Uniwersytet Łódzki
Wydział Matematyki

Agnieszka Czuba

Słabe ostre minima

Praca magisterska wykonana w Zakładzie
Metod Numerycznych pod kierunkiem
prof. dr hab. Marcina Studniarskiego

Łódź 2001

Spis treści

Wstęp 3

Rozdział 1. Pojęcie słabego ostrego minimum 4

1.1. Rodzaje minimów 4
1.2. Słabe ostre minima w teorii dźwięku 7
1.3. Słabe ostre minima w teorii gazów 11

Rozdział 2. Warunki istnienia słabego ostrego minimum w przypadkach nieograniczonego obszaru poszukiwań 13

2.1. Słabe ostre minima a pojęcie stożka normalnego 13
2.2. Słabe ostre minima a pojęcie pochodnej 17
2.2.1. Warunek wystarczający istnienia słabego ostrego minimum 19
2.2.2. Warunek konieczny i wystarczający istnienia słabego ostrego minimum 20

Rozdział 3. Warunek konieczny i wystarczający istnienia słabego ostrego minimum dla maksimum ze skończonych i ściśle różniczkowalnych funkcji 23

Rozdział 4. Warunki istnienia słabego ostrego minimum w przypadkach z ograniczeniami 27

4.1. Ograniczenia jawne obszaru poszukiwań 27
4.2. Ograniczenia obszaru poszukiwań uwikłane poprzez układ warunków 29
4.2.1. Warunek wystarczający istnienia słabego ostrego minimum pierwszego rzędu 29
4.2.2. Warunek wystarczający istnienia słabego ostrego minimum drugiego rzędu 31
4.2.3. Warunek wystarczający istnienia słabego ostrego minimum pierwszego rzędu w przypadku ograniczeń liniowych 33

Podsumowanie 36

Spis rysunków 37

Bibliografia 38

Wstęp

Do podstawowych zagadnień, jakimi zajmuje się teoria optymalizacji, należy poszukiwanie wartości minimalnych funkcji opisujących zjawiska z otaczającego nas świata. Problemem w rozwiązywaniu tych zagadnień jest sam sposób opisania zjawiska, a następnie metoda minimalizacji otrzymanej funkcji w całej jej dziedzinie, albo przy określonych ograniczeniach.

W przedstawionej pracy rozważania skupią się wokół rozwiązania zadania

$$\min\{f(x) \mid x \in C\}, \tag{1}$$

gdzie $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$, $C \subset \mathbb{R}^n$, natomiast minimum należy do zbioru minimów o szczególnych własnościach i jest nazywane słabym ostrym minimum rzędu m.

Ten specjalny rodzaj minimum zostanie przedstawiony w rozdziale pierwszym. Podany również tutaj będą przykłady funkcji posiadających słabe ostre minimum, które opisują zjawiska otaczającego nas świata.

Głównym celem pracy jest zaprezentowanie twierdzeń, na podstawie których mogą być tworzone metody poszukiwania słabych ostrych minimów dla problemu (1). W kolejnych częściach będą więc cytowane odpowiednie warunki oraz zostaną przedstawione przykłady ich praktycznego zastosowania.

Drugi rozdział będzie poświęcony poszukiwaniom słabego ostrego minimum w przypadku ogólnym, kiedy zbiór C jest całą przestrzenią $\mathbb{R}^n$. Zostaną przedstawione dwa warunki wystarczające i konieczne istnienia słabego ostrego minimum. W jednym z nich wykorzystane będzie pojęcie stożka normalnego, a w drugim zmodyfikowana pochodna kierunkowa. Omówione również będą problemy, jakie mogą pojawić się przy stosowaniu cytowanych twierdzeń. W rozdziale trzecim zostanie przedstawione twierdzenie, które pozwala pominąć te problemy w przypadku specjalnej klasy funkcji (maksimum ze skończonych i ściśle różniczkowalnych funkcji).

Ponieważ w rzeczywistości często w zadaniach optymalizacyjnych zbiór punktów dopuszczalnych jest ograniczony, w rozdziale czwartym zostaną przedstawione sposoby poszukiwania słabych ostrych minimów funkcji f w przypadku, kiedy zbiór C jest podzbiorem właściwym przestrzeni $\mathbb{R}^n$. Jeden z zacytowanych warunków będzie dotyczył przypadku, kiedy zbiór C jest podany w sposób jawny, natomiast oddzielnie rozpatrzony zostanie przypadek zbioru C opisanego przy pomocy układu warunków. Zaprezentowane zostaną również twierdzenia wykorzystujące do rozwiązania zadania (1) funkcję Lagrange'a.

Rozdział 1

Pojęcie słabego ostrego minimum

1.1. Rodzaje minimów

Wśród zagadnień teorii optymalizacji podstawowe znaczenie ma problem polegający na wybraniu z danego podzbioru przestrzeni wektorowej tego konkretnego wektora, który minimalizuje daną funkcję. Czasami jednak pojawia się dodatkowe żądanie dotyczące nie tylko obszaru poszukiwań, ale także rodzaju znalezionego minimum. Okazuje się, że funkcja może przyjmować wartość minimalną w różny sposób i w związku z tym rozważania mogą dotyczyć nie tylko pytania, gdzie funkcja osiąga minimum, ile ono wynosi, ale także, co się dzieje w jego otoczeniu.

Klasycznie przyjmujemy, że funkcja $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ osiąga minimum w punkcie $\overline{x} \in \mathbb{R}^n$, jeżeli istnieje otoczenie U punktu $\overline{x}$ takie, że:

$$f(\overline{x}) \leqslant f(x) \quad \text{dla dowolnego} \quad x \in U \tag{1.1}$$

Jeśli rozważy się dokładniej otoczenie punktu $\overline{x}$, w którym funkcja f osiąga minimum, to okaże się, że minima – ze względu na ich własności – można podzielić na kilka grup:

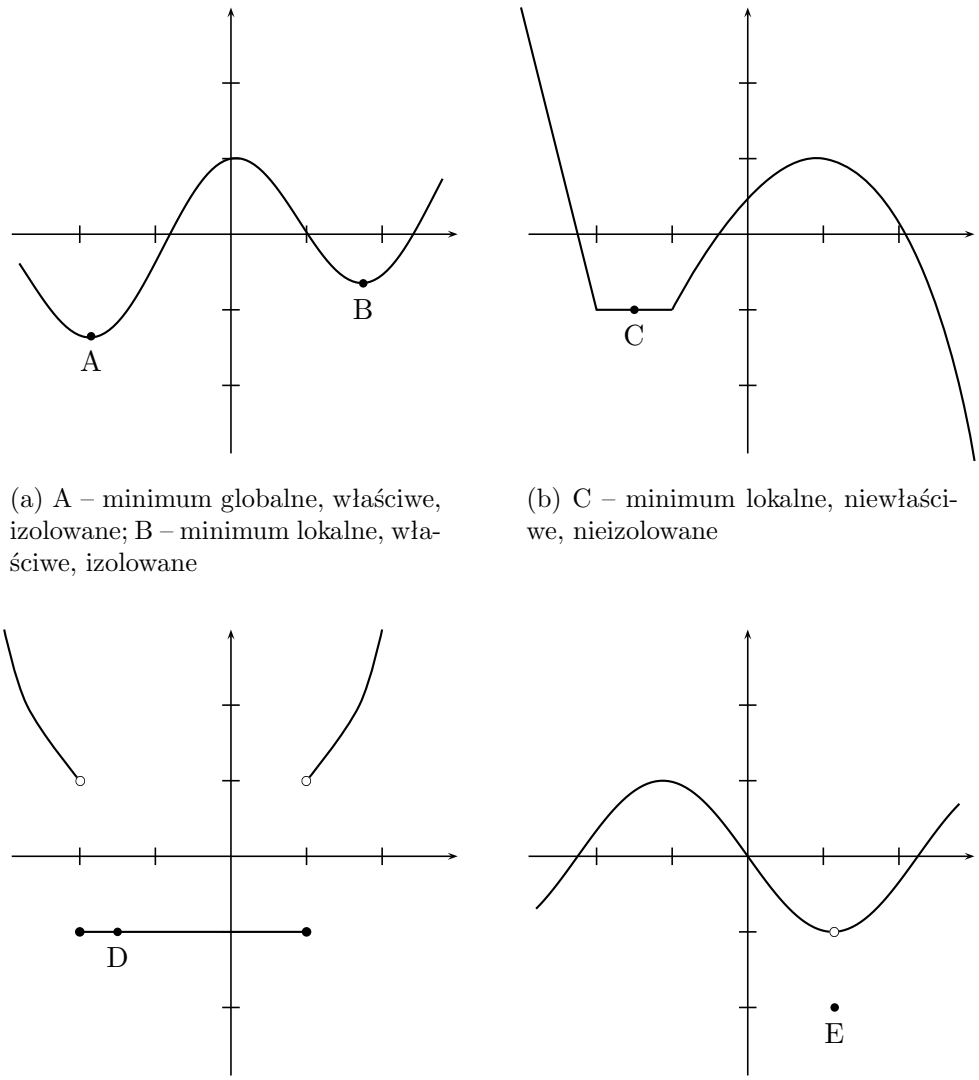
1. Jeżeli istnieje otoczenie U z warunku (1.1), to mówimy, że funkcja f osiąga w $\overline{x}$ minimum *lokalne*.
2. Jeżeli warunek (1.1) jest spełniony dla dowolnego punktu x z dziedziny funkcji f, to mówimy, że funkcja f osiąga w $\overline{x}$ minimum *globalne*.
3. Jeżeli istnieje otoczenie U punktu $\overline{x}$ takie, że:

$$f(\overline{x}) < f(x) \quad \text{dla dowolnego} \quad x \in U \setminus \{\overline{x}\}, \tag{1.2}$$

   to mówimy, że funkcja f osiąga w $\overline{x}$ minimum *właściwe*.
4. Jeśli w każdym otoczeniu U z warunku (1.1) istnieje punkt $\overline{\overline{x}} \neq \overline{x}$ taki, że $f(\overline{x}) = f(\overline{\overline{x}})$, to funkcja f osiąga w $\overline{x}$ minimum *niewłaściwe*.
5. Jeżeli w $\overline{x}$ funkcja f osiąga minimum, które jest punktem izolowanym zbioru wszystkich minimów lokalnych funkcji f, to mówimy, że funkcja f osiąga w $\overline{x}$ minimum *izolowane*.
6. Jeżeli minimum nie jest izolowane, to nazywamy je *nieizolowanym*.

W zbiorze minimów nieizolowanych można wyróżnić rozważane w tej pracy słabe ostre minima rzędu m.

Definicja 1.1. Niech funkcja $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ przyjmuje wartość stałą i skończoną na zbiorze $S \subset \mathbb{R}^n$ i niech $C \subset \mathbb{R}^n$, $\overline{x} \in S \cap C$, $m \geqslant 1$. Mówimy, że funkcja



(a) A – minimum globalne, właściwe, izolowane; B – minimum lokalne, właściwe, izolowane

(b) C – minimum lokalne, niewłaściwe, nieizolowane

(c) D – minimum globalne, niewłaściwe, nieizolowane

(d) E – minimum lokalne, właściwe, izolowane

Rysunek 1.1. Przykłady różnego rodzaju minimów

f osiąga w $\overline{x}$ słabe ostre minimum rzędu m (ang. *weak sharp minimum of order* m) dla problemu (1), jeżeli istnieje $\beta > 0$ taka, że

$$f(x) - f(\overline{x}) \geqslant \beta \cdot \operatorname{dist}^m(x, S) \quad \text{dla dowolnego} \quad x \in C, \tag{1.3}$$

gdzie $\operatorname{dist}^m(x, S)$ oznacza m-tą potęgę odległości punktu x od zbioru S:

$$\operatorname{dist}^m(x, S) = \inf\{\|y - x\|^m \mid y \in S\}.$$

Natomiast mówimy, że funkcja f osiąga w $\overline{x}$ lokalne słabe ostre minimum rzędu m (ang. *weak sharp local minimum of order* m) (dla problemu (1)), jeżeli istnieją $\beta > 0$ i $\varepsilon > 0$ takie, że:

$$f(x) - f(\overline{x}) \geqslant \beta \cdot \operatorname{dist}^m(x, S) \quad \text{dla dowolnego} \quad x \in C \cap B(\overline{x}, \varepsilon), \tag{1.4}$$

gdzie $B(\overline{x}, \varepsilon)$ oznacza kulę domkniętą o środku $\overline{x}$ i promieniu ε:

$$B(\overline{x}, \varepsilon) = \{y \in \mathbb{R}^n \mid \|\overline{x} - y\| \leqslant \varepsilon\}.$$

Przykład 1.1. Rozważmy funkcję $f : \mathbb{R}^2 \to \mathbb{R}$, $f(x, y) = |x|^m$, $m \geqslant 1$, $C = \mathbb{R}^2$, $S = \{(0, y) \mid y \in \mathbb{R}\}$. Niech $(0, \overline{y}) \in S$, gdzie $\overline{y} \in \mathbb{R}$. Wtedy $f(0, \overline{y}) = 0$. Zauważmy, że dla dowolnego $(x, y) \in \mathbb{R}^2$:

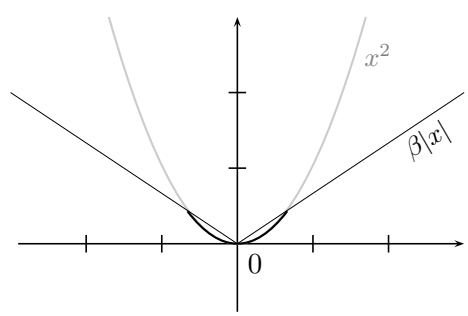
$$\begin{aligned} \operatorname{dist}^m((x, y), S) &= \inf\{\|(x, y) - (x_s, y_s)\|^m \mid (x_s, y_s) \in S\} \\ &= \|(x, y) - (0, y)\|^m = \|(x, 0)\|^m = (\sqrt{x^2})^m = |x|^m, \end{aligned} \tag{1.5}$$

$$\begin{aligned} f(x, y) - f(0, \overline{y}) &= |x|^m - 0^m = |x|^m \\ &= \operatorname{dist}^m((x, y), S) \geqslant 1 \cdot dist^m((x, y), S), \end{aligned}$$

a więc funkcja f osiąga słabe ostre minimum rzędu m w dowolnym punkcie zbioru S.

Zauważmy, że w przypadku, kiedy zbiór S jest jednoelementowy ($S = \{x\}$) i funkcja f osiąga w punkcie x minimum, nie oznacza to wcale, że funkcja f osiąga w x słabe ostre minimum rzędu m, $m \geqslant 1$.

Przykład 1.2. Rozpatrzmy przykład funkcji $f(x) = x^2$, $S = \{0\}$, $C = \mathbb{R}$. Funkcja f osiąga w $x = 0$ minimum. Czy jest to także słabe ostre minimum rzędu pierwszego?



