

1 Wstęp

1.1 Notacja

Często będziemy pisać $i \in \mathbf{m}$ zamiast $i = 1, \dots, m$.

1.2 Optymalizacja z ograniczeniami nierównościowymi

Rozpatrujemy zadanie optymalizacji wypukłej

$$\begin{aligned} & \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimize}} && f_0(\mathbf{x}) \\ & \text{subject to} && f_i(\mathbf{x}) \leq 0, \quad i \in \mathbf{m} \end{aligned} \quad (1)$$

Dla zadania tego możemy zdefiniować zbiór dopuszczalny (*feasible set*)

$$\Omega = \{\mathbf{x} \in \mathbb{R}^n \mid f_i(\mathbf{x}) \leq 0, \quad i \in \mathbf{m}\}, \quad (2)$$

który jest zbiorem wypukłym z uwagi to, że funkcje $f_i(\mathbf{x})$, $i \in \mathbf{m}$, są wypukłe. Zadania powyższego typu można rozwiązywać metodami funkcji bariery (*barrier function*). Zastosowanie funkcji bariery umożliwia sprowadzenie zadań optymalizacji z ograniczeniami nierównościowymi do sekwencji zadań bez ograniczeń do rozwiązywania których możemy zastosować poznane wcześniej metody np. metodę Newtona.

Będziemy w dalszym ciągu zakładać, że funkcje ograniczeń nierównościowych $f_i(\mathbf{x})$, $i \in \mathbf{m}$ oraz funkcja celu f_0 są dwukrotnie różniczkowalne. Dodatkowo zakładamy, że wartość optymalna p^* funkcji celu jest skończona i osiągana w punkcie optymalnym \mathbf{x}^* , oraz że zadanie jest ściśle dopuszczalne (*strictly feasible*), tzn. że istnieje pewien punkt (wektor) $\bar{\mathbf{x}} \in \mathbb{R}^n$ taki, że

$$f_i(\bar{\mathbf{x}}) < 0, \quad i \in \mathbf{m}. \quad (3)$$

1.3 Logarytmiczna funkcja bariery

Funkcję ciągłą $\varphi: \mathbb{R}^n \rightarrow \mathbb{R}$ nazywamy wypukłą funkcją barierą (*convex barrier function*) dla zbioru Ω , jeśli jest wypukła na zbiorze Ω oraz $\varphi(\mathbf{x}) \rightarrow \infty$ dla \mathbf{x} dążących do brzegu zbioru Ω . Najczęściej używa się tzw. logarytmicznej funkcji bariery (*logarithmic barrier function*)

$$\varphi(\mathbf{x}) = \begin{cases} -\sum_{i=1}^m \log(-f_i(\mathbf{x})) & \text{dla } \mathbf{x} \in \text{relint } \Omega \\ +\infty & \text{dla } \mathbf{x} \notin \text{relint } \Omega \end{cases}, \quad (4)$$

gdzie

$$\text{relint } \Omega = \{\mathbf{x} \in \mathbb{R}^n \mid f_i(\mathbf{x}) < 0, \quad i \in \mathbf{m}\}. \quad (5)$$

Zauważmy, że dla $i \in \mathbf{m}$, $k \in \mathbf{n}$, mamy

$$\frac{\partial \log(-f_i(\mathbf{x}))}{\partial x_k} = \frac{1}{f_i(\mathbf{x})} \frac{\partial f_i(\mathbf{x})}{\partial x_k}, \quad (6)$$

zatem dla $i \in \mathbf{m}$

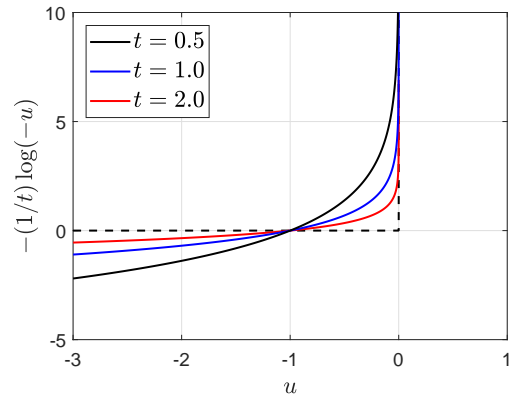
$$\nabla \log(-f_i(\mathbf{x})) = \begin{bmatrix} \frac{1}{f_i(\mathbf{x})} \frac{\partial f_i(\mathbf{x})}{\partial x_1} \\ \frac{1}{f_i(\mathbf{x})} \frac{\partial f_i(\mathbf{x})}{\partial x_2} \\ \vdots \\ \frac{1}{f_i(\mathbf{x})} \frac{\partial f_i(\mathbf{x})}{\partial x_n} \end{bmatrix} = \frac{1}{f_i(\mathbf{x})} \nabla f_i(\mathbf{x}). \quad (7)$$

Gradient funkcji φ jest zatem równy

$$\begin{aligned} \nabla \varphi(\mathbf{x}) &= -\sum_{i=1}^m \nabla \log(-f_i(\mathbf{x})) \\ &= -\sum_{i=1}^m \frac{1}{f_i(\mathbf{x})} \nabla f_i(\mathbf{x}), \end{aligned} \quad (8)$$

zaś jej hesjan jest równy

$$\begin{aligned} \nabla^2 \varphi(\mathbf{x}) &= D \nabla \varphi(\mathbf{x}) \\ &= \sum_{i=1}^m \left(\frac{\nabla f_i(\mathbf{x}) \nabla f_i(\mathbf{x})^T}{(f_i(\mathbf{x}))^2} - \frac{\nabla^2 f_i(\mathbf{x})}{f_i(\mathbf{x})} \right). \end{aligned} \quad (9)$$



Rysunek 1: Wykres funkcji $-(1/t) \log(-u)$ dla różnych wartości parametru t .

