[x_1, x_2]^T\|_\infty \leqslant 1,\ \|[y_1, y_2]^T\|_2 \leqslant \sqrt{2},\ y_1 \geqslant 0 \,\}.$$

Kolokwium poprawkowe z Optymalizacji II

(*Ściśle tajne* przed godz. 16 : 00 15 stycznia 2016.)

Proszę uważnie przeczytać treść zadań. Na ocenę bardzo duży wpływ będzie miała czytelność rozwiązań i poprawność uzasadnienia każdej odpowiedzi.

1. Iloczynem tensorowym funkcji $f\colon X \to \mathbb{R}$ i $g\colon Y \to \mathbb{R}$ jest funkcja $h = f \otimes g\colon X \times Y \to \mathbb{R}$ określona wzorem $h(\mathbf{x}, \mathbf{y}) \stackrel{\text{def}}{=} f(\mathbf{x})g(\mathbf{y})$.

   Zbadaj, które z podanych niżej warunków są dostateczne, aby iloczyn tensorowy dodatnich funkcji f i g klasy C^1 był funkcją wypukłą:

a) funkcje f i g są wypukłe,

b) funkcje $\sqrt{f}$ i $\sqrt{g}$ są wypukłe,

c) funkcje $\ln f$ i $\ln g$ są wypukłe.

2. Znajdź globalne minimum i maksimum funkcji

$$f(x, y, z) \stackrel{\text{def}}{=} (x-1)^2 + (y-1)^2 + z^3$$

   w zbiorze

$$W = \{\, (x, y, z)\colon x^2 + y^2 \leqslant (1-z)^2,\ 0 \leqslant z \leqslant 1 \,\}.$$

Kolokwium z Optymalizacji II

(*Ściśle tajne* przed godz. 12 : 15 9 grudnia 2016.)

Proszę uważnie przeczytać treść zadań. Na ocenę bardzo duży wpływ będzie miała czytelność rozwiązań i poprawność uzasadnienia każdej odpowiedzi.

1. Niech $W \subset \mathbb{R}^n$ będzie zbiorem wypukłym i niech f_0, f_1 będą różniczkowalnymi rzeczywistymi funkcjami wypukłymi określonymi w zbiorze W. Niech $g\colon W \times [0,1] \to \mathbb{R}$ będzie funkcją określoną wzorem

$$g(\boldsymbol{x}, y) = (1-y)f_0(\boldsymbol{x}) + yf_1(\boldsymbol{x}).$$

Wśród podanych niżej warunków znajdź warunki dostateczne, aby funkcja g była wypukła i wskaż najsłabszy z nich:

a) istnieje punkt $\boldsymbol{x} \in W$, taki że $Df_0(\boldsymbol{x}) = Df_1(\boldsymbol{x}) = \boldsymbol{0}^T$,

b) istnieją stałe $c > 0$, $d \in \mathbb{R}$, takie że $f_0(\boldsymbol{x}) = cf_1(\boldsymbol{x}) + d$ dla każdego $\boldsymbol{x} \in W$,

c) istnieje stała $d \in \mathbb{R}$, taka że $f_0(\boldsymbol{x}) = f_1(\boldsymbol{x}) + d$ dla każdego $\boldsymbol{x} \in W$.

2. Znajdź globalne minimum i maksimum funkcji

$$f(x,y,z) \stackrel{\text{def}}{=} (1-z)\min\{x, y, x-y\} + (x^2+y)z$$

w zbiorze

$$W = \{\, (x,y,z)\colon x^2+y^2 \leqslant 1,\ 0 \leqslant z \leqslant 1 \,\}.$$

Ćwiczenia z optymalizacji

1. Wykaż, że symetryczna macierz $A \in \mathbb{R}^{n\times n}$ jest dodatnio określona wtedy i tylko wtedy, gdy istnieją kwadratowe macierze L i U, odpowiednio trójkątna dolna i górna, które mają wszystkie współczynniki diagonalne dodatnie i takie że $A = LU$.

   Wskazówka. Dokonaj podziału blokowego

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{12}^T & A_{22}, \end{bmatrix}, \qquad L = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22}, \end{bmatrix}, \qquad U = \begin{bmatrix} U_{11} & U_{12} \\ 0 & U_{22}, \end{bmatrix}.$$

   Pokaż, że $A_{11} = L_{11}U_{11}$ i że jeśli blok A_{11} jest nieosobliwy, to można znaleźć takie bloki trójkątne L_{11} i U_{11} (np. z jedynką na całej diagonali L_{11}), a następnie skorzystaj z twierdzenia Cauchy'ego i z kryterium Sylvestra.

2. Zbadaj, czy macierz

$$A = \begin{bmatrix} 10 & 3 & 0 \\ 3 & 1 & 3 \\ 0 & 3 & 10 \end{bmatrix}$$

   jest dodatnio określona.

3. Niech $A \in \mathbb{R}^{m\times n}$, $\mathbf{b} \in \mathbb{R}^m$, $\mathbf{x} \in \mathbb{R}^n$ (m i n mogą być dowolne).

$$f(\mathbf{x}) = \|\mathbf{r}(\mathbf{x})\|_2^2 = \|\mathbf{b} - A\mathbf{x}\|_2^2.$$

   Udowodnij, że funkcja f ma minimalną wartość w punkcie $\mathbf{x}$ wtedy i tylko wtedy, gdy $A^T\mathbf{r}(\mathbf{x}) = \mathbf{0} \in \mathbb{R}^n$. Kiedy minimum jest ścisłe?

   Rozw. Niech $\mathbf{x}, \mathbf{x}' \in \mathbb{R}^n$ oraz $\mathbf{y} = A\mathbf{x}$, $\mathbf{y}' = A\mathbf{x}'$. Przypuśćmy, że wektor $\mathbf{y}$ jest obrazem wektora $\mathbf{b}$ w rzucie prostopadłym na podprzestrzeń $V \subset \mathbb{R}^m$ rozpiętą przez kolumny macierzy A. Zatem wektor $\mathbf{r}(\mathbf{x})$ jest prostopadły do wszystkich elementów tej podprzestrzeni, w tym do wszystkich kolumn macierzy A, skąd wynika, że $A^T\mathbf{r}(\mathbf{x}) = \mathbf{0}$. Jest

$$\mathbf{r}(\mathbf{x}') - \mathbf{r}(\mathbf{x}) = \mathbf{b} - A\mathbf{x}' - \mathbf{b} + A\mathbf{x} = \mathbf{y} - \mathbf{y}' \in V,$$

   a zatem $\mathbf{r}(\mathbf{x}') = \mathbf{r}(\mathbf{x}) + (\mathbf{y} - \mathbf{y}')$ oraz $(\mathbf{y} - \mathbf{y}')^T\mathbf{r}(\mathbf{x}) = 0$. Z twierdzenia Pitagorasa $\|\mathbf{r}(\mathbf{x}')\|_2^2 = \|\mathbf{r}(\mathbf{x})\|_2^2 + \|\mathbf{y} - \mathbf{y}'\|_2^2$, czyli $f(\mathbf{x}') \geqslant f(\mathbf{x})$.

   Wektor $\mathbf{x}$ jest rozwiązaniem układu równań normalnych $A^TA\mathbf{x} = A^T\mathbf{b}$, przy czym macierz A^TA jest nieosobliwa (dodatnio określona) wtedy i tylko wtedy, gdy kolumny macierzy A są liniowo niezależne (w tym przypadku musi być $m \geqslant n$). Wtedy rozwiązanie jest jednoznaczne i minimum jest ścisłe.

4. (NPP) Znajdź minimum funkcji

$$f(x_1, x_2) = 5x_1^2 + x_2^2 - 4x_1x_2 - 2x_1 + 3.$$

   Rozw.

$$Df(x_1, x_2) = [10x_1 - 4x_2 - 2, 2x_2 - 4x_1],$$
$$D^2f = \begin{bmatrix} 10 & -4 \\ -4 & 2 \end{bmatrix}.$$

   Macierz D^2f jest dodatnio określona, zatem funkcja f jest wypukła w $\mathbb{R}^2$ i ma jedno minimum.

   Znajdujemy (x_1, x_2) t.że $Df(x_1, x_2) = \mathbf{0}$:

$$\begin{cases} 10x_1 - 4x_2 = 2, \\ -4x_1 + 2x_2 = 0. \end{cases}$$

   Po wstawieniu $x_2 = 2x_1$ do pierweszego równania i rozwiązaniu dostajemy $x_1 = 1$, $x_2 = 2$.

5. (NPP) Znajdź minimum funkcji

$$f(x) = \frac{1}{4}x^4 - x^3 + x^2 + 1.$$

   Rozw. Obliczamy $f'(x) = x^3 - 3x^2 + 2x = x(x^2 - 3x + 2) = x(x-1)(x-2)$ oraz $f''(x) = 3x^2 - 6x + 2$. Badamy miejsca zerowe f':

$$\begin{aligned} x = 0 &: f''(0) = 2 && \Rightarrow \text{ minimum lokalne } f(0) = 1 \\ x = 1 &: f''(1) = -1 && \Rightarrow \text{ maksimum lokalne} \\ x = 2 &: f''(2) = 2 && \Rightarrow \text{ minimum lokalne } f(2) = 1. \end{aligned}$$

6. (NPP) Dane są punkty $\mathbf{y}^1, \ldots, \mathbf{y}^m \in \mathbb{R}^n$. Znajdź punkt $\hat{\mathbf{y}} \in \mathbb{R}^n$, do którego suma kwadratów odległości od punktów $\mathbf{y}^i$ jest najmniejsza.

   Rozw. Minimalizujemy funkcję

$$f(\mathbf{y}) = \sum_{i=1}^m \|\mathbf{y} - \mathbf{y}^i\|_2^2 = \sum_{i=1}^m \sum_{j=1}^n (y_j - y_j^i)^2.$$
$$Df(\mathbf{y}) = \left[2\sum_{i=1}^m y_j - y_j^i\right]_{j=1,\ldots,n},$$
$$D^2f(\mathbf{y}) = 2mI_{n\times n} \quad \text{— dodatnio określona.}$$

   Rozwiązaniem równania $Df(\mathbf{y}) = \mathbf{0}$ jest $\hat{y}_j = \frac{1}{m}\sum_{i=1}^m y_j^i$, czyli $\hat{\mathbf{y}} = \frac{1}{m}\sum_{i=1}^m \mathbf{y}^i$.

7. Dane są punkty $\mathbf{y}^1, \dots, \mathbf{y}^m \in \mathbb{R}^n$. Znajdź hiperpłaszczyznę w $\mathbb{R}^n$, taką że suma kwadratów odległości punktów $\mathbf{y}^i$ od tej hiperpłaszczyzny jest najmniejsza.

   Rozw. Hiperpłaszczyzna jest wyznaczona przez punkt $\mathbf{p}$ i jednostkowy wektor normalny $\mathbf{n}$; należą do niej punkty $\mathbf{x}$, takie że $\mathbf{n}^T(\mathbf{p}-\mathbf{x}) = 0$. Niech punkty $\mathbf{x}^1, \dots, \mathbf{x}^m$ będą rzutami prostopadłymi punktów $\mathbf{y}^1, \dots, \mathbf{y}^m$ na hiperpłaszczyznę; wtedy wektory $\mathbf{y}^i - \mathbf{x}^i$ mają kierunek wektora $\mathbf{n}$, a iloczyny skalarne $\mathbf{n}^T(\mathbf{y}^i - \mathbf{x}^i) = \mathbf{n}^T(\mathbf{y}^i - \mathbf{p})$ są z dokładnością do znaku równe odległościom punktów $\mathbf{y}^i$ od hiperpłaszczyzny.

   Niech $A \in \mathbb{R}^{m\times n}$ oznacza macierz, której każdy wiersz jest równy $\mathbf{n}^T$. Niech $\mathbf{b} \in \mathbb{R}^m$ będzie wektorem, którego kolejne współrzędne b_i są iloczynami skalarnymi $\mathbf{n}^T\mathbf{y}^i$, dla $i = 1, \dots, m$. Dla ustalonego wektora $\mathbf{n}$ należy znaleźć minimum funkcji

$$f(\mathbf{p}) = \sum_{i=1}^{m} \bigl(\mathbf{n}^T(\mathbf{y}^i - \mathbf{p})\bigr)^2 = \|A\mathbf{p} - \mathbf{b}\|_2^2.$$

   Z zadania 3 wiemy, że punkt $\mathbf{p}$ musi spełniać układ równań normalnych $A^TA\mathbf{p} = A^T\mathbf{b}$. Jest $A^TA = m\mathbf{n}\mathbf{n}^T$ oraz $A^T\mathbf{b} = \sum_{i=1}^m \mathbf{n}\mathbf{n}^T\mathbf{y}^i = \mathbf{n}\mathbf{n}^T\sum_{i=1}^m \mathbf{y}^i$. Oczywistym rozwiązaniem układu równań normalnych (którego macierz $n\times n$ ma rząd 1) jest punkt $\mathbf{p} = \frac{1}{m}\sum_{i=1}^m \mathbf{y}^i$.

   Mając punkt $\mathbf{p}$ wyznaczamy wektor jednostkowy $\mathbf{n}$. Oznaczamy $\mathbf{v}^i = \mathbf{y}^i - \mathbf{p}$ i określamy funkcję

$$g(\mathbf{n}) = \sum_{i=1}^{m}(\mathbf{n}^T\mathbf{v}^i)^2 = \sum_{i=1}^{m}\mathbf{n}^T\mathbf{v}^i\mathbf{v}^{iT}\mathbf{n} = \mathbf{n}^T\Bigl(\sum_{i=1}^{m}\mathbf{v}^i\mathbf{v}^{iT}\Bigr)\mathbf{n} = \mathbf{n}^TB\mathbf{n}.$$

   Macierz symetryczna B jest nieujemnie określona; pozostaje dowieść, że funkcja g osiąga minimum w zbiorze wektorów jednostkowych, jeśli wektor $\mathbf{n}$ jest jej wektorem własnym przynależnym do najmniejszej wartości własnej.

8. (NPP) Wyznacz gradient i hesjan funkcji $f\colon \mathbb{R}^n \to \mathbb{R}$:

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^TA\mathbf{x} + \mathbf{b}^T\mathbf{x} + c,$$

   z macierzą niesymetryczną $A \in \mathbb{R}^{n\times n}$, wektorem $\mathbf{b} \in \mathbb{R}^n$ i stałą c.

   Rozw. Z definicji pochodnej

$$\begin{aligned} f(\mathbf{x}+\varepsilon\mathbf{d}) - f(\mathbf{x}) &= \frac{1}{2}(\mathbf{x}+\varepsilon\mathbf{d})^TA(\mathbf{x}+\varepsilon\mathbf{d}) + \mathbf{b}^T(\mathbf{x}+\varepsilon\mathbf{d}) - \frac{1}{2}\mathbf{x}^TA\mathbf{x} - \mathbf{b}^T\mathbf{x} \\ &= \frac{1}{2}\varepsilon\mathbf{x}^TA\mathbf{d} + \frac{1}{2}\varepsilon\mathbf{d}^TA\mathbf{x} + \varepsilon\mathbf{b}^T\mathbf{d} + \frac{1}{2}\varepsilon^2\mathbf{d}^TA\mathbf{d} \\ &= \Bigl(\frac{1}{2}\mathbf{x}^TA^T + \frac{1}{2}\mathbf{x}^TA + \mathbf{b}^T\Bigr)(\varepsilon\mathbf{d}) + \frac{1}{2}\varepsilon^2\mathbf{d}^TA\mathbf{d}, \end{aligned}$$

   stąd

$$Df(\mathbf{x}) = \frac{1}{2}(A^T + A)\mathbf{x} + \mathbf{b}.$$

   Ze wzoru Taylora

$$f(\mathbf{x}+\mathbf{d}) = f(\mathbf{x}) + Df(\mathbf{x})^T\mathbf{d} + \frac{1}{2}\mathbf{d}^TD^2f(\mathbf{x})\mathbf{d} + o(\|\mathbf{d}\|^2).$$

   Porównując to z

$$f(\mathbf{x}+\mathbf{d}) = f(\mathbf{x}) + \Bigl(\frac{1}{2}(A^T + A)\mathbf{x} + \mathbf{b}\Bigr)\mathbf{d} + \frac{1}{2}\mathbf{d}^TA\mathbf{d},$$

   możemy wyciągnąć błędny wniosek, że $D^2f(\mathbf{x}) = A$. Ale macierz $D^2f(\mathbf{x})$ jest symetryczna; jedyna taka macierz, która spełnia warunek $\mathbf{d}^TB\mathbf{d} = \mathbf{d}^TA\mathbf{d}$ dla każdego $\mathbf{d} \in \mathbb{R}^n$ to macierz $B = \frac{1}{2}(A^T + A)$. Zatem $D^2f = \frac{1}{2}(A^T + A)$.

9. (NPP) Dana jest funkcja $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^TA\mathbf{x} + \mathbf{b}^T\mathbf{x}$. Zbadaj, czy jest ona wypukła albo wklęsła, jeśli macierz A jest równa

$$\begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix}, \quad \begin{bmatrix} -3 & 1 \\ 1 & -2 \end{bmatrix}, \quad \begin{bmatrix} 3 & 1 \\ 1 & -2 \end{bmatrix}, \quad \begin{bmatrix} 3 & 2 \\ 0 & 2 \end{bmatrix}.$$

10. (NPP) Dana jest przestrzeń probabilistyczna $(\Omega, \mathcal{F}, \mathbb{P})$, $\mathbb{P}(\Omega) = 1$, i zmienna losowa $\boldsymbol{\eta} \in L^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^n)$. Znajdź punkt $\hat{\mathbf{x}} \in \mathbb{R}^n$, taki że $\mathbb{E}\|\boldsymbol{\eta} - \hat{\mathbf{x}}\|^2$ jest najmniejsza.

    Rozw. Określamy

$$\begin{aligned} f(\mathbf{x}) = \mathbb{E}\|\boldsymbol{\eta} - \mathbf{x}\|^2 &= \mathbb{E}(\boldsymbol{\eta} - \mathbf{x})^T(\boldsymbol{\eta} - \mathbf{x}) = \mathbb{E}(\boldsymbol{\eta}^T\boldsymbol{\eta} - \boldsymbol{\eta}^T\mathbf{x} - \mathbf{x}^T\boldsymbol{\eta}) - \mathbf{x}^T\mathbf{x} \\ &= \mathbf{x}^T\mathbf{x} - 2(\mathbb{E}\boldsymbol{\eta})^T\mathbf{x} + \mathbb{E}\|\boldsymbol{\eta}\|^2. \end{aligned}$$

    Zatem

$$f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^TA\mathbf{x} + \mathbf{b}^T\mathbf{x} + c,$$

    gdzie $A = 2I$, $\mathbf{b} = 2\mathbb{E}\boldsymbol{\eta}$, $c = \mathbb{E}\|\boldsymbol{\eta}\|^2$. Macierz A jest symetryczna i dodatnio określona. Mamy

$$Df(\mathbf{x}) = 2I\mathbf{x} - 2\mathbb{E}\boldsymbol{\eta},$$

    stąd $Df(\mathbf{x}) = \mathbf{0}$, gdy $\mathbf{x} = \mathbb{E}\boldsymbol{\eta}$. To jest minimum globalne.

11. (NPP) Znajdź minimum funkcji $f(x_1, x_2) = 2x_1 + 3x_2$ na zbiorach $W_1 = \{(x_1, x_2)\colon x_1^2 + x_2^2 = 1\}$, $W_2 = \{(x_1, x_2)\colon x_1^2 + x_2^2 \leqslant 1\}$.

    Rozw. $Df = [2, 3]$, zatem nie może być minimum we wnętrzu W_2, musi być na brzegu. Bierzemy $x_2 = \sqrt{1 - x_1^2}$, określamy $f(x) = 2x + 3\sqrt{1 - x^2}$, $f'(x) = 2 + 3x/\sqrt{1 - x^2}$, $f'(x) = 0$ jeśli $3x = 2\sqrt{1 - x^2}$, czyli $9x^2 = 4 - 4x^2$, czyli $13x^2 = 4$. Zatem $x = \pm\frac{2}{\sqrt{13}}$.

    W punkcie $(x_1, x_2) = \big(-\frac{2}{\sqrt{13}}, -\frac{3}{\sqrt{13}}\big)$ jest minimum.

    W punkcie $(x_1, x_2) = \big(\frac{2}{\sqrt{13}}, \frac{3}{\sqrt{13}}\big)$ jest maksimum.

12. (NPP) Dana jest funkcja Peano

$$f(x_1, x_2) = (x_2^2 - x_1)(x_2^2 - 2x_1) = x_2^4 - 3x_1x_2^2 + 2x_1^2.$$

    Udowodnij, że obcięcie tej funkcji do dowolnej prostej przechodzącej przez $(0, 0)$ ma w tym punkcie minimum, ale funkcja f nie ma lokalnego minimum w tym punkcie, ani nigdzie w $\mathbb{R}^2$.

    Rozw. Jest $Df(x_1, x_2) = [4x_1 - 3x_2^2, 4x_2^3 - 6x_1x_2]$, czyli $Df(0, 0) = \mathbf{0}$ Jest to jedyne miejsce zerowe pochodnej w $\mathbb{R}^2$.

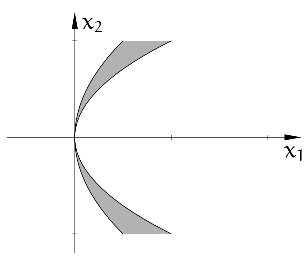
    Dla prostej $x_2 = 0$ mamy $g(x) = f(x, 0) = 2x^2$ — funkcja ta ma minimum dla $x = 0$.

    Dla prostej $x_1 = ax_2$ mamy $g(x) = f(ax, x) = (x^2 - ax)(x^2 - 2ax)$, jest $g(0) = 0$, a jeśli $|x| < |a|$, to $g(x) > 0$, jest zatem minimum na każdej takiej prostej.

    Teraz podstawiamy $x_1 = \frac{2}{3}x_2^2$: mamy

$$g(x) = f\Big(\frac{2}{3}x^2, x\Big) = \Big(x^2 - \frac{2}{3}x^2\Big)\Big(x^2 - \frac{4}{3}x^2\Big) = -\frac{1}{9}x^4,$$

    teraz dla $x = 0$ jest maksimum, nie ma minimum dla żadnego x, czyli f nie ma minimum w $\mathbb{R}^2$.



13. Znajdź przykład funkcji $f\colon \mathbb{R}^3 \to \mathbb{R}$, której obcięcie do dowolnej prostej przechodzącej przez punkt $(0, 0, 0)$ ma w tym punkcie minimum, ale funkcja f nie ma minimum.

    Odp. $f(x, y, z) = (y^2 + z^2 - x)(y^2 + z^2 - 2x)$. Ale trzeba uzasadnić.

14. Znajdź przykład funkcji $f\colon \mathbb{R}^3 \to \mathbb{R}$, której obcięcie do dowolnej płaszczyzny przechodzącej przez punkt $(0, 0, 0)$ ma w tym punkcie minimum, ale funkcja f nie ma minimum.

    Odp. $f(x, y, z) = (y - x^2)^2 + (z - x^3)^2 - \frac{1}{4}x^6$. Ale trzeba uzasadnić.

15. (EII) Niech $W = \{\, (x, y, z) \in \mathbb{R}^3\colon x^2 + y^2 + z^2 = 2,\ 2xy + 2yz + 2zx + 1 = 0 \,\}$. Znajdź minimum i maksimum funkcji $f(x, y, z) = 2x + 2y - 3z$ w zbiorze W.

    Rozw. Zbiór W jest przecięciem dwóch kwadryk: sfery o promieniu $\sqrt{2}$ i hiperboloidy obrotowej (jeszcze nie wiadomo, jakiej). Jej równanie przedstawiamy w postaci macierzowej:

$$[x, y, z]\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \\ z \end{bmatrix} + 1 = 0.$$

    Macierz A występująca w tym równaniu ma wartość własną $\lambda_1 = 2$ i przynależny do niej wektor własny $\boldsymbol{v}_1 = [1, 1, 1]^T$ oraz dwukrotną wartość własną $\lambda_{2,3} = -1$, której przestrzeń własna jest rozpięta przez wektory $\boldsymbol{v}_2 = [1, 0, -1]^T$ i $\boldsymbol{v}_3 = [-1, 2, -1]^T$. Wektory $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$ tworzą bazę ortogonalną przestrzeni $\mathbb{R}^3$, tworzymy z nich macierz ortogonalną zmiany układu współrzędnych:

$$Q = \Big[\frac{1}{\sqrt{3}}\boldsymbol{v}_1, \frac{1}{\sqrt{2}}\boldsymbol{v}_2, \frac{1}{\sqrt{6}}\boldsymbol{v}_3\Big],$$

    i wprowadzamy nowe zmienne, u, v, w. Jest $[x, y, z]^T = Q[u, v, w]^T$, czyli $[u, v, w]^T = Q^T[x, y, z]^T$. Równanie sfery w nowych zmiennych jest takie samo: $u^2 + v^2 + w^2 = 2$, natomiast równanie hiperboloidy ma postać (na podstawie wartości własnych) $2u^2 - v^2 - w^2 = -1$. Jest to więc hiperboloida jednopowłokowa.

    Podstawiając wyznaczone z równania sfery $v^2 + w^2 = 2 - u^2$ do równania hiperboloidy, dostajemy $u^2 = 1/3$, zatem punkty zbioru W mają współrzędną $u = \pm 1/\sqrt{3}$. Teraz określamy (przez zamianę zmiennych) funkcję $g(u, v, w) = f(x, y, z)$. Czyli jest $g(u, v, w) = [2, 2, -3]Q[u, v, w]^T$. Dalej,

$$[2, 2, -3]Q = \Big[\frac{1}{\sqrt{3}}, \frac{5}{\sqrt{2}}, \frac{5}{\sqrt{6}}\Big], \quad \text{zatem} \quad g(u, v, w) = \frac{u}{\sqrt{3}} + \frac{5v}{\sqrt{2}} + \frac{5w}{\sqrt{6}}.$$

Pozostaje znaleźć ekstrema funkcji (wielomianów pierwszego stopnia) $g_1(v,w) = g(1/\sqrt{3}, v, w)$ i $g_2(v,w) = g(-1/\sqrt{3}, v, w)$ na okręgu $v^2 + w^2 = 5/3$. Gradient obu tych funkcji jest taki sam: $Dg_i = [\frac{5}{\sqrt{2}}, \frac{5}{\sqrt{6}}]$.

Stąd maksimum jest dla $u = \frac{1}{\sqrt{3}}$, $v = \frac{\sqrt{5}}{2}$, $w = \frac{\sqrt{5}}{2\sqrt{3}}$, minimum jest dla $u = -\frac{1}{\sqrt{3}}$, $v = -\frac{\sqrt{5}}{2}$, $w = -\frac{\sqrt{5}}{2\sqrt{3}}$. Za pomocą macierzy Q znajdujemy odpowiednie współrzędne x, y, z punktów ekstremalnych. Maksimum i minimum są odpowiednio w punktach

$$Q\begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{\sqrt{5}}{2} \\ \frac{\sqrt{5}}{2\sqrt{3}} \end{bmatrix} = \begin{bmatrix} \frac{1}{3} + \frac{\sqrt{5}}{3\sqrt{2}} \\ \frac{1}{3} + \frac{\sqrt{5}}{3\sqrt{2}} \\ \frac{1}{3} - \frac{\sqrt{10}}{3} \end{bmatrix} \quad \text{i} \quad Q\begin{bmatrix} -\frac{1}{\sqrt{3}} \\ -\frac{\sqrt{5}}{2} \\ -\frac{\sqrt{5}}{2\sqrt{3}} \end{bmatrix} = \begin{bmatrix} -\frac{1}{3} - \frac{\sqrt{5}}{3\sqrt{2}} \\ -\frac{1}{3} - \frac{\sqrt{5}}{3\sqrt{2}} \\ -\frac{1}{3} + \frac{\sqrt{10}}{3} \end{bmatrix}.$$

16. (NPP) Rozważając problem

$$\begin{cases} xyzt \to \max, \\ x + y + z + t = 4c, \\ x, y, z, t \in \mathbb{R}_+^{\star}, \end{cases}$$

udowodnij nierówność średnich.

