ROCZNIKI POLSKIEGO TOWARZYSTWA MATEMATYCZNEGO
Seria III: MATEMATYKA STOSOWANA XXIX (1987)

Ministerstwo Nauki i Szkolnictwa Wyższego

WALERIAN DUBNICKI, KRYSTIAN ZORYCHTA (Warszawa)

Funkcje quasi- i pseudowypukłe w programowaniu nieliniowym

(*Praca wpłynęła do Redakcji 1985.05.06*)

W roku 1983 Państwowe Wydawnictwo Naukowe wydało książkę Béli Martosa *Programowanie nieliniowe, teoria i metody*. W książce tej przedstawiona jest wyczerpująco teoria funkcji quasi-wypukłych i pseudowypukłych oraz ich zastosowań w teorii i metodach programowania nieliniowego. W artykule tym pragniemy przedstawić główne wyniki teoretyczne zawarte w tej książce, gdyż naszym zdaniem problematyka ta warta jest spopularyzowania. Uwagę naszą skupimy na funkcjach quasi-wypukłych i pseudowypukłych, quasi-monotonicznych i pseudomonotonicznych oraz teorii Kuhna–Tuckera dla zadań programowania nieliniowego formułowanych w klasie takich funkcji. Nie będziemy natomiast omawiali prezentowanych przez B. Martosa metod, gdyż od czasu wydania oryginału (1975) ich rozwój poszedł dalej niż jest to uwidocznione w jego książce.

1. Zadanie programowania wypukłego. *Programowaniem matematycznym* nazywamy tę część teorii i metod optymalizacji, która obejmuje zadania polegające na maksymalizacji lub minimalizacji danej funkcji (tzw. funkcji celu) na zbiorze rozwiązań dopuszczalnych określonych za pomocą układu równań i nierówności, liniowych bądź nieliniowych. Jeśli pozostaniemy na gruncie przestrzeni euklidesowej, to zadanie programowania matematycznego będzie miało postać następującą:

$$\min_{x \in \mathcal{Q}} f_0(x), \tag{1.1}$$

gdzie

$$\mathcal{Q} = \{x = (x_1, \ldots, x_n) \in \mathcal{X}\colon f_i(x) \leqslant 0,\ i \in \overline{1, m}\} \tag{1.2}$$

oraz f_i ($i \in \overline{0, m}$) są funkcjami rzeczywistymi, określonymi na zbiorze otwartym $\mathcal{X} \subseteq \mathcal{E}^n$ (¹). Interesuje nas przy tym nie tylko minimalna (lub maksymal-

(¹) Nie jest konieczne, aby w każdym przypadku zbiór $\mathcal{X}$ musiał być otwarty, jednak przyjęcie tego założenia jest bardzo wygodne i nie osłabia wyników w sposób istotny.

na) wartość funkcji celu, ale także punkt $x \in \mathscr{Q}$, w którym ta wartość jest przyjmowana (tzw. punkt optymalny).

Zadanie, w którym wszystkie funkcje f_i są funkcjami afinicznymi (suma funkcji liniowej i stałej)

$$f_i(x) = \sum_{j=1}^{n} a_{ij} x_j - b_i, \qquad i \in \overline{0, m},$$

nazywamy *zadaniem programowania liniowego*. W przeciwnym wypadku mamy do czynienia z zadaniem programowania nieliniowego.

Ostatnie 30 lat stanowią okres intensywnego rozwoju teorii i metod programowania matematycznego, a także ich zastosowań. Rozwój ten zaczął się od programowania liniowego. Najpierw powstała jedna z najbardziej znanych i najintensywniej wykorzystywanych po wojnie metod, metoda sympleksowa, a następnie piękna i bogata w interpretacje teoria dualności. Trudno jest znaleźć dziedzinę wiedzy lub działalności ludzkiej, w której nie znalazłyby zastosowania metody programowania liniowego. Szczególnie znane są zastosowania o charakterze ekonomicznym, które w latach sześćdziesiątych przynosiły poprawę efektywności modelowanych obiektów średnio o 10 - 15%.

O wiele skromniejsze są efekty programowania liniowego w zastosowaniach technicznych, sterowaniu i projektowaniu. Doskonalenie procesów technologicznych, sterowanie obiektami różnorodnej natury, a także opracowywanie efektywnych i przyszłościowych konstrukcji wymaga uwzględniania zależności nieliniowych. W ostatnich latach coraz częściej pojawia się konieczność budowy modeli nieliniowych również dla procesów ekonomicznych. Stąd naturalna dążność do uogólnienia wyników uzyskanych dla programowania liniowego.

Pierwszy, stosunkowo łatwy krok doprowadził do programowania wypukłego, to znaczy do takich zadań, w których zbiór $\mathscr{X}$ oraz wszystkie funkcje f_i $(i \in \overline{0, m})$ są wypukłe (lub wklęsłe, jeśli nierówności w (1.2) są przeciwne oraz funkcja celu f_0 jest maksymalizowana).

Dość często zadania programowania matematycznego spotykane w technice i ekonomii formułowane są za pomocą wypukłych lub wklęsłych funkcji f_i, określających wypukłe zbiory rozwiązań dopuszczalnych i optymalnych. Dla przykładu, zależność wskaźników efektywności systemów technicznych od swoich parametrów opisywana jest zwykle funkcją wklęsłą; im wyższe są techniczne charakterystyki układu, tym trudniej zwiększyć jego efektywność. Podobnie funkcja produkcji w określonym systemie ekonomicznym jest zwykle funkcją wklęsłą wielkości wykorzystywanych zasobów. Przy dużych nakładach kapitałowych przyrost wolumenu produkcji na jednostkę nakładu jest mniejszy niż przy nakładach małych. Im bowiem skala produkcji jest większa, tym więcej środków pochłania organizacja współdziałania między elementami systemu.

Naturalną metodą badania zadań programowania matematycznego wydaje się być metoda mnożników Lagrange'a. Wykorzystanie jej wymaga jednak dodatkowego wysiłku z uwagi na nierówności, które stanowią jakość istotnie różną od równań, dla których funkcjonuje metoda klasyczna.

Zdefiniujmy więc funkcję Lagrange'a dla zadania programowania matematycznego (1.1)

$$L(x, u) = f_0(x) + \sum_{i=1}^{m} u_i f_i(x), \qquad x \in \mathcal{X},\ u_i \geqslant 0\ (i \in \overline{1, m}). \tag{1.3}$$

Ważnym pojęciem teorii jest pojęcie punktu siodłowego tej funkcji. Mówimy, że punkt $(\bar{x}, \bar{u})$ (gdzie $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_n) \in \mathcal{X}$, $\bar{u} = (\bar{u}_1, \ldots, \bar{u}_m)$, $\bar{u}_i \geqslant 0$ dla $i \in \overline{1, m}$) jest *punktem siodłowym* funkcji Lagrange'a, jeśli dla każdego $x \in \mathcal{X}$ i dla każdego $u = (u_1, \ldots, u_m)$ o nieujemnych współrzędnych spełnione są nierówności

$$L(\bar{x}, u) \leqslant L(\bar{x}, \bar{u}) \leqslant L(x, \bar{u}). \tag{1.4}$$

Prawdziwe są następujące twierdzenia:

TWIERDZENIE 1.1. *Jeśli $(\bar{x}, \bar{u})$ jest punktem siodłowym funkcji L, to $\bar{x}$ jest punktem optymalnym zadania* (1.1).

TWIERDZENIE 1.2. *Jeśli spełnione są założenia:*

(1) *zadanie* (1.1) *jest zadaniem programowania wypukłego*;

(2) *istnieje $x \in \mathcal{X}$, dla którego $f_i(x) < 0$ $(i \in \overline{1, m})$ (warunek regularności Slatera)*;

(3) *$\bar{x}$ jest punktem optymalnym zadania* (1.1);

to istnieje $\bar{u} = (\bar{u}_1, \ldots, \bar{u}_m)$, $\bar{u}_i \geqslant 0$ $(i \in \overline{1, m})$ takie, że $(\bar{x}, \bar{u})$ jest punktem siodłowym funkcji L.

Zwróćmy uwagę, że wypukłość odgrywa rolę jedynie w twierdzeniu 1.2. Oba twierdzenia funkcjonują niezależnie od tego, czy funkcje f_i są różniczkowalne. Jeśli jednak nie są różniczkowalne, to twierdzenia te nie dają efektywnego analitycznego narzędzia rozwiązywania zadań (1.1).

Ograniczając się do funkcji różniczkowalnych uzyskujemy następne dwa twierdzenia.