Rysunek 1.2. Przykład minimum funkcji, które nie jest jej słabym ostrym minimum rzędu pierwszego

Zauważmy, że dla dowolnego $x \in C$

$$\operatorname{dist}(x, S) = \inf\{|x - y| \mid y \in S\} = |x - 0| = |x|, \tag{1.6}$$

$$f(x) - f(0) = x^2. \tag{1.7}$$

Z własności funkcji kwadratowej wiemy, że dla dowolnego $\beta > 0$ istnieje $x \in C$ taki, że:

$$f(x) - f(0) = x^2 < \beta \cdot |x| = \beta \cdot \operatorname{dist}(x, S),$$

a więc funkcja f nie osiąga w punkcie $x = 0$ słabego ostrego minimum rzędu pierwszego.

Natomiast funkcja f osiąga w tym punkcie lokalne słabe ostre minimum rzędu wyższego, ponieważ:

$$f(x) - f(0) = x^2 \geqslant 1 \cdot |x|^m \quad \text{dla dowolnego } x \in [-1, 1] \subset C,\ m \geqslant 2$$

Przykład ten pokazuje, że definicja słabego ostrego minimum rzędu m w punkcie x w przypadku $S = \{x\}$ (nazywanego wówczas ścisłym minimum (ang. *strict minimum*)) nie jest równoważna definicji minimum w punkcie x.

1.2. Słabe ostre minima w teorii dźwięku

Źródłem dźwięku jest zawsze jakieś ciało drgające. Wskutek drgań źródła dźwięku w otaczającym je powietrzu wytwarzają się na przemian zagęszczenia i rozrzedzenia. Dzięki sprężystości powietrza odkształcenia te są przenoszone z miejsca na miejsce w postaci fal podłużnych. Ilość pełnych drgań wykonanych przez źródło dźwięku w jednostce czasu nazywamy częstotliwością, natomiast wysokość powstałej fali amplitudą. W praktyce częstotliwość dźwięku wiąże się z jego wysokością[1], zaś amplitudę z natężeniem i głośnością[2].

Podczas nagrywania dźwięku fala dochodząca do membrany mikrofonu wywołuje zagęszczenia i rozrzedzenia powietrza, w wyniku których membrana zaczyna się poruszać, co powoduje wytwarzanie prądu zmiennego. Prąd ten za pomocą przetwornika analogowo-cyfrowego (ang. *Analog-to-Digital Converter, ADC*) jest przetwarzany na cyfrowy zapis fali dźwiękowej. Zapis polega na tym, że pobiera się określoną ilość[3] „próbek" w jednostce czasu (tzn. mierzy się amplitudę fali w kolejnych odstępach czasowych).

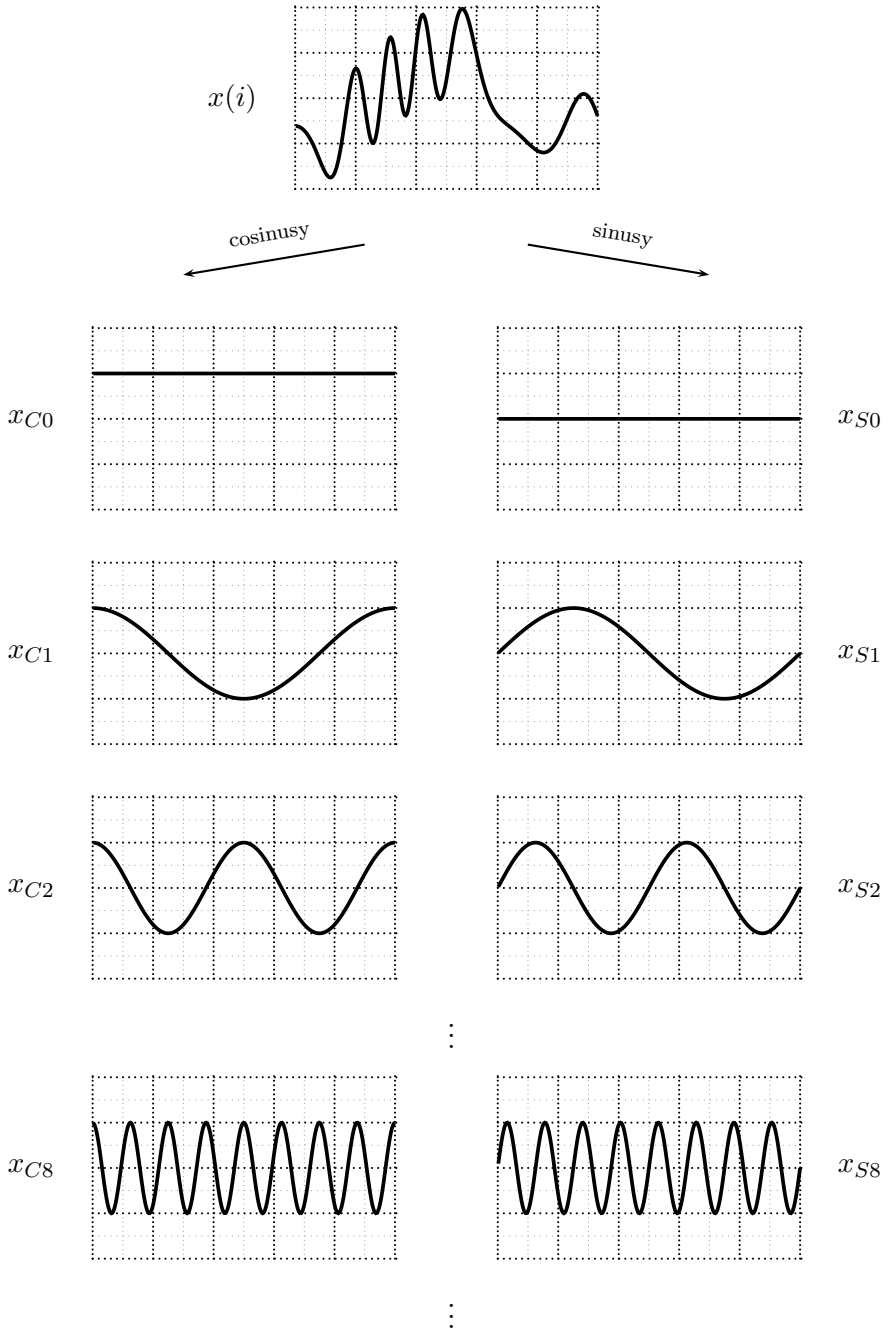
W przypadku, kiedy chcemy odtworzyć nagrany dźwięk, przetwornik cyfrowo-analogowy (ang. *Digital-to-Analog Converter, DAC*) interpoluje zapis cyfrowy dźwięku, w wyniku czego powstaje odpowiedni prąd zmienny, który dociera do głośnika i powoduje, że membrana głośnika zaczyna drgać.

W taki właśnie sposób – w wielkim uproszczeniu i skrócie – przebiega zapisywanie i odtwarzanie dźwięku. Czasami jednak pojawiają się problemy. Aby móc je dobrze przedstawić, należy nawiązać do przekształcenia, jakiemu można poddać cyfrowy zapis dźwięku – dyskretnej transformacji Fouriera (ang. *Discrete Fourier Transformation, DFT*). Do tego celu zakłada się, co nie jest do końca zgodne z rzeczywistością, że funkcja opisująca dany dźwięk jest funkcją okresową. „Wycina się" więc pewien fragment zapisu cyfrowego fali akustycznej i przyjmuje, że ten wycięty kawałek będzie okresowo powtarzał się w funkcji. W ten sposób powstały zapis można poddać transformacji Fouriera, tzn. przybliżyć powstałą w ten sposób funkcję okresową f za pomocą szeregu Fouriera. DFT da nam w wyniku przybliżenie, które będzie zawierało skończoną ilość składników. Sumowane oczywiście będą funkcje sinus i cosinus, a okresy kolejnych składników (a więc częstotliwość dźwięku) będą się zwiększać. Transformację Fouriera ilustruje przykład 1.3.

[1] im wyższa częstotliwość, tym wyższy dźwięk

[2] im wyższa amplituda, tym większe natężenie i głośność dźwięku

[3] oczywiście dokładność (wierność) zapisu jest lepsza, jeśli ilość pobranych próbek będzie duża



$$x(i) \approx \sum_{k=0}^{N/2} \frac{x_{Ck}}{N/2} \cos\left(\frac{2\pi ki}{N}\right) + \sum_{k=0}^{N/2} \frac{x_{Sk}}{N/2} \sin\left(\frac{2\pi ki}{N}\right)$$

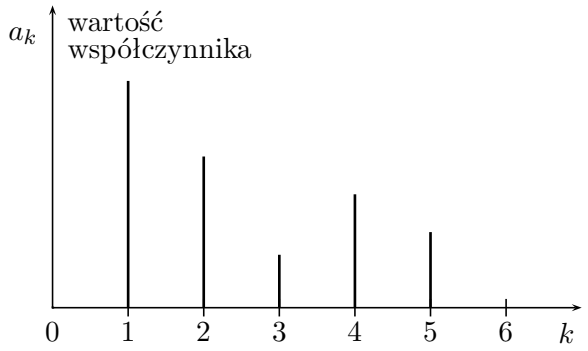
N – ilość wszystkich próbek

Rysunek 1.3. Transformacja Fouriera

Transformacja Fouriera oblicza współczynniki kolejnych częstotliwości[4]

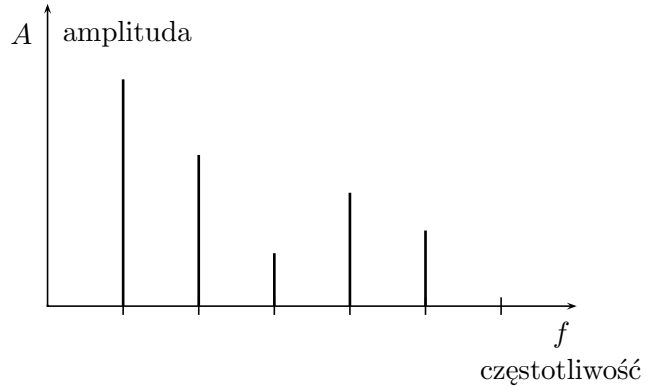
[4] w literaturze częstotliwością określana jest również funkcja sinus lub cosinus o danym okresie

w taki sposób, aby zsumowane funkcje składowe dały w wyniku jak najlepsze przybliżenie funkcji f.



Rysunek 1.4. Wynik działania DFT

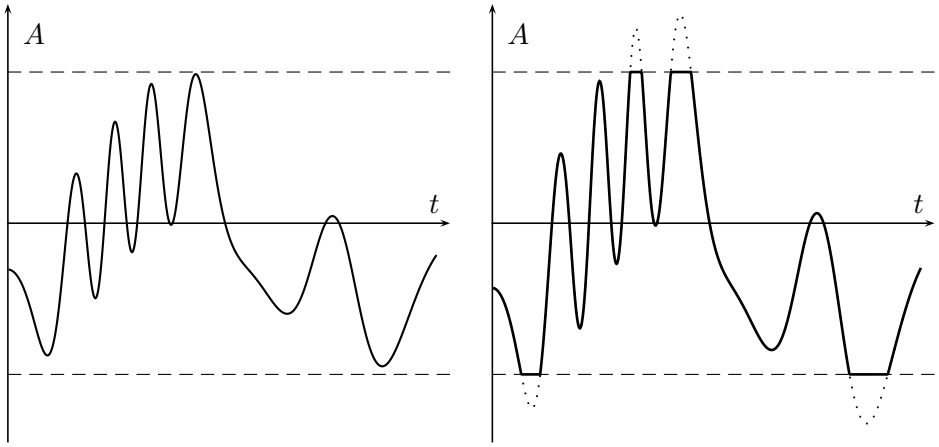
Numer współczynnika odpowiada częstotliwości dźwięku, zaś wartość współczynnika – maksymalnej amplitudzie dźwięku o danej częstotliwości. Dlatego też wynik transformacji Fouriera podaje się w postaci zależności amplitudy od częstotliwości.



Rysunek 1.5. Wynik działania DFT – zależność amplitudy od częstotliwości

Mając tak przedstawiony dźwięk możemy zrozumieć, jaki pojawia się problem w przypadku, kiedy odtwarzamy nagrany dźwięk coraz głośniej. W wyniku ograniczonych możliwości fizycznych urządzeń odtwarzających, amplituda dźwięku (a także jego częstotliwość) są ograniczone. A zatem na skutek zwiększania głośności może się okazać, że fala, jaką ma odtworzyć głośnik, przekroczyła dozwolony zakres. Dochodzi wtedy do obcięcia (ang. *clipping*) amplitudy, a więc do zmiany kształtu fali. W związku z tym zmieniają się też współczynniki częstotliwości w transformacji Fouriera.

Ponieważ współczynniki o najwyższych wartościach odpowiadają najniższym, a także najsłabiej słyszalnym częstotliwościom, więc najczęściej właśnie amplitudy tych częstotliwości są obcinane. W takich wypadkach zmodyfikowane częstotliwości muszą być odpowiednio „zastąpione” przez częstotliwości wyższe, które są dobrze słyszalne. W efekcie odbiorca dźwięku słyszy zakłócenia.



(a) fala przed zwiększeniem głośności

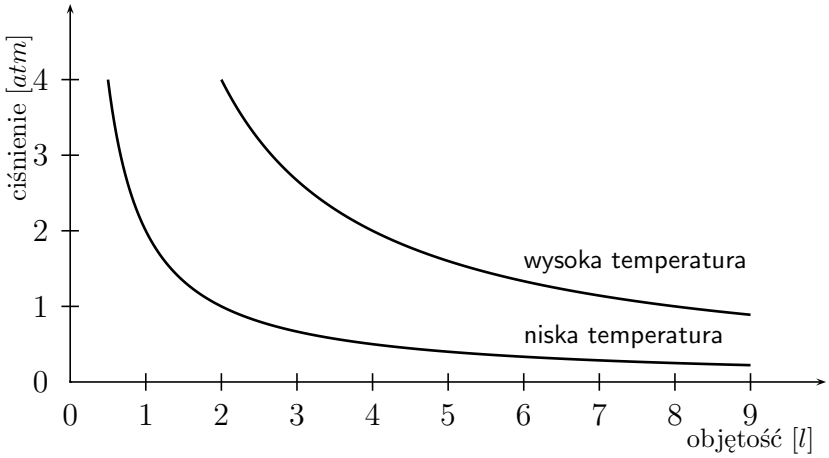
(b) fala po zwiększeniu głośności

Rysunek 1.6. Przykład obcięcia fali dźwiękowej

Jakie ma powiązania opisany wyżej przykład ze słabymi ostrymi minimami? Otóż obcięcie amplitudy powstałe w wyniku zwiększenia głośności odtwarzanego dźwięku powoduje, że funkcję g opisującą częstotliwość najniższą przed obcięciem zastępuje funkcja h, która posiada słabe ostre minima (oraz słabe ostre maksima). To właśnie istnienie słabych ostrych minimów i maksimów jest przyczyną zakłóceń, jakie słyszy odbiorca.