<u>Rozw.</u> Przyjmujemy $t = 4c - x - y - z$ i określamy funkcję

$$g(x,y,z) = 4cxyz - x^2yz - xy^2z - xyz^2,$$

której maksimum trzeba znaleźć. Obliczamy $Dg =$

$$[4cyz - 2xyz - y^2 - yz^2, 4cxz - x^2z - 2xyz - xz^2, 4cxy - x^2y - xy^2 - 2xyz].$$

Jeśli pochodna jest zerowa, to

$$\begin{aligned} 4cyz &= 2xyz + y^2z + yz^2 &\Rightarrow\quad 4c &= 2x + y + z, \\ 4cxz &= x^2z + 2xyz + xz^2 &\Rightarrow\quad 4c &= x + 2y + z, \\ 4cxy &= x^2y + xy^2 + 2xyz &\Rightarrow\quad 4c &= x + y + 2z. \end{aligned}$$

Odejmując każdą parę z powyższych trzech równań dostajemy $x - y = 0$, $y - z = 0$ i $x - z = 0$, czyli $Dg = \mathbf{0}$ wtedy, gdy $x = y = z = c$.

Hesjan funkcji g jest równy

$$\begin{bmatrix} -2yz & 4cz - 2xz - 2yz - z^2 & 4cy - 2xy - y^2 - 2yz \\ 4cz - 2xz - 2yz - z^2 & -2xz & 4cz - x^2 - 2xy - 2xz \\ 4cy - 2xy - y^2 - 2yz & 4cz - x^2 - 2xy - 2xz & -2xy \end{bmatrix}.$$

Dla $x = y = z = c$ hesjan jest równy

$$\begin{bmatrix} -2c^2 & -c^2 & -c^2 \\ -c^2 & -2c^2 & -c^2 \\ -c^2 & -c^2 & -2c^2 \end{bmatrix} = -c^2\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{bmatrix}.$$

Na podstawie kryterium Sylvestra (zobacz też zadanie 1) hesjan w punkcie (c,c,c) jest ujemnie określony, zatem w tym punkcie jest lokalne maksimum funkcji g. Ale nie ma innych punktów krytycznych i funkcja g, jako wielomian, jest ograniczona w czworościanie $\{(x,y,z)\colon x,y,z \geqslant 0, x+y+z \leqslant 4c\}$ i ma wartość 0 na jego brzegu, zatem to maksimum jest globalne. Stąd wynika nierówność

$$(xyzt)^{1/4} \leqslant \frac{x+y+z+t}{4}.$$

17. Udowodnij twierdzenie o epigrafie: funkcja $f\colon X \to \mathbb{R}$ określona na zbiorze wypukłym $X \subset \mathbb{R}^n$ jest wypukła wtedy i tylko wtedy, gdy jej epigraf jest zbiorem wypukłym.

<u>Dowód.</u> Z definicji, $\operatorname{epi} f = \{(\boldsymbol{x}, z)\colon \boldsymbol{x} \in X, z \geqslant f(\boldsymbol{x})\} \subset \mathbb{R}^{n+1}$. Niech $\boldsymbol{x}, \boldsymbol{y} \in X$, $\lambda \in [0,1]$. Dla dowolnych punktów $(\boldsymbol{x}, z), (\boldsymbol{y}, t) \in \operatorname{epi} f$ jest $z \geqslant f(\boldsymbol{x})$, $t \geqslant f(\boldsymbol{y})$ i mamy stąd oraz z wypukłości f

$$\lambda z + (1-\lambda)t \geqslant \lambda f(\boldsymbol{x}) + (1-\lambda)f(\boldsymbol{y}) \geqslant f\big((\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\big),$$

zatem $\lambda(\boldsymbol{x}, z) + (1-\lambda)(\boldsymbol{y}, t) = \big(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}, \lambda z + (1-\lambda)t\big) \in \operatorname{epi} f$.

Wynikanie w drugą stronę: dla dowolnych punktów $\boldsymbol{x}, \boldsymbol{y} \in X$ oraz $\lambda \in [0,1]$ jest $\big(\boldsymbol{x}, f(\boldsymbol{x})\big), \big(\boldsymbol{y}, f(\boldsymbol{y})\big) \in \operatorname{epi} f$ i jeśli epi f jest wypukły, to $\lambda\big(\boldsymbol{x}, f(\boldsymbol{x})\big) + (1-\lambda)\big(\boldsymbol{y}, f(\boldsymbol{y})\big) \in \operatorname{epi} f$. Ale na podstawie definicji epigrafu to oznacza, że $f\big(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\big) \leqslant \lambda f(\boldsymbol{x}) + (1-\lambda)f(\boldsymbol{y})$.

18. (NPP) Udowodnij, że jeśli zbiory $X, Y \subset \mathbb{R}^n$ są wypukłe, to $X \cap Y$ jest wypukły.

19. <u>Suma Minkowskiego</u> zbiorów $X, Y \subset \mathbb{R}^n$ jest to zbiór $X + Y = \{\boldsymbol{x} + \boldsymbol{y}\colon \boldsymbol{x} \in X, \boldsymbol{y} \in Y\}$. Wykaż, że jeśli zbiory X i Y są wypukłe, to zbiór $X + Y$ jest też wypukły.

20. Niech $X, Y \subset \mathbb{R}^n$ będą zwartymi zbiorami wypukłymi i niech funkcje $g\colon X \to \mathbb{R}$ i $h\colon Y \to \mathbb{R}$ będą wypukłe. Wykaż, że funkcja

$$f\colon X + Y \to \mathbb{R}, \quad f(\boldsymbol{z}) \stackrel{\text{def}}{=} \inf_{\boldsymbol{x} \in X, \boldsymbol{y} \in Y, \boldsymbol{x}+\boldsymbol{y}=\boldsymbol{z}} g(\boldsymbol{x}) + h(\boldsymbol{y})$$

jest wypukła. Czy i jak założenie o zwartości można osłabić? Czy istnieje związek między epigrafami funkcji f, g i h?

21. (NPP) Zbiór X jest wypukły, funkcje $f, g\colon X \to \mathbb{R}$ są wypukłe. Udowodnij, że
a) funkcja $f+g$ jest wypukła,
b) funkcja $\max(f,g)$ jest wypukła,
c) funkcja fg nie musi być wypukła. Podaj warunek dostateczny, aby była wypukła.

Rozw. a) Dla dowolnych $\boldsymbol{x}, \boldsymbol{y} \in X \subset \mathbb{R}^n$ oraz $\lambda \in [0,1]$ jest

$$\begin{aligned}\lambda(f+g)(\boldsymbol{x}) + (1-\lambda)(f+g)(\boldsymbol{y}) &= \lambda f(\boldsymbol{x}) + (1-\lambda)f(\boldsymbol{y}) + \lambda g(\boldsymbol{x}) + (1-\lambda)g(\boldsymbol{y}) \\ &\leqslant f\big(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\big) + g\big(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\big) = (f+g)\big(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\big).\end{aligned}$$

b) Epigraf funkcji $\max(f,g)$ jest przecięciem epigrafów funkcji f i g; jako przecięcie zbiorów wypukłych jest zbiorem wypukłym, przy czym jego rzutem na $\mathbb{R}^n$ jest zbiór X. Wypukłość funkcji $\max(f,g)$ jest natychmiastowym wnioskiem z twierdzenia o epigrafie.

22. (NPP) Niech $f\colon \mathbb{R} \to \mathbb{R}$ będzie funkcją wypukłą i rosnącą, a $g\colon W \to \mathbb{R}$, określona na zbiorze wypukłym $W \subset \mathbb{R}^n$, będzie wypukła. Wykaż, że $f \circ g$ jest wypukła.

Rozw. Dla dowolnych punktów $\boldsymbol{x}, \boldsymbol{y} \in W$ oraz liczby $\lambda \in (0,1)$ jest

$$g\big(\lambda x + (1-\lambda)y\big) \leqslant \lambda g(\boldsymbol{x}) + (1-\lambda)g(\boldsymbol{y}),$$

$$\begin{aligned} f\Big(g\big(\lambda x + (1-\lambda)y\big)\Big) &\underset{\substack{\uparrow \\ \text{bo } f \text{ rosnąca}}}{\leqslant} f\big(\lambda g(\boldsymbol{x}) + (1-\lambda)g(\boldsymbol{y})\big) \\ &\underset{\substack{\uparrow \\ \text{bo } f \text{ wypukła}}}{\leqslant} \lambda f\big(g(\boldsymbol{x})\big) + (1-\lambda)f\big(g(\boldsymbol{y})\big).\end{aligned}$$

23. (NPP) Udowodnij, że jeśli funkcja f jest (ściśle) wklęsła i $f>0$, to $1/f$ jest (ściśle) wypukła.

Rozw. Wystarczy zbadać, czy

$$\frac{1}{f\big((\boldsymbol{x}+\boldsymbol{y})/2\big)} \overset{?}{\leqslant} \frac{1}{2}\Big(\frac{1}{f(\boldsymbol{x})} + \frac{1}{f(\boldsymbol{y})}\Big). \tag{*}$$

Z wklęsłości f

$$\frac{f(\boldsymbol{x}) + f(\boldsymbol{y})}{2} \leqslant f\Big(\frac{\boldsymbol{x}+\boldsymbol{y}}{2}\Big), \quad \text{czyli} \quad \frac{2}{f(\boldsymbol{x}) + f(\boldsymbol{y})} \geqslant \frac{1}{f\big((\boldsymbol{x}+\boldsymbol{y})/2\big)}. \tag{**}$$

Sprawdzamy zatem, czy zachodzi nierówność

$$\frac{2}{f(\boldsymbol{x}) + f(\boldsymbol{y})} \overset{?}{\leqslant} \frac{1}{2f(\boldsymbol{x})} + \frac{1}{2f(\boldsymbol{y})} = \frac{f(\boldsymbol{y}) + f(\boldsymbol{x})}{2f(\boldsymbol{x})f(\boldsymbol{y})}.$$

Dzięki temu, że $f(\boldsymbol{x}), f(\boldsymbol{y}) > 0$ to jest równoważne

$$\big(f(\boldsymbol{x}) + f(\boldsymbol{y})\big)^2 \geqslant 4f(\boldsymbol{x})f(\boldsymbol{y}) \quad \Leftrightarrow \quad \big(f(\boldsymbol{x}) - f(\boldsymbol{y})\big)^2 \geqslant 0.$$

Rzeczywiście, ostatnia nierówność jest prawdziwa. Nierówności w (*) i (**) można zamienić na ostre.

Uwaga. Z tego, że funkcja f jest wypukła i dodatnia *nie wynika* wklęsłość funkcji $1/f$. Znajdź odpowiedni kontrprzykład.

24. (NPP) Niech $X \subset \mathbb{R}^n$ będzie zbiorem wypukłym i niech funkcja $f\colon X \to \mathbb{R}$ będzie wypukła. Niech $A \subset X$ oznacza zbiór punktów $\boldsymbol{x} \in X$, w których f osiąga minimalną wartość. Założymy, że $A \neq \emptyset$.
a) wykaż, że zbiór A jest wypukły,
b) wykaż, że każde minimum lokalne funkcji f w X jest minimum globalnym.

Rozw. a) Niech $\boldsymbol{x}, \boldsymbol{y} \in A$ oraz $\boldsymbol{a} = \lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}$ dla dowolnego $\lambda \in (0,1)$. Wtedy $\boldsymbol{a} \in X$ oraz

$$f(\boldsymbol{a}) = f\big(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\big) \leqslant \lambda f(\boldsymbol{x}) + (1-\lambda)f(\boldsymbol{y}) = f(\boldsymbol{x}) = f(\boldsymbol{y}) \leqslant f(\boldsymbol{a}).$$

Stąd musi być $\boldsymbol{a} \in A$.

b) Niech $\boldsymbol{x}^*$ będzie minimum lokalnym funkcji f. Przypuśćmy, że istnieje punkt $\overline{\boldsymbol{x}} \in X$, taki że $f(\overline{\boldsymbol{x}}) < f(\boldsymbol{x}^*)$, a więc $\boldsymbol{x}^*$ nie jest minimum globalnym. Z wypukłości f, dla każdego $\lambda \in (0,1)$ jest

$$f\big(\lambda\boldsymbol{x}^* + (1-\lambda)\overline{\boldsymbol{x}}\big) \leqslant \lambda f(\boldsymbol{x}^*) + (1-\lambda)f(\overline{\boldsymbol{x}}).$$

Ponieważ punkt $\boldsymbol{x}^*$ jest minimum lokalnym, istnieje takie $r>0$, że dla każdego $\boldsymbol{x} \in X$: $\|\boldsymbol{x} - \boldsymbol{x}^*\| \leqslant r$ jest $f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}^*)$, a zatem $\|\overline{\boldsymbol{x}} - \boldsymbol{x}^*\| > r$. Weźmy $\lambda^* = r/\|\overline{\boldsymbol{x}} - \boldsymbol{x}^*\|$; jest $\lambda^* \in (0,1)$ i mamy

$$\begin{aligned} f\big(\lambda^*\boldsymbol{x}^* + (1-\lambda^*)\overline{\boldsymbol{x}}\big) \geqslant f(\boldsymbol{x}^*) &= \lambda^* f(\boldsymbol{x}^*) + (1-\lambda^*)f(\boldsymbol{x}^*) \\ &> \lambda^* f(\boldsymbol{x}^*) + (1-\lambda^*)f(\overline{\boldsymbol{x}}).\end{aligned}$$

Z przypuszczenia wynika sprzeczność.

25. (NPP) Niech $f\colon W \to \mathbb{R}$, przy czym zbiór $W \subset \mathbb{R}^n$ jest wypukły i $f \in C^2$. Niech $D^2f(\boldsymbol{x}) \geqslant 0$ dla każdego $\boldsymbol{x} \in W$ oraz $Df(\boldsymbol{x}^*) = \boldsymbol{0}$ dla pewnego $\boldsymbol{x}^* \in W$. Wykaż, że punkt $\boldsymbol{x}^*$ jest minimum globalnym.

Rozw. Na podstawie wzoru Taylora

$$f(\boldsymbol{x}) = f(\boldsymbol{x}^*) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}^*)^T D^2 f(\boldsymbol{y})(\boldsymbol{x} - \boldsymbol{x}^*),$$

gdzie $\boldsymbol{y}$ jest pewnym punktem odcinka $\overline{\boldsymbol{x}^*\boldsymbol{x}}$, a więc elementem zbioru W. Jest $D^2f(\boldsymbol{y}) \geqslant 0$, a zatem $f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}^*)$.

26. (NPP) Znajdź przykład funkcji $f \in C^2$ która ma $D^2f|_x \geqslant 0$ dla każdego $x \in X \subset \mathbb{R}^n$, ma minimum lokalne w zbiorze X i nie ma minimum globalnego.

Wskazówka. Zbiór X nie może być wypukły.

27. (NPP) Udowodnij, że funkcja

$$f(x_1, x_2) = x_1^4 + x_1^2 - x_1x_2 + 2x_2^2$$

osiąga minimum globalne na zbiorze $X = \{\,(x_1, x_2)\colon x_1 \geqslant -1\,\}$.

Rozw. Zbiór X jest nieograniczony i domknięty. Funkcja f jest ciągła (bo to wielomian) i koercywna, bo

$$f(x_1, x_2) = x_1^4 + (x_1 - \frac{1}{2}x_2)^2 + \frac{7}{4}x_2^2 \xrightarrow[\|(x_1,x_2)\|\to\infty]{} \infty.$$

28. (NPP) Niech $W \subset \mathbb{R}^n$ oznacza zbiór wypukły otwarty i niech funkcje $g, h\colon W \to \mathbb{R}$ będą różniczkowalne. Wykaż, że jeśli zachodzi jeden z warunków:
a) g — wypukła i nieujemna, h — wklęsła i dodatnia,
b) g — wypukła i niedodatnia, h — wypukła i dodatnia,
to funkcja $f = g/h$ jest pseudowypukła.

Rozw. a) Gradient ilorazu jest opisany wzorem

$$Df|_x = \frac{Dg|_x h(x) - g(x)Dh|_x}{\big(h(x)\big)^2}.$$

Funkcja f jest pseudowypukła $\Leftrightarrow \big(Df|_{\overline{x}}(x - \overline{x}) \geqslant 0 \Rightarrow f(x) \geqslant f(\overline{x})\big)$.

Na podstawie wzoru na gradient,

$$Df|_{\overline{x}}(x - \overline{x}) \geqslant 0 \;\Leftrightarrow\; Dg|_{\overline{x}}(x - \overline{x})h(\overline{x}) \geqslant g(\overline{x})Dh|_{\overline{x}}(x - \overline{x}), \qquad (*)$$

a ponieważ $h > 0$, jest też

$$f(x) \geqslant f(\overline{x}) \;\Leftrightarrow\; g(x)h(\overline{x}) \geqslant g(\overline{x})h(x). \qquad (**)$$

Z istnienia płaszczyzny podpierającej mamy

$$h(\overline{x})g(x) \geqslant h(\overline{x})\big(g(\overline{x}) + Dg|_{\overline{x}}(x - \overline{x})\big),$$
$$g(\overline{x})h(x) \leqslant g(\overline{x})\big(h(\overline{x}) + Dh|_{\overline{x}}(x - \overline{x})\big).$$

Czyli jeśli jest spełniona nierówność

$$h(\overline{x})\big(g(\overline{x}) + Dg|_{\overline{x}}(x - \overline{x})\big) \geqslant g(\overline{x})\big(h(\overline{x}) + Dh|_{\overline{x}}(x - \overline{x})\big),$$

to zachodzi (**). Ale to wynika z (*).

b) (prawie) tak samo, jak a).

29. Znajdź zbiór punktów ekstremalnych zbioru

$$X = \{\,(x, y, z)\colon y^2 + z^2 \leqslant (x + 1)^2,\ x \geqslant 0,\ x \leqslant 1,\ z \geqslant 0\,\}.$$

30. Znajdź maksimum funkcji $f(x, y, z) = (x - 2)^2 + y + z$ w zbiorze X z poprzedniego zadania.

31. (NPP) Zbiór $W \subset \mathbb{R}^n$ jest otwarty i wypukły, funkcje $g, h\colon W \to \mathbb{R}$ są różniczkowalne. Wykaż, że jeśli funkcja g jest wypukła i niedodatnia, a h jest wklęsła i dodatnia, to funkcja $f = gh$ jest pseudowypukła.

Rozw. Z zadania 23 funkcja $1/h$ jest wypukła, jest też dodatnia. Reszta wynika z zadania 28.

32. (NPP) Niech $S \subset \mathbb{R}$ i funkcja $f\colon S \to \mathbb{R}$ będzie pseudowypukła. Wykaż, że jeśli $f'(x_0) = 0$, to x_0 jest minimum funkcji f na S.

Rozw. Z definicji pseudowypukłości, $\forall_{x \in S}\, f'(x_0)(x - x_0) \geqslant 0 \;\Rightarrow\; f(x) \geqslant f(x_0)$. Ponieważ $f'(x_0) = 0$, punkt x_0 nie może nie być minimum.

33. (NPP) Czy zbiór $W = \{\,(x_1, x_2, x_3)\colon x_1^6 + x_2^2 + x_2x_3 + x_3^2 \leqslant 1\,\}$ jest wypukły?

Rozw.

$$g(x_1, x_2, x_3) = x_1^6 + x_2^2 + x_2x_3 + x_3^2 - 1,$$
$$Dg = [6x_1^5, 2x_2 + x_3, 2x_3 + x_2],$$
$$D^2g = \begin{bmatrix} 30x_1^4 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix}.$$

Macierz D^2g jest nieujemnie określona, zatem funkcja g jest wypukła i jej zbiór poziomicowy $W = \{\,x\colon g(x) \leqslant 0\,\}$ jest wypukły.

34. (NPP) Zbiór $W \subset \mathbb{R}^n$ jest niepusty, wypukły, domknięty i ograniczony. Funkcja $f\colon W \to \mathbb{R}$ jest wklęsła. Niech $A = \{\,x\colon f(x) = \min_{y \in W} f(y)\,\}$. Udowodnij, że do zbioru A należą pewne punkty ekstremalne zbioru W.

Rozw. $A \neq \emptyset$, bo funkcja f jest ciągła na zbiorze zwartym. Niech $x = a_1z_1 + \cdots + a_kz_k \in A$, przy czym punkty $z_1, \ldots z_k$ są punktami ekstremalnymi, $a_1, \ldots, a_k \geqslant 0$ oraz $a_1 + \cdots + a_k = 1$.

Ponieważ funkcja f jest wklęsła, jest $f(x) \geqslant a_1(fz_1) + \cdots + a_kf(z_k)$, ale $f(x) \leqslant f(z_i)$ dla każdego i z definicji zbioru A. Zatem, *wszystkie* punkty z_i, takie że $a_i > 0$ (co najmniej jeden spośród $z_1, \ldots, z_k$) spełniają równość $f(z_i) = f(x)$, a więc należą do A.

35. (NPP) Rozwiąż zadanie minimalizacji z ograniczeniami:

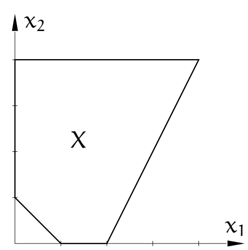
$$\begin{cases} \log_2(x_1+x_2) - 2|x_2 - 2x_1| - x_2^2 + 2x_2 \to \min, \\ X = \{\, (x_1,x_2)\colon x_1 \geqslant 0,\ x_2 \geqslant 0,\ |x_1 - x_2| + |x_1| \leqslant 4,\ x_1 + x_2 \geqslant 1 \,\} \end{cases}$$

<u>Rozw.</u>

$$\begin{aligned} g(\boldsymbol{x}) &= \log_2(x_1+x_2) - 2|x_2 - 2x_1| - x_2^2 + 2x_2, \\ Dg(\boldsymbol{x}) &= \left[\frac{1}{(x_1+x_2)\ln 2} + 4\,\mathrm{sgn}(x_2 - 2x_1),\ \frac{1}{(x_1+x_2)\ln 2} - 2\,\mathrm{sgn}(x_2-2x_1) + 2x_2 + 2\right], \\ D^2g(\boldsymbol{x}) &= \begin{bmatrix} \frac{-1}{(x_1+x_2)^2\ln 2} & \frac{-1}{(x_1+x_2)^2\ln 2} \\ \frac{-1}{(x_1+x_2)^2\ln 2} & \frac{-1}{(x_1+x_2)^2\ln 2} - 2 \end{bmatrix}. \end{aligned}$$

Na podstawie kryterium Sylvestra $D^2g < 0$, zatem funkcja g jest wklęsła.

Badamy zbiór X: Ponieważ $x_1 \geqslant 0$, nierówność $|x_1 - x_2| + |x_1| \leqslant 4$ można zastąpić przez $|x_1 - x_2| + x_1 \leqslant 4$ i dalej $2x_1 - 4 \leqslant x_2 \leqslant 4$. Zbiór X jest wypukłym pięciokątem o wierzchołkach (punktach ekstremalnych) $(0,1)$, $(1,0)$, $(2,0)$, $(0,4)$, $(4,4)$. Obliczamy wartości funkcji g w tych punktach:



$$\begin{aligned} &g(0,1) = 1, && g(2,0) = -7, && g(4,4) = -13, \\ &g(1,0) = -4, && g(0,4) = -14. \end{aligned}$$

Minimum funkcji g jest w punkcie $(0,4)$.

36. (NPP) Znajdź maksimum funkcji $f(x,y) = \sin x + \sin y - \sin(x+y)$ w zbiorze $X = \{\,(x,y)\colon x \geqslant 0,\ y \geqslant 0,\ x + y \leqslant 2\pi\}$.

<u>Rozw.</u> Maksimum istnieje, bo funkcja jest ciągła na zbiorze zwartym X.

$$Df(x,y) = [\cos x - \cos(x+y), \cos y - \cos(x+y)].$$

Gradient f jest równy $\boldsymbol{0}$, jeśli $\cos x = \cos y$, stąd $x = y$ lub $y = 2\pi - x$. Jest $\cos(x+y) = \cos x \cos y - \sin x \sin y$.

Dla $x = y$: $\cos(2x) = 2\cos^2 x - 1$, jeśli $z = \cos x$, to mamy $2z^2 - z - 1 = 0$, stąd $z = 1$ (i wtedy $x = y = 0$) lub $z = -1/2$ (i wtedy $x = y = 2\pi/3$).

Dla $y = 2\pi - x$: $\cos x = \cos y$ i $\sin x = -\sin y$. Mamy dwa rozwiązania: $x = 0, y = 2\pi$ oraz $x = 2\pi, y = 0$.

Wewnątrz zbioru X jest miejsce zerowe gradientu: $x = y = 2\pi/3$. Funkcja f ma na całym brzegu zbioru X wartość 0, bo

$$\begin{aligned} f(x,0) &= \sin x - \sin x = 0, \\ f(0,y) &= \sin y - \sin y = 0, \\ f(x, 2\pi - x) &= \sin x - \sin x + \sin(2\pi) = 0. \end{aligned}$$

W punkcie $(2\pi/3, 2\pi/3)$ jest $f(2\pi/3, 2\pi/3) = 3\sqrt{3}/2 > 0$ — to jest wartość maksymalna.

37. a) Udowodnij, że dowolna norma w $\mathbb{R}^n$ jest funkcją wypukłą.
b) Niech M oznacza wypukły, domknięty i ograniczony podzbiór $\mathbb{R}^n$, którego wnętrze jest niepuste i który spełnia warunek $\boldsymbol{x} \in M \Leftrightarrow -\boldsymbol{x} \in M$. Udowodnij, że funkcja nieujemna $f\colon \mathbb{R}^n \to \mathbb{R}$, określona wzorem

$$f(\boldsymbol{x}) = \begin{cases} \min\limits_{t\colon t\boldsymbol{x} \in M} \dfrac{1}{|t|} & \text{dla } \boldsymbol{x} \neq \boldsymbol{0}, \\ 0 & \text{dla } \boldsymbol{x} = \boldsymbol{0}, \end{cases}$$

jest normą.

38. Niech wielomian $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T A\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{x} + c$ będzie ściśle wypukłą funkcją w $\mathbb{R}^n$. Udowodnij, że minimum funkcji f w zbiorze rozwiązań układu równań liniowych $C\boldsymbol{x} = \boldsymbol{d}$, którego macierz ma niezależne liniowo wiersze, jest blokiem $\boldsymbol{x}$ w jedynym rozwiązaniu układu równań liniowych

$$\begin{cases} \frac{1}{2}(A^T + A)\boldsymbol{x} + C^T\boldsymbol{\lambda} = -\boldsymbol{b}, \\ C\boldsymbol{x} \qquad\qquad\qquad\ = \boldsymbol{d} \end{cases}$$

<u>Rozw. I</u> Jeśli $C \in \mathbb{R}^{k\times n}$, to mamy tu k regularnych funkcji ($g_i(\boldsymbol{x}) = \boldsymbol{c}_i^T\boldsymbol{x} - d_i$), które opisują ograniczenia równościowe w zadaniu minimalizacji. Pierwszy podukład, tj. $\frac{1}{2}(A^T + A)\boldsymbol{x} + C^T\boldsymbol{\lambda} = -\boldsymbol{b}$ wyraża równość $Df|_x + Dg|_x\boldsymbol{\lambda} = 0$, przy czym współrzędne wektora $\boldsymbol{\lambda} \in \mathbb{R}^k$ są mnożnikami Lagrange'a.