TWIERDZENIE 1.3 (*warunek konieczny Kuhna–Tuckera*). *Jeśli $(\bar{x}, \bar{u})$ jest punktem siodłowym funkcji L oraz f_i dla $i \in \overline{0, m}$ są różniczkowalne w $\bar{x}$, to $(\bar{x}, \bar{u})$ spełnia następujące warunki Kuhna–Tuckera*:

$$\begin{gathered} f_0'(\bar{x}) + \sum_{i=1}^{m} \bar{u}_i f_i'(\bar{x}) = 0, \\ f_i(\bar{x}) \leqslant 0, \qquad \bar{u}_i \geqslant 0, \\ \bar{u}_i f_i(\bar{x}) = 0 \quad \text{dla } i \in \overline{1, m}, \end{gathered} \tag{1.5}$$

gdzie f' oznacza gradient funkcji f.

TWIERDZENIE 1.4 (warunek dostateczny Kuhna–Tuckera). *Jeśli f_i są różniczkowalne w punkcie $\bar{x}$ oraz wypukłe, a ponadto w punkcie $(\bar{x}, \bar{u})$ spełnione są warunki* (1.5) *to $\bar{x}$ jest punktem optymalnym zadania* (1.1).

Uzyskanie twierdzeń mówiących o tym, kiedy dla danego punktu optymalnego $\bar{x}$ istnieje $\bar{u} \geqslant 0$ taki, że para $(\bar{x}, \bar{u})$ spełnia warunki Kuhna–Tuckera (1.5) wymaga wprowadzenia warunków regularności, takich jak na przykład wykorzystany w twierdzeniu 1.2 warunek Slatera.

Ta krótka teoria daje narzędzie pozwalające zbadać, czy – przy określonych założeniach – dany punkt $\bar{x}$ jest punktem optymalnym, czy też nie. Warunki Kuhna–Tuckera mogą również stanowić istotny element metod poszukiwania punktów optymalnych.

Warto sobie także uświadomić, że przyjęte założenie wypukłości funkcji f_i $(i \in \overline{0, m})$ gwarantuje następujące pozytywne cechy zadań programowania wypukłego:

(1.6) Zbiór rozwiązań dopuszczalnych jest zbiorem wypukłym.

(1.7) Każde minimum lokalne funkcji celu f_0 jest minimum globalnym.

(1.8) Dla klasy zadań programowania wypukłego istnieją metody, których koszt przy ustalonej tolerancji błędu rośnie stosunkowo wolno (liniowo) względem n (za jednostkę kosztu przyjmujemy koszt jednokrotnego obliczenia wartości wszystkich funkcji f_i oraz ich gradientów w danym punkcie x).

Z punktu widzenia metod rozwiązywania zadań programowania matematycznego są to cechy bardzo pożądane. Jeśli zadanie nie ma drugiej cechy, to pojawia się natychmiast problem przeglądu minimów lokalnych. Do metod wkracza więc czynnik kombinatoryczny, którego negatywny wpływ na efektywność metod rośnie bardzo szybko wraz ze wzrostem liczby minimów. Brak pierwszej cechy podważa już sens takich metod, które polegają na poszukiwaniu minimów na kolejnych półprostych. Iloczyn mnogościowy danej półprostej i zbioru dopuszczalnego mógłby być bowiem zbiorem niespójnym, złożonym z dużej i nieznanej liczby przedziałów rozłącznych, co zmuszałoby nas do rozpatrywania funkcji celu na każdym z tych przedziałów oddzielnie. Pozytywne skutki trzeciej cechy nie wymagają komentarza.

Powstaje pytanie, czy założenie wypukłości funkcji f_i stanowi kres, przy którym przedstawione twierdzenia oraz trzy omówione wyżej cechy pozostają w mocy? Czy i na ile można to założenie osłabić, aby nic nie utracić z uzyskanych wyników?

Odpowiedź jest pozytywna, wymaga jednak wprowadzenia nowych klas wypukłości: quasi-wypukłości i pseudowypukłości. Szczegółowemu omówieniu tych, jak i innych związanych z wypukłością funkcji pojęć oraz ich konsekwencji w teorii programowania nieliniowego poświęcona jest książka Béli Martosa *Programowanie nieliniowe, teoria i metody*. Jest to tłumaczenie kiążki, wydanej w języku angielskim w roku 1975. Książka nie jest więc nowa. Należy jednak wziąć pod uwagę, że również w tej dziedzinie istotnie nowe wyniki teoretyczne pojawiają się coraz rzadziej. Sam autor pisze w przedmowie, że w tym procesie uogólniania dochodzi niemal do punktu, w którym osłabione założenie staje się nie tylko wystarczające, ale i konieczne. Książka nie straciła więc na aktualności, a jej znaczenie dla polskiego czytelnika umacnia fakt, że jest to jedyna książka przedstawiająca tak wyczerpująco tę tematykę w języku polskim.

W dalszej części artykułu omówimy w skrócie najważniejsze elementy problematyki zawartej w książce Martosa.

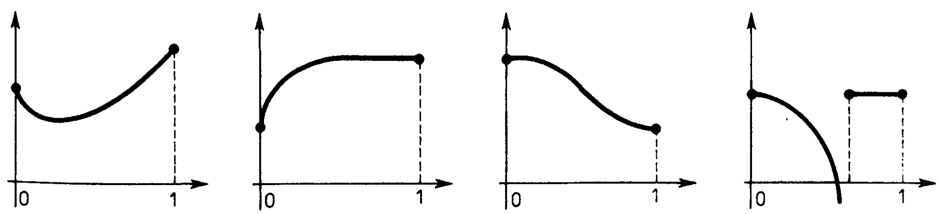
2. Funkcje quasi-wypukłe i pseudowypukłe. Rozpatrywać będziemy funkcje o wartościach skalarnych, jednak omawiane własności odnosić będziemy także do funkcji o wartościach wektorowych, jeśli poszczególne składowe funkcji mieć będą te własności. W przypadku funkcji pseudowypukłych zakładać będziemy ich różniczkowalność na odpowiednich zbiorach (tam gdzie wykorzystywany będzie gradient), natomiast nie będziemy czynić żadnych dodatkowych założeń (poza wypukłością dziedziny) przy wprowadzaniu funkcji quasi-wypukłych.

Niech f będzie funkcją o wartościach skalarnych określoną na zbiorze wypukłym $\mathcal{X} \subseteq \mathcal{E}^n$. Funkcję f nazywamy *quasi-wypukłą* na $\mathcal{X}$, jeśli jest spełniony warunek

$$(2.1)\qquad f(x_0) \leqslant \max\{f(x_1), f(x_2)\} \quad \text{dla dowolnych } x_1, x_2 \in \mathcal{X},\ x_0 \in (x_1, x_2).$$

Symbolem (x_1, x_2) oznaczamy tu odcinek bez końców łączący punkty x_1 i x_2, tzn. $(x_1, x_2) = \{(1-\lambda)x_1 + \lambda x_2\colon \lambda \in (0, 1)\}$. Funkcja jest więc quasi-wypukła, gdy na każdym odcinku przyjmuje wartości nie większe od maksymalnej spośród przyjmowanych na krańcach tego odcinka. Na rysunku 1 przedstawiono wykresy czterech funkcji quasi-wypukłych na [0, 1]. Tylko pierwsza z tych funkcji jest wypukła. Przypomnijmy, że wypukłość f na $\mathcal{X}$ definiuje się warunkiem

$$(2.2)\qquad f((1-\lambda)x_1 + \lambda x_2) \leqslant (1-\lambda)f(x_1) + \lambda f(x_2) \quad \text{dla dowolnych}$$
$$x_1, x_2 \in \mathcal{X},\ \lambda \in (0, 1).$$



Rys. 1

Funkcję f nazywamy *ściśle quasi-wypukłą* na $\mathscr{X}$, jeśli spełniony jest warunek (2.1) oraz warunek

$$f(x_1) \neq f(x_2) \Rightarrow f(x_0) < \max\{f(x_1), f(x_2)\}. \tag{2.3}$$

Żądanie spełnienia warunku (2.1) w powyższej definicji wynika z faktu, iż warunek (2.3) nie implikuje quasi-wypukłości. Łatwo przekonać się o tym, rozpatrując funkcję

$$f(x) = \begin{cases} 0 & \text{dla } x \neq 0, \\ 1 & \text{dla } x = 0. \end{cases}$$

Dowiedziono jednak, że w przypadku funkcji półciągłych z dołu implikacja ta zachodzi. Dla tej klasy funkcji, a więc w szczególności dla funkcji ciągłych, ścisłą quasi-wypukłość definiować można tylko warunkiem (2.3). Spośród czterech funkcji przedstawionych na rysunku 1 tylko pierwsza i trzecia są ściśle quasi-wypukłe na $[0, 1]$.

