Elementy optymalizacji nieliniowej

dla studentów MSEM, semestr letni 2011

Paweł Goldstein
Instytut Matematyki UW

Wykłady 1 i 2

1. Zasady zaliczania.

- 4 do 5 karkówek (20 min), zapowiedziane tydzień wcześniej - 1 pytanie z wykładu, jedno zadanie
- aby zaliczyć ćwiczenia, trzeba mieć co najmniej 50% punktów z kartkówek. Ocenę z ćwiczeń wystawia prowadzący je, przede wszystkim na podstawie kartkówek, ale biorąc pod uwagę aktywność, wynik kolokwium itp.
- na koniec egzamin. Do egzaminu w pierwszym terminie dopuszczone tylko te osoby, które zaliczyły ćwiczenia, w drugim terminie - wszyscy.

2. Wstęp i terminologia. Teoria optymalizacji to jeden z najszerzej stosowanych działów matematyki. Ma istotne znaczenie zarówno w zagadnieniach technicznych (automatyka), logistyce, jak i – a może przede wszystkim – we współczesnej ekonomii. Ten wykład, w odróżnieniu od konwersatorium z metod wspomagania podejmowania decyzji, będzie miał nieco bardziej teoretyczny charakter. Będzie tak z dwóch powodów: po pierwsze, materiał teoretyczny, jaki mają w ramach tego przedmiotu Państwo opanować, jest obszerny i nietrywialny; w zupełności wystarczy go, by zapełnić nam cały semestr. Po drugie, przedmiot ten jest pomyślany po części jako kontynuacja wykładu z analizy matematycznej – i wprowadzenie Państwa w takie jej metody, które mają związek z teorią optymalizacji i są najużyteczniejsze dla ekonomisty.

optimus , *optima* , *optimum* – *łac.* najlepszy (-a, -e)

Czym zatem zajmuje się optymalizacja? Jak sama nazwa wskazuje – optymalizowaniem, a więc szukaniem *optymalnej* wartości pewnych parametrów. Może to być na przykład taki podział portfela inwestycyjnego na akcje (A), obligacje (O) i lokaty (L) (mierzony ich procentowym udziałem w portfelu), który zmaksymalizuje spodziewany zysk, przy określonym przez inwestora maksymalnym poziomie ryzyka, na jaki jest gotów się zgodzić.

Po przykładzie tym widać, że sprawa nie jest banalna: musimy umieć mierzyć

- spodziewany zysk – funkcję $Z(A, O, L)$, której maksimum szukamy,
- związane z podziałem portfela ryzyko inwestycyjne $R(A, O, L)$ – przy jego pomocy ustalamy dopuszczalne zakresy zmiennych – parametrów A, O i L.

Na naszym przedmiocie nie będziemy się, oczywiście, zajmować tym, jak wyznaczyć funkcje Z i R; interesować nas będzie „wypreparowane”, ogólne zagadnienie matematyczne: znajdź maksimum funkcji $Z(A, O, L)$ na zbiorze parametrów opisanym warunkiem $0 \leq R(A, O, L) \leq R_o$, gdzie R_o to maksymalne ryzyko, na jakie się godzi inwestor. Zadanie narzuca nam jeszcze jeden naturalny warunek na akceptowalne wartości parametrów: $A+O+L = 100\%$ (chyba, że dopuszczamy sytuację, w której nie rozdysponujemy całego portfela – wtedy powinniśmy przyjąć warunek $0\% \leq A + O + L \leq 100\%$).

Różnymi zagadnieniami tego typu będziemy zajmować się w zasadzie przez cały semestr.

Na początek wprowadźmy trochę terminologii. Niech U będzie pewnym podzbiorem $\mathbb{R}^n$, zaś $f : U \to \mathbb{R}$ będzie funkcją określoną w punktach zbioru U.

Definicja 1. Mówimy, że funkcja f przyjmuje w $x_o \in U$ **minimum** (globalne), jeżeli dla wszystkich $y \in U$ zachodzi $f(x_o) \leq f(y)$.

Analogicznie, f przyjmuje w $x_1 \in U$ **maksimum** (globalne), jeżeli dla wszystkich $y \in U$ zachodzi $f(x_1) \geq f(y)$.

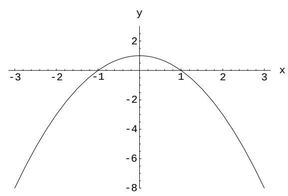
Definicja 2. Mówimy, że funkcja f przyjmuje w $x_o \in U$ **minimum lokalne**, jeżeli istnieje $\varepsilon > 0$ takie, że dla wszystkich $y \in U$ takich, że $\|x_o - y\| < \varepsilon$, zachodzi $f(x_o) \leq f(y)$.

Analogicznie, f przyjmuje w $x_1 \in U$ **maksimum lokalne**, jeżeli dla wszystkich $y \in U$, takich, że $\|x_1 - y\| < \varepsilon$, zachodzi $f(x_1) \geq f(y)$.

Przykład 1. Funkcja $f(x) = x^4 - 2x^2 + 1$ przyjmuje na $A = \mathbb{R}$ maksimum lokalne w $x = 0$, a minima lokalne w $x = -1$ i $x = 1$. Owe minima lokalne są równocześnie minimum globalnym, podczas gdy maksimum lokalne w $x = 0$ nie jest maksimum globalnym f. Skąd to wiemy? Powinni to Państwo pamiętać z wykładu analizy, poświęcimy jednak dalej trochę miejsca na przypomnienie tych metod.

Minima i maksima funkcji obejmujemy wspólną nazwą **ekstremów** lub **optimów**. Podobnie możemy mówić o **ekstremach (optimach) lokalnych**. Funkcję f, której ekstremów szukamy, nazywać będziemy czasem **funkcją celu**.

Warto zapamiętać, że terminy **minimum** i **maksimum** funkcji oznaczają wartości funkcji, nie punkty, w których wartości te są osiągane. Często mówi się, że np. funkcja $f(x) = 1 - x^2$, określona na całym $\mathbb{R}$, ma w $x_o = 0$ maksimum — i jest to prawda, ale maksimum to jest równe 1 (a nie 0). Błędem, a w każdym razie nieścisłością jest stwierdzenie, że $x_o = 0$ *jest maksimum* $f(x) = 1 - x^2$.



Wykres funkcji $f(x) = 1 - x^2$

Przypomnijmy (znane z wykładu analizy) jedno z najważniejszych twierdzeń o ekstremach funkcji ciągłych:

Twierdzenie 1 (Weierstrassa). *Funkcja ciągła przyjmuje na zbiorze zwartym swoje kresy.*

Karl Theodor Wilhelm Weierstrass (1815–1897) był twórcą współczesnej analizy matematycznej, jej terminologii, jemu też zawdzięczamy ścisłe sformułowanie jej podstaw. Jego autorstwa są definicje typu „dla każdego ε istnieje δ..."

Innymi słowy, jeżeli funkcja $f(x)$ jest ciągła, zbiór A zwarty (jeżeli $A \subset \mathbb{R}^n$, to zwartość oznacza po prostu, że A jest domknięty i ograniczony) i $\sup_{x \in A} f(x) = M$, to istnieje $x_M \in A$ taki, że $f(x_M) = M$ (a więc w rzeczywistości $M = \max_{x \in A} f(x)$. Analogicznie dla infimum i minimum: jeżeli $\inf_{x \in A} f(x) = m$, to istnieje $x_m \in A$ taki, że $f(x_m) = m$, a zatem $m = \min_{x \in A} f(x)$.