<u>Rozw. II</u> Niech $S = \frac{1}{2}(A^T + A)$; macierz S jest symetryczna i zachodzi równość $f(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T S\boldsymbol{x} + \boldsymbol{b}^T\boldsymbol{x} + c$ (zobacz zad. 8). Ścisła wypukłość funkcji f oznacza, że macierz S jest dodatnio określona. Niech $\boldsymbol{x}$ spełnia badany układ równań razem z pewnym wektorem $\boldsymbol{\lambda} \in \mathbb{R}^k$. W szczególności wektor $\boldsymbol{x}$ jest rozwiązaniem układu $C\boldsymbol{x} = \boldsymbol{d}$. Każde inne rozwiązanie tego układu ma postać $\boldsymbol{x} + \boldsymbol{z}$, gdzie $C\boldsymbol{z} = \boldsymbol{0}$.

Obliczmy

$$f(\boldsymbol{x} + \boldsymbol{z}) = f(\boldsymbol{x}) + \frac{1}{2}\boldsymbol{z}^T S\boldsymbol{z} + \boldsymbol{x}^T S\boldsymbol{z} + \boldsymbol{b}^T\boldsymbol{z}.$$

Ale jest $x^T S z = -\lambda^T C z - b^T z = -b^T z$, skąd wynika, że

$$f(x+z) = f(x) + \frac{1}{2} z^T S z.$$

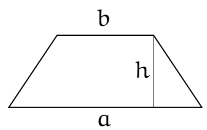
Ponieważ macierz S jest dodatnio określona, dla każdego $z \neq 0$ jest $f(x) < f(x+z)$.

Trzeba jeszcze sprawdzić, że układ ma rozwiązanie. Z pierwszego podukładu musi być $x = -S^{-1}(b + C^T\lambda)$ (S^{-1} istnieje), po wstawieniu do drugiego $d = Cx = -CS^{-1}b - CS^{-1}C^T\lambda$, czyli $CS^{-1}C^T\lambda = -CS^{-1}b - d$. Ponieważ wiersze C są liniowo niezależne, macierz $CS^{-1}C^T$ jest dodatnio określona — nieosobliwa — zatem λ oraz x są określone jednoznacznie.

39. (NPP) Metalowa belka ma przekrój w kształcie symetrycznego trapezu o ustalonym polu S. Górna i boczna powierzchnia ma być pomalowana farbą antykorozyjną. Ustal kształt przekroju tak, aby powierzchnia do pomalowania była jak najmniejsza.

Rozw. Dolny bok trapezu ma długość a, górny b, wysokość trapezu jest równa h. Należy zminimalizować funkcję

$$f(a,b,h) = b + 2\sqrt{\left(\frac{a-b}{2}\right)^2 + h^2}$$



w zbiorze miejsc zerowych funkcji

$$g(a,b,h) = \frac{a+b}{2}h - S.$$

Niech $x = a+b$, $y = a-b$, oraz $\hat{f}(x,y,h) = (x-y)/2 + \sqrt{y^2+4h^2}$, $\hat{g}(x,y,h) = xh/2 - S$. Wtedy mamy znaleźć minimum funkcji $\hat{f}$ w zbiorze miejsc zerowych funkcji $\hat{g}$. Stosujemy mnożniki Lagrange'a.

$$D\hat{f} = \left[\frac{1}{2}, -\frac{1}{2} + \frac{y}{\sqrt{y^2+4h^2}}, \frac{4h}{\sqrt{y^2+4h^2}}\right],$$
$$D\hat{g} = \left[\frac{h}{2}, 0, \frac{x}{2}\right].$$

Ponieważ dopuszczamy tylko $h > 0$ i $x > 0$, $D\hat{g}$ nie znika. Rozwiązujemy równanie

$$D\hat{f} + \lambda D\hat{g} = 0,$$

czyli układ równań

$$\frac{1}{2} + \lambda\frac{h}{2} = 0, \quad -\frac{1}{2} + \frac{y}{\sqrt{y^2+4h^2}} = 0, \quad \frac{4h}{\sqrt{y^2+4h^2}} + \lambda\frac{x}{2} = 0.$$

Z pierwszego $h = -1/\lambda$, z drugiego $y^2 = h^2 + y^2/4$, czyli $y^2 = 4h^2/3 = 4/(3\lambda^2)$, z trzeciego

$$x = \frac{-8h}{\lambda\sqrt{y^2+4h^2}} = \frac{8}{\lambda^2\sqrt{\frac{4}{3\lambda^2} + \frac{4}{\lambda^2}}} = \frac{-2\sqrt{3}}{\lambda}.$$

Z ograniczenia $\hat{g} = 0$ dostajemy $2\sqrt{3}/(2\lambda^2) = S$, czyli $\lambda = -\sqrt[4]{3}/\sqrt{S}$ (bo $\lambda < 0$). Musi być $y \geqslant 0$, stąd otrzymujemy

$$x = 2\sqrt[4]{3}\sqrt{S}, \quad y = \frac{2}{3}\sqrt[4]{3}\sqrt{S}, \quad h = \frac{\sqrt{S}}{\sqrt[4]{3}}.$$

Czy to jest minimum globalne? To jest możliwe. Dla $h \searrow 0$ jest $x \nearrow \infty$ i $\hat{f} \nearrow \infty$, podobnie dla $x \searrow 0$ jest $h \nearrow \infty$ i $\hat{f} \nearrow \infty$, czyli na brzegu nie ma minimum. Należałoby sprawdzić warunek II-go rzędu ($D^2\hat{f} \geqslant 0$).

40. (NPP) Rozwiąż zadanie minimalizacji

$$\begin{cases} x_1 + x_2 + x_3 + x_4 \to \min, \\ x_1x_2x_3x_4 = c^4, \\ x_1, x_2, x_3, x_4 \geqslant 0. \end{cases}$$

dla $c > 0$.

Rozw. $f(x_1,x_2,x_3,x_4) = x_1 + x_2 + x_3 + x_4$, $g(x_1,x_2,x_3,x_4) = x_1x_2x_3x_4 - c^4$.

$$Df = [1,1,1,1],$$
$$Dg = [x_2x_3x_4, x_1x_3x_4, x_1x_2x_4, x_1x_2x_3].$$

Jest $Dg \neq 0$, bo x_i są nieujemne i ich iloczyn jest dodatni. Równanie

$$Df + \lambda Dg = 0$$

po rozpisaniu daje

$$\begin{aligned} x_2x_3x_4 = x_1x_3x_4 \quad &\Rightarrow \quad x_1 = x_2, \\ x_2x_3x_4 = x_1x_2x_4 \quad &\Rightarrow \quad x_1 = x_3, \\ x_1x_3x_4 = x_1x_2x_3 \quad &\Rightarrow \quad x_2 = x_4, \end{aligned}$$

skąd $x_1 = x_2 = x_3 = x_4 = c$. Z równości $1 + \lambda x_2x_3x_4 = 0$ wynika $\lambda = -1/c^3$.

Określamy $L(\mathbf{x},\lambda) = f(\mathbf{x}) + \lambda g(\mathbf{x})$. Jest

$$D^2_{\mathbf{xx}}L(\mathbf{x},\lambda) = \lambda \begin{bmatrix} 0 & x_3x_4 & x_2x_4 & x_2x_3 \\ x_3x_4 & 0 & x_1x_4 & x_1x_3 \\ x_2x_4 & x_1x_4 & 0 & x_1x_2 \\ x_2x_3 & x_1x_3 & x_1x_2 & 0 \end{bmatrix},$$

$$D^2_{\mathbf{xx}}|_{((c,c,c,c),1/c^3)}L(\mathbf{x},\lambda) = -\frac{c^2}{c^3}\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} = -\frac{1}{c}\begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} =: H.$$

Bierzemy wektor $\mathbf{d} = [d_1, d_2, d_3, d_4]^T \neq \mathbf{0}$, taki że $Dg|_{((c,c,c,c),1/c^3)}\mathbf{d} = 0$; ponieważ $Dg|_{((c,c,c,c),1/c^3)} = c^3[1,1,1,1]$, musi być $d_1 + d_2 + d_3 + d_4 = 0$. Obliczamy

$$\mathbf{d}^T H\mathbf{d} = -\frac{1}{c}\mathbf{d}^T \begin{bmatrix} d_2+d_3+d_4 \\ d_1+d_3+d_4 \\ d_1+d_2+d_4 \\ d_1+d_2+d_3 \end{bmatrix} = -\frac{1}{c}\mathbf{d}^T \begin{bmatrix} -d_1 \\ -d_2 \\ -d_3 \\ -d_4 \end{bmatrix} = \frac{d_1^2+d_2^2+d_3^2+d_4^d}{c}.$$

Zatem, $\mathbf{d}^T H\mathbf{d} > 0$, co oznacza, że w punkcie $x_1 = x_2 = x_3 = x_4 = c$ jest minimum.

41. Rozwiąż zadanie 15 metodą mnożników Lagrange'a.

Rozw. Mamy

$$\begin{cases} f(x,y,z) = 2x + 2y - 3z, \\ g_1(x,y,z) = x^2 + y^2 + z^2 - 2, \\ g_2(x,y,z) = 2xy + 2yz + 2zx + 1. \end{cases}$$

Określamy

$$\begin{aligned} \tilde{g}_1(x,y,z) &= g_1(x,y,z) + g_2(x,y,z) = (x+y+z)^2 - 1, \\ \tilde{g}_2(x,y,z) &= 2g_1(x,y,z) - g_2(x,y,z) = (x-y)^2 + (y-z)^2 + (z-x)^2 - 5. \end{aligned}$$

Obliczamy

$$\begin{aligned} Df &= [2,2,-3], \\ D\tilde{g}_1 &= 2(x+y+z)[1,1,1], \\ D\tilde{g}_2 &= [2x-y-z, 2y-z-x, 2z-x-y]. \end{aligned}$$

Trzeba rozwiązać układ równań

$$\begin{cases} Df + \lambda_1 D\tilde{g}_1 + \lambda_2 D\tilde{g}_2 = \mathbf{0}, \\ \tilde{g}_1 = 0, \\ \tilde{g}_2 = 0, \end{cases}$$

czyli

$$\begin{cases} \begin{bmatrix} 2 \\ 2 \\ -3 \end{bmatrix} + \lambda_1 2(x+y+z)\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \lambda_2 \begin{bmatrix} 2x-y-z \\ 2y-z-x \\ 2z-x-y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \\ (x+y+z)^2 = 1, \\ (x-y)^2 + (y-z)^2 + (z-x)^2 = 5. \end{cases}$$

Z równania $\tilde{g}_1 = 0$ wynika $x + y + z = \pm 1$. Rozpatruję przypadek $x + y + z = +1$. Wtedy

$$\begin{bmatrix} 2 \\ 2 \\ -3 \end{bmatrix} + \lambda_1 \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} + \lambda_2 \begin{bmatrix} 3x-1 \\ 3y-1 \\ 3z-1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

skąd wynika, że

$$\begin{aligned} \lambda_2(x-y) &= 0, \\ 1 - 3\lambda_2(y-z) &= 0. \end{aligned}$$

Musi być $\lambda_2 \neq 0$, a zatem $x = y \neq z$. Podstawiając to do równań $\tilde{g}_1 = \tilde{g}_2 = 0$, dostajemy

$$2y + z = 1, \qquad 2(y-z)^2 = 5,$$

co po rozwiązaniu daje $x = y = \frac{1}{3} \pm \frac{\sqrt{5}}{3\sqrt{2}}$, $z = 1 - 2y = \frac{1}{3} \mp \frac{\sqrt{10}}{3}$.

Teraz przypadek $x + y + z = -1$. Z układu równań

$$\begin{bmatrix} 2 \\ 2 \\ -3 \end{bmatrix} + \lambda_1 \begin{bmatrix} -2 \\ -2 \\ -2 \end{bmatrix} + \lambda_2 \begin{bmatrix} 3x+1 \\ 3y+1 \\ 3z+1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

po wyeliminowaniu λ_1 otrzymujemy

$$\begin{aligned} \lambda_2(x-y) &= 0, \\ 1 + 3\lambda_2(y-z) &= 0. \end{aligned}$$

Jak poprzednio, musi być $\lambda_2 \neq 0$ oraz $x = y \neq z$. Po rozwiązaniu równań ograniczeń dostajemy $x = y = -\frac{1}{3} \mp \frac{\sqrt{5}}{3\sqrt{2}}$, $z = 1 - 2y = -\frac{1}{3} \pm \frac{\sqrt{10}}{3}$.

Jest $f\big(\frac{1}{3} + \frac{\sqrt{5}}{3\sqrt{2}}, \frac{1}{3} + \frac{\sqrt{5}}{3\sqrt{2}}, \frac{1}{3} - \frac{\sqrt{10}}{3}\big) = (1 + 5\sqrt{10})/3$ — maksimum,
$f\big(\frac{1}{3} - \frac{\sqrt{5}}{3\sqrt{2}}, \frac{1}{3} - \frac{\sqrt{5}}{3\sqrt{2}}, \frac{1}{3} + \frac{\sqrt{10}}{3}\big) = (1 - 5\sqrt{10})/3$.
$f\big(-\frac{1}{3} - \frac{\sqrt{5}}{3\sqrt{2}}, -\frac{1}{3} - \frac{\sqrt{5}}{3\sqrt{2}}, -\frac{1}{3} + \frac{\sqrt{10}}{3}\big) = -(1 + 5\sqrt{10})/3$ — minimum,
$f\big(-\frac{1}{3} + \frac{\sqrt{5}}{3\sqrt{2}}, -\frac{1}{3} + \frac{\sqrt{5}}{3\sqrt{2}}, -\frac{1}{3} - \frac{\sqrt{10}}{3}\big) = -(1 - 5\sqrt{10})/3$.

42. (EII) Niech

$$H = \{ (x, y, z) \in \mathbb{R}^3 \colon x^2 + y^2 - z^2 + 4 = 0 \}.$$

Znajdź w zbiorze H punkt, którego odległość od $(2, 4, 0)$ jest najmniejsza.

Rozw. I Zbiór H opisany równaniem $x^2 + y^2 = z^2 - 4$ jest hiperboloidą dwupowłokową. Dla $|z| > 2$ mamy równanie okręgu $x^2 + y^2 = r^2$, gdzie $r^2 = z^2 - 4$. Możemy znaleźć najbliższy punkt na każdym takim okręgu — ma on współrzędne $y = 2x > 0$. Podstawiamy $5x^2 = z^2 - 4$, stąd $z^2 = 5x^2 + 4$. Kwadrat odległości od $(2, 4, 0)$ jest równy

$$d^2 = (x - 2)^2 + (y - 4)^2 + z^2.$$

Określamy

$$f(x) = d^2 = (x - 2)^2 + (y - 4)^2 + z^2 = 10x^2 - 12x + 24.$$

Stąd $f(x) = 20x - 12$, $f(x) = 0 \;\Rightarrow\; x = \frac{6}{10}$, $y = \frac{12}{10}$ oraz $z = \pm\sqrt{\frac{29}{5}}$.

Rozw. II Metoda mnożników Lagrange'a: Określamy

$$\begin{aligned} f(x, y, z) &= (x - 2)^2 + (y - 4)^2 + z2, & Df &= [2x - 4, 2y - 8, 2z], \\ g(x, y, z) &= x2 + y2 - z2 + 4, & Dg &= [2x, 2y, -2z] \end{aligned}$$

i znajdujemy minimum funkcji f w zbiorze H miejsc zerowych funkcji g, rozwiązując układ równań

$$\begin{cases} Df + \lambda Dg &= \mathbf{0}, \\ g(x, y, z) &= 0. \end{cases}$$

43. (NPP) Rozwiąż zadanie

$$\begin{cases} -x_1x_2 - x_2x_3 - x_1x_3 \to \min, \\ x_1 + x_2 + x_3 = 3. \end{cases}$$

Rozw. Mamy tu funkcje $f(x_1, x_2, x_3) = -x_1x_2 - x_2x_3 - x_1x_3$, $g(x_1, x_2, x_3) = x_1 + x_2 + x_3 - 3$. Jest

$$\begin{aligned} Df &= [-x_2 - x_3, -x_1 - x_3, -x_1 - x_2], \\ Dg &= [1, 1, 1]. \end{aligned}$$

Ograniczenie jest regularne. Gradient funkcji Lagrange'a ma być równy $\mathbf{0}$:

$$D_x L(\boldsymbol{x}, \lambda) = [-x_2 - x_3, -x_1 - x_3, -x_1 - x_2] + \lambda[1, 1, 1] = \mathbf{0},$$

stąd $x_1 = x_2 = x_3 = 1$, $\lambda = 2$. Hesjan funkcji Lagrange'a

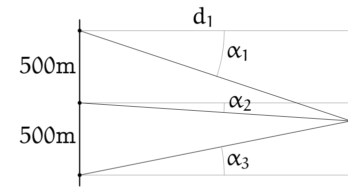
$$D^2_{xx} L(\boldsymbol{x}, \lambda) = \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & -1 \\ -1 & -1 & 0 \end{bmatrix} =: H.$$

Niezerowy wektor $\mathbf{d} \in \mathbb{R}^3$ taki, że $[1, 1, 1]\mathbf{d} = 0$, jest kombinacją liniową wektorów $\mathbf{d}_1 = [1, 0, -1]^T$ i $\mathbf{d}_2 = [0, 1, -1]^T$. Macierz $B = [\mathbf{d}_1, \mathbf{d}_2]$ umożliwia zapisanie tego w postaci $\mathbf{d} = B\boldsymbol{z}$, gdzie $\boldsymbol{z} \in \mathbb{R}^2$. Wtedy

$$\mathbf{d}^T H \mathbf{d} = \boldsymbol{z}^T B^T H B \boldsymbol{z} = \boldsymbol{z}^T C \boldsymbol{z}, \quad \text{gdzie macierz} \quad C = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

jest dodatnio określona. Kończy to dowód, że znaleźliśmy minimum.

44. (NPP) Dla wyznaczenia pozycji statku zostały z brzegu wykonane pomiary kątów: $\tilde{\alpha}_1 = 0.083$, $\tilde{\alpha}_2 = 0.024$, $\tilde{\alpha}_3 = -0.017$. Należy je skorygować tak, aby proste łączące odpowiednie punkty na brzegu ze statkiem przecinały się w jednym punkcie, a suma kwadratów poprawek była najmniejsza.



Rozw. Określamy funkcję $f(\alpha_1, \alpha_2, \alpha_3) = \sum_{i=1}^{3} (\alpha_i - \tilde{\alpha}_i)^2$, mamy znaleźć jej minimum. Ze zgodności kątów jest

$$\operatorname{tg} \alpha_1 = \frac{d_2}{d_1}, \qquad \operatorname{tg} \alpha_2 = \frac{d_2 - 500}{d_1}, \qquad \operatorname{tg} \alpha_3 = \frac{d_2 - 1000}{d_1}.$$

Stąd $\operatorname{tg} \alpha_1 - \operatorname{tg} \alpha_3 = 2(\operatorname{tg} \alpha_1 - \operatorname{tg} \alpha_2)$. Biorąc $\operatorname{tg} \alpha \approx \alpha$, dostajemy

$$g(\alpha_1, \alpha_2, \alpha_3) = \alpha_1 - 2\alpha_2 + \alpha_3 = 0.$$

Dalej, mamy

$$\begin{aligned} Df &= 2[\alpha_1 - \tilde{\alpha}_1, \alpha_2 - \tilde{\alpha}_2, \alpha_3 - \tilde{\alpha}_3], \\ Dg &= [1, -2, 1]. \end{aligned}$$

Rozwiązujemy równanie

$$\mathbf{0} = Df + \lambda Dg = [2(\alpha_1 - \tilde{\alpha}_1) + \lambda, 2(\alpha_2 - \tilde{\alpha}_2) - 2\lambda, 2(\alpha_3 - \tilde{\alpha}_3) + \lambda].$$

Warto cały układ zapisać w postaci macierzowej:

$$\begin{bmatrix} 2 & 0 & 0 & 1 \\ 0 & 2 & 0 & -2 \\ 0 & 0 & 2 & 1 \\ 1 & -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \lambda \end{bmatrix} = \begin{bmatrix} \tilde{\alpha}_1 \\ \tilde{\alpha}_2 \\ \tilde{\alpha}_3 \\ 0 \end{bmatrix}.$$

Po eliminacji α_1, α_2, α_3 z ostatniego równania dostajemy
$6\lambda = \tilde\alpha_1 - 2\tilde\alpha_2 + \tilde\alpha_3$, czyli $\lambda = \frac{1}{6}(\tilde\alpha_1 - 2\tilde\alpha_2 + \tilde\alpha_3) = 0.003$ i dalej
$\alpha_1 = \tilde\alpha_1 - \lambda = 0.08$, $\alpha_2 = \tilde\alpha_2 + 2\lambda = 0.03$ i $\alpha_3 = \tilde\alpha_3 - \lambda = -0.02$. Funkcja f jest ściśle wypukła, mamy zatem minimum. Możemy dalej obliczyć
$d_1 = 1000\mathrm{m}/(\alpha_1 - \alpha_3) = 10000\mathrm{m}$, $d_2 = d_1\alpha_1 = 800\mathrm{m}$.

45. (NPP) Rozwiąż zadanie

$$\begin{cases} f(x_1, x_2) = 2x_1 + x_2 \to \min, \\ g(x_1, x_2) = x_1^2 + x_2^2 - 1 \leqslant 0. \end{cases}$$

<u>Rozw.</u> Jest $Dg = [2x_1, 2x_2]$ — ograniczenie g jest regularne (tj. $Dg \neq \mathbf{0}$ w zbiorze miejsc zerowych funkcji g). Rozwiązujemy układ równań i nierówności

$$\begin{cases} Df + \lambda Dg = 0, \\ \lambda g \qquad = 0, \\ \lambda \qquad \geqslant 0. \end{cases}$$

Jeśli $\lambda = 0$, to $D_x L(x, \lambda) = Df \neq \mathbf{0}$, czyli nie ma rozwiązań z $\lambda = 0$.

Badamy zatem, czy jest rozwiązanie z $\lambda > 0$. Z równania $D_x L(x, \lambda) = \mathbf{0}$ jest

$$2 + 2\lambda x_1 = 0 \quad \Rightarrow \quad x_1 = -1/\lambda,$$
$$1 + 2\lambda x_2 = 0 \quad \Rightarrow \quad x_2 = -1/(2\lambda).$$

Ponieważ $\lambda > 0$, musi być $g(x_1, x_2) = 0$, czyli

$$\frac{1}{\lambda^2} + \frac{1}{4\lambda^2} = 1, \qquad \Rightarrow \qquad \lambda = \frac{\sqrt5}{2},$$

dalej $x_1 = -2/\sqrt5$, $x_2 = -1/\sqrt5$. Warunek II-go rzędu jest spełniony, bo funkcja $L(x, \lambda)$ jest ściśle wypukła jako funkcja zmiennej x.

46. (NPP) Rozwiąż zadanie

$$\begin{cases} f(x_1, x_2) = x_2 \to \min, \\ g_1(x_1, x_2) = -x_1 - x_2 + 2 \leqslant 0, \\ g_2(x_1, x_2) = x_1^2 + x_2^2 - 2 \leqslant 0. \end{cases}$$

<u>Rozw.</u> Jeśli zbiór $\{x\colon g_1(x) \leqslant 0, g_2(x) \leqslant 0\}$ jest niepusty, to minimum istnieje, bo ten zbiór jest zwarty i funkcja f jest ciągła. Określamy

$$L(x, \boldsymbol\lambda) = f(x) + \lambda_1 g_1(x) + \lambda_2 g_2(x)$$

i rozwiązujemy układ

$$\begin{cases} D_x L = [0, 1] + \lambda_1[-1, -1] + \lambda_2[2x_1, 2x_2] = \mathbf{0}, \\ \lambda_1 g_1 = 0, \\ \lambda_2 g_2 = 0, \\ \lambda_1, \lambda_2 \geqslant 0. \end{cases}$$

1. Jeśli $\lambda_1 = \lambda_2 = 0$, to brak rozwiązań.

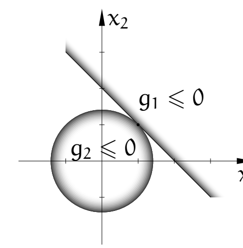
2. Jeśli $\lambda_1 = 0$, $\lambda_2 > 0$, to $2\lambda_2 x_1 = 0$, stąd $x_1 = 0$, $1 + 2\lambda_2 x_2 = 0$, stąd $x_2 = -1/(2\lambda_2)$. Ograniczenie g_2 jest aktywne, czyli $\big(1/(2\lambda_2)\big)^2 = 2$, stąd $\lambda_2 = 1/(2\sqrt2)$, $x_2 = -\sqrt2$. Jest $g_1(0, -\sqrt2) = 2 + \sqrt2 > 0$, zatem to ograniczenie jest niespełnione.

3. Jeśli $\lambda_1 > 0$, $\lambda_2 = 0$, to wychodzi sprzeczność, $\lambda_1 = 0$.

4. Jeśli $\lambda_1, \lambda_2 > 0$, to jest

$$2\lambda_2 x_1 - \lambda_1 = 0,$$
$$2\lambda_2 x_2 - \lambda_1 + 1 = 0,$$

ale z ograniczeń $x_1 + x_2 = 2$, $x_1^2 + x_2^2 = 2$ wynika $x_1 = x_2 = 1$, co prowadzi do sprzeczności, $2\lambda_1 - \lambda_1 = 0$, $2\lambda_2 - \lambda_1 + 1 = 0$.



Pozostają punkty nieregularne.
1. $g_1 = 0$, $g_2 < 0$ — brak,
2. $g_1 < 0$, $g_2 = 0$ daje $[2x_1, 2x_2] = \mathbf{0}$, stąd $x_1 = x_2 = 0$, ale wtedy $g_2 < 0$,
3. $g_1 = g_2 = 0$, wtedy $x_1 = x_2 = 1$ i gradienty $Dg_1 = [-1, -1]$ oraz $Dg_2 = [2x_1, 2x_2]$ są liniowo zależne. Zbiór punktów spełniających ograniczenia jest jednoelementowy (zobacz rysunek).

47. (NPP) Wykaż, że dla zadania

$$\begin{cases} f(x) \to \min \quad (f\colon \mathbb{R}^n \to \mathbb{R}), \\ g_i(x) \leqslant 0, \qquad i = 1, \ldots, m, \\ x \geqslant 0, \end{cases}$$

warunek I rzędu przyjmuje postać

$$\begin{cases} D_x L(x_0, \boldsymbol\lambda) \geqslant 0, \\ x_0^T D_x L(x_0, \boldsymbol\lambda) = 0, \\ \lambda_i g_i(x_0) = 0, \qquad i = 1, \ldots, m, \\ \lambda_i \geqslant 0, \qquad i = 1, \ldots, m, \end{cases} \tag{*}$$

gdzie

$$L(\boldsymbol{x},\boldsymbol{\lambda}) = f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i g_i(\boldsymbol{x}).$$

Rozw. Standardowy warunek I rzędu:

$$\begin{cases} D_x\hat{L}(\boldsymbol{x}_0,\boldsymbol{\lambda},\boldsymbol{\gamma}) = 0, & \\ \lambda_i g_i(\boldsymbol{x}_0) = 0 & i = 1,\ldots,m, \\ \gamma_i(-\boldsymbol{x}_0)_j = 0, & j = 1,\ldots,n, \\ \lambda_i, \gamma_j \geqslant 0, & i = 1,\ldots,m,\ j = 1,\ldots,n, \end{cases} \tag{**}$$

gdzie

$$\hat{L}(\boldsymbol{x},\boldsymbol{\lambda},\boldsymbol{\gamma}) = f(\boldsymbol{x}) + \sum_{i=1}^{m} \lambda_i g_i(\boldsymbol{x}) + \sum_{j=1}^{n} \gamma_j(-x_j).$$

Ma miejsce równość $D_x\hat{L}(\boldsymbol{x}_0,\boldsymbol{\lambda},\boldsymbol{\gamma}) = D_x L(\boldsymbol{x}_0,\boldsymbol{\lambda}) - \boldsymbol{\gamma}$.