Funkcję f nazywamy *quasi-wklęsłą* (*ściśle quasi-wklęsłą*) na $\mathscr{X}$, jeśli $-f$ jest funkcją quasi-wypukłą (ściśle quasi-wypukłą) na $\mathscr{X}$. Odpowiada to przyjęciu w (2.1) i (2.3) nierówności o przeciwnych zwrotach z jednoczesną zamianą maksimum na minimum. Na rysunku 1 quasi-wklęsłymi na $[0, 1]$ są funkcje druga i trzecia, ta ostatnia jest przy tym ściśle quasi-wklęsła.

Funkcję f nazywamy *quasi-monotoniczną* (*ściśle quasi-monotoniczną*) na $\mathscr{X}$, jeśli f jest quasi-wypukła i quasi-wklęsła (ściśle quasi-wypukła i ściśle quasi-wklęsła) na $\mathscr{X}$. Można wykazać, że każdy z warunków

(2.4) $\min\{f(x_1), f(x_2)\} \leqslant f(x_0) \leqslant \max\{f(x_1), f(x_2)\}$ dla dowolnych

$x_1, x_2 \in \mathscr{X}$, $x_0 \in (x_1, x_2)$,

(2.5) f jest monotoniczna na każdym odcinku domkniętym $[x_1, x_2] \subseteq \mathscr{X}$,

jest równoważny quasi-monotoniczności f na $\mathscr{X}$. Podobnie, każdy z warunków

(2.6) $\min\{f(x_1), f(x_2)\} < f(x_0) < \max\{f(x_1), f(x_2)\}$ dla dowolnych

$x_1, x_2 \in \mathscr{X}$, $x_0 \in (x_1, x_2)$,

(2.7) f jest rosnąca, malejąca lub stała na każdym odcinku $[x_1, x_2] \subseteq \mathscr{X}$,

jest równoważny ścisłej quasi-monotoniczności f na $\mathscr{X}$. Własności funkcji f, o których mowa w (2.5) i (2.7), rozumiemy tu jako własności funkcji $h(\lambda) = f\big((1-\lambda)x_1 + \lambda x_2\big)$ na przedziale $[0, 1]$. Na rysunku 1 funkcje druga i trzecia są quasi-monotoniczne na $[0, 1]$, trzecia jest przy tym ściśle quasi-monotoniczna. Zauważmy jeszcze, że każda funkcja stała na $\mathscr{X}$ jest ściśle quasi-monotoniczna na $\mathscr{X}$.

Jeśli f jest wypukła na $\mathscr{X}$, to dla dowolnych x_1, $x_2 \in \mathscr{X}$ i $\lambda \in (0, 1)$ zachodzą nierówności:

$$f\big((1-\lambda)x_1 + \lambda x_2\big) \leqslant (1-\lambda) f(x_1) + \lambda f(x_2) \leqslant \max\{f(x_1), f(x_2)\},$$

co dowodzi, że f jest quasi-wypukła na $\mathscr{X}$. Co więcej, jeśli $f(x_1) \neq f(x_2)$, to druga z tych nierówności jest ostra. Wykazaliśmy zatem następujące

TWIERDZENIE 2.1. *Funkcja wypukła na $\mathscr{X}$ jest ściśle quasi-wypukła na $\mathscr{X}$.*

W definicji ścisłej quasi-wypukłości żądaliśmy jednak quasi-wypukłości; stąd implikacje:

wypukłość $\Rightarrow$ ścisła quasi-wypukłość $\Rightarrow$ quasi-wypukłość.

Zauważmy, że tak jak quasi-wypukłość jest uogólnieniem wypukłości, tak quasi-monotoniczność jest uogólnieniem afiniczności (funkcję afiniczną zdefiniować można jako zarazem wypukłą i wklęsłą). Zilustrujmy zależności te następującym schematem:

afiniczność

$\Updownarrow$

wypukłość i wklęsłość

$\Downarrow$ $\Downarrow$

ścisła quasi-wypukłość i ścisła quasi-wklęsłość

$\Updownarrow$

ścisła quasi-monotoniczność

$\Downarrow$

quasi-monotoniczność

W przedstawionych definicjach odwoływaliśmy się do własności funkcji na odcinku zawartym w wypukłej dziedzinie $\mathscr{X}$. Funkcje quasi-wypukłe i quasi-monotoniczne można równoważnie określić poprzez zbiory typu warstwicowego. Zbiory

$$L(b) = \{x \in \mathscr{X}: f(x) \leqslant b\},$$

$$U(b) = \{x \in \mathscr{X}: f(x) \geqslant b\},$$

$$G(b) = \{x \in \mathscr{X}: f(x) = b\}$$

będziemy nazywać odpowiednio *zbiorem podwarstwicowym, nadwarstwicowym* i *warstwicowym* funkcji f.

TWIERDZENIE 2.2. *Funkcja f jest quasi-wypukła na $\mathscr{X}$ wtedy i tylko wtedy, gdy każdy jej zbiór podwarstwicowy jest wypukły.*

Własność ta często jest wykorzystywana do definiowania funkcji quasi-wypukłych i jest tą własnością, która leży u podstaw wykorzystania tych funkcji w programowaniu matematycznym. Do problematyki tej powrócimy przy okazji omawiania zbiorów dopuszczalnych. Z twierdzenia 2.2 wynika bezpośrednio następujący

WNIOSEK. *Funkcja f jest quasi-monotoniczna na $\mathscr{X}$ wtedy i tylko wtedy, gdy każdy jej zbiór podwarstwicowy i każdy jej zbiór nadwarstwicowy jest wypukły.*

Zauważmy, że każdy zbiór warstwicowy funkcji quasi-monotonicznej jest wypukły (jako przekrój zbiorów wypukłych $L(b)$ i $U(b)$). Twierdzenie odwrotne jednak nie zachodzi, funkcja

$$f(x) = \begin{cases} -x & \text{dla } x \in [0, \frac{1}{2}], \\ x & \text{dla } x \in (\frac{1}{2}, 1] \end{cases}$$

ma bowiem wypukłe zbiory warstwicowe (jednoelementowe lub puste), a nie jest quasi-monotoniczna.

W praktyce często rozstrzygać trzeba, jaki jest typ wypukłości danej funkcji. Jest to na ogół problem trudny. W prostych przypadkach wykorzystać można analizę zbiorów typu warstwicowego. Na przykład, funkcja $f(x, y) = \frac{x}{y}$ dla $x > 0$ i $y > 0$ jest quasi-monotoniczna w swej dziedzinie, bo jej zbiory podwarstwicowe i nadwarstwicowe określone są warunkami liniowymi ($x > 0$, $y > 0$, $\frac{x}{y} \leqslant b \Leftrightarrow x > 0$, $y > 0$, $x \leqslant by$), a więc są wypukłe. Częściej jednak wykorzystuje się do tego celu różnorakie twierdzenia o superponowaniu. Przytoczymy dwa z nich, pierwsze w wersji uproszczonej.

TWIERDZENIE 2.3. *Jeśli funkcja $f\colon \mathscr{X} \to \mathscr{E}^m$ jest wypukła na $\mathscr{X}$, a funkcja skalarna g jest wypukła (quasi-wypukła, ściśle quasi-wypukła) i składowo niemalejąca na* $\operatorname{conv} f(\mathscr{X})$, *to funkcja $g \circ f$ jest wypukła (quasi-wypukła, ściśle quasi-wypukła) na $\mathscr{X}$.*

TWIERDZENIE 2.4. *Jeśli funkcja $f\colon \mathscr{X} \to \mathscr{E}^1$ jest quasi-wypukła (ściśle quasi-wypukła) na $\mathscr{X}$, a funkcja skalarna g jest niemalejąca (rosnąca) na* $\operatorname{conv} f(\mathscr{X})$, *to funkcja $g \circ f$ jest quasi-wypukła (ściśle quasi-wypukła) na $\mathscr{X}$.*

Symbol $\operatorname{conv} f(\mathscr{X})$ oznacza tu powłokę wypukłą zbioru $f(\mathscr{X})$ wartości funkcji f. Fakt, że g jest w twierdzeniu 2.3 składowo niemalejąca oznacza, że g jest niemalejącą funkcją każdej swej zmiennej przy pozostałych zmiennych ustalonych. Rozważmy na przykład funkcję

$$h(x) = x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_n^{\alpha_n} \quad \text{dla } x \in \mathscr{E}_+^n.$$

Przedstawmy funkcję tę w postaci

$$h(x) = e^{\alpha_1 \ln x_1 + \alpha_2 \ln x_2 + \ldots + \alpha_n \ln x_n}$$

i rozważmy przypadek $\alpha_i \leqslant 0$ $(i \in \overline{1, n})$. Funkcja $f(x) = \alpha_1 \ln x_1 + \alpha_2 \ln x_2 + \ldots + \alpha_n \ln x_n$ jest wtedy wypukła na $\mathscr{E}_+^n$ jako suma funkcji wypukłych, a funkcja $g(t) = e^t$ jest wypukła i rosnąca na $f(\mathscr{E}_+^n) = \mathscr{E}^1$, więc na mocy twierdzenia 2.3 h jest wypukła na $\mathscr{E}_+^n$. W przypadku $\alpha_i \geqslant 0$ $(i \in \overline{1, n})$ f jest ściśle quasi-wklęsła na $\mathscr{E}_+^n$, ponieważ jest to funkcja wklęsła. Ale g jest rosnąca, więc na mocy twierdzenia 2.4, a w zasadzie jego modyfikacji dotyczącej funkcji quasi-wklęsłych i ściśle quasi-wklęsłych wnioskujemy, że h jest w tym przypadku ściśle quasi-wklęsła na $\mathscr{E}_+^n$.