Twierdzenie Weierstrassa nie podpowiada nam wprawdzie, jak tych ekstremów szukać, ale gwarantuje nam, że będą one istniały.

Łatwo zauważyć, że zarówno założenie o ciągłości f, jak i o zwartości A sa istotne:

Przykład 2. Funkcja

$$f(x) = \begin{cases} x^2 & 0 \leq x < 1 \\ 0 & 1 \leq x \leq 2 \end{cases}$$

określona na odcinku $A = [0, 2]$ nie przyjmuje w tym odcinku maksimum.

Przykład 3. Funkcja $f(x) = x^3$ nie przyjmuje na $A = \mathbb{R}$ ani minimum, ani maksimum.

Przykład 4. Funkcja $f(x) = x^2$ nie przyjmuje na odcinku $A = [0, 1]$ ani minimum, ani maksimum.

3. Powtórka z analizy: pochodne kierunkowe. Ważnym narzędziem, które będzie się co jakiś czas pojawiało w treści wykładu, są pochodne kierunkowe funkcji wielu zmiennych. Były one wprawdzie wprowadzone na wykładzie z analizy, na wszelki wypadek jednak przypomnijmy:

Definicja 3. Niech U będzie otwartym podzbiorem $\mathbb{R}^n$, $f : U \to \mathbb{R}$. **Pochodną kierunkową** funkcji f w punkcie $x_o \in U$ w kierunku wektora $\vec{v}$ nazywamy liczbę

$$f'_{\vec{v}}(x_o) = \lim_{t\to 0} \frac{f(x_o + t\vec{v}) - f(x_o)}{t}$$

(pod warunkiem, oczywiście, że granica ta istnieje).

Zauważmy, że pochodna kierunkowa jest po prostu pochodną (w zerze) funkcji jednej zmiennej, danej wzorem $g(t) = f(x_o + t\vec{v})$. Korzystając z (odpowiedniej, wielowymiarowej wersji) twierdzenia o różniczkowaniu funkcji złożonej otrzymujemy

Przez $\langle \vec{v}, \vec{w} \rangle$ oznaczam iloczyn skalarny wektorów $\vec{v}$ i $\vec{w}$.

Twierdzenie 2.

$$f'_{\vec{v}}(x_o) = \langle \nabla f(x_o), \vec{v} \rangle.$$

Proof.

$$f'_{\vec{v}}(x_o) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} g(t) = \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0} f(x_o + t\vec{v}) = \langle \nabla f(x_o), \frac{\mathrm{d}}{\mathrm{d}t}\Big|_{t=0}(x_o + t\vec{v}) \rangle = \langle \nabla f(x_o), \vec{v} \rangle.$$

$\square$

Pochodna kierunkowa $f'_{\vec{v}}(x_o)$ mierzy, jak szybko rośnie funkcja, gdy poruszamy się z x_o w kierunku wskazanym przez wektor $\vec{v}$ (ale w jednostkach zależnych od długości wektora $\vec{v}$! – zauważcie, że $f'_{2\vec{v}}(x_o) = \langle 2\vec{v}, \nabla f(x_o) \rangle = 2\langle \vec{v}, \nabla f(x_o) \rangle = 2f'_{\vec{v}}(x_o)$, pomimo, że wektory $\vec{v}$ i $2\vec{v}$ mają dokładnie ten sam kierunek i zwrot). Oczywiście, jeżeli $f'_{\vec{v}}(x_o) > 0$, to f w tym kierunku rośnie, a jeżeli $f'_{\vec{v}}(x_o) < 0$, to maleje.

Pierre de Fermat (1601-1665) - francuski matematyk – samouk, z wykształcenia prawnik i lingwista. Znany powszechnie ze względu na tzw. Wielkie Twierdzenie Fermata, udowodnione niedawno przez Wilesa, zasłużył się matematyce w wielu dziedzinach - analizie, optyce geometrycznej, teorii liczb i rachunku prawdopodobieństwa.

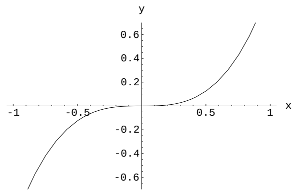
4. Zadania ekstremalne bez ograniczeń.

4.1. Twierdzenie Fermata.

Twierdzenie 3 (Fermata). *Niech U będzie* otwartym *podzbiorem* $\mathbb{R}^n$. *Jeżeli funkcja różniczkowalna $f : U \to \mathbb{R}$ przyjmuje w punkcie $x_0 \in U$ ekstremum lokalne, to $\nabla f(x_o) = 0$.*

Oczywiście jeżeli $n = 1$ (a więc f jest funkcją jednej zmiennej), ∇f to po prostu f'. Nie całkiem precyzyjne, ale upraszczające notację sformułowanie $\nabla f(x_o) = 0$ oznacza oczywiście, że gradient f jest równy wektorowi zerowemu.

Twierdzenie Fermata daje nam **warunek konieczny** (pierwszego rzędu) istnienia optimum. Łatwo można zauważyć, że nie jest to warunek dostateczny. Funkcja $f(x) = x^3$ spełnia warunek $f'(0) = 0$, gdyż $f'(x) = 3x^2$, pomimo, że nie ma w punkcie $x_o = 0$ ani minimum, ani maksimum. .Podobnie funkcja $g(x, y) = x^2 - y^2$ nie ma w $(0, 0)$ ekstremum, lecz punkt siodłowy, choć $\nabla g(x, y) = (2x, -2y)$, więc $\nabla g(0, 0) = (0, 0)$.

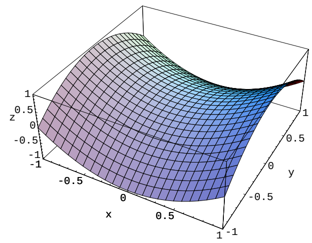


Wykres $f(x) = x^3$

Punkty, w których gradient funkcji celu jest równy zero, nazywamy **punktami stacjonarnymi**. Twierdzenie Fermata mówi, że funkcja różniczkowalna określona na zbiorze otwartym osiąga ekstrema lokalne jedynie w punktach stacjonarnych.

Oczywiście założenie, że funkcja f jest różniczkowalna w x_o, jest istotne. Ilustruje to

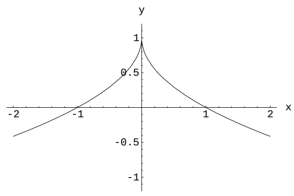
PRZYKŁAD 5. Funkcja $f(x) = 1 - \sqrt{|x|}$ przyjmuje w $x_o = 0$ maksimum lokalne, ale nie jest w tym punkcie różniczkowalna.


Wykres $g(x, y) = x^2 - y^2$

Szkic dowodu tw. Fermata:
Dla uproszczenia notacji załóżmy, że $n = 2$, zaś funkcja f ma w punkcie $(x_o, y_o) \in \mathbb{R}^2$ minimum (dowód dla większej liczby wymiarów jest dokładnie taki sam, podobnie jak dla przypadku, gdy w (x_o, y_o) jest maksimum). Dowód przeprowadzimy nie wprost: założmy, że $\nabla f(x_o, y_o) \neq (0, 0)$. Oznacza to, że co najmniej jedna z pochodnych cząstkowych $\frac{\partial f}{\partial x}(x_o, y_o)$ i $\frac{\partial f}{\partial y}(x_o, y_o)$ jest różna od zera. Załóżmy, że pierwsza z nich jest większa od zera i zauważmy, że $\frac{\partial f}{\partial x}(x_o, y_o) = \phi'(x_o)$, gdzie $\phi(x) = f(x, y_o)$. Oznacza to, że funkcja $\phi(x)$ jest w otoczeniu x_o rosnąca — a więc dla $x < x_o$, x dostatecznie bliskich x_o, mamy $f(x, y_o) = \phi(x) < \phi(x_o) = f(x_o, y_o)$, a więc funkcja f nie może mieć w (x_o, y_o) minimum.