Idea rozwiązania zadania (1) polega na zastąpieniu go przybliżeniem

$$\underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimize}} \quad f_0(\mathbf{x}) + \frac{1}{t} \varphi(\mathbf{x}) \quad (10)$$

gdzie $t > 0$ jest parametrem. Zakładamy, że punkt optymalny zadania (10) istnieje i jest jedyny, oznaczmy go \mathbf{x}_t^* , oraz że istnieje punkt startowy \mathbf{x}_0 , który spełnia

$$f_i(\mathbf{x}_0) < 0, \quad i \in \mathbf{m}. \quad (11)$$

Użycie funkcji $\varphi(\mathbf{x})$ zapobiega opuszczeniu przez rozwiązanie obszaru Ω , innymi słowy $\varphi(\mathbf{x})$ jest swego rodzaju barierą zbioru Ω , ponieważ $\varphi(\mathbf{x}) = +\infty$ dla $\mathbf{x} \notin \Omega$ oraz $\varphi(\mathbf{x})$ dąży do $+\infty$ kiedy \mathbf{x} zbliża się do brzegu zbioru Ω . (Mówiąc bardziej precyzyjnie, $\varphi(\mathbf{x}) = +\infty$ dla $\mathbf{x} \notin \text{relint } \Omega$ oraz $\varphi(\mathbf{x})$ dąży do $+\infty$ kiedy \mathbf{x} zbliża się do względnego brzegu zbioru Ω . Więcej informacji na temat względnego wnętrza zbioru i względnego brzegu zbioru można znaleźć w [1]).

Minimalizator \mathbf{x}_t^* (tzn. rozwiązanie) zadania (10) jest ściśle dopuszczalny (*strictly feasible*) dla zadania (1) tzn.

$$f_i(\mathbf{x}_t^*) < 0, \quad i \in \mathbf{m}. \quad (12)$$

Przyjmijmy oznaczenie

$$\psi_t(\mathbf{x}) = f_0(\mathbf{x}) + \frac{1}{t} \varphi(\mathbf{x}). \quad (13)$$

Z warunków optymalności pierwszego rzędu mamy

$$\nabla \psi_t(\mathbf{x}_t^*) = \nabla_{\mathbf{x}} \left[f_0(\mathbf{x}) + \frac{1}{t} \varphi(\mathbf{x}) \right] \Big|_{\mathbf{x}=\mathbf{x}_t^*} = 0, \quad (14)$$

zatem

$$\nabla f_0(\mathbf{x}_t^*) + \sum_{i=1}^m \frac{1}{-tf_i(\mathbf{x}_t^*)} \nabla f_i(\mathbf{x}_t^*) = 0. \quad (15)$$

Zdefiniujmy

$$(\lambda_t^*)_i = \frac{1}{-tf_i(\mathbf{x}_t^*)} > 0, \quad i \in \mathbf{m}. \quad (16)$$

Zauważmy, że lagranżjan dla zadania (1) jest postaci

$$L(\mathbf{x}, \boldsymbol{\lambda}) = f_0(\mathbf{x}) + \sum_{i=1}^m \lambda_i f_i(\mathbf{x}). \quad (17)$$

Zauważmy, że dla

$$\lambda_i = (\lambda_t^*)_i = \frac{1}{-tf_i(\mathbf{x}_t^*)} > 0, \quad i \in \mathbf{m} \quad (18)$$

lagranżjan ten

$$L(\mathbf{x}, \boldsymbol{\lambda}_t^*) = f_0(\mathbf{x}) + \sum_{i=1}^m (\lambda_t^*)_i f_i(\mathbf{x}) \quad (19)$$

osiąga minimum w \mathbf{x}_t^* ponieważ

$$\begin{aligned} \nabla_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}_t^*)|_{\mathbf{x}=\mathbf{x}_t^*} &= \nabla L(\mathbf{x}_t^*, \boldsymbol{\lambda}_t^*) \\ &= \nabla f_0(\mathbf{x}_t^*) + \sum_{i=1}^m (\lambda_t^*)_i \nabla f_i(\mathbf{x}_t^*) \\ &= \nabla f_0(\mathbf{x}_t^*) + \sum_{i=1}^m \frac{1}{-tf_i(\mathbf{x}_t^*)} \nabla f_i(\mathbf{x}_t^*) \\ &= \mathbf{0}. \end{aligned} \quad (20)$$

Biorąc pod uwagę, że dualna funkcja Lagrange'a

$$g(\boldsymbol{\lambda}) = \min_{\mathbf{x}} L(\mathbf{x}, \boldsymbol{\lambda}) \quad \boldsymbol{\lambda} \geq \mathbf{0} \quad (21)$$

jest dolnym ograniczeniem dla wartości optymalnej p^* zadania (1), tzn.