Przejdziemy teraz do omawiania funkcji pseudowypukłych. Niech f będzie funkcją skalarną różniczkowalną na wypukłym zbiorze $\mathscr{X} \subseteq \mathscr{E}^n$. Funkcję f nazywamy *pseudowypukłą* na $\mathscr{X}$, jeśli spełniony jest warunek

$$\bigwedge_{x_1, x_2 \in \mathscr{X}} \left(\langle f'(x_1), x_2 - x_1 \rangle \geqslant 0 \Rightarrow f(x_2) - f(x_1) \geqslant 0 \right)(^1). \tag{2.8}$$

Warunek ten geometrycznie interpretować możemy następująco: jeśli w punkcie x_1 pochodna kierunkowa jest nieujemna w kierunku wskazującym inny punkt zbioru, to wartość funkcji nie zmaleje, gdy z punktu x_1 będziemy przemieszczać się wzdłuż tego kierunku.

Zauważmy, że jeśli f jest pseudowypukła na $\mathscr{E}^n$, to każdy jej punkt stacjonarny (tzn. taki punkt x_1, że $f'(x_1) = 0$) jest punktem bezwarunkowego minimum globalnego tej funkcji. Wynika to natychmiast z definicji, w punkcie stacjonarnym x_1 jest bowiem spełniony warunek $\langle f'(x_1), x_2 - x_1 \rangle \geqslant 0$ dla dowolnych $x_2 \in \mathscr{E}^n$; stąd spełniony jest też warunek $f(x_2) \geqslant f(x_1)$ dla dowolnych $x_2 \in \mathscr{E}^n$. Odpowiednikiem tej sytuacji dla przypadku minimalizacji funkcji f na zbiorze $\mathscr{X}$, a więc dla przypadku minimalizacji warunkowej, jest następujące twierdzenie:

TWIERDZENIE 2.5. *Jeśli funkcja f jest pseudowypukła na $\mathscr{X}$ i $x_1 \in \mathscr{X}$ jest potencjalnym punktem minimalnym f na $\mathscr{X}$ (tzn. spełniona jest nierówność $\langle f'(x_1), x_2 - x_1 \rangle \geqslant 0$ dla dowolnych $x_2 \in \mathscr{X}$), to x_1 jest punktem minimum globalnego f na $\mathscr{X}$.*

Odwołując się do geometrycznej interpretacji pochodnej kierunkowej, łatwo stwierdzamy, że potencjalnym punktem minimalnym f na $\mathscr{X}$ jest taki punkt $x_1 \in \mathscr{X}$, w którym pochodna kierunkowa nie wskazuje na istnienie kierunku dopuszczalnego i poprawiającego (tzn. prowadzącego do zbioru $\mathscr{X}$ i dającego spadek wartości funkcji f). Własność funkcji pseudowypukłych wyrażona w twierdzeniu 2.5, podobnie jak własność funkcji quasi-wypukłych wyrażona w twierdzeniu 2.2, jest decydująca z punktu widzenia wykorzystania tych funkcji w programowaniu matematycznym (por. cechy (1.6) i (1.7)).

Funkcję f nazywamy *pseudowklęsłą* na $\mathscr{X}$, jeśli $-f$ jest pseudowypukła na $\mathscr{X}$. Funkcję f nazywamy *pseudomonotoniczną* na $\mathscr{X}$, jeśli f jest pseudowypukła i pseudowklęsła na $\mathscr{X}$.

Następne twierdzenie ujmuje zależności pomiędzy wypukłością, pseudowypukłością i quasi-wypukłością.

TWIERDZENIE 2.6. *Funkcja wypukła i różniczkowalna na $\mathscr{X}$ jest pseudowypukła na $\mathscr{X}$, a funkcja pseudowypukła na $\mathscr{X}$ jest ściśle quasi-wypukła na $\mathscr{X}$.*

Dla funkcji różniczkowalnych mamy zatem implikacje:

wypukłość $\Rightarrow$ pseudowypukłość $\Rightarrow$ ścisła quasi-wypukłość $\Rightarrow$ quasi-wypukłość.

---

(1) Symbolem $\langle x, y \rangle$ oznaczamy iloczyn skalarny wektorów $x, y \in \mathscr{E}^n$, tzn.

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i.$$

Odwołując się do analogicznych implikacji obejmujących różne typy wklęsłości oraz do poprzednio wprowadzonych definicji, odnotujmy jeszcze implikacje:

$$\text{afiniczność} \Rightarrow \text{pseudomonotoniczność} \Rightarrow$$
$$\Rightarrow \text{ścisła quasi-monotoniczność} \Rightarrow \text{quasi-monotoniczność}.$$

Sformułowane wyżej definicje quasi-wypukłości i pseudowypukłości mają charakter globalny, dotyczą bowiem własności funkcji f na całym zbiorze $\mathscr{X}$. Własności te można jednak rozpatrywać w sensie lokalnym, w ustalonym punkcie $\bar{x} \in \mathscr{X}$. Wtedy można nawet pominąć założenie wypukłości zbioru $\mathscr{X}$, my jednak przytoczymy odpowiednie definicje w węższym sensie, zachowując założenie wypukłości $\mathscr{X}$. I tak, funkcję f nazywamy *quasi-wypukłą* w punkcie $\bar{x} \in \mathscr{X}$ (w odniesieniu do zbioru wypukłego $\mathscr{X}$), jeśli spełniony jest warunek

$$(2.9) \qquad f\big((1-\lambda)\bar{x}+\lambda x\big) \leqslant \max\{f(\bar{x}), f(x)\} \quad \text{dla dowolnych } x \in \mathscr{X},\ \lambda \in (0, 1).$$

Podobnie, funkcję f różniczkowalną w punkcie $\bar{x} \in \mathscr{X}$ nazywamy *pseudowypukłą* w punkcie $\bar{x}$ (w odniesieniu do zbioru wypukłego $\mathscr{X}$), jeśli spełniony jest warunek

$$(2.10) \qquad \bigwedge_{x \in \mathscr{X}} \big(\langle f'(\bar{x}), x-\bar{x}\rangle \geqslant 0 \Rightarrow f(x)-f(\bar{x}) \geqslant 0\big).$$

Oczywiste jest, że f jest quasi-wypukła (pseudowypukła) na zbiorze wypukłym $\mathscr{X}$ wtedy i tylko wtedy, gdy f jest quasi-wypukła (pseudowypukła) w każdym punkcie zbioru $\mathscr{X}$.

Na zakończenie podamy twierdzenie o linearyzacji ograniczeń pseudomonotonicznych. Orzeka ono o możliwości zastąpienia warunku nieliniowego $f(x) \leqslant b$ z pseudomonotoniczną funkcją f warunkiem liniowym $\langle f'(x_0), x-x_0\rangle \leqslant 0$.

TWIERDZENIE 2.7. *Jeśli funkcja f jest pseudomonotoniczna na $\mathscr{X}$, to dla każdego $b \in f(\mathscr{X})$ oraz $x_0 \in \mathscr{G}(b)$ zachodzi równość*

$$L(b) = \{x \in \mathscr{X}\colon \langle f'(x_0), x-x_0\rangle \leqslant 0\}.$$

3. Zbiory dopuszczalne i minima. Skupimy teraz naszą uwagę na zbiorze rozwiązań dopuszczalnych zadania programowania matematycznego. Niech będzie to zbiór postaci

$$\mathscr{Q} = \{x \in \mathscr{X}\colon f(x) \leqslant b\},$$

gdzie $b \in \mathscr{E}^m$, a f jest m-wymiarową wektor funkcją określoną na $\mathscr{X} \subseteq \mathscr{E}^n$. Bardzo pożądanymi cechami tego zbioru są jego wypukłość i domkniętość. Stąd duże znaczenie następnego twierdzenia, będącego w istocie bezpośrednią konsekwencją twierdzenia 2.2 i definicji półciągłości.