Jeżeli natomiast $\frac{\partial f}{\partial x}(x_o, y_o) < 0$, to $\phi(x)$ jest malejąca, a więc dla $x > x_o$ mamy $\phi(x) < \phi(x_o)$, skąd otrzymujemy sprzeczność, analogicznie jak w poprzednim przypadku.

Jeżeli zaś $\frac{\partial f}{\partial y}(x_o, y_o) \neq 0$, to prowadzimy dowód w ten sam sposób, rozważając funkcje $\psi(y) = f(x_o, y)$. Czy potrafimy podać warunki dostateczne istnienia ekstremum? Owszem. Szczególnie łatwą maja one postać w przypadku $n = 1$:

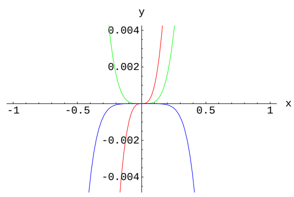

Wykres $f(x) = 1 - \sqrt{|x|}$

TWIERDZENIE 4. *Niech funkcja $f : (a, b) \to \mathbb{R}$ będzie dwukrotnie różniczkowalna w $x_o \in (a, b)$ oraz*

- $f'(x_o) = 0$ *(warunek stacjonarności),*
- $f''(x_o) \neq 0$.

Wówczas funkcja f przyjmuje w x_o ekstremum lokalne, przy czym jeżeli $f''(x_o) > 0$, to jest to minimum lokalne, a jeżeli $f''(x_o) < 0$ — maksimum lokalne.

Oczywiście łatwo można podać przykłady prostych funkcji, dla których powyższe twierdzenie nie potrafi rozstrzygnąć, czy w punkcie x_o jest ekstremum, a tym bardziej — jaki jest jego charakter. Zawodzi ono wówczas, gdy druga pochodna jest równa 0. Na przykład funkcja $f(x) = x^3$ nie ma w $x_o = 0$ ekstremum, $g(x) = x^4$ ma minimum, a $h(x) = -x^6$ — maksimum. Mamy $f'(0) = g'(0) = h'(0) = 0$, spełnione są więc warunki stacjonarności, ale $f''(0) = g''(0) = h''(0) = 0$.


Wykresy $f(x) = x^3$, $g(x) = x^4$ i $h(x) = -x^6$.

Znają Państwo zapewne wzmocnienie Twierdzenia 4, pomagające nam rozstrzygnąć powyższe przypadki:

TWIERDZENIE 5. *Niech funkcja $f : (a, b) \to \mathbb{R}$ będzie n-krotnie różniczkowana w $x_o \in (a, b)$ oraz*

- $f'(x_o) = f''(x_o) = \dots = f^{(n-1)}(x_o) = 0$ *(warunek stacjonarności),*
- $f^{(n)}(x_o) \neq 0$.

Wówczas

- *jeżeli n jest nieparzyste, to funkcja f nie przyjmuje w x_o ekstremum,*
- *jeżeli n jest parzyste, to funkcja f przyjmuje w x_o ekstremum lokalne, przy czym jeżeli $f^{(n)}(x_o) > 0$, to jest to minimum lokalne, a jeżeli $f^{(n)}(x_o) < 0$ — maksimum lokalne.*

Tak wygodne narzędzie mamy, niestety, jedynie dla funkcji jednej zmiennej...

Aby sformułować twierdzenie analogiczne do Twierdzenia 4, ale dla funkcji wielu zmiennych, musimy wprowadzić kilka nowych pojęć, rodem z algebry liniowej. W odróżnieniu od przypadku funkcji jednej zmiennej, pierwsza pochodna (gradient) nie jest już liczbą, lecz wektorem, a druga pochodna jest macierzą. Oczywiście nie sprawia nam trudności warunek stacjonarności — tak jak w Twierdzeniu Fermata, wystarczy zażądać, by $\nabla f(x_0) = 0$ [1]. Nie bardzo jednak wiadomo, co zrobić z warunkiem na drugą pochodną — jak rozumieć nierówność „macierz>0" czy „macierz<0".

Definicja 4. Mówimy, że symetryczna macierz kwadratowa M, o wymiarach $n\times n$, jest

- dodatnio określona, jeżeli dla dowolnego wektora $\vec{v} \in \mathbb{R}^n$ zachodzi $\vec{v}^T M\vec{v} > 0$,
- dodatnio półokreślona, jeżeli dla dowolnego wektora $\vec{v} \in \mathbb{R}^n$ zachodzi $\vec{v}^T M\vec{v} \geq 0$,
- ujemnie określona, jeżeli dla dowolnego wektora $\vec{v} \in \mathbb{R}^n$ zachodzi $\vec{v}^T M\vec{v} < 0$,
- ujemnie półokreślona, jeżeli dla dowolnego wektora $\vec{v} \in \mathbb{R}^n$ zachodzi $\vec{v}^T M\vec{v} \leq 0$.

Jeżeli nie zachodzi żadna z powyższych sytuacji, mówimy, że macierz jest nieokreślona.

Twierdzenie 6. *Macierz M jest dodatnio (ujemnie) określona wtedy i tylko wtedy, gdy wszystkie jej wartości własne są dodatnie (ujemne). Podobnie, M jest dodatnio (ujemnie) półokreślona wtedy i tylko wtedy, gdy wszystkie jej wartości własne są nieujemne (niedodatnie).*

Warto zauważyć, że jeżeli wyznacznik macierzy jest różny od zera, to albo jest ona (dodatnio lub ujemnie) określona, albo nie jest półokreślona (ani dodatnio, ani ujemnie - a zatem jest nieokreślona).

Szkic dowodu (dla macierzy dodatnio określonej, w pozostałych przypadkach dowód jest analogiczny): Jeżeli któraś z wartości własnych M jest niedodatnia ($\lambda_i \leq 0$), to oznaczając przez $\vec{w}_i$ odpowiadający jej unormowany wektor własny otrzymujemy $\vec{w}_i^T M\vec{w}_i = \lambda_i\|\vec{w}_i\|^2 = \lambda_i \leq 0$ — a więc sprzeczność z dodatnią określonością M. I odwrotnie: macierz symetryczną można zdiagonalizować, istnieje dla niej baza unormowanych wektorów własnych $\{\vec{w}_i\}_{i=1\dots n}$. Załóżmy, że wszystkie wartości własne λ_i macierzy M są dodatnie. Przedstawiając (dowolny) wektor $\vec{v}$ w tej bazie ($\vec{v} = v_1\vec{w}_1 + v_2\vec{w}_2 + \cdots v_n\vec{w}_n$) otrzymamy

$$\vec{v}^T M\vec{v} = \langle \vec{v}\,,\, M\vec{v}\rangle = \left\langle \sum_{i=1}^n v_i\vec{w}_i, M\left(\sum_{i=1}^n v_i\vec{w}_i\right)\right\rangle = \sum_{i=1}^n v_i^2\lambda_i\|w_i\|^2 > 0$$

Możemy teraz sformułować warunki dostateczne istnienia ekstremum dla funkcji wielu zmiennych. Jak poprzednio, zbiór U jest otwartym podzbiorem $\mathbb{R}^n$ (w szczególności może być całą przestrzenią $\mathbb{R}^n$).