1.3. Słabe ostre minima w teorii gazów

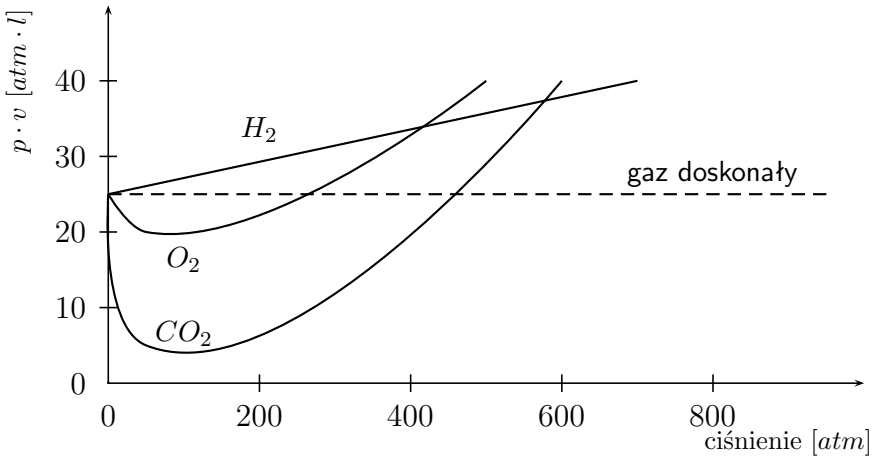
Jedno z podstawowych praw teorii gazów (pochodzące z XVII wieku prawo Boyle'a-Mariotte'a) mówi, że w stałej temperaturze objętość danej ilości gazu jest odwrotnie proporcjonalna do ciśnienia. Wykresem zależności ciśnienia od objętości w danej temperaturze jest hiperbola zwana izotermą.



Rysunek 1.7. Izotermy gazu spełniającego prawo Boyle'a-Mariotte'a

Kolejne prawo Gay-Lussaca podaje, że przy stałym ciśnieniu objętość danej masy gazu jest proporcjonalna do temperatury bezwzględnej[5].

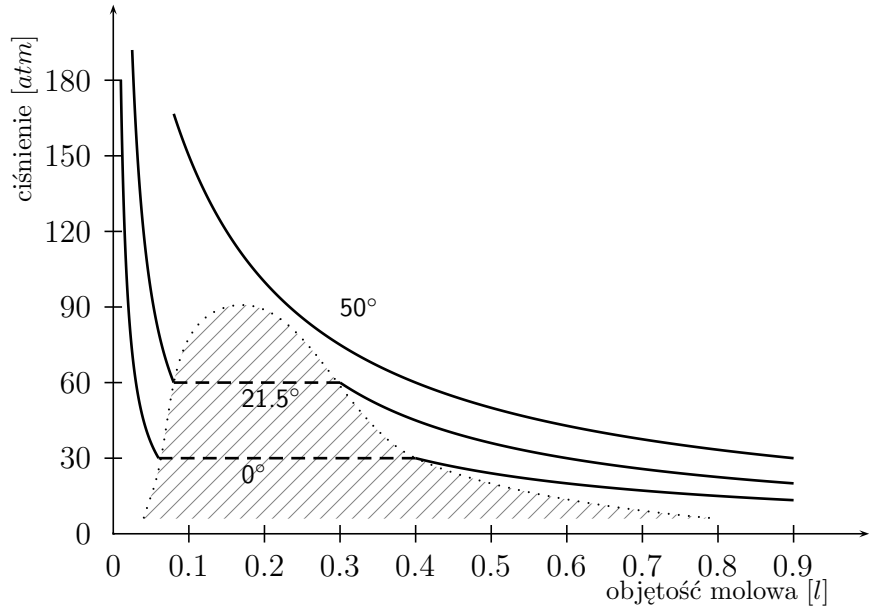
Pomiary wykonane pod wyższymi ciśnieniami, a nawet bardzo dokładne pomiary przeprowadzone pod zwykłymi ciśnieniami wykazały, że występują odchylenia od praw Boyle'a-Mariotte'a i Gay-Lussaca. Gazy zachowujące się zgodnie z tymi prawami nazywamy doskonałymi. Jednakże gazy rzeczywiste zachowują się w sposób doskonały jedynie w pewnych zakresach ciśnienia i temperatury.



Rysunek 1.8. Odchylenia od doskonałego zachowania się gazów rzeczywistych

[5] temperatura w skali Kelvina

Okazuje się, że izotermy gazów rzeczywistych dla wyższych temperatur wykazują niewielkie odchylenia od hiperbol oczekiwanych dla gazu doskonałego. Podobnie jest dla izoterm niskotemperaturowych w obszarze niskich ciśnień i dużych objęteości.



Rysunek 1.9. Izotermy i obszar skraplania CO_2

Gdy przyjrzymy się izotermie gazu rzeczywistego dla niskiej temperatury, to przy początkowym wzroście ciśnienia objętość maleje zgodnie z prawem Boyle'a-Mariotte'a. Jednak w pewnym momencie funkcja opisująca zachowanie gazu osiąga lokalne słabe ostre maksimum. Jest to moment, w którym oddziaływania między cząsteczkami gazu stają się tak silne, iż gaz zaczyna przekształcać się w ciecz. Wszystkie następne punkty izotermy, w których znajdują sie lokalne słabe ostre minima, wyznaczają obszar, w jakim ciśnienie odpowiadające równowagowej prężności pary w danej temperaturze pozostaje stałe. Gaz przekształca się w ciecz. Kiedy cała ilość gazu jest już skroplona, dalsze zwiększanie ciśnienia powoduje minimalne zmniejszenie objętości. Izoterma zaczyna gwałtownie rosnąć przy objętości zmierzającej do zera.

Wykreślając izotermy danego gazu dla kolejnych temperatur i zaznaczając miejsca, gdzie funkcje przyjmują lokalne słabe ostre maksimum oraz lokalne słabe ostre minima, można określić tzw. obszar skraplania gazu. Na rysunku jest on zaznaczony linią kropkowaną. Obszar pod tą linią odpowiada współistnieniu cieczy i pary, obszar po prawej stronie – fazie gazowej, a obszar po lewej stronie – fazie ciekłej.

Rozdział 2

Warunki istnienia słabego ostrego minimum w przypadkach nieograniczonego obszaru poszukiwań

2.1. Słabe ostre minima a pojęcie stożka normalnego

Poszukiwania metody znajdowania słabych ostrych minimów rzędu m w przypadkach ogólnych, tzn. wtedy gdy problem (1) jest rozpatrywany dla dowolnej funkcji f, a obszar poszukiwań C jest nieograniczoną przestrzenią $\mathbb{R}^n$, zakończyły się sukcesem w postaci warunku koniecznego i wystarczającego przedstawionego w [12]. Twierdzenie, o którym mowa, opiera się na badaniu znaku granicy ciągu pewnych ilorazów różnicowych funkcji f. Do budowania tych ilorazów wykorzystuje się wektory specjalnego zbioru – stożka normalnego – zależnego od zbioru S z definicji 1.1 i punktu $\overline{x}$, w którym chcemy stwierdzić istnienie słabego ostrego minimum.

Definicja 2.1. Niech S oznacza niepusty podzbiór przestrzeni $\mathbb{R}^n$, $\overline{x} \in \operatorname{cl} S$. Stożkiem normalnym[1] do zbioru S w punkcie $\overline{x}$ nazywamy zbiór:

$$\begin{aligned} N(S,\overline{x}) := \{y \mid \quad & \text{istnieją ciągi} \quad \{y_j\},\{x_j\},\{t_j\} \quad \text{i} \quad \{s_j\} \quad \text{takie, że} \\ & y_j \to y \quad \text{i} \quad x_j \to \overline{x} \quad \text{i} \quad s_j \in P(S,x_j) \quad \text{i} \\ & \{t_j\} \subset (0,+\infty) \quad \text{i} \quad y_j = \frac{x_j - s_j}{t_j}\} \end{aligned}$$

Przez $P(S,x_j)$ rozumiemy zbiór punktów domkniętego zbioru S, których odległość od punktu $x_j \in \mathbb{R}^n$ jest najmniejsza, tzn.

$$P(S,x_j) := \{w \in \operatorname{cl} S \mid \|x_j - w\| = \operatorname{dist}(x_j,S)\}.$$

Przykład 2.1. Rozpatrzmy zbiory:

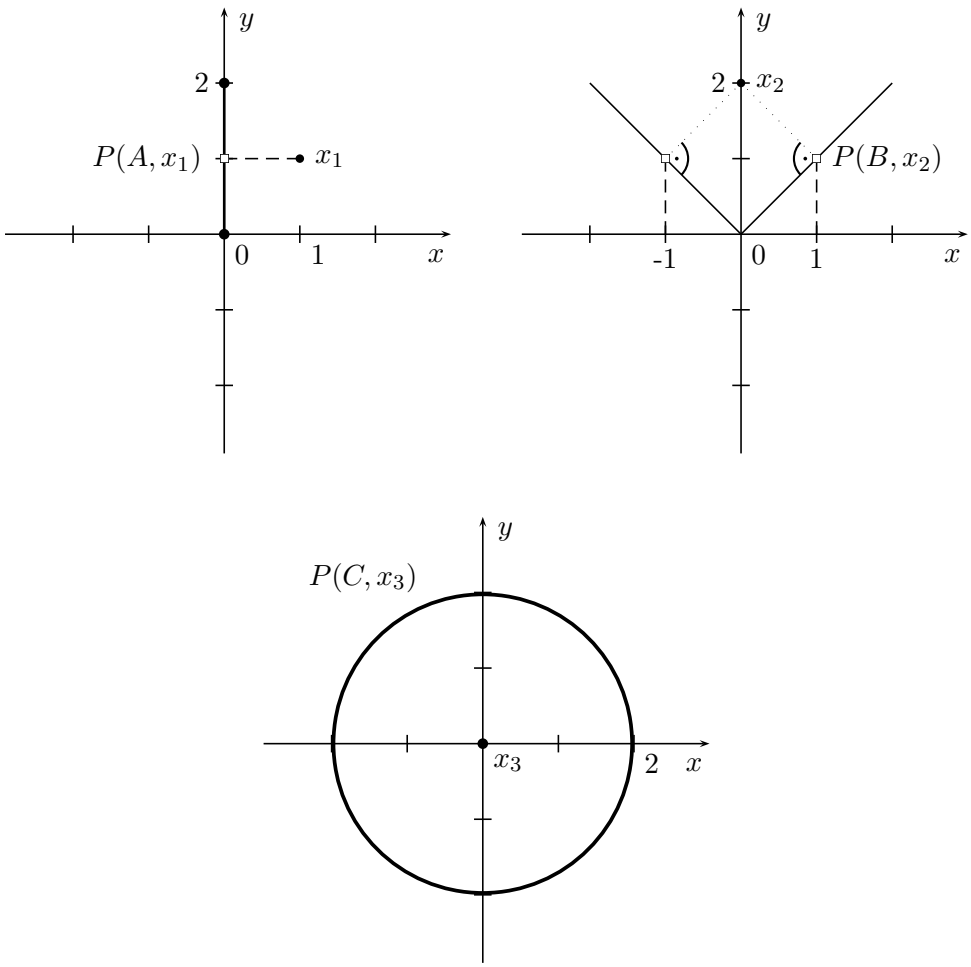
$$A = \{(0,y) \mid y \in [0,2]\}\ ,\ B = \{(x,y) \mid y = |x|\}\ ,\ C = \{(x,y) \mid x^2 + y^2 = 4\}$$

oraz punkty: $x_1 = (1,1), x_2 = (0,2), x_3 = (0,0)$.

Wówczas:

$$P(A,x_1) = \{(0,1)\}\ ,\ P(B,x_2) = \{(-1,0),(1,0)\}\ ,\ P(C,x_3) = C.$$

[1] zbiór ten nazywany jest także stożkiem Mordukhovicha



Rysunek 2.1. Przykłady punktów danego zbioru, których odległość od danego punktu jest najmniejsza

Przykład 2.2. Przykład ilustrujący definicję stożka normalnego do S w punkcie $\overline{x}$. Rozważmy zbiór $S \subset \mathbb{R}^2$:

$$S = \{(x, y) \mid (x \in [-1, 0] \quad \text{i} \quad y = 0) \quad \text{lub} \quad (x = 0 \quad \text{i} \quad y \in [-1, 0])\}$$

1. Niech x_s będzie punktem zbioru S takim, że:

$$x_s = (\overline{x}, 0),\ \overline{x} \in (-1, 0).$$

Stożkiem normalnym do zbioru S w punkcie x_s (rysunek 2.2) jest zbiór:

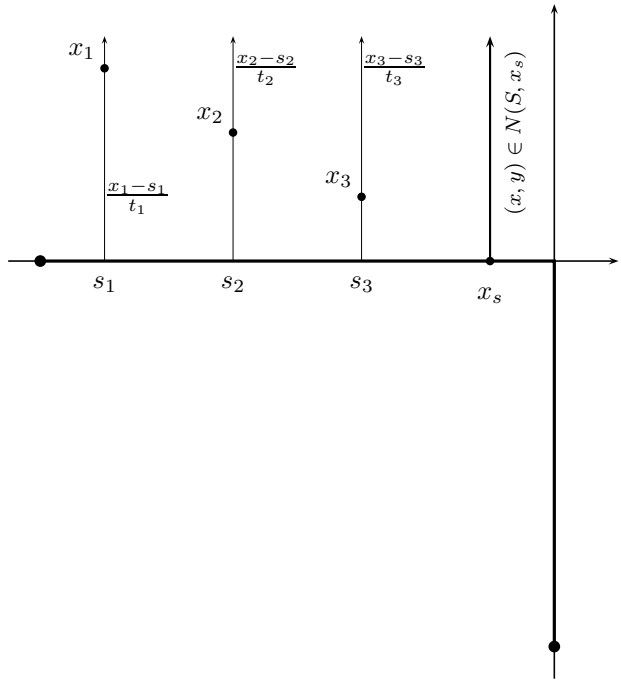
$$N(S, x_s) = \{(\overline{x}, y) \mid y \in \mathbb{R}\}.$$

2. Niech x_s będzie punktem zbioru S takim, że:

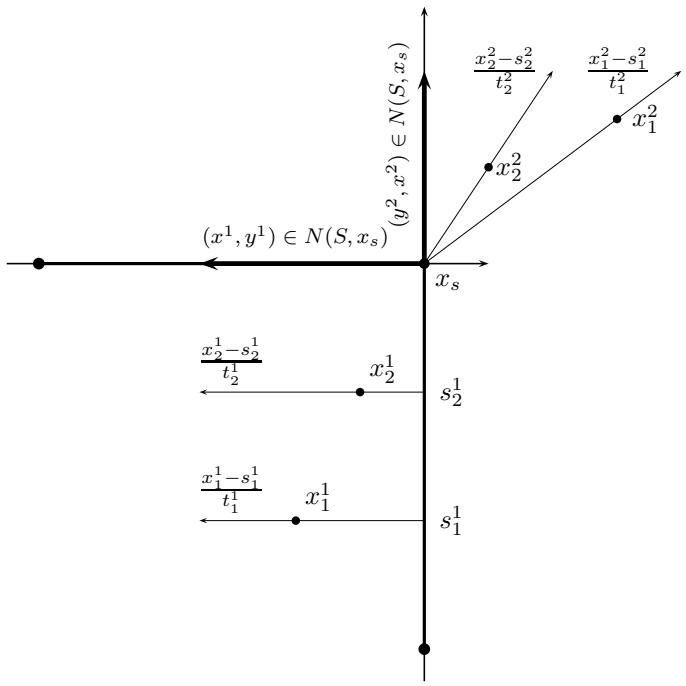
$$x_s = (0, 0)$$

Stożkiem normalnym do zbioru S w punkcie x_s (rysunek 2.3) jest zbiór:

$$\begin{aligned} N(S, x_s) = \{(x, y) \mid (x \geqslant 0 \quad \text{i} \quad y \geqslant 0) \quad \text{lub} \\ (x = 0 \quad \text{i} \quad y < 0) \quad \text{lub} \\ (x < 0 \quad \text{i} \quad y = 0)\}. \end{aligned}$$



Rysunek 2.2. Przykład elementu (x, y) stożka $N(S, x_s)$, gdzie $x_s = (\overline{x}, 0)$, $\overline{x} \in (-1, 0)$



Rysunek 2.3. Przykład elementów $(x^1, y^1), (x^2, y^2)$ stożka $N(S, x_s)$, gdzie $x_s = (0, 0)$

Na rysunkach 2.2 oraz 2.3 przedstawiono przykłady ciągów punktów $x_1^n, x_2^n, x_3^n, \ldots$ zbieżnych do x_s oraz odpowiadających im ciągów punk-

tów $s_1^n, s_2^n, s_3^n, \ldots$, które przy pomocy ciągów $t_1^n, t_2^n, t_3^n, \ldots$ tworzyły ciąg $\frac{x_1^n - s_1^n}{t_1^n}, \frac{x_2^n - s_2^n}{t_2^n}, \frac{x_3^n - s_3^n}{t_3^n}, \ldots$ zbieżny do elementu stożka $N(S, x_s)$.