TWIERDZENIE 3.1. *Jeśli zbiór $\mathscr{X}$ jest wypukły i domknięty, to $\mathscr{Q}$ jest wypukły i domknięty dla każdego $b \in \mathscr{E}^m$ wtedy i tylko wtedy, gdy f jest quasi-wypukła i półciągła z dołu na $\mathscr{X}$.*

Zwróćmy uwagę na to, że żądanie wypukłości i domkniętości $\mathcal{Q}$ dla każdego $b \in \mathscr{E}^m$ jest o wiele mocniejsze niż w praktyce. Tam bowiem b jest na ogół ustalone i $\mathcal{Q}$ może okazać się wypukły i domknięty, pomimo tego, iż niektóre spośród składowych funkcji f nie będą quasi-wypukłe i półciągłe z dołu na $\mathcal{X}$ (np. funkcja f_i ze zbędnego warunku ograniczającego $f_i(x) \leqslant b_i$). Jeśli wśród warunków ograniczających występują równania, to warunkiem wystarczającym (ale nie koniecznym) wypukłości i domkniętości $\mathcal{Q}$ dla każdego b jest quasi-monotoniczność i ciągłość funkcji występujących w tych równaniach (dla pozostałych warunków wymagania nie ulegają zmianie). Wynika to natychmiast z możliwości zastąpienia każdego równania dwoma nierównościami.

Szczególną rolę w zastosowaniach pełnią zadania z wielościennymi zbiorami rozwiązań dopuszczalnych. Kolejne twierdzenie podaje warunki na to, aby $\mathcal{Q}$ był takim właśnie zbiorem.

TWIERDZENIE 3.2. *Jeśli $\mathcal{X}$ jest zbiorem wielościennym i f jest quasi-monotoniczna i półciągła z dołu na $\mathcal{X}$, to $\mathcal{Q}$ jest zbiorem wielościennym dla każdego $b \in \mathscr{E}^m$.*

Skoro $\mathcal{Q}$ jest przy tych założeniach zbiorem wielościennym, to określić go można liniowymi warunkami ograniczającymi. Wykazano, że każdy z warunków definiujących zbiór $\mathcal{Q}$, w którym występuje funkcja quasi-monotoniczna i półciągła z dołu, zastąpić można co najwyżej dwoma warunkami liniowymi. Nie jest jednak znana procedura umożliwiająca efektywne wyznaczenie układu liniowego zastępującego układ wyjściowy. Potrafimy to zrobić tylko w szczególnym przypadku, kiedy f jest pseudomonotoniczna na $\mathcal{X}$ (f jest więc też wtedy quasi-monotoniczna i ciągła). Zgodnie z twierdzeniem 2.7 każde ograniczenie $f_i(x) \leqslant b_i$ $(i \in \overline{1, m})$ zastąpić możemy wówczas dokładnie jednym ograniczeniem liniowym $\langle f'(x_0), x - x_0 \rangle \leqslant 0$. Podkreślmy przy tym, że punkt x_0 nie musi być dobierany jako wspólny dla wszystkich ograniczeń (poszczególne ograniczenia linearyzować można niezależnie). W istocie trzeba więc dla każdego $i \in \overline{1, m}$ wyznaczyć $x_{i0} \in \mathcal{X}$ taki, że $f_i(x_{i0}) = b_i$, a następnie przyjąć

$$\mathcal{Q} = \{x \in \mathcal{X}:\ \langle f_i'(x_{i0}), x - x_{i0} \rangle \leqslant 0,\ i \in \overline{1, m}\}.$$

Przejdziemy teraz do rozpatrywania zbioru $\mathcal{Q}^*$ rozwiązań optymalnych zadania (1.1). Interesować nas będą warunki, przy których:

– $\mathcal{Q}^* \neq \emptyset$ (f_0 osiąga minimum na $\mathcal{Q}$),

– minimum lokalne f na $\mathcal{Q}$ jest minimum globalnym f na $\mathcal{Q}$,

– $\mathcal{Q}^*$ jest zbiorem wypukłym (w szczególności wielościennym),

– f_0 osiąga minimum na $\mathcal{Q}$ w punkcie ekstremalnym zbioru $\mathcal{Q}$ (w szczególności w wierzchołku zbioru wielościennego $\mathcal{Q}$) [1].

[1] $\bar{x} \in \mathcal{Q}$ jest punktem ekstremalnym (wierzchołkiem) zbioru wypukłego $\mathcal{Q}$, jeśli $\mathcal{Q} \setminus \{\bar{x}\}$ jest zbiorem wypukłym.

Warunków wystarczających na to, aby $\mathcal{Q}^* \neq \emptyset$, nie będziemy szerzej komentować. Przypomnimy tylko dwa z takich warunków:

(1) $\mathcal{Q}$ jest niepustym zbiorem zwartym i f_0 jest półciągła z dołu na $\mathcal{Q}$,

(2) $\mathcal{Q}$ jest niepustym zbiorem wielościennym i f_0 jest afiniczna i ograniczona z dołu na $\mathcal{Q}$.

Powszechnie wiadomo, że każde minimum lokalne funkcji wypukłej f_0 na zbiorze wypukłym $\mathcal{Q}$ jest jej minimum globalnym na tym zbiorze. Ogólniejsza wersja tego twierdzenia brzmi następująco.

TWIERDZENIE 3.3. *Jeśli $\mathcal{Q}$ jest wypukły i f_0 jest ściśle quasi-wypukła na $\mathcal{Q}$, to minimum lokalne f_0 na $\mathcal{Q}$ jest minimum globalnym f_0 na $\mathcal{Q}$.*

Podkreślmy, że założenia ścisłej quasi-wypukłości nie można tu osłabić, zakładając tylko quasi-wypukłość. Łatwo się o tym przekonać, rozpatrując przykład $\min\{\operatorname{sgn} x\colon x \geqslant 0\}$.

Kolejne dwa twierdzenia wynikają bezpośrednio z twierdzeń 2.2 i 3.2, jeśli uwzględnić fakt, że $\mathcal{Q}^* = \{x \in \mathcal{Q}\colon f_0(x) \leqslant f_0(\bar{x})\}$, gdzie $\bar{x} \in \mathcal{Q}^*$ ($\bar{x}$ jest dowolnym ustalonym rozwiązaniem optymalnym).

TWIERDZENIE 3.4. *Jeśli $\mathcal{Q}$ jest zbiorem wypukłym i f_0 jest quasi-wypukła na $\mathcal{Q}$, to $\mathcal{Q}^*$ jest zbiorem wypukłym.*

TWIERDZENIE 3.5. *Jeśli $\mathcal{Q}$ jest zbiorem wielościennym i f_0 jest quasi-monotoniczna i półciągła z dołu na $\mathcal{Q}$, to $\mathcal{Q}^*$ jest zbiorem wielościennym.*

Twierdzenia te nie zawierają oczywiście warunków na to, aby $\mathcal{Q}^* \neq \emptyset$ (dopuszczamy w nich możliwość $\mathcal{Q}^* = \emptyset$). Zauważmy, że w szczególności, minimalizując na zbiorze wielościennym $\mathcal{Q}$ funkcję afiniczną (a więc w programowaniu liniowym) lub, ogólniej, funkcję pseudomonotoniczną, uzyskujemy $\mathcal{Q}^*$ jako zbiór wielościenny.

Zajmiemy się teraz warunkami, przy których f_0 osiąga minimum na $\mathcal{Q}$ w punkcie ekstremalnym (wierzchołku) zbioru $\mathcal{Q}$. Zadania programowania matematycznego mające tę własność są na ogół dużo prostsze do rozwiązania, można bowiem ograniczyć się wtedy do analizy wyłącznie punktów ekstremalnych, których jest zwykle znacznie mniej niż wszystkich rozwiązań dopuszczalnych. Ważne znaczenie ma tu także to, że zbiór wypukły i zwarty jest wypukłą powłoką swych punktów ekstremalnych (wykazano nawet więcej, że każdy punkt zbioru wypukłego i zwartego w $\mathcal{E}^n$ jest wypukłą kombinacją liniową co najwyżej $n+1$ punktów ekstremalnych tego zbioru). Jeśli np. punktami ekstremalnymi wielościanu $\mathcal{Q}$ (tzn. zbioru wielościennego ograniczonego) są $x_1, x_2, \ldots, x_k$, to każdy $x \in \mathcal{Q}$ jest postaci

$$x = \sum_{i=1}^{k} \lambda_i x_i \quad \Big(\sum_{i=1}^{k} \lambda_i = 1;\ \lambda_i \geqslant 0,\ i \in \overline{1, k}\Big).$$

Stąd dla quasi-wklęsłej funkcji f mamy

$$f_0(x) \geqslant \min\{f_0(x_i)\colon i \in \overline{1, k}\} = f(x_s).$$

Wykazaliśmy zatem, że quasi-wklęsła funkcja f_0 osiąga minimum na wielościanie $\mathcal{Q}$ i to w wierzchołku tego wielościanu (x_s jest tym wierzchołkiem). Podobnie dowieść można innego twierdzenia, dotyczącego niekoniecznie zbiorów wielościennych.