Twierdzenie 7. *Niech funkcja $f : U \to \mathbb{R}$ będzie dwukrotnie różniczkowalna w $x_o \in U$ oraz*

- $\nabla f(x_o) = 0$ *(warunek stacjonarności),*

[1] Gdy piszemy, że wektor jest równy 0, mamy na myśli to, że wszystkie jego współrzędne są równe zeru. Jest to pewne nadużycie notacji, ale bardzo upraszcza i skraca zapis.

- *$D^2 f(x_o)$ jest macierzą dodatnio lub ujemnie określoną.*

Wówczas funkcja f przyjmuje w x_o ekstremum lokalne, przy czym jeżeli $D^2 f(x_o)$ jest dodatnio określona, to jest to minimum lokalne, a jeżeli $D^2 f(x_o)$ jest ujemnie określona — maksimum lokalne.

Nazwy „macierz Hessa" i „hessjan" użył po raz pierwszy J.J. Sylvester, dla upamiętnienia niemieckiego matematyka Ludwiga Ottona Hesse (1811–1874).

Przez $D^2 f(x_o)$ oznaczam oczywiście macierz złożoną z drugich pochodnych cząstkowych:

$$D^2 f(a) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(a) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(a) & \frac{\partial^2 f}{\partial x_2 \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(a) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(a) & \frac{\partial^2 f}{\partial x_n \partial x_2}(a) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(a) \end{pmatrix}$$

Macierz ta nazywana jest często macierzą Hessa, a jej wyznacznik – hessjanem funkcji f w punkcie a. Możemy również uzupełnić twierdzenie Fermata o dodatkowy warunek konieczny II rzędu:

$$\begin{pmatrix} a & b & e & f \\ c & d & g & h \\ i & j & m & n \\ k & l & o & p \end{pmatrix}$$

Wyznaczniki podmacierzy zakreślonych prostokątami to minory (2×2), przy czym wyznacznik macierzy zakreślonej średniej grubości kreską jest minorem głównym, a zakreślonej najgrubszą kreską — wiodącym minorem głównym.

TWIERDZENIE 8 (Warunek konieczny II rzędu istnienia ekstremum). *Jeżeli funkcja $f : U \to \mathbb{R}$ dwukrotnie różniczkowana w $x_o \in U$ ma w x_o minimum (maksimum) lokalne, to*

- *$\nabla f(x_o) = 0$,*
- *$D^2 f(x_o)$ jest macierzą dodatnio (ujemnie) półokreśloną.*

4.2. Kryterium Sylvestera.

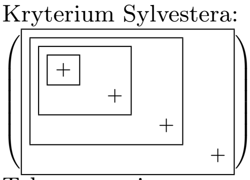
Czy aby sprawdzić jak określona jest druga pochodna musimy sprawdzać znaki jej wartości własnych? Na szczęście nie. Potrzebujemy jednak kilku pojęć.

James Joseph Sylvester (1814–1897) — wybitny angielski matematyk i prawnik, wniósł istotny wkład we współczesną algebrę, szczególnie w teorię niezmienników.

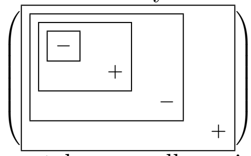
DEFINICJA 5. MINOREM macierzy M nazywamy wyznacznik dowolnej macierzy kwadratowej A otrzymanej z wycięcia z M pewnej ilości kolumn i rzędów. Jeżeli macierz M jest kwadratowa, a macierz A otrzymaliśmy z wycinania rzędów i kolumn o tych samych numerach (np. usunęliśmy z M 1,3,5 i 6-ty wiersz oraz 1,3,5 i 6-tą kolumnę), to wyznacznik A nazywamy MINOREM GŁÓWNYM, a jeżeli dodatkowo usuwane kolumny i rzędy to wszystkie o numerach powyżej pewnego $k \leq n$, to mówimy, że $\det A$ jest WIODĄCYM MINOREM GŁÓWNYM.

Okazuje się, że to, czy macierz jest dodatnio lub ujemnie określona, można sprawdzić badając znaki wiodących minorów głównych.

Kryterium Sylvestera:



Tak powinny wyglądać znaki wiodących minorów głównych dla macierzy dodatnio,



a tak — dla ujemnie określonej.

TWIERDZENIE 9 (Kryterium Sylvestera). *Macierz symetryczna jest dodatnio określona wtedy i tylko wtedy, gdy wszystkie jej wiodące minory główne są dodatnie. Podobnie, macierz jest ujemnie określona, gdy jej kolejne wiodące minory główne, licząc od wyrazu w lewym górnym rogu, mają znaki na zmianę, zaś pierwszy z nich jest ujemny.*

Niestety kryterium Sylvestera, jak przekonamy się na ćwiczeniach, zawodzi dla macierzy półokreślonych — można podać przykłady macierzy, które nie są dodatnio półokreślone, choć wszystkie ich wiodące minory główne są nieujemne. Zachodzi jednak uogólnienie kryterium Sylvestera:

TWIERDZENIE 10. *Jeżeli wszystkie minory główne macierzy są nieujemne, to jest ona dodatnio półokreślona.*

Możemy je stosować również do macierzy ujemnie półokreślonych, gdy zauważymy, że M jest ujemnie półokreślona $\Leftrightarrow -M$ jest dodatnio półokreślona.

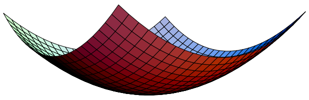
W praktyce kryterium Sylvestera stosuje się tylko wtedy, gdy rozmiary macierzy nie są zbyt duże. Obliczanie wyznaczników jest czynnością czasochłonną nie tylko dla ludzi, ale i dla komputerów i przy dużych rozmiarach macierzy zaczyna to odgrywać istotną rolę. Określoność macierzy bada się wówczas np. sprowadzając ją do postaci trójkątnej — otrzymujemy wówczas na przekątnej wartości własne macierzy, co załatwia sprawę.

4.3. Podsumowanie.

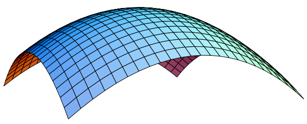
Aby znaleźć ekstrema lokalne funkcji należy

(1) wyznaczyć jej punkty krytyczne (czyli punkty, w których gradient jest równy 0)
(2) w każdym z uzyskanych punktów krytycznych zbadać, jak określona jest macierz drugiej pochodnej. W punktach, w których jest ona dodatnio (ujemnie) określona, na pewno jest minimum (maksimum), w punktach, w których jest dodatnio (ujemnie) półokreślona — może być minimum (maksimum), ale może być też punkt siodłowy, a w punktach, w których nie jest półokreślona — na pewno jest punkt siodłowy. Szczególnym przypadkiem jest sytuacja, gdy macierz Hessa jest macierzą zerową (a więc jest zarówno dodatnio, jak i ujemnie półokreślona) — nie możemy wówczas wykluczyć żadnej z trzech ewentualności.

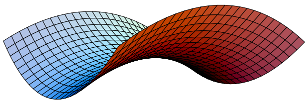
W punktach stacjonarnych, w których hessjan jest różny od zera, mamy trzy możliwe zachowania:



minimum



maksimum



i punkt siodłowy.