Twierdzenie 2.1 (Warunek konieczny i wystarczający istnienia słabego ostrego minimum). *Niech funkcja $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ przyjmuje stałą i skończoną wartość na domkniętym zbiorze $S \subset \mathbb{R}^n$ i niech $\overline{x} \in S$, $m \geqslant 1$. Funkcja f osiąga w $\overline{x}$ lokalne słabe ostre minimum rzędu m wtedy i tylko wtedy, gdy dla dowolnego $y \in N(S, \overline{x})$ takiego, że $\|y\| = 1$ oraz dla dowolnych ciągów $\{x_j\}$, $\{s_j\}$ takich, że $lim_{j\to\infty} x_j = \overline{x}$, $s_j \in P(S, x_j)$ i $lim_{j\to\infty} \frac{x_j - s_j}{\|x_j - s_j\|} = y$ spełniony jest warunek:*

$$\liminf_{j\to\infty} \frac{f(x_j) - f(s_j)}{\|x_j - s_j\|^m} > 0. \tag{2.1}$$

Metoda oparta na twierdzeniu zacytowanym powyżej, mimo że skuteczna, jest niestety bardzo skomplikowana. W związku z tym ma małe zastosowanie praktyczne. Zaczęto więc poszukiwać kolejnych sposobów rozwiązania problemu (1). Jak pokażą dalsze rozważania, nie udało się jednak znaleźć twierdzenia, które można by było stosować w tak szerokim zakresie jak twierdzenie 2.1. Okazuje się, że mniej skomplikowane metody są zarazem mniej uniwersalne.

2.2. Słabe ostre minima a pojęcie pochodnej

Klasyczna teoria optymalizacji łączy często poszukiwanie wartości ekstremalnych funkcji z jej różniczkowalnością. Do znajdowania wartości minimalnych lub maksymalnych funkcji wykorzystuje się między innymi metody posługujące się pojęciem pochodnej. Na przykład dla funkcji rzeczywistych jednej zmiennej mamy:

— metodę porównywania znaków pochodnej (metoda opiera się na twierdzeniu: „Niech f będzie funkcją różniczkowalną w pewnym otoczeniu $(x_0 - \delta, x_0 + \delta)$, $\delta > 0$ punktu x_0.
  a) Jeżeli $f'(x) < 0$ dla $x \in (x_0 - \delta, x_0)$ i $f'(x) > 0$ dla $x \in (x_0, x_0 + \delta)$, to funkcja osiąga w punkcie x_0 minimum lokalne.
  b) Jeżeli $f'(x) > 0$ dla $x \in (x_0 - \delta, x_0)$ i $f'(x) < 0$ dla $x \in (x_0, x_0 + \delta)$, to funkcja osiąga w punkcie x_0 maksimum lokalne.”);

— metodę pochodnych wyższych rzędów (metoda opiera się na twierdzeniu: „Niech f będzie funkcją rzeczywistą n-krotnie różniczkowalną w sposób ciągły w otoczeniu punktu x_0, $n > 1$ oraz:
  $f'(x_0) = f''(x_0) = \cdots = f^{(n-1)}(x_0) = 0, f^{(n)}(x_0) \neq 0$.
  a) Jeżeli n jest liczbą nieparzystą, to funkcja f nie ma w punkcie x_0 ekstremum.
  b) Jeżeli n jest liczbą parzystą, to funkcja f osiąga w punkcie x_0 ekstremum lokalne właściwe (minimum, gdy $f^{(n)}(x_0) > 0$, zaś maksimum gdy $f^{(n)}(x_0) < 0$).”);

Często najwięcej informacji o monotoniczności i wartościach ekstremalnych funkcji dostarcza jej pochodna, a ponieważ w rzeczywistości istnieje wiele układów, które nie dają się opisać za pomocą funkcji różniczkowalnych, stosuje się różne metody, które pozwalają mimo wszystko badać funkcję z wykorzystaniem pojęcia pochodnej. W tym celu albo aproksymuje się niegładkie problemy za pomocą funkcji gładkich (ale to rozwiązanie niesie ze

sobą wiele nowych trudności związanych głównie z błędami aproksymacji), albo dla niegładkich funkcji definiuje się pojęcia, które w pewnym sensie są odpowiednikami pochodnej funkcji gładkiej. Ta „odpowiedniość" polega na analogiach między twierdzeniami wykorzystującymi nowo zdefiniowane pojęcia do znalezienia punktów ekstremalnych a twierdzeniami występującymi w klasycznej teorii optymalizacji.

I tak na przykład dla niegładkich funkcji wypukłych wprowadzono pojęcie subgradientu. Subgradient w ustalonym punkcie jest to wektor, który ma tę własność, że hiperpłaszczyzna przechodząca przez ustalony punkt, generowana przez ten wektor, jest dolnym przybliżeniem funkcji. Zbiór wszystkich subgradientów to subróżniczka. Ta koncepcja pozwala stosować bardzo podobne metody, z jakich korzysta się w przypadkach gładkich.

Zmodyfikowane pojęcie pochodnej pojawia się więc również podczas rozważania słabych ostrych minimów. W celu badania istnienia lokalnego ścisłego minimum rzędu m wprowadzono następujący odpowiednik pojęcia klasycznej pochodnej kierunkowej:

$$\underline{d}^m f(x,y) := \liminf_{(t,v)\to(0^+,y)} \frac{f(x+tv)-f(x)}{t^m} \tag{2.2}$$

Natomiast w przypadku, gdy S zawiera więcej niż jeden element, pochodna kierunkowa uogólnia się w następujący sposób:

Definicja 2.2. Niech $S \subset \mathbb{R}^n$ będzie niepustym, domkniętym zbiorem, $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ przyjmuje wartości skończone na $\operatorname{bd} S$, $x \in \operatorname{bd} S$ oraz $y \in \mathbb{R}^n$. Pochodną kierunkową definiujemy jako:

$$\underline{d}_S^m f(x,y) := \liminf_{\substack{s\to_{\operatorname{bd} S} x \\ (t,v)\to(0^+,y)}} \frac{f(s+tv)-f(s)}{t^m} \tag{2.3}$$

gdzie $s\to_{\operatorname{bd} S} x$ oznacza, że $s \to x$ oraz $s \in \operatorname{bd} S$.

Należy zauważyć, że w przypadku, gdy S jest zbiorem jednoelementowym pojęcia (2.2) i (2.3) są równoważne. Warto też nadmienić, że $\underline{d}^m f(x,y)$ i $\underline{d}_S^m f(x,y)$ mogą się pokrywać w przypadkach, kiedy f jest funkcją „wystarczająco gładką" [12].

Przy formułowaniu warunków istnienia słabego ostrego minimum będziemy też używać pojęcia ścisłej różniczkowalności.

Definicja 2.3. O funkcji $\mathrm{f}\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ mówimy, że jest ściśle różniczkowalna w $\overline{x}$, jeżeli $f(\overline{x})$ jest wartością skończoną oraz istnieje taki wektor $\nabla f(\overline{x})$ (nazywany gradientem funkcji f w punkcie $\overline{x}$), że dla dowolnego $y \in \mathbb{R}^n$ spełniony jest warunek:

$$\lim_{(v,w,t)\to(\overline{x},y,0^+)} \frac{f(v+tw)-f(v)}{t} = \langle \nabla f(\overline{x}), y\rangle\,. \tag{2.4}$$

Oczywiście, istnieją zależności między ścisłą różniczkowalnością funkcji f a jej pochodnymi kierunkowymi różnego typu. Otóż, jeśli funkcja $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ przyjmuje wartości skończone na $\operatorname{bd} S$ i $\overline{x} \in \operatorname{bd} S$ oraz jeżeli f jest ściśle różniczkowalna w $\overline{x}$, to dla dowolnego $y \in \mathbb{R}^n$ zachodzi równość:

$$\underline{d}_S^1 f(\overline{x},y) = \langle \nabla f(\overline{x}), y\rangle = \underline{d}^1 f(\overline{x},y). \tag{2.5}$$

2.2.1. Warunek wystarczający istnienia słabego ostrego minimum

Pojęcie uogólnionej pochodnej kierunkowej wykorzystano do sformułowania warkunku wystarczającego istnienia lokalnego słabego ostrego minimum rzędu m.

Twierdzenie 2.2. *Niech funkcja $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ przyjmuje wartość stałą i skończoną na domkniętym zbiorze $S \subset \mathbb{R}^n$, $\overline{x} \in \operatorname{bd} S$. Jeżeli $m \geqslant 1$ oraz*

$$\underline{d}_S^m f(\overline{x}, y) > 0 \quad \textit{dla dowolnego} \quad y \in N(S, \overline{x}) \setminus \{0\}, \tag{2.6}$$

to funkcja f osiąga w $\overline{x}$ lokalne słabe ostre minimum rzędu m.

Niestety, powyższy warunek nie zawsze rozstrzyga fakt istnienia lokalnego słabego ostrego minimum.

Przykład 2.3. Rozważmy funkcję $f\colon \mathbb{R}^2 \to \overline{\mathbb{R}}$

$$f(x,y) := \begin{cases} x^m & , x \geqslant 0 \quad \text{i} \quad 0 \leqslant y \leqslant 1 \\ +\infty & , \text{ w przeciwnym wypadku} \end{cases}$$

oraz zbiór $S = \{(0,y) \mid 0 \leqslant y \leqslant 1\}$. Zauważmy, że $f(x,y) = 0$ dla dowolnego punktu $(x,y) \in S$.

Niech $(0,\overline{y}) \in S$ oraz $(x,y) \in \mathbb{R}^2$. Rozważmy dwa przypadki:

1. $x \geqslant 0$ i $y \in [0,1]$. Wtedy $f(x,y) = x^m$,

$$\operatorname{dist}^m((x,y),S) = \|(x,y) - (0,y)\| = \|(x,0)\| = (\sqrt{x^2})^m = |x|^m = x^m,$$

zatem

$$f(x,y) - f(0,\overline{y}) = x^m - 0 = x^m \geqslant 1 \cdot \operatorname{dist}^m((x,y),S).$$

2. $x < 0$ lub $y \notin [0,1]$. Wtedy $f(x,y) = +\infty$,

$$f(x,y) - f(0,\overline{y}) = +\infty - 0 = +\infty > \beta \cdot \operatorname{dist}^m((x,y),S)$$

dla dowolnego $\beta > 0$, gdyż $\operatorname{dist}^m((x,y),S)$ przyjmuje wartość skończoną.

A zatem S jest zbiorem słabych ostrych minimów funkcji f.

Niech $y \in (0,1)$. Wtedy $N(S,(0,y)) = \{(z,0) \mid z \in \mathbb{R}\}$. Sprawdźmy, ile wynosi wartość pochodnej kierunkowej w kierunku wektora $(z,0) \in N(S,(0,y))$ w punkcie $(0,y)$:

$$\begin{aligned} \underline{d}_S^m f((0,y),(z,0)) &= \liminf_{\substack{(\overline{x},\overline{y}) \to \operatorname{bd} S(0,y) \\ (x_0,y_0) \to (z,0) \\ t \to 0^+}} \frac{f(\overline{x} + tx_0, \overline{y} + ty_0) - f(\overline{x},\overline{y})}{t^m} \\ &= \begin{cases} z^m & , z \geqslant 0 \\ +\infty & , z < 0 \end{cases}, \end{aligned}$$

a więc $\underline{d}_S^m f((0,y),(z,0)) > 0$ dla dowolnego $(z,0) \in N(S(0,y)) \setminus \{(0,0)\}$.

Rozważmy teraz stożek $N(S,(0,0)) = \{(x,y) \mid y \leqslant 0\}$.
Dla $(z,0) \in N(S,(0,0))$, $z \neq 0$ mamy:

$$\underline{d}_S^m f((0,0),(z,0)) = \begin{cases} z^m & , z > 0 \\ +\infty & , z < 0 \end{cases},$$

więc pochodna kierunkowa przyjmuje wartości dodatnie, lecz jeżeli jako kierunek obierzemy wektor $(0,-1) \in N(S,(0,0))$, to:

$$\underline{d}_S^m f((0,0),(0,-1)) = 0.$$

Podobną sytuację otrzymamy w przypadku stożka
$N(S,(0,1)) = \{(x,y) \mid y \geqslant 0\}$:

$$\underline{d}_S^m f((0,1),(0,1)) = 0.$$

Warunek (2.6) nie jest zatem spełniony w punktach $(0,0)$ i $(0,1)$, mimo, że funkcja f osiąga w nich słabe ostre minima.

Udało się ustalić [12], co potwierdza powyższy przykład, że warunek (2.6) nie zachodzi w przypadku, kiedy S jest wypukłym zbiorem zawierającym więcej niż jeden element, ale mającym puste wnętrze, a $\overline{x} \in S$ (w którym rozpatrywana funkcja osiąga słabe ostre minimum) nie należy do „relatywnego wnętrza”[2] (ang. *relative interior*) zbioru S.

2.2.2. Warunek konieczny i wystarczający istnienia słabego ostrego minimum

Aby móc zacytować kolejne warunki, należy wcześniej zdefiniować pojęcie stożka stycznego[3] do zbioru.