TWIERDZENIE 3.6. *Funkcja quasi-wklęsła i półciągła z dołu na zbiorze $\mathcal{Q}$ wypukłym i zwartym osiąga minimum na $\mathcal{Q}$ w punkcie ekstremalnym zbioru $\mathcal{Q}$.*

Oczywiście założenie półciągłości z dołu, którego poprzednio nie było, gwarantuje istnienie minimum i rekompensuje osłabienie założeń o zbiorze $\mathcal{Q}$.

Powróćmy teraz jeszcze raz do rozpatrywania zbiorów wielościennych. Pewne metody minimalizacji funkcji afinicznej na zbiorze wielościennym (np. metoda sympleks) polegają na kolejnym wyznaczaniu wierzchołków sąsiednich tego zbioru tak, aby każdy następny był nie gorszy od poprzedniego. Podobną procedurę stosować można w przypadku pewnych funkcji nieliniowych. Kluczowym jest więc pytanie, kiedy wierzchołek nie gorszy (z punktu widzenia minimalizacji funkcji f_0) od wierzchołków sąsiednich jest nie gorszy od wszystkich innych wierzchołków i dalej, nie gorszy od wszystkich innych punktów zbioru. Aby sformułować odpowiednie twierdzenia, nazwijmy wierzchołek $\bar{x}$ zbioru wielościennego $\mathcal{Q}$ *wierzchołkowym minimum globalnym* (*lokalnym*) f_0 na $\mathcal{Q}$, jeśli $f_0(\bar{x}) \leqslant f_0(x)$ dla każdego (każdego sąsiedniego do $\bar{x}$) wierzchołka x zbioru $\mathcal{Q}$.

TWIERDZENIE 3.7. *Jeśli $\mathcal{Q}$ jest wielościanem i f_0 jest quasi-wklęsła na $\mathcal{Q}$, to wierzchołkowe minimum globalne f_0 na $\mathcal{Q}$ jest minimum globalnym f_0 na $\mathcal{Q}$.*

TWIERDZENIE 3.8. *Jeśli $\mathcal{Q}$ jest wielościanem i f_0 jest quasi-wklęsła i ściśle quasi-wypukła na $\mathcal{Q}$, to wierzchołkowe minimum lokalne f_0 na $\mathcal{Q}$ jest wierzchołkowym minimum globalnym f_0 na $\mathcal{Q}$ (i w konsekwencji — minimum globalnym f_0 na $\mathcal{Q}$).*

Pierwsze z tych twierdzeń było już tu wypowiedziane, w innym sformułowaniu, przed twierdzeniem 3.6. Zwróćmy uwagę, że twierdzenie 3.8 stwarza możliwość podania prostego algorytmu minimalizacji na wielościanie funkcji ściśle quasi-monotonicznej, czy też w szczególności funkcji pseudomonotonicznej. Dla tej ostatniej postępowanie jest analogiczne do klasycznej metody sympleks, możemy bowiem używać gradientu funkcji do oceny, czy dany wierzchołek jest wierzchołkowym minimum lokalnym. Dla funkcji pseudomonotonicznej, tak jak dla funkcji afinicznej, pochodna kierunkowa wskazuje kierunek ścisłej monotoniczności funkcji lub jej stałości (por. (2.7)). Stąd przy ocenie, czy uzyskany wierzchołek $\bar{x}$ jest już rozwiązaniem optymalnym, wystarczy uwzględniać tylko pochodne kierunkowe w $\bar{x}$ w kierunkach wierzchołków sąsiednich do $\bar{x}$. Dodajmy jeszcze, że twierdzenia 3.7 i 3.8 można przenieść na przypadek, w którym $\mathcal{Q}$ jest zbiorem wielościennym i dla tego przypadku również sformułować odpowiednie algorytmy. Zainteresowanego czytelnika odsyłamy do wspomnianej monografii B. Martosa, w której ta problematyka jest szczególnie godna przestudiowania.

4. Warunek dostateczny Kuhna–Tuckera. Wróćmy teraz do twierdzeń formułujących warunki konieczne i dostateczne dla rozwiązań optymalnych zadań programowania nieliniowego. Zajmiemy się tylko warunkami w klasie zadań o funkcjach różniczkowalnych, tak zwanymi *warunkami Kuhna–Tuckera.* Znane są dwa tego typu twierdzenia: warunek dostateczny optymalności i warunek konieczny optymalności.

Zanim je sformułujemy, zdefiniujemy zbiór ograniczeń aktywnych w punkcie $x \in \mathcal{Q}$ jako zbiór indeksów tych funkcji ograniczających f_i $(i \in \overline{1, m})$, które w punkcie x przyjmują wartość 0, czyli

$$(4.1) \qquad \mathcal{I}(x) = \{i \in \overline{1, m}: f_i(x) = 0\}.$$

Znaczenie zbioru $\mathcal{I}(x)$ polega na tym, że (np. przy założeniu ciągłości funkcji f_i) w najbliższym otoczeniu punktu x istotnymi nierównościami definiującymi zbiór dopuszczalny $\mathcal{Q}$ są tylko ograniczenia aktywne w punkcie x.

Pierwsze ze wspomnianych wyżej twierdzeń jest zwykle formułowane w sposób następujący:

TWIERDZENIE 4.1 (*warunek dostateczny optymalności*). *Załóżmy, że dany jest punkt dopuszczalny $\bar{x} \in \mathcal{Q}$. Jeśli dla funkcji f_i: $\mathcal{X} \to \mathcal{E}^1$ $(i \in \overline{0, m})$ spełnione są założenia:*

(1) *funkcja celu f_0 jest pseudowypukła w punkcie $\bar{x}$;*

(2) *funkcje f_i $(i \in \mathcal{I}(\bar{x}))$ reprezentujące ograniczenia aktywne w $\bar{x}$ są różniczkowalne i quasi-wypukłe w $\bar{x}$;*

(3) *istnieje wektor $\bar{u} = (\bar{u}_1, \ldots, \bar{u}_m)$ o współrzędnych nieujemnych taki, że para $(\bar{x}, \bar{u})$ spełnia równanie*

$$(4.2) \qquad f_0'(\bar{x}) + \sum_{i \in \mathcal{I}(\bar{x})} \bar{u}_i f_i'(\bar{x}) = 0,$$

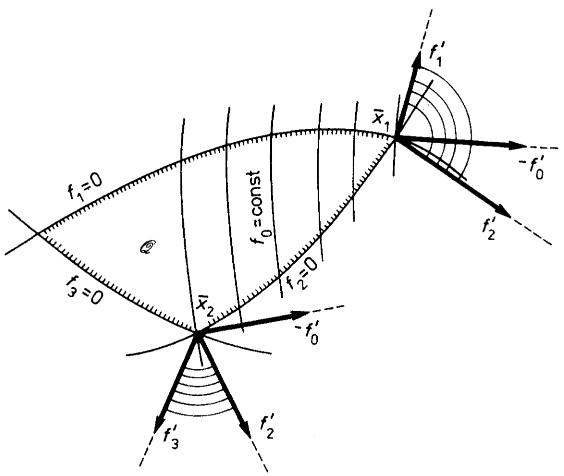
to $\bar{x}$ jest rozwiązaniem optymalnym zadania (1.1).

Twierdzenie to można traktować jako kryterium pozwalające zbadać, czy dany punkt jest optymalny (jeśli spełnione są przyjęte założenia). W tym celu należy wyznaczyć zbiór $\mathcal{I}(\bar{x})$, obliczyć gradienty $P_i = f_i'(x)$ dla $i \in \mathcal{I}(\bar{x}) \cup \{0\}$ oraz podjąć próbę rozwiązania układu równań i nierówności liniowych (np. metodami programowania liniowego):

$$(4.3) \qquad \sum_{i \in \mathcal{I}(\bar{x})} P_i u_i = -P_0, \qquad u_i \geqslant 0 \ (i \in \mathcal{I}(\bar{x})).$$

Jeśli znajdziemy rozwiązanie, to na pewno $\bar{x}$ jest rozwiązaniem optymalnym. Wniosek przeciwny na ogół nie jest prawdziwy. Świadczy o tym następujące zadanie: $\min\{x: x^2 \leqslant 0\}$. Założenia (1) i (2) twierdzenia 4.1 są spełnione. Równanie (4.2) dla $\bar{x} = 0$ przyjmuje postać: $1 + 0u = 0$, a więc jest sprzeczne. Mimo to $\bar{x} = 0$ jest rozwiązaniem optymalnym. Jeśli jednak spełnione będą jednocześnie pewne założenia gwarantujące zachodzenie warunków Kuhna–Tuckera dla każdego punktu optymalnego (por. p. 5), to będzie można mieć również pewność, że jeśli układ (4.3) jest sprzeczny, to $\bar{x}$ nie jest optymalny i na jego miejsce szukać należy innego kandydata.