(**) $\Rightarrow$ (*):
Z warunku $D_x\hat{L} = \boldsymbol{0}$ i $\boldsymbol{\gamma} \geqslant 0$ otrzymujemy $D_x L \geqslant 0$. Mnożymy $D_x\hat{L}$ z lewej strony przez $\boldsymbol{x}_0^T$:

$$0 = \boldsymbol{x}_0^T D_x\hat{L}(\boldsymbol{x}_0,\boldsymbol{\lambda},\boldsymbol{\gamma}) = \boldsymbol{x}_0^T D_x L(\boldsymbol{x}_0,\boldsymbol{\lambda}) - \underbrace{\boldsymbol{x}_0^T\boldsymbol{\gamma}}_{=0 \text{ z założenia}}.$$

(*) $\Rightarrow$ (**):
Niech $\boldsymbol{\gamma} = D_x L(\boldsymbol{x}_0,\boldsymbol{\lambda})$. Wtedy $D_x\hat{L}(\boldsymbol{x}_0,\boldsymbol{\lambda},\boldsymbol{\gamma}) = \boldsymbol{0}$. Z $D_x L(\boldsymbol{x}_0,\boldsymbol{\lambda}) \geqslant 0$ wynika $\boldsymbol{\gamma} \geqslant 0$.
Czy $\gamma_j x_j = 0$, tzn. $\big(D_x L(\boldsymbol{x}_0,\boldsymbol{\lambda})\big)_j(\boldsymbol{x}_0)_j = 0$?
To wynika z $\boldsymbol{x}_0^T D_x L(\boldsymbol{x}_0,\boldsymbol{\lambda}) = 0$ i z warunków $D_x L(\boldsymbol{x}_0,\boldsymbol{\lambda}) \geqslant 0$, $\boldsymbol{x}_0 \geqslant 0$.

48. Dla jakich wartości parametru α poniższy problem ma rozwiązanie?

$$\begin{cases} x_1 + x_2 + \alpha(x_1-1)^2 \to \min, \\ x_1 - x_2 \leqslant 0, \\ x_1 \geqslant 0. \end{cases}$$

Rozw. Postać ograniczeń umożliwia użycie sposobu z zadania 47. Określamy funkcję

$$L(\boldsymbol{x},\lambda) = x_1 + x_2 + \alpha(x_1-1)^2 + \lambda(x_1 - x_2)$$

i mamy rozwiązać układ równań i nierówności

$$\begin{cases} \frac{\partial}{\partial x_1}L(\boldsymbol{x},\lambda) \geqslant 0, \\ \frac{\partial}{\partial x_2}L(\boldsymbol{x},\lambda) = 0, \\ x_1\frac{\partial}{\partial x_1}L(\boldsymbol{x},\lambda) = 0, \\ \lambda(x_1 - x_2) = 0, \\ \lambda \geqslant 0. \end{cases} \quad \text{czyli} \quad \begin{cases} 1 + 2\alpha(x_1-1) + \lambda \geqslant 0, \\ 1 - \lambda = 0, \\ x_1(1 + 2\alpha(x_1-1) + \lambda) = 0, \\ \lambda(x_1 - x_2) = 0, \\ \lambda \geqslant 0. \end{cases}$$

Przypadek $\lambda = 0$ jest sprzeczny z $1 - \lambda = 0$.

Badamy $\lambda > 0$; musi być $\lambda = 1$ i ograniczenie $x_1 = x_2$ jest aktywne.
Z równania $x_1(1 + 2\alpha(x_1-1) + 1) = 0$ jest $x_1 = 0$ lub $x_1 = 1 - 1/\alpha$.

Jeśli $\alpha \leqslant 1$, to punkt podejrzany $x_1 = x_2 = 0$ (dla $\alpha < 1$ jest $1 - 1/\alpha < 0$, zatem nie może być $x_1 = 1 - 1/\alpha$). Funkcja f ma minimum, $f(0,0) = \alpha$.

Jeśli $\alpha > 1$, to punkty podejrzane $x_1 = x_2 = 0$ oraz $x_1 = x_2 = 1 - 1/\alpha$. Pierwszy z nich nie spełnia nierówności $\frac{\partial}{\partial x_1}L(\boldsymbol{x},\lambda) \geqslant 1$ z $\lambda = 1$. Stąd minimum jest w drugim punkcie, $f(1-1/\alpha, 1-1/\alpha) = 2 - 1/\alpha$.

49. Niech $\boldsymbol{b} \in \mathbb{R}^n$, $\boldsymbol{b} \neq \boldsymbol{0}$ oraz $p > 1$. Znajdź minimum funkcji $f(\boldsymbol{x}) = \boldsymbol{b}^T\boldsymbol{x}$ w zbiorze $\{\boldsymbol{x}\colon \|\boldsymbol{x}\|_p \leqslant 1\}$.

Rozw. Jest $\|\boldsymbol{x}\|_p = \big(\sum_{i=1}^n |x_i|^p\big)^{1/p}$;

$$D_x\big(\|\boldsymbol{x}\|_p - 1\big) = \|\boldsymbol{x}\|_p^{1-p}\big[|x_1|^{p-1}\operatorname{sgn} x_1, \ldots, |x_n|^{p-1}\operatorname{sgn} x_n\big].$$

Gradient jest różny od 0 dla każdego $\boldsymbol{x} \neq \boldsymbol{0}$.

Określamy $L(\boldsymbol{x},\lambda) = \boldsymbol{b}^T\boldsymbol{x} + \lambda\big(\|\boldsymbol{x}\|_p - 1\big)$. Warunki K-T:

$$\begin{aligned} &D_x L(\boldsymbol{x},\lambda) = \boldsymbol{0}, \quad \Rightarrow \quad \boldsymbol{b}^T + \lambda D_x\big(\|\boldsymbol{x}\|_p - 1\big) = \boldsymbol{0}, \\ &\lambda\big(\|\boldsymbol{x}\|_p - 1\big) = 0, \\ &\lambda \geqslant 0. \end{aligned}$$

Badamy $\lambda = 0$ — jest sprzeczność, bo $\boldsymbol{b} \neq \boldsymbol{0}$.

Dla $\lambda > 0$ jest $\|\boldsymbol{x}\|_p = 1$. Wtedy

$$b_i + \lambda\|\boldsymbol{x}\|^{1-p}|x_i|^{p-1}\operatorname{sgn} x_i = 0 \quad \Rightarrow \quad \operatorname{sgn} x_i = -\operatorname{sgn} b_i.$$

Możemy zatem napisać

$$|b_i| - \lambda\|\boldsymbol{x}\|_p^{1-p}|x_i|^{p-1} = 0, \quad \text{a ponieważ } \|\boldsymbol{x}\|_p = 1, \quad |x_i| = \left(\frac{|b_i|}{\lambda}\right)^{1/(p-1)}.$$

Dalej,

$$1 = \|\boldsymbol{x}\|_p^p = \sum_{i=1}^n |x_i|^p = \frac{1}{\lambda^{p/(p-1)}} \sum_{i=1}^n |b_i|^{p/(p-1)},$$

skąd

$$\lambda = \left(\sum_{i=1}^n |b_i|^{p/(p-1)}\right)^{(p-1)/p} = \|\boldsymbol{b}\|_q,$$

gdzie $q = \frac{p}{p-1}$ (czyli $\frac{1}{p} + \frac{1}{q} = 1$). Jest jeden punkt spełniający warunki K-T. Z wypukłości normy wynika, że minimum jest globalne.

50. Niech $\boldsymbol{x}^*$ będzie rozwiązaniem poprzedniego zadania. Udowodnij, że $\boldsymbol{x}^*$ jest też rozwiązaniem problemu

$$\begin{cases} \|\boldsymbol{x}\|_p \to \min, \\ \boldsymbol{b}^T\boldsymbol{x} \leqslant \boldsymbol{b}^T\boldsymbol{x}^*. \end{cases}$$

<u>Rozw.</u> Określamy funkcję $L(\boldsymbol{x}, \hat{\lambda}) = \|\boldsymbol{x}\|_p + \hat{\lambda}\boldsymbol{b}^T(\boldsymbol{x} - \boldsymbol{x}^*)$ i zapisujemy warunki

$$\begin{aligned} &D_x L(\boldsymbol{x}, \hat{\lambda}) = D_x\|\boldsymbol{x}\|_p + \hat{\lambda}\boldsymbol{b}^T = \boldsymbol{0}, \\ &\hat{\lambda}\boldsymbol{b}^T(\boldsymbol{x} - \boldsymbol{x}^*) = 0, \\ &\hat{\lambda} \geqslant 0. \end{aligned}$$

Jeśli $\hat{\lambda} = 0$, to $\boldsymbol{x} = 0$ (z warunku $D_x\|\boldsymbol{x}\|_p = \boldsymbol{0}$), ale wtedy nie jest spełniona nierówność $\boldsymbol{b}^T\boldsymbol{x} \leqslant \boldsymbol{b}^T\boldsymbol{x}^*$ (bo $\boldsymbol{b}^T\boldsymbol{x}^* < 0$).

Jeśli $\hat{\lambda} > 0$, to $\boldsymbol{b}^T(\boldsymbol{x} - \boldsymbol{x}^*) = 0$. Kula jednostkowa jest ściśle wypukła, zatem jedyny wektor $\boldsymbol{x}$ z tej kuli spełniający tę równość to $\boldsymbol{x}^*$. Wtedy $D_x L(\boldsymbol{x}^*, \hat{\lambda}) = \boldsymbol{0}$ dla $\hat{\lambda} = 1/\lambda$, λ z poprzedniego zadania. Dostateczność warunku wynika ze ściśłej wypukłości normy.

51. Rozwiąż problem

$$\begin{aligned} (x_1 - 1)^2 + x_2^2 &\to \min, \\ 2kx_1 - x_2^2 &\leqslant 0, \end{aligned}$$

z parametrem $k > 0$.

<u>Rozw.</u> Jest $D_x g = [2k, -2x_2] \neq \boldsymbol{0}$ dla każdego x_1, x_2. Określamy

$$L(\boldsymbol{x}, \lambda) = (x_1 - 1)^2 + x_2 + \lambda(2kx_1 - x_2^2).$$

Warunki K-T:

$$\begin{cases} 2(x_1 - 1) + 2k\lambda = 0, \\ 2x_2 - 2\lambda x_2 = 0, \\ \lambda(2kx_1 - x_2^2) = 0, \\ \lambda \geqslant 0. \end{cases}$$

Sprawdzamy $\lambda = 0$: $x_1 = 1$, $x_2 = 0$, $2k < 0$ — jest sprzeczność.

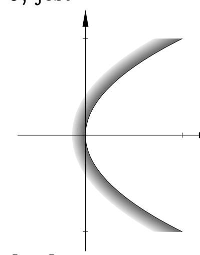
Sprawdzamy $\lambda > 0$: $x_1 = 1 - k\lambda$, $x_2 - \lambda x_2 = 0$, stąd $x_2 = 0$ lub $\lambda = 1$.

Sprawdzamy $\lambda = 1$: $x_1 = 1 - k$, $2k(1-k) = x_2^2$, stąd $x_2 = \pm\sqrt{2k(1-k)}$, musi być $k \in (0, 1]$.

Sprawdzamy $x_2 = 0$: $2kx_1 = 0$, stąd $x_1 = 0$, $\lambda = 1/k$.

$$D_x L(\boldsymbol{x}, \lambda) = [2(x_1 - 1) - 2k\lambda, 2x_2 - 2\lambda x_2],$$

$$D_{xx}^2 L(\boldsymbol{x}, \lambda) = \begin{bmatrix} 2 & 0 \\ 0 & 2(1-\lambda) \end{bmatrix}.$$



Dla $x = (0, 0)$, $k \geqslant 1$, $\lambda = 1/k$: jeśli $k > 1$, to $\lambda < 1$, $D_{xx}^2 L > 0$, jest rozwiązanie lokalne.

Dla $k = 1$ jest $\lambda = 1$, $D_{xx}^2 L = \left[\begin{smallmatrix} 2 & 0 \\ 0 & 0 \end{smallmatrix}\right]$. Bierzemy $C = \{\, \boldsymbol{d}\colon D_x g \cdot \boldsymbol{d} = 0 \,\} = \{\, \boldsymbol{d}\colon kd_1 = x_2 d_2 \,\}$. W punkcie $(0, 0)$ zbiór $C = \{\, \boldsymbol{d}\colon kd_1 = 0 \,\}$, jest $\boldsymbol{d}^T D_{xx}^2 L\boldsymbol{d} = 0$ dla każdego $\boldsymbol{d} \in C$, to nie daje rozstrzygnięcia.

Dla $\boldsymbol{x} = (0, 0)$, $k \leqslant 1$ warunek konieczny II rzędu nie jest spełniony.

Dla $\boldsymbol{x} = (1 - k, \pm\sqrt{2k(1-k)})$, $k \in (0, 1)$, $\lambda = 1$ jest $D_{xx}^2 L = \left[\begin{smallmatrix} 2 & 0 \\ 0 & 0 \end{smallmatrix}\right]$.

Należy zbadać $\boldsymbol{d}^T D_{xx}^2 L\boldsymbol{d} \overset{?}{>} 0$ dla $\boldsymbol{d} \in C$.

Przypadek $+\sqrt{2k(1-k)}$: $C = \{\, \boldsymbol{d} \neq \boldsymbol{0}\colon kd_1 = \sqrt{2k(1-k)}d_2 \,\}$, $\operatorname{sgn} d_1 = \operatorname{sgn} d_2$. Jest $\forall_{\boldsymbol{d} \in C}\ \boldsymbol{d}^T D_{xx}^2 L\boldsymbol{d} = 2d_1^2 > 0$.

Przypadek $-\sqrt{2k(1-k)}$: $\operatorname{sgn} d_1 \neq \operatorname{sgn} d_2$, ale też jest $\forall_{\boldsymbol{d} \in C}\ \boldsymbol{d}^T D_{xx}^2 L\boldsymbol{d} = 2d_1^2 > 0$.

Przypadek $\boldsymbol{x} = (0, 0)$, $k = 1$, $\lambda = 1$: $x_2^2 = 2x_1$. Minimalizujemy $(x_1 - 1)^2 + 2x_1 = x_1^2 + 1$, to ma minimum dla $x_1 = 0$.

52. (GP') Zbadaj wypukłość zbiorów
a) $\{\, (x, y, z)\colon x, y, z \geqslant 0, x + y + z = 1 \,\}$,
b) $\{\, (x, y, z)\colon |x| + |y| + z \leqslant 2 \,\}$,
c) $\{\, (x, y, z)\colon x^3y^3z^3 \geqslant 8, x, y, z \in [0, 2] \,\}$,
d) $\{\, (x, y)\colon x^3 - x + y^2 < 0, x \geqslant 0 \,\}$.

53. Wykaż, że jeśli funkcja $f\colon A \to \mathbb{R}$ określona na zbiorze wypukłym $A \subset \mathbb{R}^n$ jest wypukła, to dla każdego $x \in A$ subróżniczka $\partial f|_x$ jest zbiorem niepustym, wypukłym i domkniętym, przy czym jeśli $x \in \operatorname{int} A$, to subróżniczka jest zbiorem ograniczonym.

Rozw. Epigraf funkcji f jest zbiorem wypukłym, zatem dla $x \in A$ istnieje hiperpłaszczyzna (w $\mathbb{R}^{n+1}$) podpierająca epigraf w punkcie $\bigl(x, f(x)\bigr)$. Należą do niej punkty (x', y') spełniające równanie $y' - f(x) = \xi(x' - x)$; funkcjonał liniowy ξ jest subgradientem funkcji f w punkcie x. Zatem subróżniczka nie jest zbiorem pustym.

Weźmy $\xi_0, \xi_1 \in \partial f|_x$ oraz $t \in (0,1)$. Niech $\xi_t = (1-t)\xi_0 + t\xi_1$. Wtedy dla każdego $x' \in A$

$$\begin{aligned} f(x') - f(x) &= (1-t)\bigl(f(x') - f(x)\bigr) + t\bigl(f(x') - f(x)\bigr) \\ &\geqslant (1-t)\xi_0(x'-x) + t\xi_1(x'-x) = \xi_t(x'-x), \end{aligned}$$

a zatem $\xi_t \in \partial f|_x$, co kończy dowód wypukłości.

Dla dowolnego ciągu zbieżnego subgradientów w punkcie x, $(\xi_i)_{i\in\mathbb{N}}$, granica tego ciągu, ξ, jest subgradientem, co wynika z tego, że definiująca subgradient nierówność jest nieostra.

Niech $x \in \operatorname{int} A$. Ustalmy $\varepsilon > 0$, taki że kula o tym promieniu i środku w x jest zawarta w A. Przypuśćmy, że subróżniczka zawiera nieograniczony ciąg subgradientów $(\xi_i)_{i\in\mathbb{N}}$. Możemy wskazać w tej kuli ciąg punktów $(x_i)_{i\in\mathbb{N}}$, taki że ciąg $\bigl(\xi_i(x_i - x)\bigr)_{i\in\mathbb{N}}$ jest nieograniczony z góry. Wobec nierówności $f(x_i) \geqslant f(x) + \xi_i(x_i - x)$ ciąg $\bigl(f(x_i)\bigr)_{i\in\mathbb{N}}$ jest też nieograniczony z góry. Wobec dowolności wyboru ε funkcja f jest nieograniczona w dowolnie małym otoczeniu x, czyli jest nieciągła, a więc nie może też być wypukła.

54. (GP') Znajdź subróżniczkę funkcji $f\colon \mathbb{R}^2 \to \mathbb{R}$:
a) $f(x,y) = |x-y|$,
b) $f(x,y) = |x| + |y|$,
c) $f(x,y) = |x-y| + |x+y|$,
d) $f(x,y) = \max\{x, y, x+y\}$,
e) $f(x,y) = |2x - 3y + 1|$,
f) $f(x,y) = |2x - y| + |3x + y - 5| + 1$.

55. Niech $f\colon W \subset R^2 \to \mathbb{R}$ będzie dodatnią funkcją wypukłą w otwartym zbiorze wypukłym W i niech $g(x) = f^2(x)$.

a) Wykaż, że funkcja g jest wypukła,

b) Wykaż, że $\partial g(x) = 2f(x)\partial f(x)$.

Rozw. b) Niech $\xi \in \partial f(x)$. Wtedy dla $\mathbf{d}$ takiego że $x + \mathbf{d} \in W$ jest

$$\begin{aligned} f(x+\mathbf{d}) &\geqslant f(x) + \xi^T\mathbf{d}, \\ f^2(x+\mathbf{d}) &\geqslant f^2(x) + 2f(x)\xi^T\mathbf{d} + (\xi^T\mathbf{d})^2 \geqslant f^2(x) + 2f(x)\xi^T\mathbf{d}, \end{aligned}$$

czyli $g(x+\mathbf{d}) \geqslant g(x) + 2f(x)\xi^T\mathbf{d}$, zatem $2f(x)\partial f(x \in \partial g(x)$, a stąd $2f(x)\partial f(x) \subset \partial g(x)$.

Aby dowieść zawierania w drugą stronę, użyjemy twierdzenia opisującego związki pochodnych kierunkowych funkcji z ich subróżniczkami (ustalamy punkt x i pomijamy go w notacji niżej):

$$\frac{\partial f}{\partial \mathbf{d}} = \max_{\xi\in\partial f} \xi^T\mathbf{d}, \qquad \frac{\partial g}{\partial \mathbf{d}} = \max_{\eta\in\partial g} \eta^T\mathbf{d}.$$

Zatem, dla dowolnego wektora $\mathbf{d}$ istnieją wektory $\xi_0 \in \partial f$ oraz $\eta_1 \in \partial g$, takie że

$$\frac{\partial f}{\partial \mathbf{d}} = \xi_0^T\mathbf{d}, \qquad \frac{\partial g}{\partial \mathbf{d}} = \eta_1^T\mathbf{d}.$$

Oznaczmy $\xi_1 = \frac{1}{2f(x)}\eta_1$; przypuśćmy, że $\xi_1 \notin \partial f$, co wyraża się w ten sposób, że dla pewnego wektora $\mathbf{d}$ jest

$$f(x+\mathbf{d}) < f(x + \xi_1^T\mathbf{d}.$$

Mamy przy tym, dla odpowiednio dobranego wektora $\xi_0 \in \partial f$, równości

$$\begin{aligned} f(x+\mathbf{d}) &= f(x) + \xi_0^T\mathbf{d} + r_0, \\ f^2(x+\mathbf{d}) &= f^2(x) + 2f(x)\xi_0^T\mathbf{d} + r_1, \\ g(x+\mathbf{d}) &= g(x + 2f(x)\xi_1^T\mathbf{d} + r_2, \end{aligned}$$

gdzie $r_0 = o(\|\mathbf{d}\|)$, $r_1 = o(\|\mathbf{d}\|)$, $r_2 = o(\|\mathbf{d}\|)$. Odejmując stronami dwie ostatnie równości, dostajemy

$$2f(x)(\xi_1 - \xi_0)^T\mathbf{d} = (r_1 - r_2) = o(\|\mathbf{d}\|),$$

ale jeśli wyrażenie po lewej stronie nie jest zerem, to jest rzędu $\|\mathbf{d}\|$, czyli mamy sprzeczność.

56. Niech funkcje $f_i\colon A \to \mathbb{R}$ określone na zbiorze wypukłym A będą wypukłe i niech szereg $\sum_{i=1}^{\infty} f_i(x)$ będzie zbieżny w każdym punkcie $x \in A$. Wykaż, że funkcja f będąca sumą tego szeregu jest wypukła oraz że dla każdego $x \in A$

$$\partial f|_x = \operatorname{cl}\Bigl\{\, \xi = \sum_{i=1}^{\infty} \xi_i \colon \xi_i \in \partial f_i|_x \,\Bigr\}.$$

57. Niech $A = \{\boldsymbol{x} \in \mathbb{R}^n \colon \|\boldsymbol{x}\|_2 \leqslant 1\}$ (np. $n = 2$) i niech $(\boldsymbol{x}_i)_{i\in\mathbb{N}}$ oznacza dowolny ciąg punktów w A. Niech $f_i(\boldsymbol{x}) = 2^{-i}\|\boldsymbol{x} - \boldsymbol{x}_i\|_2$. Wykaż, że funkcja $f = \sum_{i=1}^{\infty} f_i$ jest określona w A i jeśli ciąg $(\boldsymbol{x}_i)_{i\in\mathbb{N}}$ jest gęsty w A, to funkcja f w żadnym punkcie zbioru A nie ma gradientu, ale ma w każdym punkcie niepustą subróżniczkę.

Czy taka funkcja (dla ciągu $(\boldsymbol{x}_i)_{i\in\mathbb{N}}$ gęstego w A) ma pochodną kierunkową w jakimś punkcie i w jakimś kierunku?

Wskaż minimum każdej takiej funkcji f. Czy ono jest ścisłe?

58. (GP) Sprawdź, czy punkty ze zbioru P są rozwiązaniami zadania minimalizacji funkcji $f\colon \mathbb{R}^2 \to \mathbb{R}$:
a) $f(x,y) = \max\{2x - 3y - 1, -x + y, x^2 + 2y^2 - 36\}$, $P = \{(4,3)\}$,
b) $f(x,y) = \max\{|4x - y + 1|, 2x^2 - 8\}$, $P = \{(2,9), (0,1), (3,3)\}$,
c) $f(x,y) = \max\{x^2 + y^2, 13x + 7y + 20, 3x + 3y + 20\}$,
$P = \{(-10,10), (-2,2), (-2,5)\}$.

59. (GP) Znajdź globalne minimum funkcji $f\colon \mathbb{R}^2 \to \mathbb{R}$:
a) $f(x,y) = |x - y + 1| + |2x + y - 4|$,
b) $f(x,y) = \max\{(x-1)^2, |y|\}$,
c) $f(x,y) = \max\{x^2 + y^2, x + 1\}$,
d) $f(x,y) = \max\{e^x, e^{-y}, y - x\}$,
e) $f(x,y) = \max\{(x - 1^2 + y^2 + 4, -2x + 4y + 2)\}$.

60. (GP) Znajdź minimum globalne funkcji f na zbiorze $A \subset \mathbb{R}^2$:
a) $f(x,y) = \max\{x,y\}$, $A = \{(x,y)\colon x^2 + y^2 \leqslant 1\}$,
b) $f(x,y) = |x - y + 2|$, $A = \{(x,y)\colon |x| + |y| \leqslant 2\}$,
c) $f(x,y) = \max\{x - y, x + y\}$,
$A = \{(x,y)\colon (x-1)^2 + y^2 \leqslant 4, (x+1)^2 + y^2 \leqslant 4\}$,
d) $f(x,y) = |2x + y - 2| + 1$,
$A = \{(x,y)\colon (x-y)^2 - 2 \leqslant x + y, (x+y)^2 - 2 \leqslant y - x\}$.

61. (NPP) Na rynku są dwa produkty, A i B. Agent ma początkowo po 10 jednostek tych produktów, jego funkcja użyteczności to $u(x_A, x_B) = 2\ln x_A + 3\ln x_B$. Niech p_A i p_B będą cenami A i B. Agent chce mieć takie ilości produktów A i B, które maksymalizują funkcję użyteczności przy ograniczeniu budżetowym

$$p_A x_A + p_B x_B \leqslant 10p_A + 10p_B.$$

Rozw. Funkcja $-u(x_A, x_B) = -2\ln x_A - 3\ln x_B$ jest wypukła, jej minimalizacja to maksymalizacja funkcji u. Funkcja Lagrange'a

$$L(\boldsymbol{x}, \lambda) = -2\ln x_A - 3\ln x_B + \lambda\bigl(p_A(x_A - 10) + p_B(x_B - 10)\bigr)$$

ma gradient

$$D_x L = \Bigl[-\frac{2}{x_A} + \lambda p_A, -\frac{3}{x_B} + \lambda p_B\Bigr].$$

Rozwiązujemy układ równań i nierówności

$$\begin{cases} D_x L = 0, \\ \lambda\bigl(p_A(x_A - 10) + p_B(x_B - 10)\bigr) = 0, \\ \lambda \geqslant 0. \end{cases}$$

Jeśli $\lambda = 0$, to $\frac{-2}{x_A} = 0$, to jest wykluczone.

Jeśli $\lambda > 0$, to

$$\lambda = \frac{2}{x_A p_A} = \frac{3}{x_B p_B} \quad\Rightarrow\quad 3x_A p_A = 2x_B p_B.$$

Razem z równaniem $p_A(x_A - 10) + p_B(x_B - 10) = 0$ daje to

$$\frac{2}{3} p_B x_B + p_B x_B = 10(p_A + p_B),$$

$$\text{stąd} \qquad x_B = \frac{6(p_A + p_B)}{p_B}, \qquad x_A = \frac{4(p_A + p_B)}{p_A}.$$

To jest minimum, bo funkcja $-u$ jest wypukła.