$$g(\boldsymbol{\lambda}) \leq p^*, \quad (22)$$

otrzymujemy dla $\boldsymbol{\lambda} = \boldsymbol{\lambda}_t^*$

$$\begin{aligned} p^* &\geq g(\boldsymbol{\lambda}_t^*) \\ &= L(\mathbf{x}_t^*, \boldsymbol{\lambda}_t^*) \\ &= f_0(\mathbf{x}_t^*) + \sum_{i=1}^m (\lambda_t^*)_i f_i(\mathbf{x}_t^*) \\ &= f_0(\mathbf{x}_t^*) + \sum_{i=1}^m \frac{1}{-tf_i(\mathbf{x}_t^*)} f_i(\mathbf{x}_t^*) \\ &= f_0(\mathbf{x}_t^*) - \frac{m}{t}, \end{aligned} \quad (23)$$

czyli

$$f_0(\mathbf{x}_t^*) - \frac{m}{t} \leq p^* \quad (24)$$

czyli

$$\boxed{f_0(\mathbf{x}_t^*) - p^* \leq \frac{m}{t}} \quad (25)$$

Nierówność (25) stanowi podstawowe uzasadnienie metody funkcji bariery, ponieważ oznacza, że rozwiązanie $\mathbf{x}^*(t)$ zadania bez ograniczeń (10) jest ϵ -suboptymalnym rozwiązaniem zadania z ograniczeniami (1), tzn. dla danego $\epsilon > 0$ zachodzi

$$f_0(\mathbf{x}_t^*) - p^* \leq \epsilon \quad \text{jeśli} \quad \frac{m}{t} \leq \epsilon, \quad (26)$$

skąd otrzymujemy

$$f_0(\mathbf{x}_t^*) \xrightarrow{t \rightarrow \infty} p^*. \quad (27)$$

Przykład 1. Weźmy pod uwagę następujące zadanie optymalizacji

$$\begin{aligned} &\underset{x \in \mathbb{R}}{\text{minimize}} && cx \\ &\text{subject to} && x \geq a, \\ &&& x \leq b. \end{aligned}$$

Zakładamy, że $b > a$, $c \neq 0$. Zadanie to możemy zapisać w postaci standardowej

$$\begin{aligned} &\underset{x \in \mathbb{R}}{\text{minimize}} && f_0(x) \\ &\text{subject to} && f_i(x) \leq 0, \quad i = 1, \dots, m \end{aligned}$$

gdzie $f_0(x) = cx$, $m = 2$, $f_1(x) = a - x$, $f_2(x) = x - b$.

Zdefiniujmy funkcję

$$\psi_t(x) = cx - \frac{1}{t} [\log(x - a) + \log(-x + b)]. \quad (28)$$

Pierwsza i druga pochodna funkcji $\psi_t(x)$ wyrażają się następująco

$$\psi_t'(x) = c - \frac{1}{t} \left[\frac{1}{x - a} + \frac{1}{x - b} \right], \quad (29)$$

$$\psi_t''(x) = \frac{1}{t} \left[\frac{1}{(x - a)^2} + \frac{1}{(x - b)^2} \right]. \quad (30)$$

Zauważmy, że wewnątrz przedziału (a, b) mamy $\psi_t''(x) > 0$ czyli funkcja ψ jest wypukła. Przyrównując do zera pochodną funkcji $\psi_t(x)$ otrzymujemy równanie kwadratowe

$$x^2 - \left(a + b + \frac{2}{ct} \right) x + ab + \frac{a + b}{ct} = 0. \quad (31)$$

Równanie to ma dwa pierwiastki

$$x = \frac{a + b + \frac{2}{ct} \pm \sqrt{(b - a)^2 + \left(\frac{2}{ct} \right)^2}}{2}. \quad (32)$$

Interesuje nas rozwiązanie wewnątrz przedziału (a, b) , zatem mamy

$$x_t^* = \begin{cases} \frac{a + b + \frac{2}{ct} - \sqrt{(b - a)^2 + \left(\frac{2}{ct} \right)^2}}{2} & \text{dla } c > 0 \\ \frac{a + b + \frac{2}{ct} + \sqrt{(b - a)^2 + \left(\frac{2}{ct} \right)^2}}{2} & \text{dla } c < 0. \end{cases} \quad (33)$$

Pamiętając, że $b > a$, dla $c > 0$ mamy

$$\begin{aligned} \lim_{t \rightarrow +\infty} x_t^* &= \lim_{t \rightarrow +\infty} \frac{a + b + \frac{2}{ct} - \sqrt{(b - a)^2 + \left(\frac{2}{ct} \right)^2}}{2} \\ &= \frac{a + b - |b - a|}{2} \\ &= \frac{a + b - (b - a)}{2} \\ &= a, \end{aligned} \quad (34)$$

a dla $c < 0$ mamy

$$\begin{aligned} \lim_{t \rightarrow +\infty} x_t^* &= \lim_{t \rightarrow +\infty} \frac{a + b + \frac{2}{ct} + \sqrt{(b - a)^2 + \left(\frac{2}{ct} \right)^2}}{2} \\ &= \frac{a + b + |b - a|}{2} \\ &= \frac{a + b + (b - a)}{2} \\ &= b. \end{aligned} \quad (35)$$