Definicja 2.4. Niech C będzie podzbiorem przestrzeni $\mathbb{R}^n$, $\overline{x} \in C$. Stożkiem stycznym do C w punkcie $\overline{x}$ nazywamy zbiór:

$$K(C,\overline{x}) := \{y \in \mathbb{R}^n \mid \begin{array}{ll} \text{istnieje ciąg} & (t_j, y_j) \to (0^+, y) \\ \text{taki, że} & x + t_j y_j \in C\}. \end{array} \tag{2.7}$$

W pozycji [13] podany jest warunek konieczny istnienia słabego ostrego minimum rzędu m:

Twierdzenie 2.3. *Niech funkcja $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ przyjmuje wartość skończoną i stałą na zbiorze $S \subset \mathbb{R}^n$ oraz niech punkt $x \in S$ będzie lokalnym słabym ostrym minimum rzędu m funkcji f. Wtedy istnieje $\beta > 0$ taka, że:*

$$\underline{d}^m f(\overline{x}, y) \geqslant \beta \cdot \operatorname{dist}^m(y, K(S,\overline{x})) \quad \textit{dla dowolnego} \quad y \in \mathbb{R}^n. \tag{2.8}$$

[2] Relatywnym wnętrzem podzbioru wypukłego S przestrzeni $\mathbb{R}^n$ nazywamy wnętrze zbioru S rozumianego jako podzbiór swojej otoczki afinicznej

[3] zbiór (2.7) podany w definicji 2.4 w pozycji [12] jest określany nazwą „contingent cone” (z łac. *contigare* znaczy dotykać); nazwa stożka stycznego (ang. *tangent cone*) pochodzi z [5] i została wprowadzona dla zbioru (2.7) przy założeniu, że C jest zbiorem wypukłym; w tej pracy zbiór (2.7) będzie nazywany stożkiem stycznym przy założeniu, że C jest dowolnym podzbiorem $\mathbb{R}^n$

Twierdzenie to dało możliwość sformułowania drugiego już warunku koniecznego i wystarczającego, w którym skomplikowane badanie znaku granicy ciągu pewnych ilorazów różnicowych z twierdzenia 2.1 zastąpiono badaniem znaku zmodyfikowanej pochodnej kierunkowej.

Twierdzenie 2.4 (warunek konieczny i wystarczający istnienia słabego ostrego minimum). *Niech funkcja $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ przyjmuje wartość skończoną i stałą na domkniętym zbiorze $S \subset \mathbb{R}^n$ oraz $x \in \operatorname{bd} S$. Jeżeli są spełnione warunki:*

$$K(S,\overline{x}) \cap N(S,\overline{x}) = \{0\}\,, \tag{2.9}$$

$$\underline{d}^m f(\overline{x},y) = \underline{d}^m_S f(\overline{x},y) \quad \textit{dla dowolnego} \quad y \in N(S,\overline{x}) \setminus \{0\}\,, \tag{2.10}$$

to funkcja f osiąga w punkcie $\overline{x}$ lokalne słabe ostre minimum rzędu m wtedy i tylko wtedy, gdy zachodzi warunek (2.6) postaci:

$$\underline{d}^m_S f(\overline{x},y) > 0 \quad \textit{dla dowolnego} \quad y \in N(S,\overline{x}) \setminus \{0\}.$$

Tak, jak wspomniano już wcześniej, prostszy do sprawdzenia warunek (2.6) nie daje się stosować w tych wszystkich przypadkach, w jakich warunek (2.1) był skuteczny. Poniżej zostaną rozważone problemy, na jakie napotykamy chcąc wykorzystać warunek z twierdzenia 2.4.

W przypadku, kiedy S jest zbiorem wypukłym, stożek normalny do zbioru S w punkcie $\overline{x}$ ma postać [6]:

$$N(S,\overline{x}) = \{z \in \mathbb{R}^n \mid y^T \cdot z \leqslant 0 \quad \text{dla dowolnego} \quad y \in K(S,\overline{x})\}.$$

Dla dowolnego $y \in K(S,\overline{x}) \cap N(S,\overline{x})$ mamy zatem: $\langle y,y\rangle \leqslant 0$, a więc $K(S,\overline{x}) \cap N(S,\overline{x}) = \{0\}$.

Warunek (2.9) jest zatem zawsze spełniony, kiedy S jest zbiorem wypukłym.

W przypadku zbiorów niewypukłych warunek ten może, choć nie musi zachodzić, o czym świadczą dwa poniższe przypadki.

Wróćmy do zbioru S rozważanego w przykładzie 2.2. W tym przypadku dla $\overline{x} = (0,0)$ mamy:

$$\begin{aligned} K(S,\overline{x}) &= \{(x,y) \mid (x \leqslant 0 \quad \text{i} \quad y = 0) \qquad \text{lub} \quad (x = 0 \quad \text{i} \quad y \leqslant 0)\},\\ N(S,\overline{x}) &= \{(x,y) \mid (x \geqslant 0 \quad \text{i} \quad y \geqslant 0) \qquad \text{lub} \quad (x = 0 \quad \text{i} \quad y \leqslant 0)\\ &\qquad\qquad\qquad\qquad\qquad\qquad \text{lub} \quad (x < 0 \quad \text{i} \quad y = 0)\}, \end{aligned}$$

a więc $K(S,\overline{x}) \cap N(S,\overline{x}) \supset S$ i warunek (2.9) nie zachodzi.

Jeżeli natomiast rozważymy zbiór niewypukły $S = \{(x,x^2) \mid x \in [-1,1]\}$ oraz punkt $\overline{x} = (0,0)$, to:

$$\begin{aligned} K(S,\overline{x}) &= \{(0,0)\},\\ N(S,\overline{x}) &= \{(0,y) \mid y \in \mathbb{R}\}, \end{aligned}$$

a więc $K(S,\overline{x}) \cap N(S,\overline{x}) = \{(0,0)\}$ i warunek (2.9) zachodzi.

Wynika stąd, że twierdzenie 2.4 można na pewno stosować w przypadku, kiedy zbiory S są wypukłe, o ile zachodzi warunek (2.10)

Przykład 2.4. Powróćmy do przykładu 2.3. Otrzymano tam, że:

$$\underline{d}_S^m f((0,0),(0,-1)) = 0.$$

Zauważmy, że:

$$\begin{aligned}\underline{d}^m f((0,0),(0,-1)) &= \liminf_{\substack{(\overline{x},\overline{y})\to(0,-1)\\ t\to 0^+}} \frac{f(0+t\cdot\overline{x}, 0+t\cdot\overline{y}) - f((0,0))}{t^m}\\ &= \liminf_{\substack{(\overline{x},\overline{y})\to(0,-1)\\ t\to 0^+}} \frac{f(t\cdot\overline{x}, t\cdot\overline{y})}{t^m} = +\infty,\end{aligned}$$

a więc

$$\underline{d}_S^m f((0,0),(0,-1)) \neq \underline{d}^m f((0,0),(0,-1)).$$

Podobnie:

$$\underline{d}_S^m f((0,1),(0,1)) = 0.$$

oraz

$$\begin{aligned}\underline{d}^m f((0,1),(0,1)) &= \liminf_{\substack{(\overline{x},\overline{y})\to(0,1)\\ t\to 0^+}} \frac{f(0+t\cdot\overline{x}, 1+t\cdot\overline{y}) - f((0,1))}{t^m}\\ &= \liminf_{\substack{(\overline{x},\overline{y})\to(0,1)\\ t\to 0^+}} \frac{f(t\cdot\overline{x}, 1+t\cdot\overline{y})}{t^m} = +\infty,\end{aligned}$$

a więc

$$\underline{d}_S^m f((0,1),(0,1)) \neq \underline{d}^m f((0,1),(0,1)).$$

Warunek (2.10) nie jest spełniony, a więc nie można rozstrzygnąć za pomocą warunku (2.6), czy funkcja f osiąga w punktach (0,0), (0,1) słabe ostre minimum rzędu m.

Warunek (2.6) jest znacznie łatwiejszy do sprawdzania niż warunek twierdzenia 2.1. Niestety, twierdzenie 2.4 zachodzi tylko w przypadkach, kiedy spełnione są założenia (2.9) i (2.10).

Jak pokazano już wcześniej, warunek (2.9) zachodzi zawsze dla zbiorów S wypukłych, natomiast (2.10) jest spełniony na pewno, kiedy funkcja f jest ściśle różniczkowalna w $\overline{x}$. Nie można zatem stosować twierdzenia 2.4 w przypadku słabych ostrych minimów rzędu pierwszego nawet dla najprostszych funkcji niegładkich. A zatem zyskując prostszy warunek w twierdzeniu 2.4 straciliśmy ogólność i uniwersalność, jaką dawało nam twierdzenie 2.1.

Rozdział 3

Warunek konieczny i wystarczający istnienia słabego ostrego minimum dla maksimum ze skończonych i ściśle różniczkowalnych funkcji

W rozdziale drugim przedstawiono twierdzenia, które formułowano z myślą o ich zastosowaniu dla jak największej ilości klas funkcji f. Okazuje się jednak, że dla niektórych rodzajów funkcji, łatwiejsze do sprawdzenia twierdzenie 2.4 jest niemożliwe do zastosowania. Dzieje się tak nawet w przypadkach bardzo prostych, ale jednak niegładkich funkcji, kiedy szukamy lokalnego słabego ostrego minimum pierwszego rzędu ($m = 1$).

Rozważmy teraz specjalną klasę niegładkich funkcji $f\colon \mathbb{R}^n \to \mathbb{R}$ typu:

$$f(x) := \max\{f_i(x) \mid i \in I\}, \tag{3.1}$$

gdzie $I = \{1, \ldots, p\}$ jest skończonym zbiorem indeksów, zaś funkcje $f_i\colon \mathbb{R}^n \to \mathbb{R}$, $i \in I$ są ściśle różniczkowalne.

Jest to klasa funkcji, które nie spełniają założenia (2.10) twierdzenia 2.4.

Przykład 3.1. Rozpatrzmy funkcję $f\colon \mathbb{R}^2 \to \mathbb{R}$ daną wzorem:

$$f(x_1, x_2) = \max\{x_1, x_2, -x_2\}. \tag{3.2}$$

Niech $S = \{(x_1, 0) \mid x_1 \leqslant 0\}$, $\overline{x} = (0, 0) \in \operatorname{bd} S$, $(x_1, x_2) \in \mathbb{R}^2$. Wtedy $f(x, y) = 0$ dla dowolnego $(x, y) \in S$ oraz

$$f(x_1, x_2) - f(0, 0) = f(x_1, x_2) = \max\{x_1, x_2, -x_2\}.$$

Zachodzą dwa przypdaki

a) jeśli $x_1 \leqslant 0$, to:

$$\operatorname{dist}((x_1, x_2), S) = \|(x_1, x_2) - (x_1, 0)\| = \|(0, x_2)\| = \sqrt{x_2^2} = |x_2| \quad \text{oraz}$$

$$f(x_1, x_2) - f(0, 0) = \max\{x_1, x_2, -x_2\} \geqslant |x_2| = 1 \cdot \operatorname{dist}((x_1, x_2), S).$$

b) jeśli $x_1 > 0$, to:

$$\begin{aligned}\operatorname{dist}((x_1, x_2), S) &= \|(x_1, x_2) - (0, 0)\| = \sqrt{x_1^2 + x_2^2} \leqslant \sqrt{x_1^2} + \sqrt{x_2^2} \\ &= |x_1| + |x_2| = x_1 + |x_2| \leqslant 2 \cdot \max\{x_1, x_2, -x_2\} \quad \text{oraz}\end{aligned}$$

$$f(x_1, x_2) - f(0, 0) = \max\{x_1, x_2, -x_2\} \geqslant \frac{1}{2} \operatorname{dist}((x_1, x_2), S).$$

Funkcja f osiąga zatem słabe ostre minimum rzędu pierwszego w punkcie $\overline{x} = (0,0)$.

Sprwawdźmy, czy można zastosować dla funkcji f twierdzenie 2.4. Założenie (2.9) jest spełnione gdyż $K(S,\overline{x}) = S$, zaś $N(S,\overline{x}) = \{(x_1,x_2)\,|\,x_1 \geqslant 0\}$, zatem $K(S,\overline{x}) \cap N(S,\overline{x}) = \{(0,0)\}$.

Natomiast założenie (2.10) nie zachodzi. Istotnie, niech $y = (0,1)$. Wtedy

$$\begin{aligned}
\underline{d}^1 f(\overline{x},y) &= \liminf_{\substack{(v_1,v_2)\to(0,1)\\ t\to 0^+}} \frac{f(0+tv_1,0+tv_2)-f(0,0)}{t} \\
&= \liminf_{\substack{(v_1,v_2)\to(0,1)\\ t\to 0^+}} \frac{\max\{tv_1,tv_2,-tv_2\}}{t} \\
&= \liminf_{\substack{(v_1,v_2)\to(0,1)\\ t\to 0^+}} \frac{t\cdot\max\{v_1,v_2,-v_2\}}{t} \\
&= \liminf_{\substack{(v_1,v_2)\to(0,1)\\ t\to 0^+}} v_2 = 1,
\end{aligned}$$

$$\underline{d}^1_S f(\overline{x},y) = \liminf_{\substack{(v_1,v_2)\to(0,1)\\ t\to 0^+\\ (s_1,s_2)\to_{\mathrm{bd}\,S}(0,0)}} \frac{f(s_1+tv_1,s_2+tv_2)-f(s_1,s_2)}{t} = 0\,,$$

zatem $\underline{d}^1 f(\overline{x},y) \neq \underline{d}^1_S f(\overline{x},y)$ dla $y \in N(S,\overline{x}) \setminus \{(0,0)\}$. Stąd wynika, iż nie można posłużyć się twierdzeniem 2.4 w przypadku danej funkcji.

Przykład 3.1 pokazuje, że podczas minimalizacji polegającej na poszukiwaniu lokalnych słabych ostrych minimów pierwszego rzędu dla funkcji typu (3.1) nie możemy stosować twierdzenia 2.4. Dla tej klasy funkcji został sformułowany oddzielny warunek konieczny i wystarczający [10]:

Twierdzenie 3.1. *Załóżmy, że funkcja f typu (3.1) przyjmuje wartość stałą na domkniętym zbiorze $S \subset \mathbb{R}^n$ oraz $\overline{x} \in S$. Wówczas równoważne są warunki:*

a) funkcja f osiąga w $\overline{x}$ lokalne słabe ostre minimum pierwszego rzędu,

b) dla każdego $y \in N(S,\overline{x})$, $\|y\| = 1$ istnieje stała $\delta > 0$ taka, że dla każdego $x \in B(\overline{x},\delta) \setminus S$ i $s \in P(S,x)$ spełniających warunek

$$\left\| \frac{x-s}{\|x-s\|} - y \right\| \leqslant \delta$$

istnieje indeks $i \in I(s)$ taki, że $\langle \nabla f_i(\overline{x}), y\rangle > 0$, gdzie przez $I(s)$ rozumiemy zbiór $I(s) := \{i \in I \mid f_i(s) = f(s)\}$.

Przykład 3.2. Powróćmy do funkcji rozważanej w przykładzie 3.1. Funkcja f ma postać: $f(x_1,x_2) = \max\{f_1(x_1,x_2), f_2(x_1,x_2), f_3(x_1,x_2)\}$, gdzie $f_1(x_1,x_2) = x_1$, $f_2(x_1,x_2) = x_2$, $f_3(x_1,x_2) = -x_2$. Pokażemy, że w punkcie $\overline{x} = (0,0)$ spełniony jest warunek (b) twierdzenia 3.1.