Wykłady 3 i 4

5. Zbiory wypukłe.

DEFINICJA 6. Mówimy, że zbiór A jest *wypukły*, jeżeli dla dowolnych punktów $x, y \in A$ cały odcinek łączący x z y leży w A.

Oczywiście definicja ta jest wygodna, gdy mamy zbiór A narysowany, lub łatwo możemy go sobie wyobrazić; by móc sprawdzać wypukłość rachunkowo lub z niej w rachunkach korzystać potrzebujemy charakteryzacji odcinka $[x, y]$. Łatwo zauważyć, że dowolny punkt odcinka $[x, y]$ można zapisać jako $z = \alpha x + (1-\alpha)y$, gdzie $\alpha \in [0, 1]$ (jeżeli pominiemy ograniczenie na wartości α, otrzymamy całą prostą przechodzącą przez punkty x i y). Oczywiście, gdy $\alpha = 0$ lub 1, otrzymujemy odpowiedni koniec odcinka, biorąc zaś $\alpha = 1/2$ — jego środek.

Zbiór A jest zatem wypukły, jeżeli dla dowolnych $x, y \in A$ i $\alpha \in [0, 1]$ punkt $\alpha x + (1-\alpha)y$ należy do A.

Często można (również w literaturze) znaleźć nieco prostszy (lecz błędny) warunek: zbiór A jest wypukły, jeżeli dla $x, y \in A$ *środek* odcinka $[x, y]$ należy do A. Warunek ten jest w rzeczywistości trochę słabszy, niż podany przeze mnie, na przykład zbiór liczb wymiernych $\mathbb{Q} \subset \mathbb{R}$ spełnia go (środek odcinka o końcach w liczbach wymiernych też jest liczbą wymierną), choć oczywiście nie jest wypukły.

TWIERDZENIE 11. *Część wspólna (dowolnie licznej, skończonej lub nie) rodziny zbiorów wypukłych jest zbiorem wypukłym.*

PROOF. Oznaczmy naszą rodzinę zbiorów przez $\{A_i \ : \ i \in I\}$, gdzie I jest pewnym zbiorem indeksów. Część wspólna tej rodziny to $B = \bigcap_{i \in I} A_i =$

$\{z\ :\ z \in A_i$ dla wszystkich $i \in I\}$. Naszym celem jest udowodnienie, że

dla dowolnych $x, y \in B$ oraz $\alpha \in [0, 1]$ mamy $\alpha x + (1 - \alpha)y \in B$.

Ustalmy x, y i α spełniające powyższe założenia ($x, y \in B$, $\alpha \in [0, 1]$). Zbiór B jest częścią wspólną zbiorów A_i, więc jeżeli x i $y \in B$, to dla wszystkich $i \in I$ mamy $x, y \in A_i$. Zbiory A_i są wypukłe, więc dla każdego $i \in I$ mamy $\alpha x + (1-\alpha)y \in A_i$, a to oczywiście oznacza, że $\alpha x + (1-\alpha)y \in B$. □

Załóżmy teraz, że zbiór A jest domknięty. Oczywiście, z definicji środki odcinków, których końce leżą w A, też należą do A. Nie każdy jednak punkt zbioru A jest środkiem takiego odcinka (niezdegenerowanego do pojedynczego punktu).

Przyjrzyjmy się odcinkowi $[a, b] \subset \mathbb{R}$. Każdy punkt z odcinka otwartego (a, b) można przedstawić jako środek odcinka $[z - \varepsilon, z + \varepsilon]$, gdzie $\varepsilon = \min\{|z - a|/2, |z - b|/2\}$. Nie da się jednak znaleźć takiego odcinka ani dla a, ani dla b.

Definicja 7. Niech A będzie domkniętym zbiorem wypukłym. Punkt $z \in A$, którego nie da się przedstawić jako środek odcinka, którego oba końce leżą w A, nazywamy *punktem ekstremalnym* zbioru A (czasem używa się też nazwy *punkt skrajny*). Zbiór wszystkich punktów ekstremalnych zbioru A oznaczamy ex$A$. Warunek na bycie punktem ekstremalnym możemy zapisać więc tak:

$$z \in \mathrm{ex}A \iff \left(z = \frac{x+y}{2},\ x, y \in A \Rightarrow x = y = z \right)$$

Powstaje natychmiast pytanie: czy każdy domknięty zbiór wypukły ma jakieś punkty ekstremalne? Bo, jak widać na przykładzie odcinka $(a, b) \subset \mathbb{R}$, otwarty zbiór wypukły może nie mieć ani jednego. Również cała przestrzeń — wypukła, a jakże, i do tego domknięta — nie ma żadnych punktów ekstremalnych. Zachodzi natomiast ważne

Twierdzenie 12 (Minkowskiego). *Każdy wypukły i zwarty zbiór $A \subset \mathbb{R}^n$ ma punkty ekstremalne.*

Hermann Minkowski (1864–1909) — niemiecki matematyk poch. żydowskiego, nauczyciel Alberta Einsteina i współpracownik Dawida Hilberta, wniósł duży wkład w teorię liczb oraz w matematyczny formalizm szczególnej teorii względności.

W rzeczywistości twierdzenie Minkowskiego mówi nieco więcej: nie tylko ex$A$ jest niepusty, ale, co więcej, A najmniejszym zwartym i wypukłym podzbiorem $\mathbb{R}^n$ zawierającym ex$A$.

Zanim przejdziemy do dowodu, zauważmy dwa użyteczne fakty: Niech $f : \mathbb{R}^n \to \mathbb{R}^m$ będzie funkcją liniową, zaś $A \subset \mathbb{R}^n$ i $B \subset \mathbb{R}^m$ — zbiorami wypukłymi.

- $f(A)$ jest wypukłym podzbiorem $\mathbb{R}^m$;
- $f^{-1}(B)$ jest wypukłym podzbiorem $\mathbb{R}^n$.

Dowód — na ćwiczeniach.

Szkic dowodu twierdzenia Minkowskiego. Dowód poprowadzimy przez indukcję ze względu na wymiar przestrzeni $\mathbb{R}^n$, której podzbiorem jest A.

Zauważmy najpierw, że jeżeli $n = 1$, to zbiór A jest odcinkiem domkniętym, a jego punktami ekstremalnymi są końce odcinka — czyli w tym przypadku twierdzenie zachodzi.

Załóżmy teraz, że dowolny zwarty i wypukły podzbiór przestrzeni $(n-1)$-wymiarowej ma punkty ekstremalne. Niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie rzutowaniem na pierwszą współrzędną, czyli

$$f(x_1, x_2, \ldots, x_n) = x_1.$$

Zbiór $f(A) \subset \mathbb{R}$ jest zbiorem zwartym (bo zwartość zachowywana jest przez przekształcenia ciągłe) i wypukłym (bo rzutowanie jest liniowe) — a więc jest odcinkiem domkniętym $[a, b]$. Przyjrzyjmy się zbiorowi $f^{-1}(b) \cap A$. Jest to zbiór wypukły i zwarty, zawarty w $(n-1)$-wymiarowej podprzestrzeni przestrzeni $\mathbb{R}^n$ — a więc na mocy założenia indukcyjnego ma punkt ekstremalny z. Wykażemy, że z jest punktem ekstremalnym zbioru A.