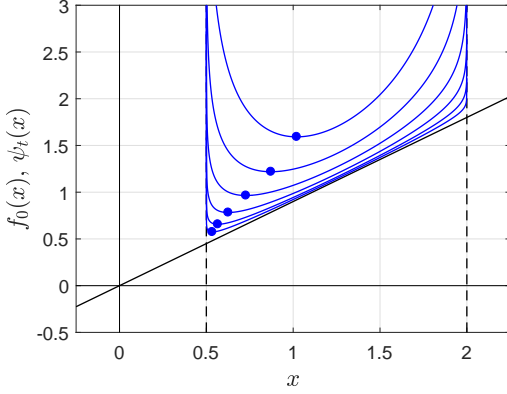
Na Rys. 2-4, dla $a = 0.5$, $b = 2$, $c = 0.9$, przedstawiono wykresy funkcji $f_0(x)$, $\psi_t(x)$, $\psi'_t(x)$ i $\psi''_t(x)$, dla wybranych wartości parametru t (Tab. 1), w Tab. 1 podano również wartości rozwiązań zadania

$$\underset{x \in \mathbb{R}}{\text{minimize}} \quad \psi_t(x) \quad (36)$$

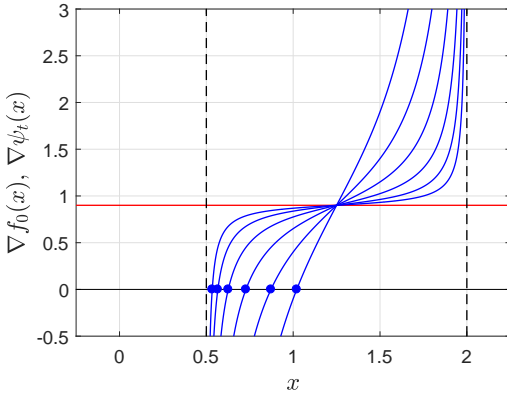
dla tych wartości parametru t .

Tabela 1

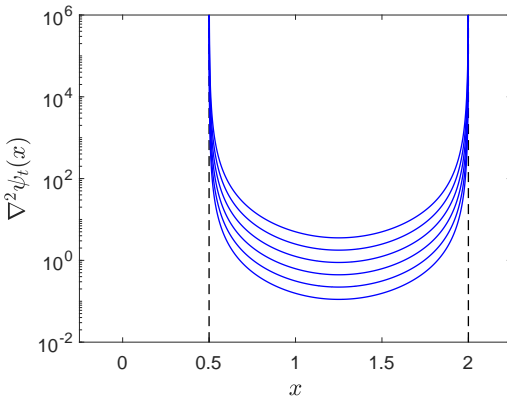
t	1	2	4	8	16	32
x_t^*	1.0206	0.8722	0.7280	0.6261	0.5662	0.5339



Rysunek 2: Wykresy funkcji $f_0(x)$ (kolor czerwony), $\psi_t(x) = cx - \frac{1}{t} [\log(x-a) + \log(-x+b)]$ (kolor niebieski) dla wartości parametru t z Tab. 1 oraz $a = 0.5$, $b = 2$, $c = 0.9$. Niebieskimi kropkami zaznaczono minima funkcji $\psi_t(x)$.



Rysunek 3: Wykres pochodnej (gradientu) $\psi'_t(x)$ (kolor niebieski) oraz pochodnej (gradientu) $f'_0(x)$ (kolor czerwony), dla wartości parametru t z Tab. 1 oraz $a = 0.5$, $b = 2$, $c = 0.9$. Niebieskimi kropkami zaznaczono miejsca zerowe $\psi'_t(x)$ odpowiadające minimum funkcji $\psi_t(x)$.



Rysunek 4: Wykres drugiej pochodnej (hesjanu) $\psi''_t(x) = \frac{1}{t} \left[\frac{1}{(x-a)^2} + \frac{1}{(x-b)^2} \right]$ dla wartości parametru t z Tab. 1 oraz $a = 0.5$, $b = 2$, $c = 0.9$.

1.4 Sekwencyjna metoda bariery

Mogłoby się wydawać, że najprościej przyjąć

$$t \geq \frac{m}{\epsilon} \quad (37)$$

i rozwiązać (10) używając np. metody Newtona. W praktyce takie podejście może okazać się mało skuteczne, nie tylko dla tego, że punkt startowy \mathbf{x}_0 może być położony daleko od punktu optymalnego \mathbf{x}^* , ale przede wszystkim dlatego, że funkcja (13) jest na ogół źle uwarunkowana dla dużych t (tzn. hesjan funkcji ψ_t zmienia się szybko w pobliżu brzegu zbioru Ω), co powoduje, że metoda Newtona może potrzebować bardzo wielu iteracji aby zbiec do \mathbf{x}_t^* . Ponadto, jeśli f_0 jest funkcją lionową, to dla dużych wartości t hesjan

$$\nabla^2 \left(f_0(\mathbf{x}) + \frac{1}{t} \varphi(\mathbf{x}) \right) \quad (38)$$

dla \mathbf{x} , które nie znajdują się blisko brzegu obszaru dopuszczalnego, będzie macierzą bliską osobliwej, co może powodować trudności numeryczne przy stosowaniu metody Newtona (Rys. 4). W związku z tym, w praktyce, rozwiązuje się sekwencję zadań (10) zaczynając od niedużych wartości t i zwiększając je stopniowo, aż do spełnienia warunku

$$\frac{m}{t} \leq \epsilon. \quad (39)$$

Postępowanie takie można przedstawić w postaci algorytmu podanego w ramce na końcu sekcji.