Niech $y = (y_1,y_2) \in N(S,\overline{x})$, $\|y\| = 1$. Może zachodzić jeden z trzech przypadków:

a) $y_1 \geqslant 0$ i $y_2 > 0$;
wtedy $2 \in I(s)$ i $\langle \nabla f_2(\overline{x}), y\rangle = \langle (0,1),(y_1,y_2)\rangle = y_2 > 0$.
a) $y_1 \geqslant 0$ i $y_2 < 0$;
wtedy $3 \in I(s)$ i $\langle \nabla f_3(\overline{x}), y\rangle = \langle (0,-1),(y_1,y_2)\rangle = -y_2 > 0$.
c) $y_1 = 1$ i $y_2 = 0$;
wtedy $1 \in I(s) = I(\overline{x}) = \{1,2,3\}$ i $\langle \nabla f_1(\overline{x}), y\rangle = \langle (1,0),(1,0)\rangle = 1 > 0$.
A zatem dla funkcji f zdefiniowanej w przykładzie 3.1 (wzór (3.2)) twierdzenie 3.1 zachodzi.

Warunek (b) twierdzenia 3.1 jest skomplikowany, jednak nie można go zastąpić prostszym warunkiem postaci:

$$\max\{\langle \nabla f_i(\overline{x}), y\rangle \mid i \in I(\overline{x})\} > 0 \quad \text{dla dowolnego} \quad y \in N(S,\overline{x}) \setminus \{0\}. \tag{3.3}$$

Problem, jaki napotkamy przy warunku (3.3), przedstawiono w poniższym przykładzie.

Przykład 3.3. Niech funkcja $f\colon \mathbb{R}^n \to \mathbb{R}$ będzie określona jako:

$$f(x_1,x_2) := \max\{f_1(x_1,x_2), f_2(x_1,x_2), f_3(x_1,x_2), f_4(x_1,x_2)\},$$

gdzie $f_1(x_1,x_2) := 0$, $f_2(x_1,x_2) := x_1^3$, $f_3(x_1,x_2) := x_2^3$, $f_4(x_1,x_2) := x_1 + x_2$. Zdefiniujmy również zbiór S:

$$S := \{(x_1,x_2) \mid x_1 \leqslant 0, x_2 \leqslant 0\},$$

na którym funkcja f przyjmuje stałą wartość równą zero

Weźmy punkt $\overline{x} = (0,0) \in S$. Mamy wtedy:

$$\begin{aligned} I(\overline{x}) &= \{1,2,3,4\}, \\ N(S,\overline{x}) &= \{(y_1,y_2) \mid y_1 \geqslant 0, y_2 \geqslant 0\}, \end{aligned}$$

$$\begin{aligned} \max\{\langle \nabla f_i(\overline{x}), (y_1,y_2)\rangle \mid i \in I(\overline{x})\} &\geqslant \langle \nabla f_4(\overline{x}), (y_1,y_2)\rangle \\ &= \langle (1,1),(y_1,y_2)\rangle = y_1 + y_2 > 0 \end{aligned} \tag{3.4}$$

dla dowolnego $(y_1,y_2) \in N(S,\overline{x}) \setminus \{(0,0)\}$.

Warunek (3.3) jest zatem spełniony w punkcie $\overline{x}$. Sprawdźmy, czy funkcja f osiąga w $\overline{x}$ lokalne słabe ostre minimum pierwszego rzędu. W każdym dostatecznie bliskim otoczeniu punktu $\overline{x}$ znajdziemy punkt $x = (x_1,x_2)$ taki, że $x_1 < 0$ i $0 < x_2 < -x_1$. Mamy wtedy: $x_2^3 > 0 > x_1 + x_2$ oraz $f(x) = x_2^3$.

Przypuśćmy, że funkcja f osiąga w $\overline{x}$ lokalne słabe ostre minimum rzędu pierwszego. Wtedy istnieje stała $\beta > 0$ taka, że dla punktów x zdefiniowanych powyżej spełniony jest warunek:

$$\begin{aligned} f(x) - f(\overline{x}) = x_2^3 \geqslant \beta \cdot \operatorname{dist}(x,S) &= \beta \cdot \|(x_1,x_2) - (x_1,0)\| \\ &= \beta \cdot \|(0,x_2)\| = \beta \cdot |x_2| = \beta \cdot x_2. \end{aligned}$$

Dzieląc obie strony nierówności przez x_2 otrzymamy

$$x_2^2 \geqslant \beta$$

i przy $x_2 \to 0^+$ mamy:

$$\beta \leqslant 0,$$

co jest sprzeczne z naszym założeniem.

Zatem funkcja f nie osiąga w $\overline{x}$ lokalnego słabego ostrego minimum pierwszego rzędu, mimo iż warunek (3.3) jest spełniony.

Sprawdźmy, jak w przypadku punktu $\overline{x}$ zadziała twierdzenie 3.1.

Niech $y = (0, 1)$. Oczywiście $y \in N(S, \overline{x})$ i $\|y\| = 1$. Dla dowolnego $\delta > 0$ istnieje punkt $x = (x_1, x_2) \in B(\overline{x}, \delta)$ taki, że:

a) $x_1 < 0$ i $0 < x_2 < -x_1$,

b) $s = (x_1, 0) \in P(S, \overline{x})$, $I(s) = \{1, 3\}$.

c)

$$\left\| \frac{x - s}{\|x - s\|} - y \right\| = \left\| \frac{(x_1, x_2) - (x_1, 0)}{\|(x_1, x_2) - (x_1, 0)\|} - (0, 1) \right\| = \left\| \frac{(0, x_2)}{x_2} - (0, 1) \right\|$$
$$= \left\| \frac{(0, x_2) - (0, x_2)}{x_2} \right\| = 0 \leqslant \delta,$$

d) $\langle \nabla f_1(\overline{x}), y \rangle = \langle \nabla f_3(\overline{x}), y \rangle = \langle (0, 0), (0, 1) \rangle = 0$, a więc $\langle \nabla f_i(\overline{x}), y \rangle \leqslant 0$ dla dowolnego $i \in I(S)$.

Zgodnie z prawami rachunku zdań i z twierdzenia 3.1 wynika, iż funkcja f nie osiąga w $\overline{x}$ lokalnego słabego ostrego minimum pierwszego rzędu.

Powyższy przykład pokazuje, że warunek (3.3) nie jest wystarczającym dla istnienia lokalnego słabego ostrego minimum pierwszego rzędu. Nie można nim więc zastąpić bardziej skomplikowanego warunku (b) twierdzenia 3.1.

Rozdział 4

Warunki istnienia słabego ostrego minimum w przypadkach z ograniczeniami

Do tej pory przedstawiono rozważania na temat znajdowania rozwiązań problemu (1) przy założeniu, że zbiór C był całą przestrzenią $\mathbb{R}^n$. Okazuje się jednak, że w zależności od zagadnienia, jakie opisuje funkcja f, zbiór C wektorów dopuszczalnych może być podzbiorem właściwym $\mathbb{R}^n$ podanym w sposób jawny lub w sposób uwikłany, poprzez układ warunków. Poniższe rozważania będą dotyczyły metod znajdowania słabego ostrego minimum funkcji f w przypadkach zawężonego obszaru poszukiwań.

4.1. Ograniczenia jawne obszaru poszukiwań

W tej części pracy podczas formułowania warunków wystarczających istnienia słabego ostrego minimum będziemy posługiwać się zmodyfikowanym pojęciem stożka stycznego, w którego definicji uwzględniony jest zbiór S.

Definicja 4.1. Stożkiem stycznym do zbioru $C \subset \mathbb{R}^n$ w punkcie $x \in C$ uwzględniającym zbiór S nazywamy zbiór:

$$\begin{aligned} K_S(C,x) := \{y \mid \quad & \text{istnieją ciągi} \quad \{t_j\},\{y_j\},\{x_j\} \quad \text{takie, że} \\ & t_j \to 0^+ \,,\, y_j \to y \,,\, x_j \to_{\mathrm{bd}\, S} x \quad \text{i} \quad x_j + t_j y_j \in C\} \end{aligned} \tag{4.1}$$

Oczywiście $K_S(C,x) = K(C,x)$ w przypadku, kiedy $S = \{x\}$. Natomiast dla zbiorów S złożonych z więcej niż jednego elementu zbiory $K_S(C,x)$ oraz $K(C,x)$ mogą się różnić.

Przykład 4.1. Rozważmy zbiory: $C = \{(x^1,x^2) \mid x^2 = |x^1|\}$ oraz $S = \{(x^1,x^2) \mid x^2 = -|x^1|\}$. Ustalmy $\overline{x} = (0,0) \in C$. Stożkiem stycznym do zbioru C w punkcie $\overline{x}$ będzie zbiór:

$$K(C,\overline{x}) = \{(x^1,x^2) \mid x^2 = \left|x^1\right|\} = C.$$

Jeżeli jednak uwzględnimy zbiór S, to otrzymamy stożek postaci:

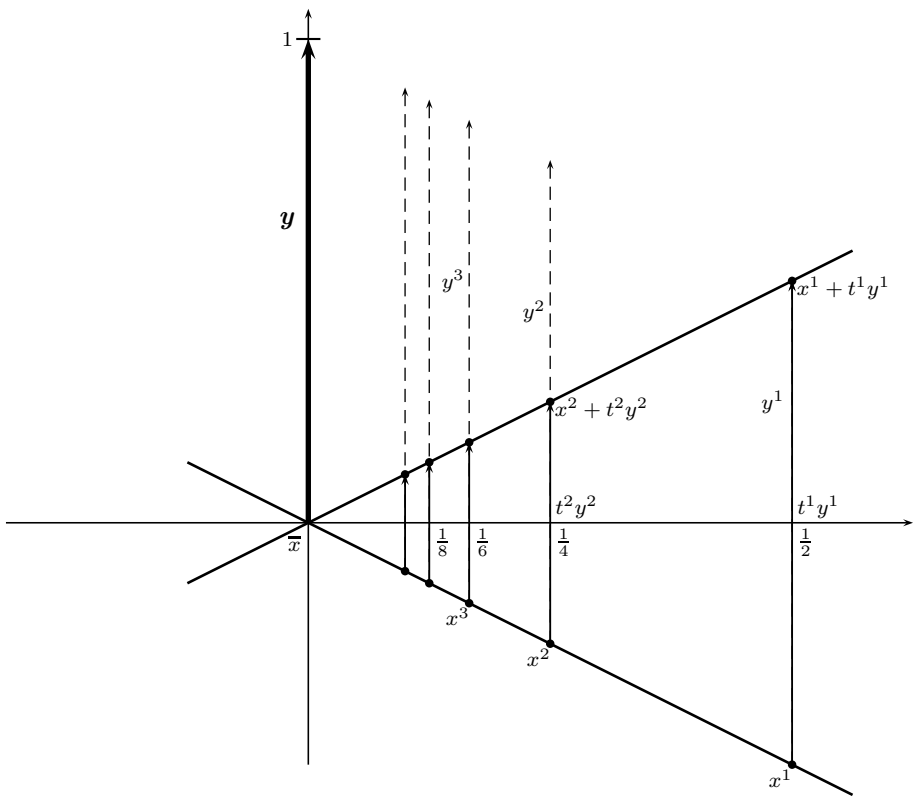
$$K_S(C,\overline{x}) = \{(x^1,x^2) \mid x^2 \geqslant \left|x^2\right|\},$$

a więc $K(C,\overline{x}) \neq K_S(C,\overline{x})$.

Na rysunku 4.1 pokazano, że punkt $(0,1)$ jest elementem zbioru $K_S(C,\overline{x})$. Istnieją bowiem ciągi $\{(x_j^1,-x_j^1)\} = \{(\frac{1}{2j},-\frac{1}{2j})\} \to_{\operatorname{bd} S} (0,0)$, $\{(y_j^1,y_j^2)\} = \{(0,1)\} \to (0,1), \{t_j\} = \{\frac{1}{j}\} \to 0^+$ takie, że

$$\begin{aligned}(x_j^1,-x_j^1) + t_j\cdot(y_j^1,y_j^2) &= (x_j^1 + t_j y_j^1, -x_j^1 + t_j y_j^2)\\ &= \left(\frac{1}{2j} + \frac{1}{j}\cdot 0, -\frac{1}{2j} + \frac{1}{j}\cdot 1\right)\\ &= \left(\frac{1}{2j}, -\frac{1}{2j} + \frac{1}{j}\right) = \left(\frac{1}{2j},\frac{1}{2j}\right) \in C\end{aligned}$$

dla dowolnego $j \in \mathbb{N}$.



Rysunek 4.1. Przykład elementu y zmodyfikowanego stożka $K_S(C,\overline{x})$

Używając zmodyfikowanego pojęcia stożka stycznego możemy sformułować warunek wystarczający istnienia lokalnego słabego ostrego minimum rzędu m w przypadku, kiedy zbiór rozwiązań dopuszczalnych jest podzbiorem przestrzeni $\mathbb{R}^n$.

Twierdzenie 4.1. *[12] Niech $C \subset \mathbb{R}^n$ oraz niech funkcja $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ przyjmuje skończoną i stałą wartość na domkniętym zbiorze S, zaś $\overline{x} \in \operatorname{bd} S \cap C$. Jeśli $\underline{d}_S^m f(\overline{x},\cdot)$ jest funkcją właściwą oraz spełniony jest warunek*

$$\underline{d}_S^m f(\overline{x},y) > 0 \quad \textit{dla dowolnego} \quad y \in (N(S,\overline{x}) \cap K_S(C,\overline{x})) \setminus \{0\}, \tag{4.2}$$

to funkcja f osiąga w $\overline{x}$ lokalne słabe ostre minimum rzędu m[1].

4.2. Ograniczenia obszaru poszukiwań uwikłane poprzez układ warunków

Rozważmy problem poszukiwania:

$$\min\{f(x) \mid x \in C\}, \tag{4.3}$$

gdzie $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$, $C \subset \mathbb{R}^n$,

$$C := \{x \in \mathbb{R}^n \mid g_i(x) \leqslant 0\,,\ i \in J\,,\ h_i(x) = 0\,,\ i \in L\,,\ x \in \Omega\}, \tag{4.4}$$

gdzie Ω jest niepustym podzbiorem przestrzeniu $\mathbb{R}^n$, $J := \{1, \ldots, p\}$, $L := \{1, \ldots, q\}$, $g_i : \mathbb{R}^n \to \overline{\mathbb{R}}$ dla $i \in J$, $h_i\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ dla $i \in L$.

Poniżej zostaną przedstawione warunki istnienia lolalnych słabych ostrych minimów pierwszego lub drugiego rzędu dla problemu (4.3). Wcześniej jednak zdefiniujmy pojęcia, jakie pojawią się w cytowanych twierdzeniach.

Niech Γ będzie domkniętym podzbiorem zbioru C i niech funkcja f przyjmuje skończoną i stałą wartość na zbiorze Γ ($f(x) = \alpha$ dla dowolnego $x \in \Gamma$).