Gdyby tak nie było, z byłby środkiem odcinka $[x, y]$, którego końce x i y leżałyby w A. Punkty x i y nie mogą równocześnie oba leżeć w zbiorze $f^{-1}(b) \cap A$, gdyż z jest jego punktem ekstremalnym — a więc co najmniej jeden z punktów x, y leży w $f^{-1}([a, b))$. Załóżmy bez straty ogólności, że jest to x — wiemy zatem, że $f(x) < b$, no i oczywiście $f(y) \le b$. Teraz łatwo otrzymujemy sprzeczność

$$b = f(z) = f\left(\frac{x+y}{2}\right) = \frac{1}{2}f(x) + \frac{1}{2}f(y) \le \frac{1}{2}f(x) + \frac{b}{2} < \frac{b}{2} + \frac{b}{2} = b,$$

co dowodzi, że rzeczywiście z jest punktem ekstremalnym A. □

Ważną własność punktów ekstremalnych daje poniższe twierdzenie:

Twierdzenie 13. *Niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją liniową, zaś A — zbiorem wypukłym zwartym. Oznaczmy $a = \min_{x \in A} f(x)$, $b = \max x \in Af(x)$. Istnieją wówczas punkty $v, w \in \mathrm{ex}A$ takie, że $f(v) = a$, $f(w) = b$.*

Zauważmy, że funkcja f, na mocy tw. Weierstrassa, przyjmuje w A zarówno minimum, jak i maksimum. Twierdzenie powyższe mówi, że szukając największej, jak i najmniejszej wartości funkcji liniowej na zbiorze A wystarczy zbadać wartości f w punktach ekstremalnych A. Jest to szczególnie użyteczne, gdy — co się często zdarza w zastosowaniach — zbiór A jest zadany nierównościami na funkcjach liniowych, czyli jest wielościanem. W takim przypadku $\mathrm{ex}A$ składa się z wierzchołków tego wielościanu, jest zatem zbiorem skończonym i można po prostu policzyć wartości we wszystkich wierzchołkach, a następnie wybrać największą i najmniejszą z nich.

Proof. Dowód jest bardzo prostą modyfikacją dowodu twierdzenia Minkowskiego. Jak wówczas, zbiór $f(A)$ jest zwartym i wypukłym podzbiorem $\mathbb{R}$, a więc jest odcinkiem domkniętym — oczywiście odcinkiem $[a, b]$. Zbiór $f^{-1}(\{a\}) \cap A$ jest zwarty i wypukły (dlaczego?) — a więc ma punkt ekstremalny. Oznaczmy go przez v. Wykażemy, że $v \in \mathrm{ex}A$. Dowodzi się tego tak samo, jak w tw. Minkowskiego: gdyby $v \notin \mathrm{ex}A$, to istniałyby punkty x i $y \in A$ takie, że $v = (x+y)/2$, przy czym nie mogłyby one oba równocześnie należeć do $f^{-1}a \cap A$, gdyż przeczyłoby to temu, że v jest punktem ekstremalnym tego ostatniego zbioru. Niech więc na przykład $x \notin f^{-1}a \cap A$ — oznacza to, że $f(x) > a$. Oczywiście $f(y) \ge a$. No, ale to oznacza, że

$$a = f(v) = f\left(\frac{x+y}{2}\right) = \frac{1}{2}f(x) + \frac{1}{2}f(y) \ge \frac{1}{2}f(x) + \frac{a}{2} > \frac{a}{2} + \frac{a}{2} = a.$$

Otrzymaliśmy sprzeczność. Dowód, że istnieje $w \in \mathrm{ex}A$ takie, że $f(w) = b$ przebiega analogicznie. □

Zauważmy, że na to, by $f(A)$ było odcinkiem $[a, b]$ wystarcza, by funkcja f była ciągła — wynika to natychmiast z tw. Weierstrassa i własności Darboux. Funkcja f przyjmuje na A swoje kresy, więc $f(A) \subset [a, b]$, $a, b \in f(A)$, a z własności Darboux funkcja f przyjmie wszystkie wartości pomiędzy a i b.

6. Funkcje wypukłe. Niech A, jak poprzednio, będzie zbiorem wypukłym.

Definicja 8. Mówimy, że funkcja $f : A \subset \mathbb{R}^n \to \mathbb{R}$ jest

- *wypukła*, jeżeli dla dowolnych $x, y \in A$ i $\alpha \in [0, 1]$
$$f(\alpha x + (1-\alpha)y) \le \alpha f(x) + (1-\alpha)f(y)$$
- *ściśle wypukła*, jeżeli dla dowolnych $x, y \in A$, $x \ne y$ i $\alpha \in (0, 1)$
$$f(\alpha x + (1-\alpha)y) < \alpha f(x) + (1-\alpha)f(y)$$

- *wklęsła*, jeżeli dla dowolnych $x, y \in A$ i $\alpha \in [0, 1]$
$$f(\alpha x + (1-\alpha)y) \geq \alpha f(x) + (1-\alpha)f(y)$$
- *ściśle wklęsła*, jeżeli dla dowolnych $x, y \in A$, $x \neq y$ i $\alpha \in (0, 1)$
$$f(\alpha x + (1-\alpha)y) > \alpha f(x) + (1-\alpha)f(y)$$

Definicja 9. Zbiór $\{(x, y) \in A \times \mathbb{R} \ : \ y \geq f(x)\}$ nazywamy *nadwykresem* lub *epigrafem* funkcji f.

Oczywiście nadwykres to zbiór punktów leżących *na i nad wykresem* funkcji f.

Nadwykres funkcji wypukłej jest zbiorem wypukłym (dlaczego?). Dla funkcji wklęsłej wypukły jest „podwykres", czyli zbiór
$$\{(x, y) \in A \times \mathbb{R} \ : \ y \leq f(x)\}.$$

Johan Ludwig William Valdemar Jensen (1859–1925), duński matematyk i inżynier

Twierdzenie 14 (Nierówność Jensena). *Jeżeli funkcja* $f : A \subset \mathbb{R}^n \to \mathbb{R}$ *jest wypukła, to dla dowolnych liczb* $\alpha_1, \alpha_2, \ldots, \alpha_n \in [0, 1]$ *spełniających warunek* $\alpha_1 + \alpha_2 + \cdots + \alpha_n = 1$ *oraz dowolnych* $x_1, x_2, \ldots, x_n \in A$ *zachodzi nierówność*
$$f(\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n) \leq \alpha_1 f(x_1) + \alpha_2 f(x_2) + \cdots + \alpha_n f(x_n).$$
Jeżeli funkcja f *jest ściśle wypukła, to równość w miejscu* $\leq$ *może zajść tylko jeżeli spełniony jest jeden z dwóch poniższych warunków:*
- *wszystkie* α_i *są równe 0, poza jedną (która musi wówczas być równa 1),*
- $x_1 = x_2 = \cdots = x_n$.

Uwaga: jeżeli f jest wklęsła, to powyższe twierdzenie zachodzi bez istotnych zmian, tylko nierówność jest w przeciwną stronę. Dotyczy to również przypadku funkcji ściśle wklęsłych — równość w nierówności Jensena zachodzi dla nich tylko, gdy spełniony jest jeden z dwóch podanych w powyższym twierdzeniu warunków.

Indukcyjny dowód nierówności Jensena widzieli Państwo w zeszłym semestrze, więc go pominę.