Każdą iterację k w powyższym algorytmie nazywamy krokiem centrującym (*centering step*) lub iteracją zewnętrzną (*outer iteration*), punkt \mathbf{x}_k^* nazywamy k -tym punktem centralnym (*central point*), zaś krzywą, którą tworzą minima \mathbf{x}_k^* funkcji (13) dla kolejnych wartości k , nazywamy ścieżką centralną (*central path*), zawiera się ona w zbiorze dopuszczalnym Ω . Z tego powodu sekwencyjna metoda bariery należy do tzw. metod punktu wewnętrznego (*interior-point methods*).

Każdy krok centrujący wymaga pewnej liczby iteracji wewnętrznych, które są iteracjami algorytmu Newtona, potrzebnymi do wyznaczenia \mathbf{x}_k^* z zadaną dokładnością.

Sekwencyjna metoda bariery (ang. *sequential barrier method*)

dane: ściśle dopuszczalny \mathbf{x}_0 , t_{init} , $\gamma > 1$, $\epsilon > 0$

1. Podstawiamy $k = 0$, $t = t_{\text{init}}$.
2. Używając metody Newtona z tłumieniem rozwiązujemy zadanie

$$\underset{\mathbf{x}}{\text{minimize}} \quad f_0(\mathbf{x}) + \frac{1}{t} \varphi(\mathbf{x})$$

dla punktu startowego \mathbf{x}_k , oznaczamy rozwiązanie optymalne tego zadania przez \mathbf{x}_k^* .

3. Aktualizujemy $\mathbf{x}_{k+1} \leftarrow \mathbf{x}_k^*$.
4. Jeśli $m/t \leq \epsilon$, zwracamy \mathbf{x}_k^* i kończymy algorytm.
5. Aktualizujemy $t \leftarrow \gamma t$, $k \leftarrow k + 1$ i przechodzimy punktu 2.

1.4.1 Liczba iteracji metody SBM

Jak już wspomniano, podstawienie $t_{\text{init}} \geq m/\epsilon$ zagwarantowałyby zbieżność omówionej metody w jednym kroku centrującym (iteracji zewnętrznej), jednak mogłoby znacznie zwiększyć liczbę iteracji wewnętrznych, dlatego zwiększa się stopniowo wartość t podstawiając $t_{k+1} = \gamma t_k$ gdzie t_k oznacza wartość t w k -tym kroku centrującym. Liczba iteracji zewnętrznych (kroków centrujących) potrzebnych do znalezienia rozwiązania zadaną dokładnością ϵ wynosi

$$\left\lceil \frac{\log(m\epsilon^{-1}/t_{\text{init}})}{\log \gamma} + 1 \right\rceil \quad (40)$$

gdzie $\lceil x \rceil$ oznacza funkcję sufit (ang. *ceiling*), tzn. najmniejszą liczbę całkowitą nie mniejszą od x . Wyprowadzenie wzoru (40) można znaleźć w [1].

1.5 Wyznaczanie ściśle dopuszczalnego punktu początkowego

Metoda bariery wymaga znajomości ściśle dopuszczalnego (*strictly feasible*) punktu startowego \mathbf{x}_0 , który można wyznaczyć rozwiązując pomocnicze zadanie optymalizacji, nazywane często zadaniem pierwszej fazy (*phase I problem*). Punkt \mathbf{x} dla którego zachodzi

$$f_i(\mathbf{x}) < 0, \quad i \in \mathbf{m} \quad (41)$$

można znaleźć rozwiązując zadanie optymalizacji

$$\begin{aligned} & \underset{\mathbf{x} \in \mathbb{R}^n, s \in \mathbb{R}}{\text{minimize}} && s \\ & \text{subject to} && f_i(\mathbf{x}) \leq s, \quad i \in \mathbf{m}, \end{aligned} \quad (42)$$

dla której zawsze możemy łatwo znaleźć ściśle dopuszczalny punkt startowy, wybierając dowolny punkt $\tilde{\mathbf{x}}_0$ (np. $\tilde{\mathbf{x}}_0 = \mathbf{0}$) a następnie dobierając wartość s_0 tak, że

$$s_0 > \max\{f_1(\tilde{\mathbf{x}}_0), \dots, f_m(\tilde{\mathbf{x}}_0)\}, \quad (43)$$

np. kładąc

$$s_0 = 1 + \max\{f_1(\tilde{\mathbf{x}}_0), \dots, f_m(\tilde{\mathbf{x}}_0)\}. \quad (44)$$

Rozwiązując zadanie pomocnicze (42), np. metodą SBM, otrzymujemy punkt optymalny $(\tilde{\mathbf{x}}^*, s^*)$, wówczas możemy wyróżnić trzy przypadki.

1. Jeśli $s^* < 0$ to znaczy, że $f_i(\tilde{\mathbf{x}}^*) \leq s^* < 0$, zatem $\tilde{\mathbf{x}}^*$ jest punktem ściśle dopuszczalnym dla zadania (1)
2. Jeśli $s^* = 0$ to znaczy, że zadanie wyjściowe nie ma ściśle dopuszczalnego rozwiązania.
3. Jeśli $s^* > 0$ to znaczy, że zadanie wyjściowe nie ma rozwiązania (*infeasible*)

W praktyce jeśli $s^* < -\epsilon$ dla pewnego „rozsądnego” $\epsilon > 0$, to $\tilde{\mathbf{x}}^*$ jest ściśle dopuszczalny (*strictly feasible*). Warunek $s^* < 0$ można przyjąć jako kryterium stopu przy rozwiązywaniu (42).