Zdefiniujmy także zbiory:

$$\begin{aligned}
I(x) &:= \{i \in J \mid g_i(x) = 0\} \quad \text{dla} \quad x \in C\,,\\
I^*(\overline{x}) &:= \{i \in I(\overline{x}) \mid \text{istnieje } \delta > 0 \text{ taka, że dla dowolnego } y \in B(\overline{x}, \delta) \cap \Gamma\\
&\qquad \text{mamy: } i \in I(y)\} \quad \text{dla } \overline{x} \in \Gamma \quad ,\\
S(\overline{x}) &:= \{x \mid g_i(x) = 0 \quad \text{dla} \quad i \in I^*(\overline{x}),\\
&\qquad h_i(x) = 0 \quad \text{dla} \quad i \in L,\ f(x) = \alpha\} \qquad \text{dla } \overline{x} \in \Gamma.
\end{aligned}$$

4.2.1. Warunek wystarczający istnienia słabego ostrego minimum pierwszego rzędu

W pozycji [12] przedstawiono twierdzenie, które pozwala uzyskać rozwiązanie problemu (4.3) w postaci lokalnego słabego ostrego minimum pierwszego rzędu.

Twierdzenie 4.2. *Niech $\overline{x} \in \Gamma$ oraz $S \subset S(\overline{x})$ będzie zbiorem domkniętym takim, że $\overline{x} \in \operatorname{bd} S$. Niech ponadto funkcje h_i będą ściśle różniczkowalne w punkcie $\overline{x}$ ($i \in L$). Jeżeli funkcja $\underline{d}^1_S f(\overline{x}, \cdot)$ jest właściwa oraz spełniony jest warunek*

$$\begin{aligned}
\underline{d}^1_S f(\overline{x}, y) > 0 \ \ &\textit{dla każdego } y \in (N(S, \overline{x}) \cap K_S(\Omega, \overline{x})) \setminus \{0\}\\
&\textit{takiego, że } \underline{d}^1_S g_i(\overline{x}, y) \leqslant 0 \ \textit{dla } i \in I^*(\overline{x})\\
&\textit{oraz } \langle \nabla h_i(\overline{x}), y\rangle = 0 \ \textit{dla } i \in L,
\end{aligned} \tag{4.5}$$

to funkcja f osiąga w punkcie $\overline{x}$ lokalne słabe ostre minimum pierwszego rzędu.

[1] Funkcję $f\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ nazywamy właściwą, jeśli zbiór $\operatorname{epi} f := \{(x, r) \mid f(x) \leqslant r\}$ (zwany epiwykresem funkcji f) jest niepusty i f nie przyjmuje wartości $-\infty$

Bardzo ważnym problemem w praktycznym wykorzystaniu cytowanego twierdzenia jest wybór zbioru S. Ilustruje to poniższy przykład.

Przykład 4.2. Rozważmy funkcję $f\colon \mathbb{R}^2 \to \overline{\mathbb{R}}$ określoną następująco: $f(x,y) = x$. Niech $J = \{1,2\}$, $L = \emptyset$, $\Omega = \mathbb{R}^2$, $g_1(x,y) = -x$, $g_2(x,y) = -y$. Będziemy poszukiwać słabego ostrego minimum pierwszego rzędu w zbiorze C zdefiniowanym jako:

$$C = \{(x,y) \in \mathbb{R}^2 \mid g_i(x,y) \leqslant 0 \quad \text{dla} \quad i \in J\}.$$

Zauważmy, że:

$$C = \{(x,y) \in \mathbb{R}^2 \mid -x \leqslant 0\,,\ -y \leqslant 0\} = \{(x,y) \in \mathbb{R}^2 \mid x \geqslant 0\,,\ y \geqslant 0\}.$$

Niech $S = \{(0,y) \mid y \geqslant 0\}$. Oczywiście $f(x,y) = 0$ dla dowolnego $(x,y) \in S$. Niech $(0,\overline{y}) \in S$, gdzie $\overline{y} \geqslant 0$. Wtedy dla każdego $(x,y) \in C$ mamy:

$$f(x,y) - f(0,\overline{y}) = x - 0 = x$$

oraz

$$\operatorname{dist}((x,y),S) = \|(x,y) - (0,y)\| = \sqrt{x^2} = |x| = x,$$

a więc istnieje $\beta = 1 > 0$ taka, że:

$$f(x,y) - f(0,\overline{y}) \geqslant \beta \cdot \operatorname{dist}((x,y),S).$$

S jest więc zbiorem słabych ostrych minimów pierwszego rzędu funkcji f w zbiorze C.

Sprawdźmy, czy warunek (4.5) jest spełniony dla zbioru S zdefiniowanego wcześniej.

1. Niech $\overline{x} = (0,y)$, $y > 0$. Oczywiście $\overline{x} \in \operatorname{bd} S$. Ponadto:
   - $K_S(\Omega,\overline{x}) = \mathbb{R}^2$;
   - $N(S,\overline{x}) = \{(z,0) \mid z \in \mathbb{R}\}$;
   - $I(\overline{x}) = \{1\}, I^*(\overline{x}) = \{1\}$;
   - $S(\overline{x}) = \{0\} \times \mathbb{R}$, a więc $S \subset S(\overline{x})$;
   -

$$\begin{aligned}
\underline{d}^1_S g_1(\overline{x},(z,0)) &= \liminf_{\substack{(x_s,y_s)\to_{\operatorname{bd} S}\overline{x}\\ t\to 0^+\\ (y_1,y_2)\to(z,0)}} \frac{g_1(x_s+ty_1, y_s+ty_2) - g_1(x_s,y_s)}{t}\\
&= \liminf_{\substack{(x_s,y_s)\to_{\operatorname{bd} S}\overline{x}\\ t\to 0^+\\ (y_1,y_2)\to(z,0)}} \frac{-x_s - ty_1 + x_s}{t}\\
&= \liminf_{\substack{(x_s,y_s)\to_{\operatorname{bd} S}\overline{x}\\ t\to 0^+\\ (y_1,y_2)\to(z,0)}} (-y_1) = -z < 0 \quad \text{dla} \quad z > 0
\end{aligned}$$

A zatem dla $(z,0)$, $z>0$ otrzymujemy:

$$\begin{aligned}\underline{d}_S^1 f(\overline{x},(z,0)) &= \liminf_{\substack{(x_s,y_s)\to \operatorname{bd} S\overline{x}\\ t\to 0^+\\ (y_1,y_2)\to(z,0)}} \frac{f(x_s+ty_1,y_s+ty_2)-f(x_s,y_s)}{t}\\ &= \liminf_{\substack{(x_s,y_s)\to \operatorname{bd} S\overline{x}\\ t\to 0^+\\ (y_1,y_2)\to(z,0)}} \frac{x_s+ty_1-x_s}{t}\\ &= \liminf_{\substack{(x_s,y_s)\to \operatorname{bd} S\overline{x}\\ t\to 0^+\\ (y_1,y_2)\to(z,0)}} y_1 = z>0.\end{aligned}$$

Warunek (4.5) jest więc spełniony dla $\overline{x}=(0,y)$, $y>0$.

2. Niech $\overline{x}=(0,0)$, $y>0$. Oczywiście $\overline{x}\in \operatorname{bd} S$. Ponadto:
   - $K_S(\Omega,\overline{x})=\mathbb{R}^2$;
   - $N(S,\overline{x})=\{(x,y)\mid y\leqslant 0\}$;
   - $I(\overline{x})=\{1,2\}, I^*(\overline{x})=\{1\}$;
   - $S(\overline{x})=\{0\}\times\mathbb{R}$, a więc $S\subset S(\overline{x})$;
   - Dla punktu $(0,-1)\in(N(S,\overline{x})\cap K_S(\Omega,\overline{x}))\setminus\{0\}$ mamy:

$$\underline{d}_S^1 g_1(\overline{x},(0,-1)) = \liminf_{\substack{(x_s,y_s)\to \operatorname{bd} S\overline{x}\\ t\to 0^+\\ (y_1,y_2)\to(0,-1)}} (-y_1) = 0\leqslant 0.$$

Dla $(0,-1)$ otrzymujemy

$$\underline{d}_S^1 f(\overline{x},(0,-1)) = \liminf_{\substack{(x_s,y_s)\to \operatorname{bd} S\overline{x}\\ t\to 0^+\\ (y_1,y_2)\to(0,-1)}} y_1 = 0.$$

Warunek (4.5) nie jest więc spełniony dla $\overline{x}=(0,0)$.

Sprawdźmy, jak zmieni się sytuacja, kiedy S zdefiniujemy jako: $S=\{0\}\times\mathbb{R}$.

1. Dla $\overline{x}=(0,y)$, $y>0$ otrzymamy takie same wyniki jak w przypadku pierwszym rozpatrywanym wcześniej.
2. Dla $\overline{x}=(0,0)$ mamy: $I(\overline{x})=\{1,2\}$, $I^*(\overline{x})=\{1\}$, $S(\overline{x})=S$,
   $N(S,\overline{x})=\{(z,0)\mid z\in\mathbb{R}\}$
   $\underline{d}_S^1 g_1(\overline{x},(z,0))=-z<0$ dla $z>0$.
   Dla $(z,0)$, $z>0$ otrzymujemy:
   $\underline{d}_S^1 f(\overline{x},(z,0))=z>0$.

A więc warunek (4.5) jest spełniony dla każdego punktu $\overline{x}=(0,y)$, $y\geqslant 0$.

Okazuje się zatem, że wybór zbioru S ma wpływ na możliwość określenia, czy w danym punkcie istnieje słabe ostre minimum pierwszego rzędu. A zatem skuteczność twierdzenia 4.2 zależy od wskazania zbioru S.

4.2.2. Warunek wystarczający istnienia słabego ostrego minimum drugiego rzędu

W przypadkach, kiedy interesuje nas rozwiązanie problemu (4.3) w postaci lokalnego słabego ostrego minimum drugiego rzędu, możemy posłużyć

się warunkiem wystarczającym, w którym badany jest znak pochodnej kierunkowej funkcji Lagrange'a, tj. funkcji postaci:

$$L(x) := f(x) + \sum_{i\in J} \lambda_i g_i(x) + \sum_{i\in L} \mu_i h_i(x)$$

gdzie $\lambda_i \geqslant 0$, $i \in J$ oraz $\mu_i \in \mathbb{R}$ dla $i \in L$.

Aby móc sformułować warunek potrzebujemy zdefiniować jeszcze zbiór kierunków:

$$\begin{aligned} D^*(\overline{x}) := \{v \mid \ \langle \nabla f(\overline{x}), v\rangle \leqslant 0\,,\ \langle \nabla g_i(\overline{x}), v\rangle \leqslant 0 \quad \text{dla} \quad i \in I^*(\overline{x}),\\ \langle \nabla h_i(\overline{x}), v\rangle = 0 \quad \text{dla} \quad i \in L\} \end{aligned}$$

oraz zbiór mnożników:

$$\begin{aligned} \Lambda^*(\overline{x}) := \{(\lambda,\mu) \in \mathbb{R}^p \times \mathbb{R}^q \mid \nabla L(\overline{x}) = 0\,,\ \lambda_i \geqslant 0 \quad \text{dla} \quad i \in J\,,\\ \lambda_i = 0 \quad \text{dla} \quad i \notin I^*(\overline{x})\}. \end{aligned}$$

Twierdzenie 4.3. *Rozpatrzmy problem (4.3). Niech $\Omega = \mathbb{R}^n$. Załóżmy, że $S \subset \Gamma$ jest zbiorem domkniętym oraz f, g_i dla $i \in J$ i h_i dla $i \in L$ są ściśle różniczkowalne w $\overline{x} \in \operatorname{bd} S$. Jeżeli istnieją mnożniki $(\lambda,\mu) \in \Lambda^*(\overline{x})$ takie, że spełniony jest warunek:*

$$\underline{d}_S^2 L(\overline{x}, y) > 0 \quad \textit{dla dowolnego} \quad y \in \Big(N(S,\overline{x}) \cap D^*(\overline{x})\Big) \setminus \{0\}\,, \tag{4.6}$$

to funkcja f osiąga w $\overline{x}$ lokalne słabe ostre minimum drugiego rzędu (dla problemu (4.3)).

Przykład 4.3. Rozpatrzmy funkcję $f(x,y) = x^2$. Niech $J = \{1,2\}$, $L = \emptyset$, $g_1(x,y) = -x$, $g_2(x,y) = -y$. Tak jak w przykładzie 4.2 obszarem rozwiązań dopuszczalnych problemu (4.3) jest zbiór

$$\begin{aligned} C =& \{(x,y) \in \mathbb{R}^2 \mid g_i(x,y) \leqslant 0 \quad \text{dla} \quad i \in J\}\\ =& \{(x,y) \in \mathbb{R}^2 \mid x \geqslant 0\,,\ y \geqslant 0\}. \end{aligned}$$

Niech $S = \{(0,y) \mid y \geqslant 0\}$. Oczywiście, $f(x,y) = 0$ dla $(x,y) \in S$.

Niech $(0,\overline{y}) \in S$, gdzie $\overline{y} \geqslant 0$. Wtedy dla dowolnego $(x,y) \in C$ otrzymujemy:

$$f(x,y) - f(0,\overline{y}) = x^2 - 0 = x^2$$

oraz

$$\operatorname{dist}^2((x,y),S) = \|(x,y) - (0,y)\|^2 = x^2,$$

a więc istnieje $\beta = 1 > 0$ taka, że

$$f(x,y) - f(0,\overline{y}) \geqslant \beta \cdot \operatorname{dist}^2((x,y),S) \quad \text{dla} \quad (x,y) \in C.$$

Zatem S jest zbiorem słabych ostrych minimów drugiego rzędu funkcji f w zbiorze C.