Zwróćmy jeszcze uwagę, że równanie (4.2) ma interesującą interpretację o charakterze analityczno-geometrycznym, a mianowicie: antygradient $-f_0'$ funkcji celu w punkcie $\bar{x}$ daje się przedstawić w postaci kombinacji liniowej z nieujemnymi współczynnikami gradientów (w punkcie $\bar{x}$) funkcji ograniczających f_i, reprezentujących ograniczenia aktywne. Innymi słowy antygradient $-f_0'(\bar{x})$ należy do stożka wypukłego rozpiętego na gradientach $f_i'(\bar{x})$ ograniczeń aktywnych. Jeśli weźmiemy pod uwagę, że gradient $f'(\bar{x})$ jest prostopadły do brzegu zbioru warstwicowego $\mathscr{G}(f(\bar{x}))$ w punkcie $\bar{x}$, to sytuację tę możemy zilustrować jak na rysunku 2.



Rys. 2. $\bar{x}_1$ — punkt optymalny, $\bar{x}_2$ — punkt nieoptymalny

Bez trudu można zauważyć, że twierdzenie 4.1 jest uogólnieniem twierdzenia 1.4 w tym sensie, że założenia zostały osłabione. Mianowicie, wypukłość funkcji f_i została zastąpiona lokalną pseudowypukłością f_0 i lokalną quasi-wypukłością f_i (i to tylko dla ograniczeń aktywnych). Zauważmy, że założenia te pozostawiają nas nadal w klasie zadań posiadających trzy pozytywne cechy (1.6)-(1.8) wymienione w p. 1. Wprawdzie lokalny charakter tych założeń zachowuje te cechy jedynie w pobliżu rozwiązania $\bar{x}$, jeśli jednak przyjmiemy te założenia w sensie globalnym, to wspomniane cechy przeniosą się w całej rozciągłości.

5. Warunek konieczny Kuhna–Tuckera; regularność. Twierdzenie 1.3 wyraża łatwy do dowiedzenia fakt, że jeśli para $(\bar{x}, \bar{u})$ jest punktem siodłowym funkcji Lagrange'a oraz funkcje f_i $(i \in \overline{0, m})$ są różniczkowalne w $\bar{x}$, to spełnione są warunki Kuhna–Tuckera (1.5). Inaczej to wygląda, gdy pytamy, kiedy dla danego punktu optymalnego $\bar{x}$ istnieje $\bar{u} = (\bar{u}_1, \ldots, \bar{u}_m)$ o nieujemnych współrzędnych i taki, że para $(\bar{x}, \bar{u})$ spełnia warunki Kuhna–Tuckera. Sygnałem trudności w odpowiedzi na to pytanie jest twierdzenie 1.2, w którym po to, aby punkt optymalny można było uzupełnić do punktu siodłowego, trzeba było założyć wypukłość funkcji f_i $(i \in \overline{0, m})$ łącznie z warunkiem Slatera. W ten sposób doszliśmy do twierdzenia mówiącego, przy jakich założeniach zachodzą warunki Kuhna–Tuckera.

Twierdzenie 5.1 (Warunek konieczny optymalności). *Jeśli spełnione są założenia:*

(1) *$\mathscr{X} \subseteq \mathscr{E}^n$ jest niepusty i otwarty;*

(2) *$\bar{x} \in \mathscr{Q}$ jest minimum lokalnym f_0 na $\mathscr{Q}$;*

(3) *funkcje f_0 oraz f_i dla $i \in \mathscr{I}(\bar{x})$ są różniczkowalne w punkcie $\bar{x}$, a dla $i \notin \mathscr{I}(\bar{x})$ ciągłe w $\bar{x}$;*

(4) *spełniony jest jeden z warunków regularności,*

to istnieją $\bar{u}_i \geqslant 0$ $(i \in \mathscr{I}(\bar{x}))$, dla których spełnione jest równanie

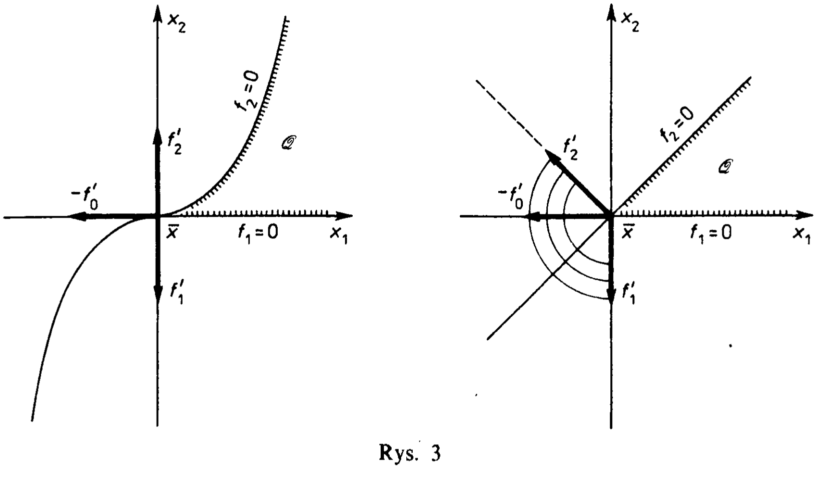
$$f_0'(\bar{x}) + \sum_{i \in \mathscr{I}(\bar{x})} \bar{u}_i f_i'(\bar{x}) = 0.$$

Do najbardziej znanych warunków regularności należą:

(C1) (*warunek liniowej niezależności wektorów normalnych*). *Gradienty $f_i'(\bar{x})$ $(i \in \mathscr{I}(\bar{x}))$ aktywnych w punkcie $\bar{x}$ funkcji ograniczających są liniowo niezależne.*

(C2) (*warunek Slatera*). *$\mathscr{X}$ jest otwarty i wypukły, funkcje f_i przy $i \in \mathscr{I}(\bar{x})$ są pseudowypukłe w $\bar{x}$ oraz istnieje punkt $x \in \mathscr{X}$ taki, że $f_i(x) < 0$ dla $i \in \mathscr{I}(\bar{x})$.*

Warunki regularności mają zwykle charakter lokalny i odnoszą się do danego punktu $\bar{x} \in \mathscr{Q}$. Warunek Slatera bardziej znany jest jednak w sformułowaniu: funkcje ograniczeń f_i są wypukłe na $\mathscr{X}$ oraz istnieje $x \in \mathscr{X}$, dla którego $f_i(x) < 0$ dla wszystkich $i \in \overline{1, m}$. Przy okazji pragniemy zwrócić uwagę na fakt, że z wypukłości funkcji f_i na zbiorze wypukłym otwartym $\mathscr{X}$ wynika ich ciągłość na tym zbiorze, a zatem tak sformułowany warunek



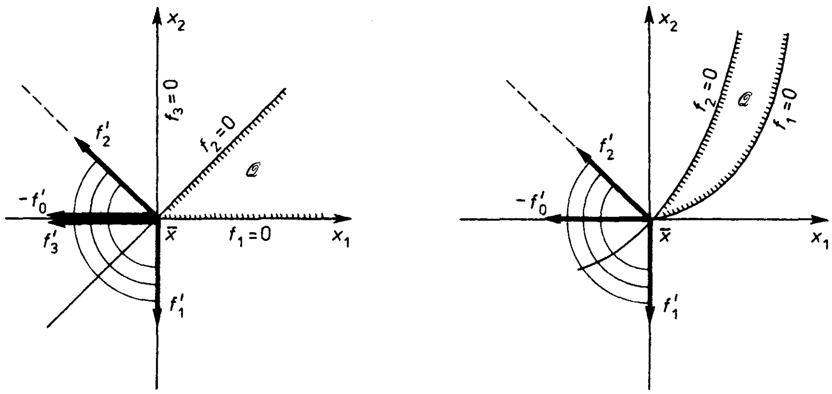
Rys. 3

Slatera implikuje, że wnętrze zbioru $\mathcal{Q}$ jest niepuste. Nie zachodzi jednak wynikanie odwrotne. Nie należy więc utożsamiać (jak to się czasem zdarza) warunku Slatera z warunkiem $\operatorname{int} \mathcal{Q} \neq \emptyset$, gdyż istota warunku Slatera nie polega na tym, że zbiór dopuszczalny ma niepuste wnętrze.