Przykład 2. Chcemy wyznaczyć $\mathbf{x} \in \mathbb{R}^n$ taki, że $\mathbf{Ax} < \mathbf{b}$, dla $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\mathbf{b} \in \mathbb{R}^m$. W tym celu formułujemy zadanie optymalizacji

$$\begin{aligned} & \underset{\mathbf{x} \in \mathbb{R}^n, s \in \mathbb{R}}{\text{minimize}} && s \\ & \text{subject to} && \mathbf{Ax} - \mathbf{b} \leq \mathbf{1}s \end{aligned} \quad (45)$$

gdzie

$$\mathbf{1} = [1 \quad \dots \quad 1]^T. \quad (46)$$

Zauważmy, że dla $\mathbf{x} = \mathbf{0}$ oraz $1 + \max(-\mathbf{b})$ nierówność w (45) jest spełniona w sposób ostry, tzn.

$$\mathbf{Ax} - \mathbf{b} < \mathbf{1}s \quad (47)$$

zatem przyjmując oznaczenia

$$\tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{x} \\ s \end{bmatrix} \in \mathbb{R}^{n+1}, \quad \tilde{\mathbf{A}} = [\mathbf{A} \quad -\mathbf{1}] \in \mathbb{R}^{m \times (n+1)}, \quad (48)$$

$$\tilde{\mathbf{b}} = \mathbf{b} \in \mathbb{R}^m, \quad \tilde{\mathbf{c}} = \begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix} \in \mathbb{R}^{n+1} \quad (49)$$

możemy zapisać zadanie (45) w postaci

$$\begin{aligned} & \underset{\tilde{\mathbf{x}} \in \mathbb{R}^{n+1}}{\text{minimize}} && \tilde{\mathbf{c}}^T \tilde{\mathbf{x}} \\ & \text{subject to} && \tilde{\mathbf{A}} \tilde{\mathbf{x}} \leq \tilde{\mathbf{b}}, \end{aligned} \quad (50)$$

dla którego możemy podać ściśle dopuszczalny punkt startowy

$$\tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{0} \\ 1 + \max(-\mathbf{b}) \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ 1 - \min(\mathbf{b}) \end{bmatrix}. \quad (51)$$

Znając ściśle dopuszczalny punkt startowy możemy metodą SBM rozwiązać zadanie (50) i tym samym, korzystając z tego, że

$$\tilde{\mathbf{x}} = \begin{bmatrix} \mathbf{x} \\ s \end{bmatrix} \in \mathbb{R}^{n+1}, \quad (52)$$

znaleźć punkt $\mathbf{x} \in \mathbb{R}^n$ taki, że $\mathbf{Ax} < \mathbf{b}$.

2 Rozwiązywanie zadań LP metodą SBM

2.1 Logarytmiczna funkcja bariery dla zadania LP

Dla zadania LP

$$\begin{aligned} & \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimize}} && \mathbf{c}^T \mathbf{x} \\ & \text{subject to} && \mathbf{a}_i^T \mathbf{x} \leq b_i, \quad i \in \mathbf{m} \end{aligned} \quad (53)$$

mamy

$$\varphi(\mathbf{x}) = - \sum_{i=1}^m \log(b_i - \mathbf{a}_i^T \mathbf{x}), \quad (54)$$

$$\nabla \varphi(\mathbf{x}) = \sum_{i=1}^m \frac{\mathbf{a}_i}{b_i - \mathbf{a}_i^T \mathbf{x}}, \quad (55)$$

$$\nabla^2 \varphi(\mathbf{x}) = \sum_{i=1}^m \frac{\mathbf{a}_i \mathbf{a}_i^T}{(b_i - \mathbf{a}_i^T \mathbf{x})^2}. \quad (56)$$

2.2 Nierówności liniowe i wielościany

Zbiór punktów $\mathbf{x} \in \mathbb{R}^n$ spełniających liniową nierówność $\mathbf{a}^T \mathbf{x} \leq b$ dla $b \in \mathbb{R}$ jest domkniętą półprzestrzenią (ang. *closed half-space*). Wektor \mathbf{a} jest normalny do brzegu tej półprzestrzeni i wskazuje na zewnątrz. Układ m liniowych nierówności

$$\mathbf{a}_i^T \mathbf{x} \leq b_i, \quad i \in \mathbf{m}, \quad (57)$$

wyznacza obszar w \mathbb{R}^n , który jest przecięciem m półprzestrzeni i który nazywamy wielościanem (ang. *polyhedron*). Wielościan jest zbiorem wypukłym (jako przecięcie zbiorów wypukłych). W zależności od układu nierówności

$$\mathbf{a}_i^T \mathbf{x} \leq b_i, \quad i \in \mathbf{m}, \quad (58)$$

wielością może być ograniczony lub nieograniczony. Wielością ograniczoną nazywamy wielokomórką (ang. *polytope*). Przyjmując oznaczenia