61! (NPP) Przypuśćmy, że na rynku działa drugi agent, którego funkcja użyteczności to $\hat{u}(\hat{x}_A, \hat{x}_B) = \ln \hat{x}_A + \ln \hat{x}_B$. Drugi agent ma początkowo odpowiednio 10 i 5 jednostek produktów. Jeśli w gospodarce są tylko ci dwaj agenci, to jak powinny się kształtować ceny p_A, p_B, aby obaj osiągnęli stan optymalny?

Rozw. Dla drugiego agenta rozwiązujemy podobnie,

$$\begin{aligned} D_x \hat{L}(\hat{\boldsymbol{x}}, \hat{\lambda}) = \Bigl[\frac{-1}{\hat{x}_A} + \hat{\lambda} p_A, \frac{-1}{\hat{x}_B} + \hat{\lambda} p_B\Bigr] &= 0, \\ \hat{x}_A p_A = \hat{x}_B p_B&, \\ p_A(\hat{x}_A - 10) + p_B(\hat{x}_B - 5) &= 0, \\ p_B \hat{x}_B + p_A \hat{x}_A = 10p_A + 5p_B&, \end{aligned}$$

otrzymując

$$\hat{x}_B = \frac{10p_A + 5p_B}{2p_B}, \qquad \hat{x}_A = \frac{10p_A + 5p_B}{2p_A}.$$

Zadowolenie obydwu agentów oznacza, że jest równowaga popytu i podaży przy optymalnych zasobach (*supply-demand equilibrium for optimal reserve*):

$$\begin{cases} x_A + \hat{x}_A = 10 + 10, \\ x_B + \hat{x}_B = 10 + 5, \end{cases}$$

Stąd

$$\frac{4(p_A + p_B)}{p_A} + \frac{10p_A + 5p_B}{2p_A} = 20, \qquad \frac{6(p_A + p_B)}{p_B} + \frac{10p_A + 5p_B}{2p_B} = 15,$$

i dalej

$$\frac{18p_A + 13p_B}{2p_A} = 20, \qquad \frac{22p_A + 17p_B}{2p_B} = 15.$$

Z obu równań wynika $22p_A - 13p_B = 0$, zatem $p_A = \frac{13}{22}p_B$. Nie są ważne bezwzględne ceny (które zależą od waluty), tylko ich proporcja.

62. (NPP) Zbiory $A \subset \mathbb{R}^n$ i $C \in \mathbb{R}^m$ są wypukłe. Funkcja $f\colon A \times C \to \mathbb{R}$ jest quasi-wypukła. Udowodnij, że funkcja $h(\boldsymbol{x}) = \inf_{y \in C} f(\boldsymbol{x}, \boldsymbol{y})$ jest quasi-wypukła.

Rozw. Niech $\boldsymbol{y}_\varepsilon^x$ oznacza dowolny punkt taki, że $f(\boldsymbol{x}, \boldsymbol{y}_\varepsilon^x) - \varepsilon \leqslant h(\boldsymbol{x})$, $\boldsymbol{x}_1, \boldsymbol{x}_2 \in A$, $\lambda \in (0,1)$, $\boldsymbol{x} = \lambda\boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2$. Jest

$$\begin{aligned}\max\{h(\boldsymbol{x}_1), h(\boldsymbol{x}_2)\} &\geqslant \max\{f(\boldsymbol{x}_1, \boldsymbol{y}_\varepsilon^{x_1}), f(\boldsymbol{x}_2, \boldsymbol{y}_\varepsilon^{x_2})\} - 2\varepsilon \\ &\geqslant f\big(\lambda\boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2, \lambda\boldsymbol{y}_\varepsilon^{x_1} + (1-\lambda)\lambda\boldsymbol{y}_\varepsilon^{x_2}\big) - 2\varepsilon.\end{aligned}$$

Dla każdego $\boldsymbol{y}_1, \boldsymbol{y}_2 \in C$ jest $h(\boldsymbol{x}) \leqslant f(\boldsymbol{x}, \lambda\boldsymbol{y}_1 + (1-\lambda)\boldsymbol{y}_2)$, zatem

$$f\big(\lambda\boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2, \lambda\boldsymbol{y}_\varepsilon^{x_1} + (1-\lambda)\lambda\boldsymbol{y}_\varepsilon^{x_2}\big) - 2\varepsilon \geqslant h\big(\lambda\boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2\big) - 2\varepsilon.$$

Ponieważ ε może być dowolny dodatni, wynika stąd teza.

63. (NPP) Zbiór $S \subset \mathbb{R}^n$ jest wypukły, funkcja $f\colon S \to \mathbb{R}$ jest klasy C^1. Wykaż, że funkcja f jest quasi-wypukła wtedy i tylko wtedy, gdy

$$\forall_{x,y \in S} f(\boldsymbol{y}) \leqslant f(\boldsymbol{x}) \Leftrightarrow Df(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x}) \leqslant 0.$$

Rozw. $\Rightarrow$: dla $\lambda \in (0,1)$

$$\frac{f\big(\boldsymbol{x} + \lambda(\boldsymbol{y} - \boldsymbol{x})\big) - f(\boldsymbol{x})}{\lambda} \leqslant 0,$$

bo $f\big(\boldsymbol{x} + \lambda(\boldsymbol{y} - \boldsymbol{x})\big) = f\big(\lambda\boldsymbol{y} + (1-\lambda)\boldsymbol{x}\big) \leqslant \max\{f(\boldsymbol{y}), f(\boldsymbol{x})\} = f(\boldsymbol{x})$. Zatem

$$Df(\boldsymbol{x})(\boldsymbol{y} - \boldsymbol{x}) = \lim_{\lambda \searrow 0} \frac{f\big(\boldsymbol{x} + \lambda(\boldsymbol{y} - \boldsymbol{x})\big) - f(\boldsymbol{x})}{\lambda} \leqslant 0.$$

$\Leftarrow$: Niech $\boldsymbol{x}_1, \boldsymbol{x}_2 \in S$, $\lambda \in (0,1)$, $\tilde{\boldsymbol{x}} = \lambda\boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2$. Załóżmy, że $f(\tilde{\boldsymbol{x}}) \geqslant \max\{f(\boldsymbol{x}_1), f(\boldsymbol{x}_2)\}$. Stąd

$$\left.\begin{aligned} Df(\tilde{\boldsymbol{x}})(\boldsymbol{x}_1 - \tilde{\boldsymbol{x}}) \leqslant 0, \\ Df(\tilde{\boldsymbol{x}})(\boldsymbol{x}_2 - \tilde{\boldsymbol{x}}) \leqslant 0, \end{aligned}\right\} \Rightarrow Df(\tilde{\boldsymbol{x}})(\boldsymbol{x}_1 - \tilde{\boldsymbol{x}}) = 0 = Df(\tilde{\boldsymbol{x}})(\boldsymbol{x}_2 - \tilde{\boldsymbol{x}}).$$

Czyli albo $f(\tilde{\boldsymbol{x}}) \leqslant \max\{f(\boldsymbol{x}_1), f(\boldsymbol{x}_2)\}$, albo $Df(\tilde{\boldsymbol{x}})(\boldsymbol{x}_1 - \tilde{\boldsymbol{x}}) = 0 = Df(\tilde{\boldsymbol{x}})(\boldsymbol{x}_2 - \tilde{\boldsymbol{x}})$ dla $\lambda \in (0,1)$.

Załóżmy, że istnieje $\lambda^* \in (0,1)$, taka że $f(\tilde{\boldsymbol{x}}^*) > \max\{f(\boldsymbol{x}_1, \boldsymbol{x}_2)\} =: \alpha$. Niech $g(\lambda) = f\big(\boldsymbol{x}_2 + \lambda(\boldsymbol{x}_1 - \boldsymbol{x}_2)\big)$. Wtedy $g(0) = f(\boldsymbol{x}_2)$, $g(\lambda^*) = f(\tilde{\boldsymbol{x}}^*)$.

Rozważmy podzbiór odcinka $\overline{\boldsymbol{x}_1\boldsymbol{x}_2}$, na którym $g(\lambda) \geqslant \alpha$. Ten podzbiór zawiera odcinek $\overline{\boldsymbol{x}^*, \tilde{\boldsymbol{x}}^*}$; ponieważ $g(\lambda^*) = f(\tilde{\boldsymbol{x}}^*) > \alpha$, a funkcja f jest klasy C^1, taki odcinek istnieje. Dla $\boldsymbol{x} \in \overline{\boldsymbol{x}^*, \tilde{\boldsymbol{x}}^*}$ jest $Df(\tilde{\boldsymbol{x}})(\boldsymbol{x}_1 - \boldsymbol{x}) = 0 = Df(\tilde{\boldsymbol{x}})(\boldsymbol{x}_2 - \boldsymbol{x})$, czyli $Df(\boldsymbol{x}) \perp \overline{\boldsymbol{x}_1, \boldsymbol{x}_2}$. Z twierdzenia o wartości średniej

$$f(\boldsymbol{x}^*) - f(\tilde{\boldsymbol{x}}^*) = Df(\boldsymbol{z})(\boldsymbol{x}^* - \tilde{\boldsymbol{x}}^*)$$

dla pewnego $\boldsymbol{z} \in \overline{\boldsymbol{x}^*, \tilde{\boldsymbol{x}}^*}$. Czyli wszędzie tam, gdzie $g(\lambda) \geqslant \alpha$, jest $g(\lambda) = g(\lambda^*) > \alpha$, co przeczy ciągłości funkcji $g(\lambda)$.

64. (NPP) Funkcja $f\colon W \to \mathbb{R}$ jest określona na zbiorze wypukłym $W \subset \mathbb{R}^n$. Udowodnij, że
 1) f jest quasi-liniowa $\Rightarrow$ zbiór $\{\boldsymbol{x} \in W\colon f(\boldsymbol{x}) = \alpha\}$ jest wypukły dla każdego $\alpha \in f(W)$.
 2) $W \subset \mathbb{R}$, f jest quasi-liniowa $\Leftrightarrow$ funkcja f jest monotoniczna.
 3) $f\colon \mathbb{R}^n \to \mathbb{R}$ jest ciągła, quasi-liniowa $\Leftrightarrow$ $f(\boldsymbol{x}) = g(\boldsymbol{a}^T\boldsymbol{x})$ dla pewnej funkcji $g\colon \mathbb{R} \to \mathbb{R}$ ciągłej i monotonicznej oraz pewnego wektora $\boldsymbol{a} \in \mathbb{R}^n$.

Rozw. 1) Niech $A = \{\boldsymbol{x} \in W\colon f(\boldsymbol{x}) = \alpha\}$, $\boldsymbol{x}, \boldsymbol{y} \in A$. Zatem $f(\boldsymbol{x}) = f(\boldsymbol{y}) = \alpha$, dla $\lambda \in (0,1)$ mamy

$$\alpha = \min\{f(\boldsymbol{x}), f(\boldsymbol{y})\} \leqslant f\big(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\big) \leqslant \max\{f(\boldsymbol{x}), f(\boldsymbol{y})\} = \alpha,$$

czyli $f\big(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\big) = \alpha$, zatem $\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y} \in A$.

2) Zbiór wypukły $W \subset \mathbb{R}$ jest przedziałem ograniczonym lub nieograniczonym. $\Rightarrow$

$$\min\{f(x), f(y)\} \leqslant f\big(\lambda x + (1-\lambda)y\big) \leqslant \max\{f(x), f(y)\}.$$

Niech $x_1, x_2, x_3 \in W$, $x_1 < x_2 < x_3$. Wtedy istnieje $\lambda \in (0,1)$, taka że $x_2 = \lambda x_1 + (1-\lambda)x_3$. Jeśli $f(x_1) = f(x_3) = \alpha$, to na podstawie punktu a) jest $f(x_2) = \alpha$. Jeśli $f(x_1) < f(x_3)$, to z quasi-liniowości jest

$$\begin{aligned}\min\{f(x_1), f(x_3)\} = f(x_1) \leqslant f(x_2) = f\bigl(\lambda x_1 + (1-\lambda)x_3\bigr) \\ \leqslant \max\{f(x_1), f(x_3)\} = f(x_3),\end{aligned}$$

zatem f jest niemalejąca. W taki sam sposób z $f(x_1) > f(x_3)$ wynika, że f jest nierosnąca.

$\Leftarrow$ oczywiste.

3) $\Leftarrow$ Niech funkcja g będzie niemalejąca, $\boldsymbol{a} \in \mathbb{R}^n$ oraz $f(\boldsymbol{x}) = g(\boldsymbol{a}^T\boldsymbol{x})$. Niech $\boldsymbol{x}, \boldsymbol{y} \in W$, $\boldsymbol{a}^T\boldsymbol{x} \leqslant \boldsymbol{a}^T\boldsymbol{y}$. Wtedy dla $\lambda \in (0,1)$

$$\begin{aligned}\min\{f(\boldsymbol{x}),f(\boldsymbol{y})\} = f(\boldsymbol{x}) = g(\boldsymbol{a}^T\boldsymbol{x}) \leqslant f\bigl(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\bigr) = \\ g\bigl(\lambda\boldsymbol{a}^T\boldsymbol{x} + (1-\lambda)\boldsymbol{a}^T\boldsymbol{y}\bigr) \leqslant g(\boldsymbol{a}^T\boldsymbol{y}) = f(\boldsymbol{y}) = \max\{f(\boldsymbol{x}), f(\boldsymbol{y})\}.\end{aligned}$$

Podobne rachunki są dla $\boldsymbol{a}^T\boldsymbol{x} \geqslant \boldsymbol{y}$ lub dla funkcji g nierosnącej.

$\Rightarrow$ a) Niech $A_1 \subset A_2 \subset \cdots$ będzie wstępującą rodziną zbiorów wypukłych. Wtedy $\bigcup_n A_n$ jest zbiorem wypukłym.

b) Niech

$$B_\alpha = \{\boldsymbol{x}\colon f(\boldsymbol{x}) < \alpha\} = \bigcup_n \{\boldsymbol{x}\colon f(\boldsymbol{x}) \leqslant \alpha - 1/n\},$$

$$C_\alpha = \{\boldsymbol{x}\colon f(\boldsymbol{x}) > \alpha\} = \bigcup_n \{\boldsymbol{x}\colon f(\boldsymbol{x}) \geqslant \alpha + 1/n\}.$$

Wypukłość zbiorów B_α i C_α wynika odpowiednio z quasi-wypukłości i quasi-wklęsłości funkcji f. Z twierdzenia o oddzielaniu istnieje niezerowy wektor $\boldsymbol{c} \in \mathbb{R}^n$, taki że $\forall_{\boldsymbol{x}\in B_\alpha, \boldsymbol{y}\in C_\alpha}\ \boldsymbol{c}^T\boldsymbol{x} \leqslant \boldsymbol{c}^T\boldsymbol{y}$.

c) Z ciągłości funkcji f zbiór $A_\alpha = \{\boldsymbol{x}\colon f(\boldsymbol{x}) = \alpha\}$, jeśli jest niepusty, to jest domknięty, wypukły i nieograniczony, ponieważ zawiera hiperpłaszczyznę oddzielającą B_α i C_α. Jeśli więc zbiory B_α i C_α są niepuste, to nie mogą istnieć dwa niezależne liniowo wektory, $\boldsymbol{c}_1$ i $\boldsymbol{c}_2$, będące wektorami normalnymi hiperpłaszczyzn oddzielających te zbiory. Gdyby istniały, to zbiór $A_\alpha = \{\boldsymbol{x}\colon f(\boldsymbol{x}) = \alpha\}$ byłby niewypukły. Stąd wynika, że zbiór A_α, jeśli jest niepusty, to albo jest hiperpłaszczyzną oddzielającą zbiory B_α i C_α, albo jest zbiorem wszystkich punktów między dwiema takimi hiperpłaszczyznami.

d) Wszystkie hiperpłaszczyzny zawarte we wszystkich zbiorach A_α są równoległe, ponieważ w przeciwnym razie zbiory te nie byłyby rozłączne. Niech H oznacza równoległą do nich hiperpłaszczyznę przechodzącą przez $\mathbf{0}$. Jeśli $\boldsymbol{x} \in A_\alpha$ (czyli $A_\alpha \neq \emptyset$), to $\boldsymbol{x} + H \subset A_\alpha$.

e) Niech $\boldsymbol{a} \in \mathbb{R}^n$ będzie wektorem jednostkowym prostopadłym do H. Wtedy $\forall_{\boldsymbol{x}\in\mathbb{R}^n}\ \exists_{t\in\mathbb{R}}\ \boldsymbol{x} \in t\boldsymbol{a} + H$, bo $\mathbb{R}^n = H \oplus \mathrm{lin}\{\boldsymbol{a}\}$. Ponadto $t = \boldsymbol{a}^T\boldsymbol{x}$, bo $\boldsymbol{x} = t\boldsymbol{a} + \boldsymbol{h}_1$, gdzie $\boldsymbol{h}_1 \perp \boldsymbol{a}$, zatem $\boldsymbol{a}^T\boldsymbol{x} = t\boldsymbol{a}^T\boldsymbol{a} + \boldsymbol{a}^T\boldsymbol{h}_1 = t$.

f) Niech $g(t) = f(t\boldsymbol{a})$. Wtedy funkcja g jest quasi-liniowa, a zatem monotoniczna i ciągła. Ponadto $f(\boldsymbol{x}) = f\bigl((\boldsymbol{a}^T\boldsymbol{x})\boldsymbol{a}\bigr) = g(\boldsymbol{a}^T\boldsymbol{x})$, ponieważ $\boldsymbol{x}$ należy do pewnego zbioru A_α, czyli jest postaci $t\boldsymbol{a} + \boldsymbol{h}_1$, i $f(\boldsymbol{x}) = f(t\boldsymbol{a})$.

65. (NPP) a) Czy suma funkcji pseudowypukłych musi być pseudowypukła?

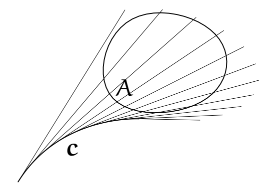
b) Czy suma funkcji pseudowypukłej i wypukłej musi być pseudowypukła?

c) Czy jeśli f i $g\colon \mathbb{R} \to \mathbb{R}$ są pseudowypukłe i $f'(x_0) = g'(x_0) = 0$ dla pewnego x_0, to $f + g$ musi być pseudowypukła?

d) Czy suma funkcji quasi-wypukłych musi być pseudowypukła?

e) Czy suma funkcji quasi-liniowej i wypukłej musi być quasi-wypukła?

66. Niech funkcja $g\colon [a,b] \to \mathbb{R}$ będzie monotoniczna i niech $\boldsymbol{c}\colon [a,b] \to \mathbb{R}^2$ będzie regularną gładką parametryzacją krzywej bez punktów przegięć. Niech $A \subset \mathbb{R}^2$ oznacza zbiór wypukły pokryty rozłącznymi półprostymi stycznymi do krzywej $\boldsymbol{c}$ (krzywa jest rozłączna z A). Wykaż, że funkcja $f\colon A \to \mathbb{R}$, która przyjmuje we wszystkich punktach prostej stycznej do $\boldsymbol{c}$ w punkcie $\boldsymbol{c}(t)$ wartość $\boldsymbol{g}(t)$, jest quasi-liniowa.



Jak to można uogólnić? a) na $\mathbb{R}^2$, b) na $\mathbb{R}^n$, $n \geqslant 2$.

67. (NPP) Niech zbiór $X \subset \mathbb{R}^n$ będzie niepusty otwarty, $f\colon X \to \mathbb{R}$, $g_i\colon X \to \mathbb{R}$ dla $i = 1, \ldots, m$, $h_j\colon X \to \mathbb{R}$ dla $j = 1, \ldots, l$. Jest zadanie minimalizacji

$$\begin{cases} f(\boldsymbol{x}) \to \min, \\ g_i(\boldsymbol{x}) \leqslant 0, \quad i = 1, \ldots, m, \\ h_j(\boldsymbol{x}) = 0, \quad j = 1, \ldots, l, \\ \boldsymbol{x} \in X. \end{cases}$$

Niech $\bar{\boldsymbol{x}}$ będzie punktem, w którym są spełnione warunki K-T:

$$Df(\bar{\boldsymbol{x}}) + \sum_{i\in I(\bar{\boldsymbol{x}})} \mu_i Dg_i(\bar{\boldsymbol{x}}) + \sum_{j=1}^{l} \lambda_j Dh_j(\bar{\boldsymbol{x}}) = \mathbf{0},$$

$$\mu_i g_i(\bar{\boldsymbol{x}}) = 0, \quad i = 1, \ldots, m,$$

$$\mu_i \geqslant 0, \quad i = 1, \ldots, m.$$

a) Udowodnij, że jeśli funkcja f jest pseudowypukła w $\bar{x}$, a funkcja Φ określona wzorem

$$\Phi(x) = \sum_{i=1}^{m} \mu_i g_i(\bar{x}) + \sum_{j=1}^{l} \lambda_j h_j(\bar{x})$$

jest quasi-wypukła w $\bar{x}$, to $\bar{x}$ jest rozwiązaniem globalnym zadania.

b) Udowodnij, że jeśli $x \to L(x, \mu, \lambda)$ jest pseudowypukła w $\bar{x}$, to $\bar{x}$ jest rozwiązaniem globalnym.

c) Wykaż, że założenia a) nie implikują założeń b) i odwrotnie.

Rozw. a) Niech $x \in X$, quasi-wypukłość funkcji Φ w $\bar{x}$ oznacza $D_x\Phi(\bar{x})(x - \bar{x}) \leqslant 0$, stąd i z warunku K-T wynika, że $Df(\bar{x})(x - \bar{x}) \geqslant 0$. Zatem, z pseudowypukłości Φ jest $f(x) \geqslant f(\bar{x})$, czyli w $\bar{x}$ minimum jest globalne.

b) Niech $x \in X$. Z pseudowypukłości funkcji L

$$D_x L(\bar{x}, \mu, \lambda)(x - \bar{x}) \geqslant 0 \quad \Rightarrow \quad L(x, \mu, \lambda) \geqslant L(\bar{x}, \mu, \lambda) = f(\bar{x}).$$

Ale

$$L(x, \mu, \lambda) = f(x) + \sum_{i=1}^{m} \mu_i \underbrace{g_i(\bar{x})}_{\leqslant 0} + \sum_{j=1}^{l} \lambda_j \underbrace{h_j(\bar{x})}_{=0} \leqslant f(x).$$

c) $\Rightarrow$:

$$\begin{cases} -x \to \min, \\ -\frac{1}{x} + 1 \leqslant 0, \\ x > 0. \end{cases}$$

Funkcja $-x$ jest pseudowypukła, $1 - 1/x$ jest quasiwypukła, bo monotoniczna.

$$L(x, \mu) = -x + \mu\left(1 - \frac{1}{x}\right),$$

$$D_x L(x, \mu) = -1 + \mu\frac{1}{x^2} = 0 \quad \Rightarrow \quad \bar{x} = 1, \mu = 1 \ (\mu = 0 \text{ daje sprzeczność.})$$

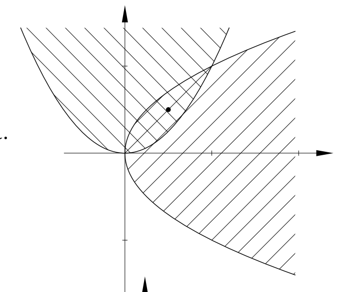
Funkcja $L(x) = -x + (1 - 1/x)$ nie jest pseudowypukła w $x = 1$, bo $DL(1) = 0$, czyli dla każdego $x > 0$ powinno być $L(x) > L(a) = -1$, tymczasem dla $x \to 0$ jest $L(x) \to -\infty$.

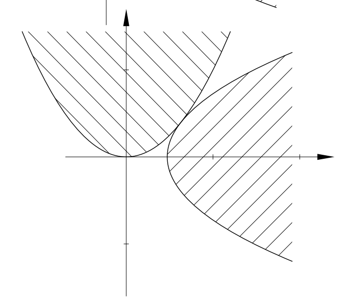
$\Leftarrow$: Z pseudowypukłości L nie można wnioskować o składnikach sumy, czyli funkcjach f i Φ.

68. (NPP) Sprawdź, dla jakich wartości parametru $a \in \mathbb{R}$ poniższe ograniczenia spełniają warunek Slatera:

$$\begin{cases} x^2 - y \leqslant 0, \\ y^2 - x - a \leqslant 0. \end{cases}$$

Rozw. Obie funkcje są pseudowypukłe jako sumy różniczkowalnych funkcji wypukłych.





1. Jeśli $a = 0$, to mamy nierówności $x^2 - y < 0$, $y^2 - x < 0$, punkt $(1/2, 1/2)$ spełnia warunki.
2. Jeśli $a \neq 0$, to ma być $x^2 < y$, $y^2 < x + a$. Aby znaleźć wartość graniczną dla a zamieniamy nierówności na równania i dobieramy a tak, aby zbiór rozwiązań był jednopunktowy.

Czyli $x^2 = y$, $y^2 = x + a$, stąd $h(x) = x^4 - x - a = 0$ — chcemy, aby pierwiastek miał krotność większą niż 1. Zatem $h'(x) = 4x^3 - 1 = 0$, stąd $x = \sqrt[3]{1/4}$, zatem $a = x^4 - x = \sqrt[3]{1/4}(1/4 - 1) = -3\sqrt[3]{1/4}/4$. Warunek Slatera jest spełniony z dowolnym punktem należącym do wnętrza zbioru dopuszczalnego, jeśli to wnętrze jest niepuste, co ma miejsce dla $a > -3\sqrt[3]{1/4}/4$. Dla $a < -3\sqrt[3]{1/4}/4$ zbiór dopuszczalny jest pusty. Sytuacja graniczna jest pokazana na rysunku.

69. Zbiór dopuszczalny $W \subset \mathbb{R}^2$ jest opisany w taki sposób:

$$\begin{aligned} &W = \{(x, y)\colon g_i(x, y) \leqslant 0, i = 1, \dots, 5\}, \\ &g_1(x, y) = x^2 + y^2 - 1, \\ &g_2(x, y) = 1 - (x - 1)^2 - y^2, \\ &g_3(x, y) = -1 - x, \\ &g_4(x, y) = -1 - x - y, \\ &g_5(x, y) = y - 1. \end{aligned}$$

Podziel zbiór W na podzbiory, których punkty jednakowo spełniają, albo nie spełniają, znane warunki dostateczne (afiniczności, liniowej niezależności lub Slatera) tego, aby stożek kierunków stycznych do W był stożkiem kierunków zlinearyzowanych. Dla warunku Slatera wskaż odpowiedni punkt (punkty) występujące w tym warunku.

Wskazówka. Przedstaw zbiór W na rysunku.

70. Znajdź minimum funkcji $f(x,y) = x^2 - x + y^2 - \sqrt{3}y$ w zbiorze W z poprzedniego zadania.

71. (GP) Funkcje $f, g\colon (0,\infty)^n \to \mathbb{R}$ są określone wzorami

$$f(x_1,\dots,x_n) = \alpha_1 \ln x_1 + \dots + \alpha_n \ln x_n, \qquad g(x_1,\dots,x_n) = x_1^{\alpha_1} \cdot \ldots \cdot x_n^{\alpha_n}.$$

Badając wypukłość/wklęsłość funkcji f pokaż, że jeśli $\alpha_1,\dots,\alpha_n \leqslant 0$, to funkcja g jest quasi-wypukła, a jeśli $\alpha_1,\dots,\alpha_n \geqslant 0$, to funkcja g jest quasi-wklęsła (zobacz zadanie 22).