Funkcja wypukła określona na odcinku otwartym jest ciągła, nie musi jednak być różniczkowalna (przykłady — na ćwiczeniach). Dla funkcji różniczkowalnych mamy łatwy test na to, czy są wypukłe:

Twierdzenie 15. *Funkcja różniczkowalna na odcinku* (a, b) *jest na nim*
- wypukła*, gdy* f' *jest na* (a, b) *niemalejąca,*
- ściśle wypukła*, gdy* f' *jest na* (a, b) *rosnąca,*
- wklęsła*, gdy* f' *jest na* (a, b) *nierosnąca,*
- ściśle wklęsła*, gdy* f' *jest na* (a, b) *malejąca.*

Są to warunki pierwszego rzędu *na wypukłość/wklęsłość. Jeżeli funkcja* f *jest dwukrotnie różniczkowalna na* (a, b)*, możemy podać następujące* warunki drugiego rzędu*:*
- wypukła*, gdy* f'' *jest na* (a, b) *nieujemna,*
- ściśle wypukła*, gdy* f'' *jest na* (a, b) *nieujemna i pomiędzy dowolnymi różnymi punktami* $x, y \in (a, b)$ *znajdziemy* ξ *takie, że* $f''(\xi) > 0$,
- wklęsła*, gdy* f'' *jest na* (a, b) *niedodatnia,*
- ściśle wklęsła*, gdy* f'' *jest na* (a, b) *niedodatnia i pomiędzy dowolnymi różnymi punktami* $x, y \in (a, b)$ *znajdziemy* ξ *takie, że* $f''(\xi) < 0$.

Jak sformułować te warunki dla funkcji wielu zmiennych?

Twierdzenie 16 (Warunki pierwszego rzędu (na wypukłość/wklęsłość funkcji wielu zmiennych)).
Funkcja $f : A \subset \mathbb{R}^n \to \mathbb{R}$, *określona na zbiorze wypukłym* A, *jest*

- wypukła *wtedy i tylko wtedy, gdy dla dowolnych* $x, y \in A$
$$f(y) \geq f(x) + \nabla f(x)^T (y - x)$$
- ściśle wypukła *wtedy i tylko wtedy, gdy dla dowolnych* $x, y \in A$, $x \neq y$,
$$f(y) > f(x) + \nabla f(x)^T (y - x)$$

Analogiczne nierówności (tylko w przeciwną stronę) zachodzą dla funkcji wklęsłych i ściśle wklęsłych.

Zauważmy, że wyrażenie $f(x) + \nabla f(x)^T (y - x)$ to pierwsze dwa wyrazy rozwinięcia Taylora wokół punktu x, czyli najlepsze przybliżenie funkcji f funkcją liniową w otoczeniu punktu x. Wykres takiej funkcji liniowej to po prostu hiperpłaszczyzna styczna to wykresu f w punkcie x. Powyższe twierdzenie uogólnia więc znany Państwu fakt, że wykres funkcji wypukłej leży *nad* styczną doń, poprowadzoną w dowolnym punkcie dziedziny, przy czym dla funkcji wypukłej wykres f może mieć z ową styczną całe wspólne odcinki (skrajnym przykładem jest funkcja liniowa, której wykres pokrywa się ze styczną doń w każdym punkcie), podczas gdy dla funkcji ściśle wypukłej jedynym punktem wspólnym stycznej i wykresu jest punkt, w którym styczna została poprowadzona.

Proof. Udowodnimy tylko pierwszy podpunkt — dla funkcji wypukłej. Mamy do udowodnienia dwie rzeczy:

a) jeżeli f jest wypukła, to $f(y) \geq f(x) + \nabla f(x)^T (y - x)$,
b) jeżeli $f(y) \geq f(x) + \nabla f(x)^T (y - x)$, to f jest wypukła.

Najpierw część **a)**. Skoro funkcja f jest wypukła, to
$$f((1-t)x + ty) = f(x + t(y - x)) \leq (1-t)f(x) + tf(y).$$
Przerzucając $(1-t)f(x)$ na lewą stronę nierówności, a następnie dzieląc przez t, otrzymujemy
$$\frac{f(x + t(y - x)) - (1-t)f(x)}{t} \leq f(y)$$
Po zamianie stron nierówności i drobnych przekształceniach
$$f(y) \geq f(x) + \frac{f(x + t(y - x)) - f(x)}{t}$$
Chcemy teraz obliczyć granice obu stron przy $t \to 0$. Ułamek po prawej stronie nierówności dąży z definicji do pochodnej kierunkowej $f'_{(y-x)}(x)$, ta zaś jest równa $\nabla f(x)^T (y - x)$. To kończy dowód a).

Teraz część **b)** dowodu. Wiemy, że nierówność $f(y) \geq f(x) + \nabla f(x)^T (y - x)$ zachodzi dla dowolnych punktów $x, y \in A$. Oznaczmy $z = \alpha x + (1 - \alpha) y$. Oczywiście $z \in A$, zachodzą więc nierówności
$$\begin{aligned} f(x) &\geq f(z) + \nabla f(z)^T (x - z), \\ f(y) &\geq f(z) + \nabla f(z)^T (y - z). \end{aligned}$$
Pomnóżmy pierwszą nierówność stronami przez α, drugą przez $1 - \alpha$ i dodajmy stronami (nierówności wolno dodawać stronami, ale **nie wolno** ich odejmować). Z lewej strony otrzymamy $\alpha f(x) + (1 - \alpha) f(y)$. Z prawej
$$\begin{aligned} \alpha(f(z) + &\nabla f(z)^T (x - z)) + (1 - \alpha)(f(z) + \nabla f(z)^T (y - z)) \\ &= (\alpha + (1 - \alpha)) f(z) + \nabla f(z)^T (\alpha (x - z) + (1 - \alpha)(y - z)) \\ &= f(z) + \nabla f(z)^T (\alpha x + (1 - \alpha) y - z) = f(z) \\ &= f(\alpha x + (1 - \alpha) y). \end{aligned} \tag{1}$$

Dostajemy zatem, że dla dowolnych $x, y \in A$

$$\alpha f(x) + (1-\alpha)f(y) \geq f(\alpha x + (1-\alpha)y),$$

co dowodzi, że funkcja f jest wypukła. □

Najważniejszym wnioskiem z Twierdzenia 16 jest

TWIERDZENIE 17. *Niech f będzie różniczkowalną funkcją wypukłą określoną na zbiorze wypukłym A. Jeżeli dla pewnego $x \in A$ zachodzi $\nabla f(x) = 0$, to*

$$f(x) = \min_{y \in A} f(y).$$

Innymi słowy - jeżeli różniczkowalna funkcja wypukła ma punkt krytyczny, to automatycznie osiąga w nim minimum. Dowód jest bardzo prosty: z Twierdzenia 16 wiemy, że dla dowolnego $y \in A$ zachodzi $f(y) \geq f(x) + \nabla f(x)^T(y-x)$, ale $\nabla f(x)^T = 0$, więc $f(y) \geq f(x)$.

Analogicznie możemy udowodnić, że jeżeli różniczkowalna funkcja wklęsła określona na zbiorze wypukłym ma punkt krytyczny, to automatycznie osiąga w nim maksimum.

Ważna uwaga: przy powyższych warunkach (pierwszego, czy też drugiego rzędu) nie możemy zrezygnować z założenia, że zbiór A, na którym funkcja jest określona, jest zbiorem wypukłym. Zauważmy, że funkcja $f(x) = \frac{1}{x^2}$ spełnia w całej swojej dziedzinie warunek $f''(x) > 0$, choć oczywiście nie jest wypukła.