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \quad (59)$$

możemy napisać układ nierówności

$$\mathbf{a}_i^T \mathbf{x} \leq b_i, \quad i \in \mathbf{m}, \quad (60)$$

w postaci macierzowo-wektorowej

$$\mathbf{A}\mathbf{x} \leq \mathbf{b}, \quad (61)$$

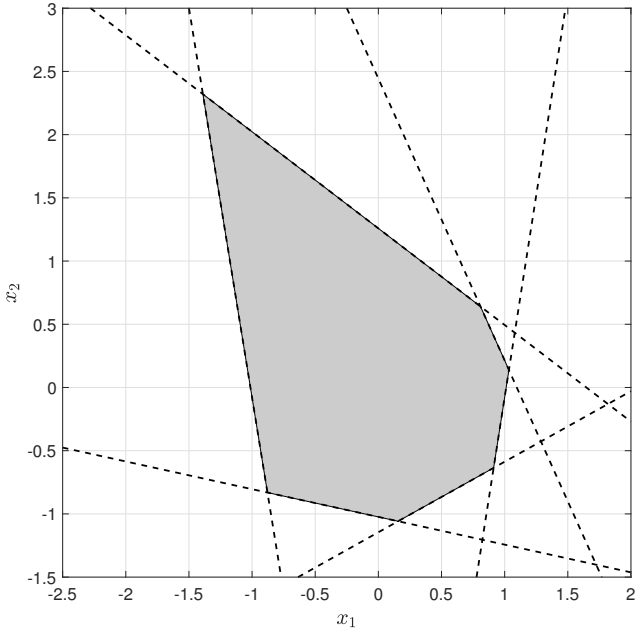
gdzie nierówność jest rozumiana jako nierówność odpowiednich elementów (ang. *component-wise inequality*), tzn. dla

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_m \end{bmatrix}, \quad \mathbf{w} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix}, \quad (62)$$

napis

$$\mathbf{v} \leq \mathbf{w} \quad (63)$$

oznacza, że $v_i \leq w_i, i \in \mathbf{m}$.



Rysunek 5: Wielokomórka opisana zależnościami (64) i (65). Wielokomórka jest przecięciem (częścią wspólną), półpłaszczyzn opisanych nierównościami $\mathbf{a}_i^T \mathbf{x} \leq b_i$, gdzie \mathbf{a}_i^T jest i -tym wierszem macierzy \mathbf{A} , natomiast b_i jest i -tym elementem wektora \mathbf{b} . Innymi słowy wielokomórka jest opisana nierównością $\mathbf{A}\mathbf{x} \leq \mathbf{b}$. Linie przerywane są wykresami prostych (ogólnie hiperpłaszczyzn) postaci $\mathbf{a}_i^T \mathbf{x} \leq b_i$.

Przykład 3. Weźmy pod uwagę układ nierówności

$$\mathbf{A}\mathbf{x} \leq \mathbf{b}, \quad (64)$$

gdzie

$$\mathbf{A} = \begin{bmatrix} 0.4873 & -0.8732 \\ 0.6072 & 0.7946 \\ 0.9880 & -0.1546 \\ -0.2142 & -0.9768 \\ -0.9871 & -0.1601 \\ 0.9124 & 0.4093 \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}. \quad (65)$$

Mamy więc 6 nierówności liniowych, z których każda definiuje pewną półprzestrzeń, natomiast ich układ definiuje wielością (Rys. 2.). Wielokomórkę reprezentowaną przez przecięcie półprzestrzeni nazywamy \mathcal{H} -wielokomórką (ang. \mathcal{H} -polytope).

3 Zadania

Zadanie 1. Napisać w środowisku Matlab funkcję, która dla zadanych macierzy \mathbf{A} i wektora \mathbf{b} będzie zwracać wektor \mathbf{x} taki, że $\mathbf{A}\mathbf{x} < \mathbf{b}$.

Zadanie 2. Napisać w środowisku Matlab skrypt do rozwiązywania zadania LP

$$\begin{aligned} & \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimize}} && \mathbf{c}^T \mathbf{x} \\ & \text{subject to} && \mathbf{A}\mathbf{x} \leq \mathbf{b} \end{aligned}$$

metodą SBM przy założeniu, że znany jest ściśle dopuszczalny punkt startowy (przyjąć $\mathbf{x}_0 = 0$), gdzie $\mathbf{c} = [-0.5 \ 0.5]^T$, zaś \mathbf{A}, \mathbf{b} są dane przez (65).

Wykonać odpowiednie wykresy (jak na Rys. 5 i Rys. 6, można skorzystać z polecenia `fill, colorbar`). Porównać otrzymany wynik z rozwiązaniem znalezionym za pomocą funkcji `linprog`.