72. (NPP) Rozwiąż metodą dualną zadanie

$$\begin{cases} x_1 \to \min, \\ x_1^2 + x_2^2 \leqslant 2, \\ (x_1,x_2) \in X = \{\,(x_1,x_2)\colon x_1 \geqslant 1\,\}. \end{cases}$$

Rozw.

$$L(x_1,x_2,\mu) = x_1 + \mu(x_1^2 + x_2^2 - 2),$$

$$L_D(\mu) = \inf_{x_1 \geqslant 1, x_2 \in \mathbb{R}} L(x_1,x_2,\mu) = \begin{cases} 1 & \text{dla } \mu = 0, \\ 1-\mu & \text{dla } \mu > 0, \end{cases}$$

$$L_P(x_1,x_2) = \sup_{\mu \geqslant 0} L(x_1,x_2,\mu) = \begin{cases} x_1 & \text{dla } x_1^2 + x_2^2 - 2 \leqslant 0, \\ \infty & \text{dla } x_1^2 + x_2^2 - 2 > 0. \end{cases}$$

Szukamy takich $(\overline{x}_1,\overline{x}_2) \in X$ i $\overline{\mu} \in \mathbb{R}_+$, aby $L_P(\overline{x}_1,\overline{x}_2) = L_D(\overline{\mu})$. Jest $L_P \geqslant 1$, $L_D \leqslant 1$, czyli to jest możliwe tylko wtedy, gdy $L_D(\overline{\mu}) = 1 = L_D(\overline{x}_1,\overline{x}_2)$, stąd $\overline{\mu} = 0$, $\overline{x}_1 = 1$, $\overline{x}_2^2 \leqslant 1$. Z twierdzenia 10.1 elementy zbioru $\{\,(x_1,x_2)\colon x_1 = 1, |x_2| \leqslant 1\,\}$ są rozwiązaniami globalnymi. Z twierdzenia 10.2 nie ma innych rozwiązań, bo zadanie jest wypukłe.

73. (NPP) Wykaż, że funkcja dualna jest wklęsła.

Rozw.

$$L_D(\boldsymbol{\mu}) = \inf_{x \in X} L(\boldsymbol{x},\boldsymbol{\mu}) = \inf_{x \in X}\Bigl(f(\boldsymbol{x}) + \sum_{i=1}^m \mu_i g_i(\boldsymbol{x})\Bigr)$$
$$= \inf_{a,b}\Bigl\{\, a + \sum_{i=1}^m \mu_i b_i\colon \exists_x a = f(\boldsymbol{x}), b_i = g_i(\boldsymbol{x}) \Bigr\}.$$

Funkcje $a + \sum_{i=1}^m b_i\mu_i$ są afiniczne, a zatem wklęsłe, więc infimum jest też funkcją wklęsłą.

74. (NPP) Rozwiąż metodą dualną zadanie

$$\begin{cases} \frac{1}{2}\sum_{i=1}^n x_i^2 \to \min, \\ \sum_{i=1}^n x_i = 1, \\ 0 \leqslant x_i \leqslant u_i, \end{cases}$$

gdzie $0 \leqslant u_1 \leqslant u_2 \leqslant \dots \leqslant u_n$, $\sum_{i=1}^n u_i \geqslant 1$.

Rozw. $X = \{\,\boldsymbol{x} \in \mathbb{R}^n\colon 0 \leqslant x_i \leqslant u_i\,\}$.

$$L(\boldsymbol{x},\lambda) = \frac{1}{2}\sum_{i=1}^n x_i^2 + \lambda\Bigl(1 - \sum_{i=1}^n x_i\Bigr),$$

$$L_D(\lambda) = \inf_{x \in X} L(\boldsymbol{x},\lambda) = \inf\Bigl(\frac{1}{2}\sum_{i=1}^n (x_i - \lambda)^2 + \lambda - \frac{n}{2}\lambda^2\Bigr)$$
$$= \frac{1}{2}\sum_{i=1}^n (u_i - \lambda)^2|_{\{u_i \leqslant \lambda\}} + \lambda - \frac{n}{2}\lambda^2.$$

Szukamy $\sup_{\lambda \geqslant 0} L_D(\lambda)$. Ponieważ funkcja $L_D(\lambda)$ jest wklęsła, należy znaleźć punkt, w którym jej pochodna zmienia znak, to będzie maksimum.
Uwaga: funkcja L_D jest różniczkowalna, natomiast ma nieciągłą drugą pochodną.

$$L_D'(\lambda) = -\sum_{i=1}^n (u_i - \lambda)|_{\{u_i < \lambda\}} + 1 - n\lambda = \sum_{i=1}^n (\lambda - u_i)_+ + 1 - n\lambda.$$

Kiedy $L = \sum_{i=1}^n (\lambda - u_i)_+ \geqslant n\lambda - 1 = P$? Dla $\lambda = 0$ lewa strona jest równa 0, prawa -1, zatem $L > P$. Dla $\lambda > u_n$ jest $L = n\lambda - \sum_{i=1}^n u_i$, $P = n\lambda - 1$, $L \leqslant P$, bo $\sum_{i=1}^n u_i \geqslant 1$. Gdzieś „po drodze” pochodna zmienia znak, czyli istnieje punkt $\overline{\lambda}$, w którym funkcja L_D osiąga maksimum. Znajdujemy $\overline{\boldsymbol{x}} = \arg\min_{x \in X} L(\boldsymbol{x},\overline{\lambda})$. Punkt $(\overline{\boldsymbol{x}},\overline{\lambda})$ jest punktem siodłowym funkcji L, czyli rozwiązaniem globalnym. Punkt $\overline{\boldsymbol{x}}$ ma postać $\overline{\boldsymbol{x}} = (u_1,\dots,u_k,\underbrace{\overline{\lambda},\dots,\overline{\lambda}}_{n-k})$.

75. (NPP) Znajdź zadanie dualne do

$$\begin{cases} \boldsymbol{c}^T\boldsymbol{x} \to \min, & \boldsymbol{c} \in \mathbb{R}^n, \\ A\boldsymbol{x} = \boldsymbol{b}, & A \in \mathbb{R}^{m\times n}, \boldsymbol{b} \in \mathbb{R}^m, \\ \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{x} \geqslant 0. \end{cases}$$

Rozw.

$$L(\boldsymbol{x},\boldsymbol{\mu},\boldsymbol{\lambda}) = \boldsymbol{c}^T\boldsymbol{x} + \boldsymbol{\lambda}^T(A\boldsymbol{x} - \boldsymbol{b}) - \boldsymbol{\mu}^T\boldsymbol{x} = (\boldsymbol{c} - \boldsymbol{\mu} + \boldsymbol{A}^T\boldsymbol{\lambda})^T\boldsymbol{x} - \boldsymbol{\lambda}^T\boldsymbol{b}.$$

Stąd

$$L_D(\mu,\lambda)=\inf_{x\in\mathbb{R}^n} L(x,\mu,\lambda)=\begin{cases}-\lambda^T\mathbf{b}, & \text{dla } \mathbf{c}-\mu+A^T\lambda=0,\\ -\infty & \text{w przeciwnym przypadku.}\end{cases}$$

Zadanie dualne:

$$\begin{cases}\lambda^T\mathbf{b}\to\max,\\ \mathbf{c}-\mu+A^T\lambda=0,\\ \lambda\geqslant 0\end{cases}\quad\Rightarrow\quad\begin{cases}\lambda^T\mathbf{b}\to\max,\\ \mathbf{c}+A^T\lambda\geqslant 0,\end{cases}$$

76. (NPP) Znajdź zadanie dualne do

$$\begin{cases}\mathbf{c}^T x\to\min,\\ Ax\leqslant\mathbf{b}.\end{cases}$$

Rozw.

$$L(x,\mu)=\mathbf{c}^T x+\mu^T(Ax-\mathbf{b})=(\mathbf{c}+A^T\mu)^T x-\mu^T\mathbf{b}.$$

$$L_D(\mu)=\inf_x L(x,\mu)=\begin{cases}-\mu^T\mathbf{b} & \text{dla } \mathbf{c}+A^T\mu=0,\\ -\infty & \text{w przeciwnym przypadku.}\end{cases}$$

Zadanie dualne:

$$\begin{cases}-\mu^T\mathbf{b}\to\max,\\ \mathbf{c}+A^T\mu=0,\\ \mu\geqslant 0.\end{cases}$$

77. (NPP) Znajdź zadanie dualne do

$$\begin{cases}x^T x\to\min,\\ Ax=\mathbf{b}.\end{cases}$$

Rozw.

$$L(x,\lambda)=x^T x+\lambda^T(Ax-\mathbf{b}),$$
$$L_D(\lambda)=\inf_x L(x,\lambda)=L(x^*,\lambda).$$

Funkcja L jest wypukła ze względu na x, więc ma minimum. Znajdujemy je z warunku $D_x L(x,\lambda)=0$.

$$D_x L(x,\lambda)=2x+A^T\lambda\quad\Rightarrow\quad x^*=-\frac{1}{2}A^T\lambda.$$

$$L_D(\lambda)=\frac{1}{4}(A^T\lambda)^T A^T\lambda+\lambda^T A\left(-\frac{1}{2}A^T\lambda\right)-\lambda^T\mathbf{b}=-\frac{1}{4}\lambda^T AA^T\lambda-\lambda^T\mathbf{b}.$$

Zadanie dualne:

$$\begin{cases}\frac{1}{4}\lambda^T AA^T\lambda-\lambda^T\mathbf{b}\to\max,\\ \lambda\in\mathbb{R}^n.\end{cases}$$

78. (NPP) Znajdź zadanie dualne do

$$\begin{cases}\frac{1}{2}x^T Hx+\mathbf{d}^T x\to\min,\\ Ax\leqslant\mathbf{b},\end{cases}$$

gdzie $\mathbf{d}\in\mathbb{R}^n$, $A\in\mathbb{R}^{m\times n}$, $\mathbf{b}\in\mathbb{R}^m$, $H\in\mathbb{R}^{n\times n}$ jest macierzą symetryczną i dodatnio określoną, $x\in\mathbb{R}^n$.

Rozw.

$$L(x,\mu)=\frac{1}{2}x^T Hx+\mathbf{d}^T x+\mu^T(Ax-\mathbf{b})=\frac{1}{2}x^T Hx+(A^T\mu)^T x-\mu^T\mathbf{b}+\mathbf{d}^T x,$$
$$L_D(\mu)=\inf_x L(x,\mu)=L(x^*,\mu).$$

Funkcja $L(x,\mu)$ jest wypukła ze względu na x, więc ma minimum:

$$D_x L(x,\mu)=Hx+(A^T\mu+\mathbf{d})=0\quad\Rightarrow\quad x^*=-H^{-1}(A^T\mu+\mathbf{d}).$$

Stąd

$$\begin{aligned}L_D(\mu)&=\frac{1}{2}(A^T\mu+\mathbf{d})^T H^{-1}(A^T\mu+\mathbf{d})-\mu^T AH^{-1}(A^T\mu+\mathbf{d})\\ &\quad-\mathbf{d}^T H^{-1}(A^T\mu+\mathbf{d})-\mu^T\mathbf{b}\\ &=-\frac{1}{2}(A^T\mu+\mathbf{d})^T H^{-1}(A^T\mu+\mathbf{d})-\mu^T\mathbf{b}.\end{aligned}$$

Zadanie dualne:

$$\begin{cases}-\frac{1}{2}(A^T\mu+\mathbf{d})^T H^{-1}(A^T\mu+\mathbf{d})-\mu^T\mathbf{b}\to\max,\\ \mu\geqslant 0.\end{cases}$$

79. Znajdź zadanie dualne do

$$\begin{cases}x^T Hx+\mathbf{e}^T x\to\min,\\ Ax\leqslant\mathbf{b},\\ Cx=\mathbf{d},\end{cases}$$

w którym macierz $H\in\mathbb{R}^{n\times n}$ jest symetryczna i dodatnio określona, suma zbiorów wierszy macierzy $A\in\mathbb{R}^{m\times n}$ i $C\in\mathbb{R}^{k\times n}$ jest liniowo niezależna, $\mathbf{e}\in\mathbb{R}^n$, $\mathbf{b}\in\mathbb{R}^m$, $\mathbf{d}\in\mathbb{R}^k$.

Rozw. Przyjmujemy $X=\mathbb{R}^n$ i określamy funkcję

$$L(x,\mu,\lambda)=x^T Hx+\mathbf{e}^T x+\mu^T(Ax-\mathbf{b})+\lambda^T(Cx-\mathbf{d}),$$

a na jej podstawie

$$L_D(\mu,\lambda)=\inf_{x\in\mathbb{R}} L(x,\mu,\lambda).$$

Dzięki wypukłości funkcji L możemy wyznaczyć $\boldsymbol{x}$:

$$D_x L(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\lambda}) = 2H\boldsymbol{x} + \boldsymbol{e} + A^T\boldsymbol{\mu} + C^T\boldsymbol{\lambda} = \boldsymbol{0},$$

skąd wynika

$$\boldsymbol{x} = -\frac{1}{2}H^{-1}(\boldsymbol{e} + A^T\boldsymbol{\mu} + C^T\boldsymbol{\lambda}).$$

Po podstawieniu do funkcji L otrzymujemy

$$\begin{aligned} L_D(\boldsymbol{\mu}, \boldsymbol{\lambda}) = {} & \frac{1}{4}(\boldsymbol{e} + A^T\boldsymbol{\mu} + C^T\boldsymbol{\lambda})^T H^{-1}(\boldsymbol{e} + A^T\boldsymbol{\mu} + C^T\boldsymbol{\lambda}) \\ & - \frac{1}{2}\boldsymbol{e}^T H^{-1}(\boldsymbol{e} + A^T\boldsymbol{\mu} + C^T\boldsymbol{\lambda}) - \frac{1}{2}\boldsymbol{\mu}^T\big(AH^{-1}(\boldsymbol{e} + A^T\boldsymbol{\mu} + C^T\boldsymbol{\lambda}) - \boldsymbol{b}\big) \\ & - \frac{1}{2}\boldsymbol{\lambda}^T\big(CH^{-1}(\boldsymbol{e} + A^T\boldsymbol{\mu} + C^T\boldsymbol{\lambda}) - \boldsymbol{d}\big). \end{aligned}$$

Po uporządkowaniu możemy przedstawić funkcję L_D w zapisie macierzowym

$$\begin{aligned} L_D(\boldsymbol{\mu}, \boldsymbol{\lambda}) = {} & -\frac{1}{4}[\boldsymbol{\mu}^T, \boldsymbol{\lambda}^T] \begin{bmatrix} AH^{-1}A^T & AH^{-1}C^T \\ CH^{-1}A^T & CH^{-1}C^T \end{bmatrix} \begin{bmatrix} \boldsymbol{\mu} \\ \boldsymbol{\lambda} \end{bmatrix} \\ & - \frac{1}{2}[\boldsymbol{e}^T H^{-1}A^T - \boldsymbol{b}^T, \boldsymbol{e}^T H^{-1}C^T - \boldsymbol{d}^T] \begin{bmatrix} \boldsymbol{\mu} \\ \boldsymbol{\lambda} \end{bmatrix} - \frac{1}{4}\boldsymbol{e}^T H^{-1}\boldsymbol{e}. \end{aligned}$$

Zadanie dualne:

$$\begin{cases} L_D(\boldsymbol{\mu}, \boldsymbol{\lambda}) \to \max, \\ \boldsymbol{\mu} \geqslant \boldsymbol{0}, \\ \boldsymbol{\lambda} \in \mathbb{R}^k. \end{cases}$$

80. Rozwiąż zadanie minimalizacji

$$\begin{cases} \|\boldsymbol{x}\|_\infty \to \min, \\ \|\boldsymbol{x} - \boldsymbol{x}_0\|_2^2 \leqslant r^2, \end{cases}$$

dla ustalonego wektora $\boldsymbol{x}_0 \in \mathbb{R}^n$ oraz liczby $r > 0$ i znajdź zadanie dualne.

Rozw. Dla dowolnego $\boldsymbol{x} \in \mathbb{R}^n$ jest $\|\boldsymbol{x}\|_\infty \leqslant \|\boldsymbol{x}\|_2$. Jeśli zatem $\|\boldsymbol{x}_0\|_2 \leqslant r$, to mamy rozwiązanie trywialne $\boldsymbol{x} = \boldsymbol{0}$. Przypuśćmy, że $\|\boldsymbol{x}_0\|_2 > r$.

Określamy funkcję

$$L(\boldsymbol{x}, \mu) = \|\boldsymbol{x}\|_\infty + \mu\big(\|\boldsymbol{x} - \boldsymbol{x}_0\|_2^2 - r^2\big).$$

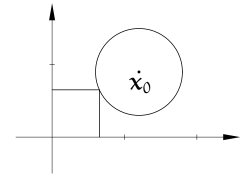
Funkcja ta dla $\mu \geqslant 0$ jest wypukła ze względu na $\boldsymbol{x}$, a zatem ma minimum, ale nie jest różniczkowalna. Dla ustalonego wektora $\boldsymbol{x}_0 = [x_{01}, \ldots, x_{0n}]^T$ określmy zbiory indeksów

$$I = \{\, i\colon |x_{0i}| \leqslant \|\boldsymbol{x}_0\|_\infty - r\,\}, \quad J = \{\, i\colon |x_{0i}| > \|\boldsymbol{x}_0\|_\infty - r\,\}.$$

Oczywiście, $I \cup J = \{1, \ldots, n\}$, przy czym zbiór I może, ale nie musi być pusty. Zbiór J ma co najmniej jeden element.

Jeśli $I = \emptyset$, to weźmy $\boldsymbol{x} = a[\operatorname{sgn} x_{01}, \ldots, \operatorname{sgn} x_{0n}]^T$ dla pewnego $a > 0$, wtedy jest $\|\boldsymbol{x}\|_\infty = a$. Funkcja dualna ma postać

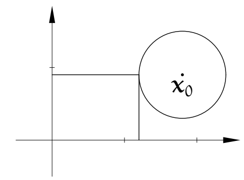
$$L_D(\mu) = a + \mu\Big(\sum_{i=1}^n (a \operatorname{sgn} x_{i0} - x_{i0})^2 - r^2\Big) = a + \mu\Big(\sum_{i=1}^n (a - |x_{0i}|)^2 - r^2\Big).$$



Należy znaleźć jej maksimum dla $\mu \geqslant 0$, ale ono istnieje wtedy, gdy wyrażenie w nawiasie jest niedodatnie. Powinno być równe 0 — wtedy ograniczenie jest aktywne, stąd mamy do rozwiązania równanie kwadratowe (z niewiadomą a)

$$na^2 - 2\Big(\sum_{i=1}^n |x_{0i}|\Big)a + \sum_{i=1}^n x_{0i}^2 - r^2 = 0,$$

czyli $na^2 - 2\|\boldsymbol{x}_0\|_1 a + \|\boldsymbol{x}_0\|_2^2 - r^2 = 0.$



Przypadek $I \neq \emptyset$ zostawiam do samodzielnego zbadania (wskazówka na rysunku obok).

81. Rozwiąż (wykonując odpowiednie rachunki) zadania minimalizacji jak w poprzednim zadaniu, dla $n = 2$ oraz a) $\boldsymbol{x}_0 = [1, 0.5]^T$, $r = 1$, b) $\boldsymbol{x}_0 = [2, 0.5]^T$, $r = 1$.

82. (NPP) Transformatą Fenchela-Legendre'a funkcji $f\colon X \to \mathbb{R}$ nazywamy funkcję $f^*\colon \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ daną wzorem

$$f^*(\boldsymbol{y}) = \sup_{x \in X}\big(\boldsymbol{y}^T\boldsymbol{x} - f(\boldsymbol{x})\big).$$

Znajdź f^* dla

a) $X = \mathbb{R}$, $f(x) = \frac{1}{2}x^2$.

Rozw.

$$f^*(y) = \sup_x\big(yx - x^2/2\big) = \sup_x(y^2/2 - y^2/2 + yx - x^2/2) = \frac{1}{2}\sup_x\big(y^2 - (y - x)^2\big) = \frac{1}{2}y^2.$$

b) $X = \mathbb{R}^n$, $f(\boldsymbol{x}) = \frac{1}{2}\|\boldsymbol{x}\|_2^2$.

Rozw.

$$f^*(y) = \sup_x \left(y^T x - \frac{1}{2}x^T x/2\right) = \frac{1}{2}\sup_x(y^T y - y^T y + y^T x - x^T x) =$$
$$\frac{1}{2}\sup_x \left(y^T y - (y-x)^T(y-x)\right) = \frac{1}{2}\|y\|_2^2.$$

c) $X = \mathbb{R}$, $f(x) = e^x$.

Rozw. $f^*(y) = \sup_x(yx - e^x)$. Dla $y > 0$ szukamy maksimum funkcji $g(x) = (yx - e^x)$; jest $g'(x) = y - e^x$. Z $g'(x) = 0$ dostajemy $x = \ln y$, czyli

$$\sup_x(yx - e^x) = y\ln y - y.$$

dla $y = 0$ $\sup_x(-e^x) = 0$,

dla $y < 0$ $\sup_x(yx - e^x) = +\infty$ (dla $x \to -\infty$)

Zatem,

$$f^*(y) = \begin{cases} y\ln y - y & \text{dla } y > 0, \\ 0 & \text{dla } y = 0, \\ +\infty & \text{dla } y < 0. \end{cases}$$

d) $X = \mathbb{R}^n$, $f(x) = \|x\|_p$, $p > 1$.

Rozw.

$$f^*(y) = \sup_x(y^T x - \|x\|_p) = \sup_{\substack{\|x\|_p \leqslant 1 \\ \lambda \geqslant 0}} \left(\lambda(y^T x - \|x\|_p)\right).$$

Wystarczy rozwiązać zadanie $y^T x - 1 \to \max$ dla $y \neq 0$. Weźmy $\|x\|_p \leqslant 1$.

$$L(x, \mu) = (1 - y^T x) + \mu(\|x_p\| - 1).$$

Rozwiązujemy układ równań

$$D_x L = \left[-y_i + \mu\|x\|_p^{(1-p)/p}|x_i|^{p-1}\operatorname{sgn} x_i\right]_i = 0,$$

Przyjęcie $\mu = 0$ daje sprzeczność z $y \neq 0$, zatem $\mu > 0$. Stąd $\operatorname{sgn} y_i = \operatorname{sgn} x_i$, $|y_i| - \mu\|x\|_p^{(1-p)/p}|x_i|^{p-1} = 0$, a dalej $|x_i| = |y_i|^{1/(p-1)}/\mu^{1/(p-1)}$, bo $\|x\|_p = 1$. Zatem

$$y^T x = \sum_{i=1}^n \frac{|y_i|^{p/(p-1)}}{\mu^{1/(p-1)}},$$
$$1 = \|x\|_p^p = \sum_{i=1}^n \left|\frac{y_i}{\mu}\right|^{p/(p-1)} = \frac{1}{\mu^{p/(p-1)}}\sum_{i=1}^n |y_i|^{p/(p-1)}.$$

Dla $\frac{1}{p} + \frac{1}{q} = 1$ jest $q = \frac{p}{p-1} = 1 + \frac{1}{p-1}$. Z $\mu^q = \sum_{i=1}^n |y_i|^q$ wynika $\mu = \|y\|_q$. Czyli

$$y^T x = \frac{\|y\|_q^q}{\mu^{q-1}} = \frac{\|y\|_q^q}{\|y\|_q^{q-1}} = \|y\|_q.$$

Ostatecznie

$$f^*(y) = \sup_{\lambda \geqslant 0}\left(\lambda(\|y\|_q^{q-1} - 1)\right) = \begin{cases} 0 & \text{dla } \|y\|_q \leqslant 1, \\ +\infty & \text{dla } \|y\|_q > 1. \end{cases}$$

83. (NPP) Udowodnij, że problem dualny do

$$\begin{cases} f(x) \to \min, & \\ Ax \leqslant b, & b \in \mathbb{R}^m,\ A \in \mathbb{R}^{m\times n}, \\ Cx = d, & d \in \mathbb{R}^l,\ C \in \mathbb{R}^{l\times n}, \\ x \in X & \end{cases}$$

ma postać

$$\begin{cases} -b^T\lambda - d^T\mu - f^*(-A^T\lambda - C^T\mu) \to \max, \\ \lambda \geqslant 0, \\ \lambda \in \mathbb{R}^m,\ \mu \in \mathbb{R}^l. \end{cases}$$

Rozw.

$$L_D = \inf_{x\in X} f(x) + \lambda^T(Ax - b) + \mu^T(Cx - d),$$

przy czym $Cx = d \Leftrightarrow Cx - d \leqslant 0 \wedge d - Cx \leqslant 0$ (przedstawienie równości jako koniunkcji dwóch nierówności jest potrzebne do poszukiwania minimum). Zatem,

$$\begin{aligned} L_D = {} & \inf_{x\in X}\left(f(x) + \lambda^T(Ax - b) + \mu^T(Cx - d)\right) = \\ & -b^T\lambda - d^T\mu + \inf_{x\in X}\left(f(x) + \lambda^T Ax + \mu^T Cx\right) = \\ & -b^T\lambda - d^T\mu - \sup_{x\in X}\left(-\lambda^T Ax - \mu^T Cx - f(x)\right) = \\ & -b^T\lambda - d^T\mu - f^*\left(-A^T\lambda - C^T\mu\right). \end{aligned}$$

84. (NPP) Udowodnij, że funkcja $F\colon \mathbb{R}^2 \times \mathbb{R}_+^2 \to \mathbb{R}$ dana wzorem

$$F(x_1, x_2, \lambda_1, \lambda_2) = x_1^2 + e^{2x_2} - 4x_1 + \lambda_1(e^{x_2} - x_1) + \lambda_2(5 - x_2)$$

ma punkt siodłowy.

Rozw. Należy znaleźć rozwiązanie zadania

$$\begin{cases} f(x_1, x_2) = x_1^2 + x_2^2 - 4x_1 \to \min, \\ g_1(x_1, x_2) = e^{x_2} - x_1 \leqslant 0, \\ g_2(x_1, x_2) = 5 - x_2 \leqslant 0. \end{cases}$$

Funkcja F jest funkcją Lagrange'a, $L(x_1, x_2, \lambda_1, \lambda_2) = F(x_1, x_2, \lambda_1, \lambda_2)$. Rozwiązujemy układ

$$D_x L = [2x_1 - 4 - \lambda_1, 2e^{2x_2} + \lambda_1 e^{x_2} - \lambda_2] = \mathbf{0}.$$

Kolejno badamy przypadki

1. $\lambda_1 = \lambda_2 = 0 \Rightarrow x_1 = 2,\ e^{2x_2} = 0$, sprzeczność.
2. $\lambda_1 = 0,\ \lambda_2 > 0 \Rightarrow x_1 = 2,\ 2e^{2x_2} = \lambda_2\ x_2 = 5 \Rightarrow \lambda_2 = 2e^{10}$, ale $e^{x_2} - x_1 = e^5 - 2 > 0$, sprzeczność z $g_1(x_1, x_2) \leqslant 0$.
3. $\lambda_1 > 0,\ \lambda_2 = 0 \Rightarrow 2e^{2x_2} + \lambda_1 e_2^x = 0$, sprzeczność.
4. $\lambda_1 > 0,\ \lambda_2 > 0 \Rightarrow 5 - x_2 = 0,\ e^{x_2} - x_1 = 0$, stąd $x_2 = 5,\ x_1 = e^5$, $\lambda_1 = 2x_1 - 4 = 2e^5 - 4 > 0,\ \lambda_2 = 2e^{2x_2} + \lambda_1 e^{x_2} = 2e^{10} + (2e^5 - 4)e^5 > 0$. Czy to jest minimum? TAK, bo funkcja $f(x_1, x_2)$ jest wypukła, g_1, g_2 też są wypukłe, zatem punkt $(e^5, 5, 2e^5 - 4, 4e^5(e^5 - 1))$ jest siodłowy.