Następujące przykłady oraz rys. 3 i 4 ilustrują działanie warunków (C1), (C2). We wszystkich tych przykładach $\bar{x} = (\bar{x}_1, \bar{x}_2) = (0, 0)$ oraz $\mathscr{X} = \mathscr{E}^2$.

(P1) $$\begin{aligned} f_0 &= x_1 && \to \min, \\ f_1 &= -x_2 && \leqslant 0, \\ f_2 &= x_2 - x_1^3 && \leqslant 0; \end{aligned}$$

(P2) $$\begin{aligned} f_0 &= x_1 && \to \min, \\ f_1 &= -x_2 && \leqslant 0, \\ f_2 &= x_2 - x_1 && \leqslant 0; \end{aligned}$$



Rys. 4

(P3) $$\begin{aligned} f_0 &= x_1 && \to \min, \\ f_1 &= -x_2 && \leqslant 0, \\ f_2 &= x_2 - x_1 && \leqslant 0, \\ f_3 &= -x_1 && \leqslant 0; \end{aligned}$$

(P4) $$\begin{aligned} f_0 &= x_1 && \to \min, \\ f_1 &= x_1^2 - x_2 && \leqslant 0, \\ f_2 &= x_2 - (x_1+1)^2 + 1 && \leqslant 0. \end{aligned}$$

Wyniki podsumujemy w tabeli.

	$\mathscr{I}(\bar{x})$	$-f_0'$	f_i', $i \in \mathscr{I}(\bar{x})$	(C1)	(C2)	$\bar{u}$
(P1)	$\{1, 2\}$	$(-1, 0)$	$(0, -1)$ $(0, 1)$	–	–	nie istnieje
(P2)	$\{1, 2\}$	$(-1, 0)$	$(0, -1)$ $(-1, 1)$	+	+	$\bar{u}_1 = 1$ $\bar{u}_2 = 1$
(P3)	$\{1, 2, 3\}$	$(-1, 0)$	$(0, -1)$ $(-1, 1)$ $(-1, 0)$	–	+	$\bar{u}_1 = 1$ $\bar{u}_2 = 1$ $\bar{u}_3 = 0$
(P4)	$\{1, 2\}$	$(-1, 0)$	$(0, -1)$ $(-2, 1)$	+	–	$\bar{u}_1 = 0.5$ $\bar{u}_2 = 0.5$

Jak na to wskazują powyższe przykłady, warunki (C1) i (C2) są niezależne w tym sensie, że żaden z nich nie jest konsekwencją drugiego. Tylko w jednym z nich, w warunku Slatera, występują pojęcia ze sfery wypukłości. Warunek (C1) nie nawiązuje do tych pojęć. W pozostałych założeniach twierdzenia 5.1 pojęcia wypukłości również nie występują. Uzasadniona zatem będzie sugestia, że istota warunku koniecznego Kuhna–Tuckera nie wiąże się tak silnie z wypukłością, jak warunek dostateczny wyrażony w twierdzeniu 4.1.

Martos w swej książce [1] posługuje się innymi warunkami regularności. Aby je sformułować, wprowadzimy następujące oznaczenia:

$$\mathcal{J}^1(\bar{x}) = \{i \in \mathcal{J}(\bar{x}) \colon f_i \textit{ jest pseudowklęsła w } \bar{x}\},$$

$$\mathcal{J}^2(\bar{x}) = \mathcal{J}(\bar{x}) \setminus \mathcal{J}^1(\bar{x}).$$

Są to zatem zbiory wskaźników ograniczeń aktywnych, które odpowiednio są i nie są pseudowklęsłe.

(C3) (*uogólniony warunek Slatera*). *Funkcje* f_i $(i \in \mathcal{J}(\bar{x}))$ *ograniczeń aktywnych są pseudowypukłe w* $\bar{x}$ *oraz istnieje punkt* $x \in \mathcal{X}$ *taki, że* $f_i(x) < 0$ *dla* $i \in \mathcal{J}^2(\bar{x})$.

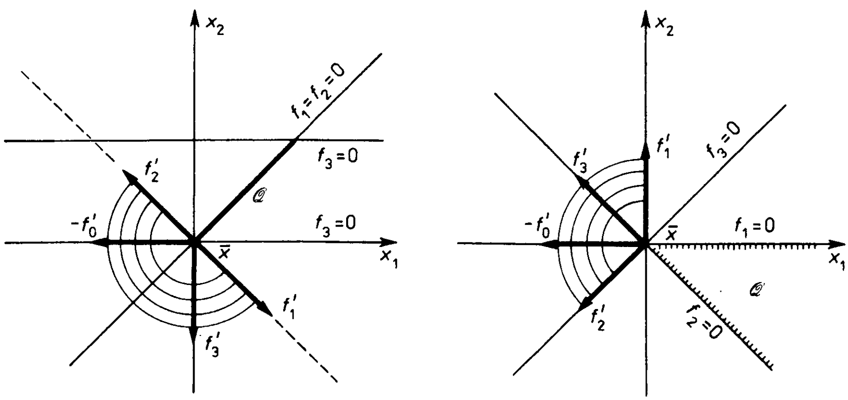
Uogólnienie warunku Slatera (C2) polega więc na tym, że jeśli wśród aktywnych funkcji ograniczających są funkcje pseudomonotoniczne w $\bar{x}$, to dla nich istnienie punktu x takiego, że $f_i(x) < 0$ $(i \in \mathcal{J}^1(\bar{x}))$ nie jest niezbędne.

(C4) (*warunek Mangasariana*). *Funkcje* f_i *są różniczkowalne w* $\bar{x}$ *dla* $i \in \mathcal{J}(\bar{x})$ *oraz istnieje wektor* $p \in \mathcal{E}^n$ *taki, że*

$$\langle p, f_i'(\bar{x})\rangle \geq 0 \quad \textit{dla } i \in \mathcal{J}^1(\bar{x}),$$

$$\langle p, f_i'(\bar{x})\rangle > 0 \quad \textit{dla } i \in \mathcal{J}^2(\bar{x}).$$

TWIERDZENIE 5.2. *Uogólniony warunek Slatera implikuje warunek Mangasariana.*



Rys. 5

Kolejne dwa przykłady oraz rys. 5 są ilustracją warunków (C3) i (C4).

(P5)
$$\begin{aligned} f_0 &= x_1 && \to \min, \\ f_1 &= x_1 - x_2 && \leqslant 0, \\ f_2 &= x_2 - x_1 && \leqslant 0, \\ f_3 &= (x_2-1)^2 - 1 && \leqslant 0; \end{aligned}$$

(P6)
$$\begin{aligned} f_0 &= x_1 && \to \min, \\ f_1 &= -(x_2-1)^2+1 && \leqslant 0, \\ f_2 &= -x_1 - x_2 && \leqslant 0, \\ f_3 &= x_2 - x_1 && \leqslant 0. \end{aligned}$$

	$\mathscr{I}(\bar{x})$	$-f_0'$	f_i'	$\mathscr{I}^1(\bar{x})$	$\mathscr{I}^2(\bar{x})$	(C3)	(C4)	p	$\bar{u}$
(P5)	$\{1, 2, 3\}$	$(-1, 0)$	$(1, -1)$ $(-1, 1)$ $(0, -2)$	$\{1, 2\}$	$\{3\}$	+	+	$(-1, -1)$	$\bar{u}_1 = 0$ $\bar{u}_2 = 1$ $\bar{u}_3 = 0.5$
(P6)	$\{1, 2, 3\}$	$(-1, 0)$	$(0, 2)$ $(-1, -1)$ $(-1, 1)$	$\{1, 2, 3\}$	$\emptyset$	−	+	$(-1, 1)$	$\bar{u}_1 = 0$ $\bar{u}_2 = 0.5$ $\bar{u}_3 = 0.5$

Zarówno uogólniony warunek Slatera, jak i warunek Mangasariana wykorzystują pseudowypukłość, pseudowklęsłość, a w konsekwencji i pseudomonotoniczność funkcji f_i, co czyni je szczególnie przydatnymi w klasie zadań opisywanych za pomocą takich funkcji. Jest zatem naturalne, że te właśnie warunki wybrał Martos dla przedstawienia w swej książce teorii Kuhna–Tuckera.

Praca cytowana

[1] Béla Martos, *Programowanie nieliniowe, teoria i metody*, PWN, Warszawa 1983.