1. Zbudujmy funkcję Lagrange'a, która będzie spełniać warunek (4.6) dla punktu $(0,y) \in S$, $y > 0$.
Ponieważ $I(0,y) = I^*(0,y) = \{1\}$, więc $\lambda_2 = 0$. Funkcja Lagrange'a ma zatem postać $L(x,y) = x^2 - \lambda_1 x$. Ponieważ $\nabla L(0,y) = (-\lambda_1, 0)$, więc $(\lambda_1, \lambda_2) \in \Lambda^*(0,y)$ wtedy i tylko wtedy, gdy $\lambda_1 = \lambda_2 = 0$, więc $L(x,y) = f(x,y) = x^2$. Zauważmy dalej, że $\nabla f(0,y) = (0,0)$ oraz $\nabla g_1(0,y) = (-1,0)$, zatem $D^*(0,y) = \{(v_1,v_2) \mid v_1 \geqslant 0\}$. Ponadto $N(S,(0,y)) = \{(z,0) \mid z \in \mathbb{R}\}$.
Niech $(y_1,y_2) \in \Big(N(S,(0,y)) \cap D^*(0,y)\Big) \setminus \{(0,0)\} = \{(v_1,0) \mid v_1 > 0\}$. Wtedy $(y_1,y_2) = (y_1,0)$, $y_1 > 0$.

$$\begin{aligned}\underline{d}_S^2 L((0,y),(y_1,0)) &= \liminf_{\substack{(x_s,y_s)\to \operatorname{bd} S(0,y)\\ (x_v,y_v)\to(y_1,0)\\ t\to 0^+}} \frac{L(x_s+tx_v, y_s+ty_v) - L(x_s,y_s)}{t^2}\\ &= \liminf_{\substack{(x_s,y_s)\to \operatorname{bd} S(0,y)\\ (x_v,y_v)\to(y_1,0)\\ t\to 0^+}} \frac{t^2\cdot x_v^2}{t^2} = y_1^2 > 0\end{aligned}$$

2. Zbudujmy funkcję Lagrange'a, która będzie spełniać warunek (4.6) dla punktu $(0,0) \in S$.
Ponieważ $I(0,0) = \{1,2\}$ i $I^*(0,0) = \{1\}$, więc $\lambda_2 = 0$ i funkcja Lagrange'a ma postać $L(x,y) = x^2 - \lambda_1 x$. Ponieważ $\nabla L(0,0) = (-\lambda_1, 0)$, zatem $D^*(0,0) = \{(v_1,v_2) \mid v_1 \geqslant 0\}$. Różnicę stanowi postać stożka normalnego do zbioru S w punkcie $(0,0)$, $N(S,(0,0)) = \{(x,y) \mid y \leqslant 0\}$. Zatem dla

$$\begin{aligned}(y_1,y_2) \in &\Big(N(S,(0,0)) \cap D^*(0,0)\Big) \setminus \{(0,0)\}\\ &= \{(x,y) \mid x \geqslant 0\,,\ y \leqslant 0\} \setminus \{(0,0)\}\end{aligned}$$

mamy:

$$\begin{aligned}\underline{d}_S^2 L((0,y),(y_1,y_2)) &= \liminf_{\substack{(x_s,y_s)\to \operatorname{bd} S(0,y)\\ (x_v,y_v)\to(y_1,y_2)\\ t\to 0^+}} \frac{L(x_s+tx_v, y_s+ty_v) - L(x_s,y_s)}{t^2}\\ &= \liminf_{\substack{(0,y_s)\to \operatorname{bd} S(0,y)\\ (x_v,y_v)\to(y_1,y_2)\\ t\to 0^+}} \frac{t^2\cdot x_v^2}{t^2} = y_1^2 > 0\end{aligned}$$

Zatem warunek (4.6) jest spełniony dla każdego punktu rozpatrywanego zbioru S.

4.2.3. Warunek wystarczający istnienia słabego ostrego minimum pierwszego rzędu w przypadku ograniczeń liniowych

Jeżeli rozważywmy zmodyfikowany zbiór mnożników postaci:

$$\begin{aligned}\Lambda(\overline{x}) := \{(\lambda,\mu) \in \mathbb{R}^p \times \mathbb{R}^q \mid \nabla L(\overline{x}) = 0\,,\ \lambda_i \geqslant 0\,,\\ \lambda_i g_i(\overline{x}) = 0 \quad \text{dla} \quad i \in J\}\end{aligned}$$

oraz pewien podzbiór indeksów:

$$M(\overline{x}) := \{i \in I(\overline{x}) \mid \lambda_i > 0\} \quad \text{dla} \quad (\lambda, \mu) \in \Lambda(\overline{x}) \,,$$

to możemy przystąpić do przedstawienia warunku wystarczającego istnienia lokalnego słabego ostrego minimum pierwszego rzędu dla problemu (4.3) rozpatrywanego przy następujących założeniach:

- $\Omega = \mathbb{R}^n$;
- $g_i(x) = \langle a_i, x\rangle - b_i$ dla $i \in J$, $h_i(x) = \langle c_i, x\rangle - d_i$ dla $i \in L$, gdzie a_i, $c_i \in \mathbb{R}^n$ zaś b_i, $d_i \in \mathbb{R}$;
- funkcja f przyjmuje stałą wartość na zbiorze

$$\begin{aligned} S = \{x \in \mathbb{R}^n \mid \ & \langle a_i, x\rangle = b_i \quad \text{dla} \quad i \in J_1 \,, \\ & \langle c_i, x\rangle = d_i \quad \text{dla} \quad i \in L\}, \quad \text{gdzie} \quad J_1 \subset J. \end{aligned} \tag{4.7}$$

Dla tak sformułowanego problemu (4.3) mamy:

Twierdzenie 4.4. *Rozważmy problem (4.3), w którym zbiór C uwzględnia założenia podane wyżej. Przypuśćmy, że funkcja f przyjmuje stałą wartość na zbiorze S postaci (4.7). Jeżeli są spełnione następujące warunki:*

- *$\overline{x} \in \Gamma = S \cap C$ jest takim punktem, że $I^*(\overline{x}) = J_1$,*
- *a_i dla $i \in J_1$ oraz c_i dla $i \in L$ są wektorami liniowo niezależnymi,*
- *funkcja f jest ściśle różniczkowalna w $\overline{x}$,*
- *istnieje $(\lambda, \mu) \in \Lambda(\overline{x})$ taka, że $M(\overline{x}) = J_1$,*

to funkcja f osiąga w $\overline{x}$ lokalne słabe ostre minimum pierwszego rzędu dla problemu (4.3).

Przykład 4.4. Rozważmy problem optymalizacji funkcji w zbiorze ograniczeń liniowych:

$$\min\{x^2 - 2y - 3z \mid x + 2y + 3z \leqslant 6 \,,\, x \geqslant 0 \,,\, y \geqslant 0 \,,\, z \geqslant 0\}.$$

Wówczas

$$\begin{aligned} f(x,y,z) &= x^2 - 2y - 3z \,, \\ g_1(x,y,z) &= x + 2y + 3z - 6 = \langle (1,2,3), (x,y,z)\rangle - 6 \,, \\ g_2(x,y,z) &= -x = \langle (-1,0,0), (x,y,z)\rangle - 0 \,, \\ g_3(x,y,z) &= -y = \langle (0,-1,0), (x,y,z)\rangle - 0 \,, \\ g_4(x,y,z) &= -z = \langle (0,0,-1), (x,y,z)\rangle - 0. \end{aligned}$$

Niech $J_1 = \{1, 2\}$ i

$$\begin{aligned} S &= \{(x,y,z) \mid g_i(x,y,z) = 0 \quad \text{dla} \quad i \in J_1\} \\ &= \{(x,y,z) \mid x + 2y + 3z = 6 \,,\, x = 0\}. \end{aligned}$$

Niech $(\overline{x}, \overline{y}, \overline{z}) \in S$. Wtedy $\overline{x} = 0$ i $\overline{y} = \frac{6-3\overline{z}}{2}$, $f(\overline{x}, \overline{y}, \overline{z}) = 0^2 - 2 \cdot \frac{6-3\overline{z}}{2} - 3\overline{z} = -6 + 3\overline{z} - 3\overline{z} = -6$. Funkcja f przyjmuje wartość stałą na zbiorze S.

Niech ponadto $(\overline{x}, \overline{y}, \overline{z}) \in C$. Wtedy $\overline{x} = 0$, $\overline{y} = \frac{6-3\overline{z}}{2}$, $\overline{y} \geqslant 0$ i $\overline{z} \geqslant 0$.

Z faktu, że $I(\overline{x}, \overline{y}, \overline{z}) = \{1, 2\}$ i $I^*(\overline{x}, \overline{y}, \overline{z}) = \{1, 2\}$ oraz z postaci zbioru S wynika, że $I^*(\overline{x}, \overline{y}, \overline{z}) = J_1$.

Ponadto wektory $a_1 = (1, 2, 3)$ i $a_2 = (-1, 0, 0)$ są liniowo niezależne oraz funkcja f jest ściśle różniczkowalna w $(\overline{x}, \overline{y}, \overline{z})$.

Funkcja Lagrange'a ma postać:

$$\begin{aligned} L(x, y, z) &= f(x, y, z) + \lambda_1 g_1(x, y, z) + \lambda_2 g_2(x, y, z) \\ &\quad + \lambda_3 g_3(x, y, z) + \lambda_4 g_4(x, y, z) \\ &= x^2 - 2y - 3z + \lambda_1(x + 2y + 3z - 6) - \lambda_2 x - \lambda_3 y - \lambda_4 z. \end{aligned}$$

$$\begin{aligned} \nabla L(\overline{x}, \overline{y}, \overline{z}) &= [\lambda_1 - \lambda_2\ ,\ -2 + 2\lambda_1 - \lambda_3\ ,\ -3 + 3\lambda_1 - \lambda_4] = [0, 0, 0] \Leftrightarrow \\ &\Leftrightarrow (\lambda_1 = \lambda_2\ ,\ \lambda_3 = 2\lambda_1 - 2\ ,\ \lambda_4 = 3\lambda_1 - 3)\ , \end{aligned}$$

a więc zbiór mnożników ma postać:

$$\begin{aligned} \Lambda(\overline{x}, \overline{y}, \overline{z}) = \{(\lambda_1, \lambda_2, \lambda_3, \lambda_4) \mid\ & \lambda_1 = \lambda_2\ ,\ \lambda_3 = 2\lambda_1 - 2\ ,\ \lambda_4 = 3\lambda_1 - 3\ , \\ & \lambda_i \geqslant 0 \quad \text{dla} \quad i \in J\ , \\ & \lambda_i g_i(\overline{x}, \overline{y}, \overline{z}) = 0 \quad \text{dla} \quad i \in J\}. \end{aligned}$$

Zauważmy, że $(1, 1, 0, 0) \in \Lambda(\overline{x}, \overline{y}, \overline{z})$. Wtedy $M(\overline{x}, \overline{y}, \overline{z}) = \{1, 2\}$, a więc $M(\overline{x}, \overline{y}, \overline{z}) = J_1$. Zgodnie z twierdzeniem 4.4 problem (4.3) ma rozwiązanie w postaci lokalnego słabego ostrego minimum pierwszego rzędu osiąganego w punkcie $(\overline{x}, \overline{y}, \overline{z})$.

Podsumowanie

W przedstawionej pracy zaprezentowane zostały warunki istnienia słabego ostrego minimum dla problemu (1). Pierwszy zacytowany warunek 2.1, w którym wykorzystano pojęcie stożka normalnego, jest skomplikowany i trudny w zastosowaniu praktycznym. Jednak, jak pokazały dalsze rozważania, można go zastosować dla wielu funkcji, które nie spełaniają założeń innych prostszych warunków. Twierdzenie 2.1 jest zatem najbardziej uniwersalnym spośród wszystkich zacytowanych w pracy warunków.

Problem z uniwersalnością zastosowania pojawia się przy metodzie opartej na twierdzeniu 2.2 wykorzystującym pojęcie zmodyfikowanej pochodnej kierunkowej. Jak się okazuje, warunki twierdzenia 2.2 są zawsze spełnione dla zbiorów S wypukłych i funkcji f ściśle różniczkowalnych w $\overline{x}$. W pozostałych przypadkach (a jest ich bardzo dużo, obejmują bowiem wszystkie funkcje posiadające słabe ostre minimum pierwszego rzędu) twierdzenie to często nie może być stosowane.

Kolejne z cytowanych twierdzeń 3.1 dotyczy specjalnej klasy funkcji. Jak pokazuje przykład 3.3 warunek tego twierdzenia nie daje się uprościć, lecz w przypadku maksimum ze skończonych i ściśle różniczkowalnych funkcji może być stosowany zamiast twierdzenia 2.1.

Następne twierdzenie 4.1 wykorzystujące również zmodyfikowaną pochodną kierunkową dotyczny przypadku minimalizacji funkcji f przy ograniczonym obszarze poszukiwań. Jeśli ograniczenie zbioru C jest podane w postaci układu warunków i jeśli poszukujemy słabego ostrego minimum pierwszego rzędu, to można stosować do rozwiązania zadania (1) twierdzenie 4.2 wykorzystujące pojęcie stożka stycznego do zbioru C, uwzględniającego zbiór S. Podany został przykład obrazujący jak ważny w stosowaniu twierdzenia 4.2 jest wybór zbioru S.

Ostatnie dwa twierdzenia 4.3 i 4.4 wykorzystują do poszukiwania słabego ostrego minimum odpowiednio rzędu drugiego lub pierwszego funkcję Lagrange'a. Zaprezentowane zostały przykłady wykorzystania tych twierdzeń w praktyce.

Spis rysunków

1.1. Przykłady różnego rodzaju minimów 5
1.2. Przykład minimum funkcji, które nie jest jej słabym ostrym minimum rzędu pierwszego 6
1.3. Transformacja Fouriera 8
1.4. Wynik działania DFT 9
1.5. Wynik działania DFT – zależność amplitudy od częstotliwości 9
1.6. Przykład obcięcia fali dźwiękowej 10
1.7. Izotermy gazu spełniającego prawo Boyle'a-Mariotte'a 11
1.8. Odchylenia od doskonałego zachowania się gazów rzeczywistych . . . 11
1.9. Izotermy i obszar skraplania CO_2 12

2.1. Przykłady punktów danego zbioru, których odległość od danego punktu jest najmniejsza 14
2.2. Przykład elementu (x, y) stożka $N(S, x_s)$, gdzie $x_s = (\overline{x}, 0)$, $\overline{x} \in (-1, 0)$ 15
2.3. Przykład elementów $(x^1, y^1), (x^2, y^2)$ stożka $N(S, x_s)$, gdzie $x_s = (0, 0)$ 15

4.1. Przykład elementu y zmodyfikowanego stożka $K_S(C, \overline{x})$ 28

Bibliografia

[1] G. M. Barrow: *Chemia fizyczna*, PWN, Warszawa 1971

[2] J. V. Burke, M. C. Ferris: *Weak sharp minima in mathematical programming*, SIAM J. Control Optim. 31 (1993), pp. 1340-1359

[3] J. Céa: *Optymalizacja: teoria i algorytmy*, PWN, Warszawa 1976

[4] D. G. Luenberger: *Teoria optymalizacji*, PWN, Warszawa 1974

[5] M. M. Mäkelä, P. Neittaanmäki: *Nonsmooth optimization: Analysis and Algorithms with Applications to Optimal Control*, World Scientific, 1992

[6] B. S. Mordukhovich: *Generalized differential calculus for nonsmooth and set-valued mappings*, J. Math. Anal. Appl, 183 (1994), pp. 250-288

[7] R. T. Rockafellar: *Convex Analysis*, Princeton University Press, 1970

[8] Monty Ross: *Power Amplifier Clipping and its Effects on Loudspeaker Reliability*, Rane Corporation 1991

[9] Steven W. Smith: *The Scientist and Engineer's Guide to Digital Signal Processing*, California Technical Publishing, San Diego, California 1999

[10] M. Studniarski: *On weak sharp minima for special class of nonsmooth functions*, Discussiones Mathematicae: Differential Inclusions, Control and Optimization, 20 (2000), pp. 195-207

[11] M. Studniarski, M. Studniarska: *New characterizations of weak sharp and strict local minimizers in nonlinear programming*, Preprint 1999/15, Wydział Matematyki Uniwersytetu Łódzkiego

[12] M. Studniarski, D. E. Ward: *Weak sharp minima: characterizations and sufficient conditions*, SIAM J. Control Optim. 38 (1999), pp. 219-236

[13] D. E. Ward: *Characterizations of strict local minima and necessary conditions for weak sharp minima*, J. Optim. Theory Appl., 80 (1994), pp. 551-571