85. (NPP) Dany jest problem optymalizacyjny

$$\begin{cases} f(x_1, x_2) = x_2 \to \min, \\ x_1 \geqslant 1, \\ x_1^2 + x_2^2 \leqslant 1. \end{cases}$$

Rozwiąż problem pierwotny i sformułuj i rozwiąż problem dualny.

Rozw. Problem pierwotny można rozwiązać geometrycznie: funkcja f ma minimum w punkcie $(1, 0)$, jest $f(1, 0) = 0$.

Problem dualny: niech $X = \{ (x_1, x_2) \colon x_1^2 + x_2^2 \leqslant 1 \}$. Określamy funkcję $L(\boldsymbol{x}, \mu) = x_2 + \mu(1 - x_1)$ oraz

$$L_D(\mu) = \inf_{(x_1, x_2) \in X} x_2 + \mu(1 - x_1), \quad \mu \geqslant 0.$$

Jest $D_x L = [-u, 1] \neq \mathbf{0}$, zatem ekstremum jest na brzegu zbioru X. Z $x_1^2 + x_2^2 = 1$ wynika $x_1 = \sqrt{1 - x_2^2}$. Minimalizujemy funkcję

$$g(x) = \bigl(x + \mu(1 - \sqrt{1 - x^2})\bigr),$$

rozwiązując równanie $g'(x) = 1 + \mu \frac{x}{\sqrt{1-x^2}} = 0$. Stąd

$$\frac{x}{\sqrt{1 - x^2}} = -\frac{1}{\mu} \quad \Rightarrow \quad \frac{x^2}{1 - x^2} = \frac{1}{\mu^2} \quad \Rightarrow \quad \frac{1 - x^2}{1 - x^2} - \frac{1}{1 - x^2} = -\frac{1}{\mu}^2 \quad \Rightarrow$$
$$1 + \frac{1}{\mu^2} = \frac{1}{1 - x^2} \quad \Rightarrow \quad \frac{\mu^2 + 1}{\mu^2} = \frac{1}{1 - x^2} \quad \Rightarrow \quad 1 - x^2 = \frac{\mu^2}{\mu^2 + 1} \quad \Rightarrow$$
$$x^2 = 1 - \frac{\mu^2}{\mu^2 + 1} = \frac{1}{\mu^2 + 1}.$$

Stąd $x_2 = \pm \frac{1}{\sqrt{1+\mu^2}}$, $x_1 = \mu/\sqrt{1 + \mu}$, $x_2 = -\frac{1}{\sqrt{1+\mu^2}}$, bo to daje minimum. Wtedy

$$\begin{aligned} L_D(\mu) &= -\frac{1}{\sqrt{1 + \mu^2}} + \mu\left(1 - \frac{\mu}{\sqrt{1 + \mu^2}}\right) = \frac{1}{\sqrt{1 + \mu^2}}(\mu\sqrt{1 + \mu^2} - (1 + \mu \\ &= \mu - \sqrt{1 + \mu^2} < 0. \end{aligned}$$

Jest $\lim_{\mu \to \infty} L_D(\mu) = 0 = \sup_{\mu \geqslant 0} L_D(\mu)$. Zatem nie istnieje $\arg\max L_D(\mu)$.

86. (NPP) Dany jest problem optymalizacyjny

$$\begin{cases} 3x_1 + 2x_2 + x_3 \to \max, \\ 2x_1 + x_2 - x_3 \leqslant 2, \\ x_1 + 2x_2 \leqslant 4, \\ x_3 \leqslant 3, \\ x_1, x_2, x_3 \geqslant 0. \end{cases}$$

Znajdź i rozwiąż zadanie dualne dla

$$X_1 = \{ (x_1, x_2, x_3) \colon 2x_1 + x_2 - x_3 \leqslant 2,\ x_1, x_2, x_3 \geqslant 0 \},$$
$$X_2 = \{ (x_1, x_2, x_3) \colon x_1 + 2x_2 \leqslant 4,\ x_1, x_2, x_3 \geqslant 0 \}.$$

Rozw. $\boxed{\boxed{X_1:}}$

$$\begin{aligned} L(\boldsymbol{x}, \boldsymbol{\mu}) &= -3x_1 - 2x_2 - x_3 + \mu_1(x_1 + 2x_2 - 4) + \mu_2(x_3 - 3) \\ &= x_1(\mu_1 - 3) + x_2(2\mu_1 - 2) + x_3(\mu_2 - 1) - 4\mu_1 - 3\mu_2. \end{aligned}$$

Jeśli $\mu_2 < 1$, μ_1 — dowolne, to

$$L_D(\boldsymbol{\mu}) = \inf_{x \in X_1} L(\boldsymbol{x}, \boldsymbol{\mu}) = -\infty \quad (x_3 \to +\infty, x_1 = x_2 = 0).$$

Jeśli $\mu_2 \geqslant 1$ oraz $\mu_1 \geqslant 3$, to

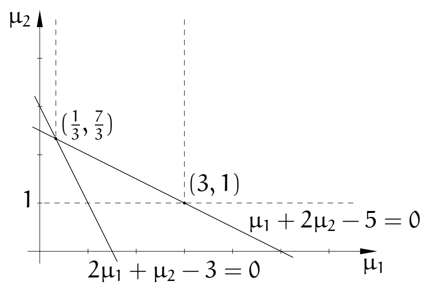
$$L_D(\boldsymbol{\mu}) = \inf_{x \in X_1} L(\boldsymbol{x}, \boldsymbol{\mu}) = -4\mu_1 - 3\mu_2 \quad (x_1 = x_2 = x_3 = 0).$$

Jeśli $\mu_2 \geqslant 1$ oraz $\mu_1 < 3$, to minimum jest osiągane w płaszczyźnie $2x_1 + x_2 - x_3 = 2$, tj. $x_3 = 2x_1 + x_2 - 2$. Mamy

$$\begin{aligned} L(\mathbf{x}, \boldsymbol{\mu}) &= x_1(\mu_1 - 3) + x_2(2\mu_1 - 2) + (2x_1 + x_2 - 2)(\mu_2 - 1) - 4\mu_1 - 3\mu_2 \\ &= x_1(\mu_1 + 2\mu_2 - 5) + x_2(2\mu_1 + \mu_2 - 3) - 4\mu_1 - 5\mu_2 + 2. \end{aligned}$$

Wtedy

$$\begin{aligned} &\mu_1 + 2\mu_2 - 5 < 0 \quad \Rightarrow \quad L_D(\boldsymbol{\mu}) = -\infty \quad (x_2 = 0,\ 2x_1 - x_3 = 2\,x_1 \to \infty), \\ &2\mu_1 + \mu_2 - 3 < 0 \quad \Rightarrow \quad L_D(\boldsymbol{\mu}) = -\infty \quad (x_1 = 0,\ x_2 - x_3 = 2,\ x_2 \to \infty). \end{aligned}$$



Jeśli $\mu_1 + 2\mu_2 - 5 \geqslant 0 \wedge 2\mu_1 + \mu_2 - 3 \geqslant 0$, to wartość minimalna jest w płaszczyźnie $x_3 = 0$, tj. $2x_1 + x_2 = 2$.

$$\begin{aligned} L(\mathbf{x}, \boldsymbol{\mu}) &= x_1(\mu_1 + 2\mu_2 - 5) + (2 - 2x_1)(2\mu_1 + \mu_2 - 3) - 4\mu_1 - 5\mu_2 + 2 \\ &= x_1(\mu_1 + 2\mu_2 - 5) + x_1(-4\mu_1 - 2\mu_2 + 6) + 4\mu_1 + 2\mu_2 - 6 - 4\mu_1 - 5\mu_2 + 2 \\ &= x_1(-3\mu_1 + 1) - 3\mu_2 - 4. \end{aligned}$$

Dla $\mu_1 \leqslant \frac{1}{3}$ jest $-3\mu_1 + 1 \geqslant 0$, $L_D(\boldsymbol{\mu}) = -3\mu_2 - 4$ $(x_1 = 0)$.
Dla $\mu_1 > \frac{1}{3}$ jest $-3\mu_1 + 1 < 0$, $L_D(\boldsymbol{\mu}) = -3\mu_1 - 3\mu_2 - 3$ $(x_1 = 1,\ x_2 = 0)$.

$L_D(\mu) = -4\mu_1 - 3\mu_2$ dla $x_1 = x_2 = x_3 = 0$.

Dla $\mu_1 \leqslant 1/3$ jest $-3\mu_2 - 4 < -4\mu_1 - 3\mu_2$, $L_D(\boldsymbol{\mu}) = -3\mu_2 - 4$.
Dla $1/3 < \mu_1 < 3$ jest $-3\mu_1 - 3\mu_2 - 3 < 4\mu_1 - 3\mu_2$, $L_d(\boldsymbol{\mu}) = -3\mu_1 - 3\mu_2 - 3$.
Dla $\mu_1 \geqslant 3$ jest $L_D(\boldsymbol{\mu}) = -4\mu_1 - 3\mu_2$.

Obliczamy $\sup_\mu L_D(\boldsymbol{\mu})$.

$\mu_1 \leqslant 1/3$, $L_D = -3\mu_2 - 4$:

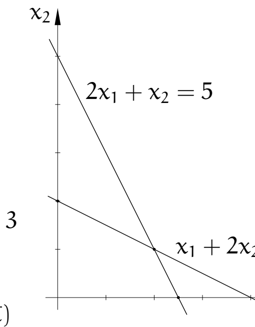
$$\begin{cases} \mu_1 + 2\mu_2 = 5 \\ 2\mu_1 + \mu_2 = 3 \end{cases} \Rightarrow \mu_1 = 1/3, \quad \mu_2 = 7/3$$

$\sup L_D = -7 - 4 = -11$.

$3 > \mu_1 > 1/3$, $L_D = -3\mu_1 - 3\mu_2 - 3$:

$$\begin{aligned} &\mu_1 + 2\mu_2 = 5 \quad \Rightarrow \quad \mu_1 = 5 - 2/\mu_2 \quad \Rightarrow \\ &\qquad L_D = -3(5 - 2\mu_2) - 3\mu_2 - 3 = -15 + 6\mu_2 - 3\mu_2 - 3 = 3\mu_2 - 18. \end{aligned}$$

Stąd $L_D(7/3) = 7 - 18 = -11$.
$\mu_1 \geqslant 3$, $L_D = -4\mu_1 - 3\mu_2$: $\sup L_D = -12 - 3 = -15$.
Zatem $\sup L_D(\boldsymbol{\mu}) = -11$ jest osiągane w punkcie $\overline{\boldsymbol{\mu}} = (1/3, 7/3)$. To oznacza, że oba ograniczenia są aktywne. Muszą być spełnione równości $x_1 + 2x_2 = 4$, $x_3 = 3$. Obliczamy



$$\begin{aligned} L(\mathbf{x}, \overline{\boldsymbol{\mu}}) &= -3x_1 - 2x_2 - x_3 = -3(4 - 2x_2) - 2x_3 - 3 \\ &= -12 + 6x_2 - 2x_2 - 3 = 4x_2 - 15. \end{aligned}$$

Ma być $2x_1 + x_2 \leqslant 5$, czyli $x_1 \in [1, 2]$. Minimum $L(\mathbf{x}, \overline{\boldsymbol{\mu}})$ jest w punkcie $x_2 = 1$, stąd $\overline{\mathbf{x}} = (2, 1, 3)$.

$\boxed{X_2:}$

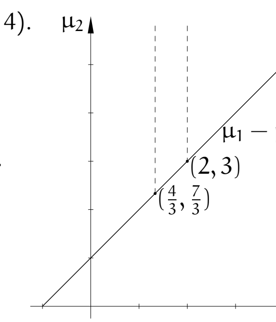
$$\begin{aligned} L(\mathbf{x}, \boldsymbol{\mu}) &= -3x_1 - 2x_2 - x_3 + \mu_1(2x_1 + x_2 - x_3 - 2) + \mu_2(x_3 - 3) = \\ &= x_1(2\mu_1 - 3) + x_2(\mu_2 - 2) + x_3(-\mu_1 + \mu_2 - 1) - 2\mu_1 - 3\mu_2. \end{aligned}$$

$-\mu_1 + \mu_2 - 1 < 0 \Rightarrow \mu_1 - \mu_2 > -1$, $L_D \to -\infty$ $(x_3 \to \infty)$.

$-\mu_1 + \mu_2 - 1 \geqslant 0 \Rightarrow 2\mu_1 - 2 < 2$, stąd $\mu_1 < 3/2$. Minimum jest osiągane na prostej $x_1 + 2x_2 = 4 \Rightarrow x_1 = 4 - 2x_2$.

$$\begin{aligned} &(4 - 2x_2)(2\mu_1 - 3) + x_2(\mu_1 - 3\mu_2) = \\ &\qquad 8\mu_1 - 12 - 4\mu_1 x_2 + 6x_2 + x_2(\mu_1 - 2) - 2\mu_1 - 3\mu_2 = \\ &\qquad x_2(-3\mu_1 + 4) + 6\mu_1 - 3\mu_2 - 12. \end{aligned}$$

$$\begin{aligned} &4 - 3\mu_1 < 0 \quad \Rightarrow \quad \mu_1 > 4/3 \quad \Rightarrow \\ &\qquad L_D = -6\mu_1 + 8 + 6\mu_1 - 3\mu_2 - 12 = -3\mu_2 - 4 \quad (\text{dla } x_2 = 2,\ x_1 = 0) \\ &4 - 3\mu_1 \geqslant 0 \quad \Rightarrow \quad \mu_1 \leqslant 4/3 \quad \Rightarrow \\ &\qquad L_D = 6\mu_1 - 3\mu_2 - 12 \quad (\text{dla } x_2 = 0, x_1 = 4). \end{aligned}$$



$2\mu_1 - 3 \geqslant 0$, $\mu_1 - 2 < 0$

$$\begin{aligned} &3/2 < \mu_1 < 2 \quad \Rightarrow \quad x_1 = 0, x_2 = 2, \\ &\qquad L_D = 2\mu_1 - 4 - 2\mu_1 - 3\mu_2 = -3\mu_2 - 4. \end{aligned}$$

$2\mu_1 - 3 \geqslant 0$, $\mu_1 - 2 \geqslant 0$

$$\mu_1 \geqslant 2 \quad \Rightarrow \quad x_1 = x_2 = 0,\ L_D = -2\mu_1 - 3\mu_2.$$

<u>Obliczamy $\sup_\mu L_D(\mu)$.</u>

$\mu_1 < 4/3$, $L_D(\mu) = 6\mu_1 - 3\mu_2 - 12$.

$$\mu_1 = \mu_2 - 1, \quad L_D = 6(\mu_2 - 1) - 3\mu_2 - 12 = 6\mu_2 - 6 - 3\mu_2 - 12 = 3\mu_2 - 18.$$
$$L_D = 3 \cdot 7/3 - 18 = 7 - 18 = -11, \quad \mu_1 = 4/3, \quad \mu_2 = 7/3.$$

$4/3 \leqslant \mu_1 < 2$, $L_D = -3\mu_2 - 4$, $L_D = -3 \cdot 7/3 - 4 = -7 - 4 = -11$. $\mu_1 = 4/3$, $\mu_2 = 7/3$.
$\mu_1 \geqslant 2$, $L_D = -2\mu_1 - 3\mu_2$, $\mu_1 = \mu_2 - 1$

$$L_D = -2(\mu_2 - 1) - 3\mu_2 = -5\mu_2 + 2,$$
$$\sup L_D = -15 + 2 = -13, \quad \mu_1 = 2\ \mu_2 = 3.$$

Mamy zatem $\sup L_D(\mu) = -11$ w punkcie $\overline{\mu} = (4/3, 7/3)$.

Oba ograniczenia są aktywne, tj. $2x_1 + x_2 - x_3 = 2$, $x_3 = 3$. Dalej, $2x_1 + x_2 = 5$, $x_3 = 3$, $x_2 = 5 - 2x_1$, $x_1 + 2x_2 \leqslant 4$, $x_1 = (5 - x_2)/2$.

$$L(x, \overline{\mu}) = -3/2 \cdot (5 - x_2) - 2x_2 - 3 = -\frac{1}{2}x_2 - \frac{21}{2}.$$

Jest $x_2 \in [0, 1]$, stąd $\min L_D(x, \overline{\mu}) = -11$ dla $x_2 = 1, x_1 = 2$. Rozwiązanie $\overline{x} = (2, 1, 3)$.

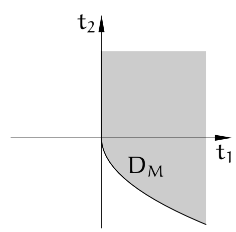
87. (NPP) Dla problemu

$$\begin{cases} x^2 + 2y^2 \to \min, \\ x^2 + y^2 \leqslant 0, \\ x \leqslant 0 \end{cases}$$

znajdź a) dziedzinę funkcji perturbacji D_M, b) wektor wrażliwości.

<u>Rozw.</u> a) Dla jakich t_1, t_2 układ

$$\begin{cases} x^2 + y^2 \leqslant t_1, \\ x \leqslant t_2 \end{cases}$$

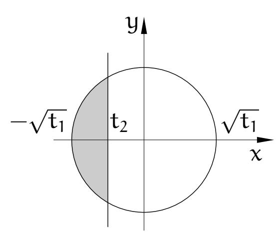


ma rozwiązanie? Musi być $t_1 \geqslant 0$, $t_2 \geqslant -\sqrt{t_1}$. Zatem,

$$D_M = \{ (t_1, t_2) \colon t_1 \geqslant 0,\ \sqrt{t_1} + t_2 \geqslant 0 \}.$$

b) Niech $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$.

$$L(x, \boldsymbol{\lambda}) = x^2 + 2y^2 + \lambda_1(x^2 + y^2) + \lambda_2 x.$$
$$D_x L = [2x + 2\lambda_1 x + \lambda_2, 4y + 2\lambda_1 y].$$



Stąd układ równań i nierówności

$$\begin{cases} 2x + 2\lambda_1 x + \lambda_2 = 0, \\ 4y + 2\lambda_1 y = 0, \\ \lambda_1(x^2 + y^2) = 0, \\ \lambda_2 x = 0, \\ \lambda_1, \lambda_2 \geqslant 0. \end{cases}$$

Z $x^2 + y^2 \leqslant 0$ wynika $x = y = 0$, stąd $\lambda_2 = 0$, λ_1 może być dowolne dodatnie. Punkt nie jest regularny. Sprawdzenie z definicji

$$0 = L(0, 0, \lambda_1, 0) \leqslant L(x, y, \lambda_1, 0) = x^2 + 2y^2 + \lambda_1(x^2 + y^2).$$

Z twierdzenia 11.4 jeśli $L(x^*, \boldsymbol{\lambda}) \leqslant L(x^*, \boldsymbol{\lambda}^*) \leqslant L(x, \boldsymbol{\lambda}^*)$, to λ^* jest wektorem wrażliwości. Jest $x^* = (0, 0)$, $\boldsymbol{\lambda}^* = (\lambda_1, \lambda_2)$, λ_1 — dowolne dodatnie, $\lambda_2 = 0$.

88. (NPP) Dla zadania

$$\begin{cases} x^2 + y^2 \to \min, \\ x + y \leqslant 0 \end{cases}$$

znajdź wektor wrażliwości i funkcję perturbacji.

<u>Rozw.</u> Rozważamy zadanie $x^2 + y^2 \to \min\ x + y \leqslant t$. Określamy funkcję

$$L_t = x^2 + y^2 + \lambda(x + y - t).$$

Mamy stąd układ

$$\begin{cases} 2x + \lambda = 0, \\ 2y + \lambda = 0, \\ \lambda(x + y - t) = 0, \\ \lambda \geqslant 0 \end{cases} \quad \Rightarrow \quad x = y,\ \lambda = -2x,\ -2x(2x - t) = 0,$$

stąd $x = 0$ lub $x = t/2$. Są dwa punkty podejrzane, $(0, 0)$ oraz $(t/2, t/2)$. Czy punkt $(0, 0)$ jest dopuszczalny? <u>Tak</u>, jeśli $t \geqslant 0$ i wtedy $\lambda = 0$. Czy punkt $(t/2, t/2)$ jest dopuszczalny? $\lambda = -t$, czyli jest dopuszczalny dla $t < 0$.

Funkcja perturbacji

$$M(t) = \begin{cases} 0 & \text{dla } t \geqslant 0, \\ \frac{1}{2}t^2 & \text{dla } t < 0. \end{cases}$$

Wektor wrażliwości

$$\lambda = \begin{cases} 0 & \text{dla } t \geqslant 0, \\ -t & \text{dla } t < 0. \end{cases}$$

89. (NPP) Dla problemu

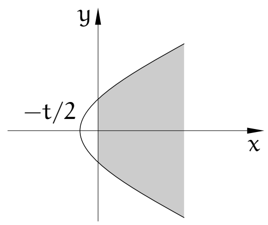
$$\begin{cases} e^{-y} \to \min, \\ \sqrt{x^2+y^2} - x \leqslant 0, \end{cases}$$

znajdź funkcję perturbacji.

<u>Rozw.</u> Ograniczenie oryginalne zastępujemy zaburzonym: $\sqrt{x^2+y^2} - x \leqslant t$.

1. Dla $t = 0$ jest $0 \leqslant \sqrt{x^2+y^2} - x \leqslant 0$, stąd $y = 0$ i $e^{-y} = 1$. $M(0) = 1$.

2. Dla $t > 0$ jest $\sqrt{x^2+y^2} - x \leqslant t$, stąd $x^2 + y^2 \leqslant t^2 + 2t^x + x^2$, czyli $y^2 \leqslant t^2 + 2tx$. Istnieje ciąg punktów dopuszczalnych (x_n, y_n), taki że $y_n \to +\infty$, zatem $e^{-y_n} \to 0$, czyli $\inf_{y\in D} e^{-y} = 0$. Dla $t > 0$ jest $M(t) = 0$, funkcja M jest nieciągła:



$$M(t) = \begin{cases} 1 & \text{dla } t = 0, \\ 0 & \text{dla } t > 0. \end{cases}$$

90. (NPP) Niech S oznacza zwarty zbiór wypukły i niech funkcja $f\colon S \to \mathbb{R}$ będzie ściśle wypukła, klasy C^2 i taka, że istnieje funkcja $\alpha(\mathbf{x}) > 0$, dla której $\mathbf{d}^T D^2 f(\mathbf{x})\mathbf{d} \geqslant \alpha(\mathbf{x})\|\mathbf{d}\|_2^2$ dla każdego $\mathbf{d}$. Udowodnij, że istnieje $m > 0$, takie że

$$\forall_{\mathbf{x},\mathbf{y}\in S} f(\mathbf{y}) \geqslant f(\mathbf{x}) - \frac{1}{m}\|Df(\mathbf{x})\|^2.$$

Jeśli $\mathbf{y}$ jest punktem minimalnym funkcji f w zbiorze S i $Df(\mathbf{y}) = \mathbf{0}$, to daje to warunek stopu algorytmu gradientowego lub Newtona.

<u>Rozw.</u> Ponieważ zbiór S jest zwarty, $\min_{\mathbf{x}\in S} \alpha(\mathbf{x}) = a > 0$. Wtedy

$$\begin{aligned} f(\mathbf{y}) &\geqslant f(\mathbf{x}) + Df(\mathbf{x})(\mathbf{y}-\mathbf{x}) + \frac{a}{2}\|\mathbf{y}-\mathbf{x}\|^2 \\ &\geqslant \inf_{\mathbf{z}\in\mathbb{R}^n} \big(f(\mathbf{x}) + Df(\mathbf{x})(\mathbf{z}-\mathbf{x}) + \frac{a}{2}\|\mathbf{z}-\mathbf{x}\|^2\big) = \inf_{\mathbf{z}\in\mathbb{R}^n} g(\mathbf{z}). \end{aligned}$$

Rozwiązujemy zdanie minimalizacji funkcji $g(\mathbf{z})$. Funkcja ta jest klasy C^1 i wypukła. Jest

$$D_{\mathbf{z}} g = Df(\mathbf{x}) + a(\mathbf{z}-\mathbf{x}) = \mathbf{0},$$

zatem punkt minimalny $\mathbf{z}$ funkcji g spełnia równość $\mathbf{z} - \mathbf{x} = -\frac{1}{a}Df(\mathbf{x})^T$. Stąd

$$f(\mathbf{y}) \geqslant f(\mathbf{x}) - \frac{1}{a}\|Df(\mathbf{x})\|^2 + \frac{a}{2}\frac{\|Df(\mathbf{x})\|^2}{a^2} = f(\mathbf{x}) - \frac{1}{2a}\|Df(\mathbf{x})\|^2$$

Jeśli $Df(\mathbf{y}) = \mathbf{0}$, to można dostać lepsze oszacowanie

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{y}) + \frac{1}{2}(\mathbf{x}-\mathbf{y})^T D^2 f(\tilde{\mathbf{y}})(\mathbf{x}-\mathbf{y}), \\ f(\mathbf{y}) &= f(\mathbf{x}) + Df(\mathbf{x})(\mathbf{y}-\mathbf{x}) + \frac{1}{2}(\mathbf{y}-\mathbf{x})^T D^2 f(\tilde{\mathbf{x}})(\mathbf{y}-\mathbf{x}). \end{aligned}$$

Podstawiając pierwsze do drugiego, dostajemy

$$0 = Df(\mathbf{x})(\mathbf{y}-\mathbf{x}) + \frac{1}{2}(\mathbf{y}-\mathbf{x})^T\big(D^2 f(\tilde{\mathbf{x}}) + D^2 f(\tilde{\mathbf{y}})\big)(\mathbf{y}-\mathbf{x}).$$

Stąd

$$Df(\mathbf{x})(\mathbf{x}-\mathbf{y}) = \frac{1}{2}(\mathbf{y}-\mathbf{x})^T\big(D^2 f(\tilde{\mathbf{x}}) + D^2 f(\tilde{\mathbf{y}})\big)(\mathbf{y}-\mathbf{x})$$

i dalej

$$\|\mathbf{y}-\mathbf{x}\|\|Df(\mathbf{x})\| \geqslant a\|\mathbf{y}-\mathbf{x}\|^2,$$

zatem

$$\frac{\|Df(\mathbf{x})\|}{\|\mathbf{y}-\mathbf{x}\|} \geqslant a.$$

Korzystając z wypukłości, $f(\mathbf{y}) \geqslant f(\mathbf{x}) + Df(\mathbf{x})(\mathbf{y}-\mathbf{x})$,

$$f(\mathbf{x}) - f(\mathbf{y}) \leqslant Df(\mathbf{x})(\mathbf{x}-\mathbf{y}) \leqslant \|\mathbf{x}-\mathbf{y}\|\|Df(\mathbf{x})\| \leqslant \frac{\|Df(\mathbf{x})\|^2}{a}.$$

Czyli $f(\mathbf{y}) \geqslant f(\mathbf{x}) - \frac{\|Df(\mathbf{x})\|^2}{a}$ — wystarczy przyjąć $m = a$.

91. (NPP) Udowodnij, że jeśli $\mathbf{d}$ jest kierunkiem spadku funkcji f w punkcie $\mathbf{x}$, to (przy założeniu, że Df spełnia warunek Lipschitza) istnieje $\delta > 0$, taka że dla $\alpha \in (0, \delta)$ ma miejsce nierówność $f(\mathbf{x} + \alpha\mathbf{d}) < f(\mathbf{x})$.