Wskazówka 1: Dla zadania LP metoda SBM ma postać

Sekwencyjna metoda bariery dla zadania LP

dane: ściśle dopuszczalny \mathbf{x}_0 , t_{init} , $\gamma > 1$, $\epsilon > 0$

1. Podstawiamy $k = 0$, $t = t_{\text{init}}$.
2. Używając metody Newtona z tłumieniem rozwiązujemy zadanie

$$\underset{\mathbf{x}}{\text{minimize}} \quad \psi_t(\mathbf{x}) \quad (66)$$

dla punktu startowego \mathbf{x}_k , gdzie

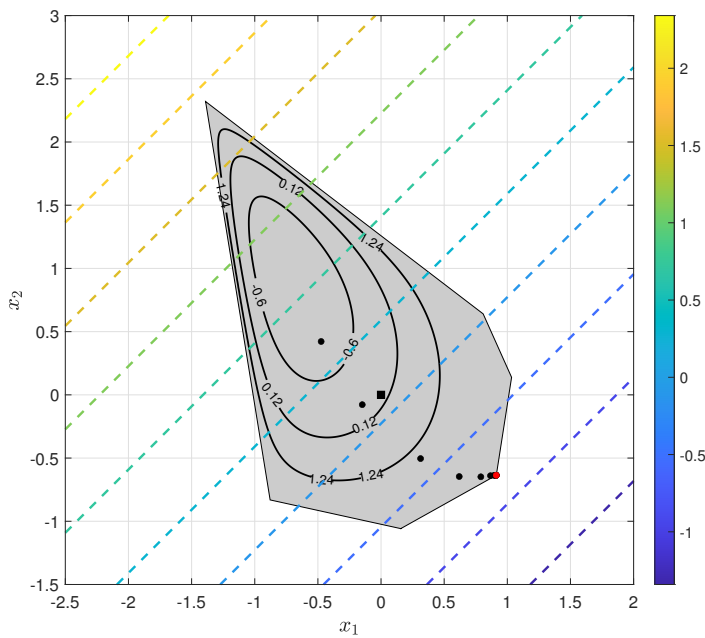
$$\begin{aligned} \psi_t(\mathbf{x}) &= f_0(\mathbf{x}) + \frac{1}{t} \varphi(\mathbf{x}) \\ &= \mathbf{c}^T \mathbf{x} - \frac{1}{t} \sum_{i=1}^m \log(b_i - \mathbf{a}_i^T \mathbf{x}), \end{aligned}$$

$$\begin{aligned} \nabla \psi_t(\mathbf{x}) &= \nabla f_0(\mathbf{x}) + \frac{1}{t} \nabla \varphi(\mathbf{x}) \\ &= \mathbf{c} + \frac{1}{t} \sum_{i=1}^m \frac{\mathbf{a}_i}{b_i - \mathbf{a}_i^T \mathbf{x}}, \end{aligned}$$

$$\begin{aligned} \nabla^2 \psi_t(\mathbf{x}) &= \nabla^2 f_0(\mathbf{x}) + \frac{1}{t} \nabla^2 \varphi(\mathbf{x}) \\ &= \frac{1}{t} \sum_{i=1}^m \frac{\mathbf{a}_i \mathbf{a}_i^T}{(b_i - \mathbf{a}_i^T \mathbf{x})^2}. \end{aligned}$$

Rozwiązanie optymalne zadania (66) oznaczamy przez \mathbf{x}_k^* .

3. Aktualizujemy $\mathbf{x}_{k+1} \leftarrow \mathbf{x}_k^*$.
4. Jeśli $m/t \leq \epsilon$, zwracamy \mathbf{x}_k^* i kończymy algorytm.
5. Aktualizujemy $t \leftarrow \gamma t$, $k \leftarrow k + 1$ i przechodzimy punktu 2.



Rysunek 6: Wykres obszaru dopuszczalnego, punktów centralnych oraz poziomic funkcji bariery dla Zadania 1. Czarnym kwadratem zaznaczono punkt startowy x_0 . Czarnymi kropkami kolejne minima iteracji zewnętrznych. Liniami przerywanymi zaznaczono poziomicę funkcji celu. Rozwiązaniem jest punkt wierzchołkowy $x_{\text{opt}} = [0.9126, -0.6359]^T$ (zaznaczony czerwoną kropką).

Wskazówka 2: Przyjmując oznaczenia

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix} \in \mathbb{R}^{m \times n}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad (67)$$

możemy wyrażenia

$$\sum_{i=1}^m \log(b_i - \mathbf{a}_i^T \mathbf{x}), \quad \sum_{i=1}^m \frac{\mathbf{a}_i}{b_i - \mathbf{a}_i^T \mathbf{x}}, \quad \sum_{i=1}^m \frac{\mathbf{a}_i \mathbf{a}_i^T}{(b_i - \mathbf{a}_i^T \mathbf{x})^2} \quad (68)$$

zapisać w środowisku Matlab następująco

```
sum(log(b-A*x))
```

```
A'*(1./(b-A*x))
```

```
(diag(1./(b-A*x))*A)'*(diag(1./(b-A*x))*A)
```

Wskazówka 3: Jak już wspomniano, rysując wykres polikomórki 2D można skorzystać z polecenia `fill`. Wierzchołkami polikomórki opisanej zależnościami (64) i (65) są kolumny macierzy V

$$V = \begin{bmatrix} 0.1562 & 0.9127 & 1.0338 & 0.8086 & -1.3895 & -0.8782 \\ -1.0580 & -0.6358 & 0.1386 & 0.6406 & 2.3203 & -0.8311 \end{bmatrix} \quad (69)$$

Wskazówka 4: Można przyjąć $t_0 = 1$, $\gamma = 2.5$.

Zadanie 3. Rozbudować skrypt z Zadania 2 w taki sposób, aby ściśle dopuszczalny punkt startowy był wyznaczany automatycznie (metodą opisaną w paragrafie 1.5).

Zadanie 4. Napisać funkcję w środowisku Matlab, do rozwiązywania metodą SBM zadań LP postaci

$$\begin{aligned} & \underset{\mathbf{x} \in \mathbb{R}^n}{\text{minimize}} && \mathbf{c}^T \mathbf{x} \\ & \text{subject to} && \mathbf{A} \mathbf{x} \leq \mathbf{b} \end{aligned}$$

dla dowolnych danych wejściowych tzn. \mathbf{A} , \mathbf{b} i \mathbf{c} .

Literatura

- [1] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. Cambridge University Press, New York, NY, USA, 2004.