<u>Rozw.</u> Jest $Df(\mathbf{x})\mathbf{d} < 0$. Niech $g(t) = f(\mathbf{x} + t\mathbf{d})$. Wtedy

$$g'(t) - g'(0) = \big(Df(\mathbf{x}+t\mathbf{d}) - f(\mathbf{x})\big)\mathbf{d} \leqslant \|Df(\mathbf{x}+t\mathbf{d}) - Df(\mathbf{x})\|\|\mathbf{d}\| \leqslant Lt\|\mathbf{d}$$

Stąd

$$\begin{aligned} g(t) = g(0) + \int_0^t g'(t)\,du \leqslant g(0) + \int_0^t g'(0) + L\|\mathbf{d}\|^2 u\,du = \\ g(0) + g'(0)t + \frac{1}{2}L\|\mathbf{d}\|^2 t^2 \equiv g_2(t). \end{aligned}$$

Niech

$$\tau = \operatorname*{arg\,min}_{t\geqslant 0} g_2(t) = -\frac{g'(0)}{L\|\mathbf{d}\|^2},$$

bo $g_2(\tau) = 0 \Rightarrow g'(0) + L\|\mathbf{d}\|^2 t = 0$. Jest $\tau > 0$, bo $g'(0) < 0$. Stąd

$$g(\tau) \leqslant g_2(\tau) = g(0) - \frac{(g'(0))^2}{2L\|\mathbf{d}\|^2},$$

czyli $f(x) + \tau\mathbf{d}) \leqslant f(\mathbf{x}) - \frac{(Df(x)\mathbf{d})^2}{2L\|\mathbf{d}\|^2} < f(\mathbf{x})$. Ponieważ w przedziale $(0, \tau)$ funkcja $g_2(t)$ nie ma ekstremów, dla $t \searrow 0$ wartość funkcji $g_2(t)$ rośnie do $g_2(0) = g(0) = f(\mathbf{x})$. Można zatem przyjąć $\delta = \tau$.

92. (NPP) Użyj metody Newtona do znalezienia minimum funkcji

$$f(x) = \begin{cases} 4x^3 - 3x^4 & x \geqslant 0, \\ 4x^3 + 3x^4 & x < 0, \end{cases}$$

startując z punktów $x = 0.4$ i $x = 0.6$ i wyjaśnij, co się dzieje.

<u>Rozw.</u> Funkcja f jest klasy C^3. Metoda Newtona:

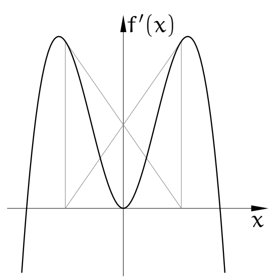
$$x_{k+1} = x_k - \frac{f'(x_k)}{f''(x_k)}$$

w tym przypadku prowadzi do wzoru

$$x_{k+1} = x_k - \begin{cases} \frac{x(1-x)}{2-3x} & x \geqslant 0, \\ \frac{x(1+x)}{2+3x} & x < 0. \end{cases}$$

Wyniki są następujące:

k	x_k	x_k
0	0.40000000	0.6000000
1	0.10000000	−0.6000000
2	0.04705882	0.6000000
3	0.02293373	−0.6000000
4	0.01133069	0.6000000
5	0.00563269	−0.6000000
6	0.00280835	0.6000000
7	0.00140219	−0.6000000
8	0.00070060	0.6000000



Wolna zbieżność metody Newtona wynika z tego, że pochodna drugiego rzędu w punkcie $x = 0$ jest równa 0. Punkt początkowy 0.6 jest poza kulą zbieżności.

93. Zbadaj inny przykład: $f(x) = |x|^{3/2}$ i przekonaj się, że jest on perfidniejszy (przyjmij dowolny punkt startowy $x_0 \neq 0$).

94. Niech H będzie macierzą symetryczną $n \times n$. Wykaż, że wektory własne przynależne do różnych wartości własnych są a) prostopadłe do siebie, b) sprzężone względem macierzy H.

Wykaż, że istnieje baza ortogonalna przestrzeni $\mathbb{R}^n$ zóżona z wektorów własnych dowolnej macierzy symetrycznej.

95. (NPP) Dany jest układ równań nieliniowych $g_i(\mathbf{x}) = 0$, $i = 1, \ldots, n$. Pokaż, jak można znaleźć rozwiązanie za pomocą metody optymalizacji.

<u>Rozw.</u> Należy znaleźć minimum funkcji

$$f(\mathbf{x}) = \sum_{i=1}^{n} |g_i(\mathbf{x})|^p$$

dla $p \geqslant 1$.

Wykonaj obliczenie dla przykładu

$$\begin{cases} 2(x-4)^4 + (2x-y)^2 - 4 &= 0, \\ x^2 - y + 1 &= 0. \end{cases}$$

96. (NPP) Niech $f(x, y) = \frac{1}{2}(x^2 + \alpha y^2)$, $\alpha > 1$. Zbadaj zbieżność metody najszybszego spadku użytej do znajdowania minimum funkcji f.

<u>Rozw.</u> Jest

$$Df = [x, \alpha y], \qquad D^2 f = \begin{bmatrix} 1 & 0 \\ 0 & \alpha \end{bmatrix}.$$

Wskaźnik uwarunkowania macierzy symetrycznej A w normie drugiej, $\mathrm{cond}_2 A = \max_i |\lambda_i| / \min_i |\lambda_i|$, dla hesjanu funkcji f jest równy α. Poziomice funkcji f są elipsami; iloraz długości półosi poziomej i pionowej każdej z nich jest równy $\sqrt{\alpha}$.

W metodzie najszybszego spadku przyjmujemy punkt początkowy $\mathbf{x}_0 = (x_0, y_0)$ — przypuśćmy, że obie jego współrzędne są różne od zera — i konstruujemy kolejne punkty, $\mathbf{x}_i = (x_i, y_i)$. Wektor o kierunku najszybszego spadku $\mathbf{d}_i = -Df(x_i, y_i)^T = [-x_i, -\alpha y_i]^T$ używamy do określenia funkcji

$$g_i(t) = f(\mathbf{x}_i + t\mathbf{d}_i) = \frac{1}{2}\big((x_i - tx_i)^2 + \alpha(y_i - t\alpha y_i)^2\big) = \\ \frac{1}{2}\big((x_i^2 + \alpha^3 y_i^2)t^2 - 2(x_i^2 + \alpha^2 y_i^2)t + (x_i^2 + \alpha y_i^2)\big).$$

Trójmian kwadratowy $at^2 - 2bt + c$ ze współczynnikiem $a > 0$ osiąga minimum dla $t = b/a$, zatem funkcja g_i ma minimum dla

$$t_1 = \frac{x_i^2 + \alpha^2 y_i^2}{x_i^2 + \alpha^3 y_i^2} = \frac{k^2 + \alpha^2}{k^2 + \alpha^3},$$

gdzie $k = x_i/y_i$. Na tej podstawie przyjmujemy

$$(x_{i+1}, y_{i+1}) = y_i \cdot (k - kt_1, 1 - \alpha t_1) = y_i \cdot \left(k\alpha^2 \frac{\alpha - 1}{k^2 + \alpha^3}, k^2 \frac{1 - \alpha}{k^2 + \alpha^3}\right),$$

a stąd $x_{i+1}/y_{i+1} = -\alpha^2/k$. Punkt (x_{i+2}, y_{i+2}) otrzymujemy, znajdując minimum funkcji

$$g_{i+1}(t) = f(\mathbf{x}_{i+1} + t\mathbf{d}_{i+1}),$$

gdzie $\mathbf{d}_{i+1} = [-x_{i+1}, -\alpha y_{i+1}]^T$ jest wektorem najszybszego spadku w punkcie $\mathbf{x}_i$. Mamy stąd

$$t_2 = \frac{x_{i+1}^2 + \alpha^2 y_{i+1}^2}{x_{i+1}^2 + \alpha^3 y_{i+1}^2} = \frac{(-\alpha^2/k)^2 + \alpha^2}{(-\alpha^2/k)^2 + \alpha^3} = \frac{\alpha^2 + k^2}{\alpha^2 + k^2\alpha}$$

i dalej

$$(x_{i+2}, y_{i+2}) = y_{i+1} \cdot \left((-\alpha^2/k) - (-\alpha^2/k)t_2, 1 - \alpha t_2\right) = \\ y_{i+1} \cdot \left(-\alpha k \frac{\alpha - 1}{\alpha + k^2}, \alpha \frac{1 - \alpha}{\alpha + k^2}\right)$$

Stąd wynika, że $x_{i+2}/y_{i+2} = k$. Ale obliczmy dalej

$$(x_{i+2}, y_{i+2}) = y_i k^2 \frac{1 - \alpha}{k^2 + \alpha^3}\left(-\alpha k \frac{\alpha - 1}{\alpha + k^2}, \alpha \frac{1 - \alpha}{\alpha + k^2}\right) = \\ \frac{k^2\alpha(\alpha - 1)^2}{(k^2 + \alpha^3)(k^2 + \alpha)}(x_i, y_i) = c(k)(x_i, y_i).$$

Rozważamy najgorszy możliwy wybór k: taki, dla którego czynnik $c(k)$ jest największy. Znajdziemy maksimum przez obliczenie miejsca zerowego pochodnej tego czynnika względem k; przyrównując licznik wyrażenia opisującego pochodną do zera, dostajemy

$$\begin{aligned} 2k\alpha(\alpha - 1)^2(k^2 + \alpha^3)(k^2 + \alpha) &= k^2\alpha(\alpha - 1)^2 \cdot 2k(2k^2 + \alpha + \alpha^3), \\ (k^2 + \alpha^3)(k^2 + \alpha) &= k^2(2k^2 + \alpha + \alpha^3), \\ k^4 + k^2\alpha^3 + k^2\alpha + \alpha^4 &= 2k^4 + k^2\alpha + k^2\alpha^3, \\ \alpha^4 &= k^4, \end{aligned}$$

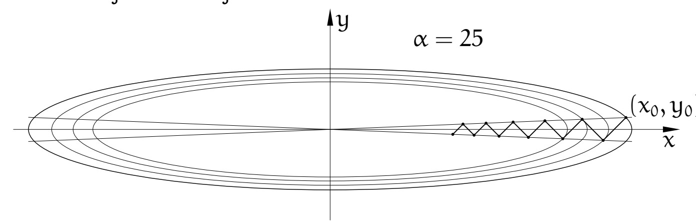
stąd $|k| = \alpha$ i wtedy

$$c(\pm\alpha) = \frac{\alpha^3(\alpha - 1)^2}{\alpha^5 + 2\alpha^4 + \alpha^3} = \frac{(\alpha - 1)^2}{(\alpha + 1)^2}.$$

Z tak wybranym punktem początkowym dla każdego $i \in \mathbb{N}$ mamy

$$(x_{i+1}, y_{i+1}) = \left(\frac{\alpha - 1}{\alpha + 1}x_i, -\frac{\alpha - 1}{\alpha + 1}y_i\right) \quad \Rightarrow \quad \|\mathbf{x}_{i+1}\| = \frac{\alpha - 1}{\alpha + 1}\|\mathbf{x}_i\|.$$

Dla $\alpha \to \infty$ ten ułamek dąży do 1, co oznacza bardzo powolną zbieżność ciągu $(\mathbf{x}_i)_i$ do $\mathbf{0}$. Przypomnijmy, że w tym zadaniu liczba α jest wskaźnikiem uwarunkowania hesjanu funkcji f.



97. Znajdź optymalny kształt beczki, określony przez 3 parametry: wysokość h, promień dna i pokrywy r_0 i promień w środku wysokości r_1; beczka jest bryłą obrotową, tworząca jej bocznej powierzchni jest parabolą. Optymalizacja polega na takim dobraniu (dodatnich) parametrów r_0, r_1, h, aby przy ustalonej pojemności V powierzchnia beczki B była jak najmniejsza.

Rozw. Tworząca beczki jest wykresem funkcji

$$R(x) = r_1 - \frac{4(r_1 - r_0)}{h^2}x^2, \quad x \in [-h/2, h/2].$$

Objętość bryły obrotowej położonej między płaszczyznami $x = -h/2$ i $x = h/2$, której tworząca jest opisana równaniem $y = R(x)$ (gdzie $R(x) \geqslant 0$) jest całką

$$V = \int_{-h/2}^{h/2} \pi R(x)^2\,dx.$$

Pole powierzchni obrotowej, której tworząca jest opisana przez funkcję różniczkowalną $R(x)$ jest równe

$$A = \int_{-h/2}^{h/2} 2\pi R(x)\sqrt{1 + R'(x)^2}\,dx.$$

Dla beczki, z uwagi na parzystość funkcji $R(x)$, mamy

$$V = 2\pi \int_0^{h/2} \left(r_1 - \frac{4(r_1 - r_0)}{h^2}x^2\right)^2 dx,$$

$$B = 2\pi r_0^2 + 4\pi \int_0^{h/2} \left(r_1 - \frac{4(r_1 - r_0)}{h^2}x^2\right)\sqrt{1 + \frac{64(r_1 - r_0)^2}{h^4}x^2}\,dx$$

(pierwszy składnik we wzorze na pole powierzchni beczki jest polem dna i pokrywy). Pierwszą z tych całek można obliczyć dosyć łatwo, bo funkcja podcałkowa jest wielomianem:

$$V = \frac{\pi}{15}\left(3r_0^2 + 4r_0r_1 + 8r_1^2\right)h.$$

Trudniejsza sprawa jest z polem powierzchni; wprawdzie można znaleźć wzór opisujący funkcję pierwotną funkcji podcałkowej, ale jest on zbyt skomplikowany. Dlatego posłużymy się wzorem przybliżonym, czyli kwadraturą:

$$B \approx f(r_0, r_1, h) = 2\pi r_0^2 + 4\pi \frac{h}{2} \sum_{i=1}^{n} A_i R_i \sqrt{1 + R_i'^2},$$

gdzie $A_1, \dots A_n$ to współczynniki kwadratury, zaś $R_i = R(x_i)$ i $R_i' = R'(x_i)$ to wartości funkcji opisującej tworzącą i jej pochodną w węzłach kwadratury $x_1, \dots, x_n$ (należących do przedziału całkowania $[0, h/2]$ — uzyskujemy je dla ustalonych liczb $t_1, \dots, t_n \in [0, 1]$, biorąc $x_1 = t_i h/2$. Liczby $A_1, \dots, A_n$ są współczynnikami kwadratury z węzłami $t_1, \dots, t_n$ odpowiedniej dla przedziału całkowania $[0, 1]$; konkretną kwadraturę wybierzemy później). Mamy

$$R_i = R_i(r_0, r_1, h) = r_1 - \frac{4(r_1 - r_0)}{h^2} x_i^2 = r_1 - (r_1 - r_0)t_i^2,$$
$$R_i' = R_i'(r_0, r_1, h) = -\frac{8(r_1 - r_0)}{h^2} x_i = -\frac{4(r_1 - r_0)t_i}{h}.$$

Teraz dla problemu optymalizacyjnego, w którym dana jest pojemność $V > 0$,

$$\begin{cases} f(r_0, r_1, h) \to \min, \\ g(r_0, r_1, h) - V = 0, \\ r_0, r_1, h > 0, \end{cases}$$

i w którym $g(r_0, r_1, h) = \frac{\pi}{15}(3r_0^2 + 4r_0 r_1 + 8r_1^2)h$, określamy funkcję Lagrange'a

$$L(r_0, r_1, h; \lambda) = f(r_0, r_1, h) + \lambda\big(g(r_0, r_1, h) - V\big)$$

(tu pomijamy ograniczenia nierównościowe) i piszemy układ równań

$$\begin{cases} Df(r_0, r_1, h) + \lambda Dg(r_0, r_1, h) = [0, 0, 0], \\ g(r_0, r_1, h) - V = 0. \end{cases} \qquad (*)$$

Możemy łatwo obliczyć

$$Dg(r_0, r_1, r_2) = \frac{\pi}{15}\big[(6r_0 + 4r_1)h,\ (4r_0 + 16r_1)h,\ 3r_0^2 + 4r_0 r_1 + 8r_1^2\big].$$

Gradient funkcji f ma współrzędne

$$\frac{\partial f}{\partial r_0} = 4\pi r_0 + 2\pi h \sum_{i=1}^{n} A_i \frac{\partial}{\partial r_0}\Big(R_i \sqrt{1 + R_i'^2}\Big),$$

$$\frac{\partial f}{\partial r_1} = 2\pi h \sum_{i=1}^{n} A_i \frac{\partial}{\partial r_1}\Big(R_i \sqrt{1 + R_i'^2}\Big),$$
$$\frac{\partial f}{\partial h} = 2\pi \sum_{i=1}^{n} A_i \bigg(R_i \sqrt{1 + R_i'^2} + h\frac{\partial}{\partial h}\Big(R_i \sqrt{1 + R_i'^2}\Big)\bigg).$$

Potrzebne wyżej pochodne są opisane wzorami

$$\frac{\partial}{\partial s}\Big(R_i \sqrt{1 + R_i'^2}\Big) = \sqrt{1 + R_i'^2}\,\frac{\partial R_i}{\partial s} + \frac{R_i R_i'}{\sqrt{1 + R_i'^2}} \frac{\partial R_i'}{\partial s},$$

przy czym za s podstawiamy kolejno r_0, r_1, h. Z kolei, do powyższych wzorów podstawimy

$$\frac{\partial R_i}{\partial r_0} = t_i^2, \qquad \frac{\partial R_i}{\partial r_1} = 1 - t_i^2, \qquad \frac{\partial R_i}{\partial h} = 0,$$
$$\frac{\partial R_i'}{\partial r_0} = \frac{4t_i}{h}, \qquad \frac{\partial R_i'}{\partial r_1} = -\frac{4t_i}{h}, \qquad \frac{\partial R_i'}{\partial h} = \frac{4(r_1 - r_0)t_i}{h^2}.$$

Układ (*) zapiszemy w postaci

$$F(\boldsymbol{x}) = \boldsymbol{0},$$

z funkcją F, której argumentem jest wektor $\boldsymbol{x} = (r_0, r_1, h, \lambda)$ i której wartość (w $\mathbb{R}^4$) jest transpozycją gradientu funkcji Lagrange'a; zatem

$$F(r_0, r_1, h, \lambda) = \begin{bmatrix} \big(Df(r_0, r_1, h) + \lambda Dg(r_0, r_1, h)\big)^T \\ g(r_0, r_1, h) - V \end{bmatrix}.$$

Rozwiązanie chcemy znaleźć metodą Newtona, do niej potrzebujemy pochodnej funkcji F, czyli hesjanu funkcji Lagrange'a L (argumenty r_0, r_1, h funkcji dla skrótu są pominięte):

$$DF = D^2L = \begin{bmatrix} D^2f + \lambda D^2g & (Dg)^T \\ Dg & 0 \end{bmatrix}$$

Hesjan funkcji g jest łatwy do znalezienia:

$$D^2g(r_0, r_1, h) = \frac{\pi}{15}\begin{bmatrix} 6h & 4h & 6r_0 + 4r_1 \\ 4h & 16h & 4r_0 + 16r_1 \\ 6r_0 + 4r_1 & 4r_0 + 16r_1 & 0 \end{bmatrix}.$$

Z funkcją f jest większy kłopot, ale pracowicie piszemy

$$\frac{\partial^2 f}{\partial r_0^2} = 4\pi + 2\pi h \sum_{i=1}^{n} A_i \frac{\partial^2}{\partial r_0^2}\Big(R_i \sqrt{1 + R_i'^2}\Big),$$

$$\frac{\partial^2 f}{\partial r_0 \partial r_1} = 2\pi h \sum_{i=1}^{n} A_i \frac{\partial^2}{\partial r_0 \partial r_1}\Bigl(R_i\sqrt{1+R_i'^2}\Bigr),$$

$$\frac{\partial^2 f}{\partial r_0 \partial h} = 2\pi \sum_{i=1}^{n} A_i \biggl(\frac{\partial}{\partial r_0}\Bigl(R_i\sqrt{1+R_i'^2}\Bigr) + h\frac{\partial^2}{\partial r_0 \partial h}\Bigl(R_i\sqrt{1+R_i'^2}\Bigr)\biggr),$$

$$\frac{\partial^2 f}{\partial r_1^2} = 2\pi h \sum_{i=1}^{n} A_i \frac{\partial^2}{\partial r_1^2}\Bigl(R_i\sqrt{1+R_i'^2}\Bigr),$$

$$\frac{\partial^2 f}{\partial r_1 \partial h} = 2\pi \sum_{i=1}^{n} A_i \biggl(\frac{\partial}{\partial r_1}\Bigl(R_i\sqrt{1+R_i'^2}\Bigr) + h\frac{\partial^2}{\partial r_1 \partial h}\Bigl(R_i\sqrt{1+R_i'^2}\Bigr)\biggr),$$

$$\frac{\partial^2 f}{\partial h^2} = 2\pi \sum_{i=1}^{n} A_i \biggl(2\frac{\partial}{\partial h}\Bigl(R_i\sqrt{1+r_i'^2}\Bigr) + h\frac{\partial^2}{\partial h^2}\Bigl(R_i\sqrt{1+r_i'^2}\Bigr)\biggr).$$

Tu potrzebujemy pochodnych drugiego rzędu odpowiednich wyrażeń, mamy

$$\frac{\partial^2}{\partial s \partial u}\Bigl(R_i\sqrt{1+R_i'^2}\Bigr) = \frac{1}{\sqrt{1+R_i'^2}}\biggl(R_i'\Bigl(\frac{\partial R_i}{\partial s}\frac{\partial R_i'}{\partial u} + \frac{\partial R_i}{\partial u}\frac{\partial R_i'}{\partial s}\Bigr) + \frac{R_i}{1+R_i'^2}\frac{\partial R_i'}{\partial s}\frac{\partial R_i'}{\partial u} + (1+R_i'^2)\frac{\partial^2 R_i}{\partial s \partial u} + R_i R_i'\frac{\partial^2 R_i'}{\partial s \partial u}\biggr). \qquad (\square)$$

Za symbole s i u podstawimy zmienne r_0, r_1 i h, potrzebujemy w sumie 6 takich par. Do powyższego wzoru mamy podstawiać pochodne pierwszego rzędu funkcji R_i i R_i' znalezione wcześniej oraz pochodne drugiego rzędu:

$$\frac{\partial^2 R_i}{\partial s \partial u} = 0, \qquad \frac{\partial^2 R_i'}{\partial r_0^2} = \frac{\partial^2 R_i'}{\partial r_0 \partial r_1} = \frac{\partial^2 R_i'}{\partial r_1^2} = 0,$$

$$\frac{\partial^2 R_i'}{\partial r_0 \partial h} = -\frac{4t_i}{h^2}, \qquad \frac{\partial^2 R_i'}{\partial r_1 \partial h} = \frac{4t_i}{h^2}, \qquad \frac{\partial^2 R_i'}{\partial h^2} = \frac{8(r_0 - r_1)t_i}{h^3}.$$

Wzór $(\square)$ możemy uprościć, korzystając z tego, że wszystkie pochodne drugiego rzędu funkcji R_i są równe 0:

$$\frac{\partial^2}{\partial s \partial u}\Bigl(R_i\sqrt{1+R_i'^2}\Bigr) = \frac{1}{\sqrt{B_i}}\biggl(R_i'\Bigl(\frac{\partial R_i}{\partial s}\frac{\partial R_i'}{\partial u} + \frac{\partial R_i}{\partial u}\frac{\partial R_i'}{\partial s}\Bigr) + \frac{R_i}{B_i}\frac{\partial R_i'}{\partial s}\frac{\partial R_i'}{\partial u} + R_i R_i'\frac{\partial^2 R_i'}{\partial s \partial u}\biggr). \qquad (\square')$$

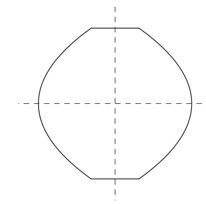
gdzie $B_i = 1 + R_i'^2$. W niektórych spośród sześciu przypadków szczególnych tego wzoru nastąpią dalsze uproszczenia, ponieważ pewne pochodne funkcji R_i' też są zerowe, a także $\frac{\partial R_i}{\partial h} = 0$.

Ponieważ wszystkie funkcje, które mamy scałkować, są gładkie (klasy C^∞), można użyć dosyć dowolnej kwadratury. Do obliczeń numerycznych niżej

wybrałem złożoną kwadraturę Simpsona. Po ustaleniu nieparzystej liczby węzłów $n \geqslant 5$, przyjmuję $x_i = (i-1)/(n-1)$ dla $i = 1, \ldots, n$ oraz $A_1 = A_n = \frac{1}{3(n-1)}$, $A_i = \frac{4}{3(n-1)}$ dla i parzystego i $A_i = \frac{2}{3(n-1)}$ dla i nieparzystego innego niż 1 i n.

W tym eksperymencie przyjąłem $n = 81$ (chwilowo, do zrobienia pozostaje oszacowanie błędu aproksymacji kwadratury i wnikliwsze dobranie n). Startując z punktu początkowego $\boldsymbol{x}_0 = (0.2, 0.6, 1.2, -3)$, metoda Newtona wygenerowała ciąg pokazany w tabelce:

krok	r_0	r_1	h	λ
0	0.200000	0.600000	1.200000	−3.000000
1	0.203561	0.635486	1.240343	−3.266381
2	0.197854	0.633788	1.245819	−3.243044
3	0.197649	0.633769	1.246140	−3.242951
4	0.197649	0.633769	1.246140	−3.242951



otrzymując rozwiązanie z dobrą dokładnością (residuum mniejsze niż 10^{-10}) po pięciu krokach.

Na podstawie kryterium drugiego rzędu możemy sprawdzić, że to jest minimum lokalne: w tym punkcie hesjan funkcji Lagrange'a i gradient funkcji ograniczenia są równe

$$D_{xx}^2 L = \begin{bmatrix} 0.7280431 & -2.0152708 & 0.1050026 \\ -2.0152708 & -0.7798431 & -1.6469667 \\ 0.1050026 & -1.6469667 & 0.1562213 \end{bmatrix},$$

$$Dg = [0.3091230, 0.9080949, 0.2554367].$$

Tworzymy macierz Y, której kolumny $\boldsymbol{y}_1$, $\boldsymbol{y}_2$ są prostopadłe do Dg, np. $\boldsymbol{y}_1 = [0.9080949, -0.3091230, 0]^T$ i $\boldsymbol{y}_2 = [0, 0.2554367, -0.9080949]^T$, a następnie obliczamy macierz

$$B = Y^T D_{xx}^2 L Y = \begin{bmatrix} 1.6572768 & -0.9548002 \\ -0.9548002 & 0.8420059 \end{bmatrix}$$

i sprawdzamy, że jest to macierz dodatnio określona (na przykład, znajdując jej wartości własne, 0.2114650 i 2.2878176).

Obliczone pole powierzchni beczki jest równe 4.8644260, przy pojemności $V = 1$. Przekrój optymalnej beczki jest na rysunku wyżej.

<u>Uwaga:</u> Otrzymane rozwiązanie odpowiada pewnemu minimum lokalnemu. Póki się nie udowodni, że to jest (albo nie jest) minimum globalne, można mieć tylko taką nadzieję. Silną poszlaką jest to, że pole powierzchni kuli o objętości 1 jest równe $\sqrt[3]{36\pi} \approx 4.8359759$, czyli bardzo niewiele mniej nż pole powierzchni znalezionej beczki.