TWIERDZENIE 18 (Warunki IIgo rzędu na wypukłość (wklęsłość)). *Niech $A \subset \mathbb{R}^n$ będzie niepustym i otwartym zbiorem wypukłym, a $f : A \to \mathbb{R}$ będzie funkcją dwukrotnie różniczkowalną. Wówczas następujące warunki są równoważne:*

a) *funkcja f jest wypukła (odp. wklęsła),*
b) *macierz Hessa $D^2 f$ jest* ***dodatnio (odp. ujemnie) półokreślona*** *we wszystkich punktach A.*

PROOF. Najpierw udowodnimy, że jeżeli funkcja jest wypukła, to $D^2 f(x)$ jest dodatnio półokreślona dla każdego $x \in A$.

Skoro zbiór A jest otwarty, do dla każdego $x \in A$ i $v \in \mathbb{R}^n$, $v \neq 0$, istnieje $\varepsilon > 0$ takie, że $y = x + \varepsilon v \in A$. Wiemy (warunek pierwszego rzędu na wypukłość), że

$$f(y) \geq f(x) + \langle \nabla f(x), (y-x) \rangle.$$

Możemy również rozwinąć funkcję f w szereg Taylora stopnia 2 wokół x:

$$f(y) = f(x) + \nabla f(x)(y-x) + (y-x)^T D^2 f(x)(y-x) + o(\|y-x\|^2).$$

Przypomnijmy, że $o(h)$ z definicji oznacza wyrażenie takie, że

$$\lim_{h \to 0} \frac{o(h)}{h} = 0$$

Gdy w miejsce y wstawimy $x + \varepsilon v$, otrzymamy

$$f(x+\varepsilon v) \geq f(x) + \varepsilon \langle \nabla f(x), v \rangle$$
$$f(x+\varepsilon v) = f(x) + \varepsilon \langle \nabla f(x), v \rangle + \frac{1}{2}\varepsilon^2 v^T D^2 f(x) v + o(\varepsilon^2 \|v\|^2).$$

Mamy zatem

$$f(x) + \varepsilon \langle \nabla f(x), v \rangle + \frac{1}{2}\varepsilon^2 v^T D^2 f(x) v + o(\varepsilon^2 \|v\|^2) \geq f(x) + \varepsilon \langle \nabla f(x), v \rangle.$$

Po odjęciu od obu stron $f(x) + \varepsilon\langle\nabla f(x)\,,\,v\rangle$ i podzieleniu przez $\frac{1}{2}\varepsilon^2$ dostajemy

$$v^T D^2 f(x)v + 2\|v\|^2 \frac{o(\varepsilon^2\|v\|^2)}{\varepsilon^2\|v\|^2} \geq 0,$$

a po przejściu z ε do 0 (wówczas i $\varepsilon^2\|v\|^2 \to 0$)

$$v^T D^2 f(x)v \geq 0.$$

Nierówność ta ma zachodzić dla dowolnego $v \in \mathbb{R}^n$, $v \neq 0$, a więc — z definicji — macierz $D^2 f$ jest dodatnio półokreślona.

Dowód w drugą stronę jest bardzo prosty. Załóżmy, że macierz $D^2 f(x)$ jest dodatnio półokreślona dla dowolnego $x \in A$ (a więc dla dowolnego $v \in \mathbb{R}^n$ zachodzi $v^T D^2 f(x)v \geq 0$). Wypiszmy raz jeszcze wzór Taylora dla f, ale tym razem stopnia 1, wypisując postać reszty:

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

gdzie $\tilde{x}$ jest pewnym punktem pośrednim pomiędzy x a y ($\tilde{x} = \theta x + (1-\theta)y$ dla pewnego $\theta \in (0,1)$). Oczywiście z wypukłości A wynika, że $\tilde{x} \in A$. Skoro $D^2 f$ jest dodatnio półokreślona we wszystkich punktach A, to $\frac{1}{2}(y-x)^T D^2 f(\tilde{x})(y-x) \geq 0$, a zatem

$$f(y) \geq f(x) + \nabla f(x)(y-x),$$

czyli funkcja f spełnia warunek pierwszego rzędu na wypukłość, jest więc wypukła. □

Przykład zastosowania nierówności Jensena

Funkcja $f(x) = \ln x$ jest ściśle wklęsła dla $x > 0$, gdyż $f'(x) = \frac{1}{x}$ jest na tej półprostej funkcją malejącą. Wypiszmy nierówność Jensena przyjmując $\alpha_1 = \alpha_2 = \cdots = \alpha_n = \frac{1}{n}$:

$$\ln\left(\frac{1}{n}x_1 + \frac{1}{n}x_2 + \cdots + \frac{1}{n}x_n\right) \geq \frac{1}{n}\ln x_1 + \frac{1}{n}\ln x_2 + \cdots + \frac{1}{n}\ln x_n.$$

dla dowolnych $x_1, \ldots, x_n > 0$, przy czym równość zachodzi jedynie wtedy, gdy $x_1 = x_2 = \cdots = x_n$. Lewa strona jest oczywiście równa $\ln\frac{x_1+x_2+\cdots+x_n}{n}$, prawą zaś możemy przekształcić, korzystając z własności logarytmu:

$$\frac{1}{n}\ln x_1 + \frac{1}{n}\ln x_2 + \cdots + \frac{1}{n}\ln x_n = \frac{1}{n}\ln(x_1 \cdot x_2 \cdots x_n) = \ln\sqrt[n]{x_1 \cdot x_2 \cdots x_n}.$$

Logarytm naturalny jest funkcją rosnącą ($\ln a \geq \ln b \Rightarrow a \geq b$, podobnie, gdy nierówność jest ostra), otrzymujemy zatem

$$\frac{x_1 + x_2 + \cdots + x_n}{n} \geq \sqrt[n]{x_1 \cdot x_2 \cdots x_n}.$$

Lewa strona nierówności to średnia arytmetyczna, a prawa — geometryczna liczb $x_1, \ldots, x_n$. Równość między nimi zachodzi wtedy i tylko wtedy, gdy wszystkie te liczby są równe.

Wykład 5

7. Zadania ekstremalne z ograniczeniami równościowymi.

Wracamy do podstawowego zadania teorii optymalizacji – do poszukiwania wartości ekstremalnych funkcji. Dotąd rozwiązywaliśmy zadania minimalizacji lub maksymalizacji funkcji na całej przestrzeni $\mathbb{R}^n$, bez żadnych dodatkowych ograniczeń. Problemy tego typu pojawiają się bardzo rzadko, ich źródłem są najczęściej pewne uproszczenia problemów z ograniczeniami (ignorujemy ograniczenia lub zastępujemy je tzw. funkcjami kary). Znacznie ważniejsze są zagadnienia typu

$$\boxed{\text{znajdź minimalną wartość } f(x) \text{ dla } x\text{-ów spełniających pewne warunki.}}$$

Oczywiście warunki na x mogą być najróżniejszego typu — na przykład w zagadnieniach ekonomicznych żądamy bardzo często, by różne wielkości (poziom produkcji, ceny itp) były nieujemne. W tym rozdziale zajmiemy się tzw. ograniczeniami równościowymi, a więc zadaniami typu

$$\boxed{\text{znajdź minimalną wartość } f(x) \text{ dla } x \text{ takich, że } g_1(x) = 0, \ldots, g_k(x) = 0.}$$

Warto pamiętać znaną z wykładu analizy matematycznej zasadę, że gradient funkcji w punkcie wskazuje kierunek, w którym funkcja najszybciej rośnie (lub też, gdy patrzymy na $-\nabla f$, w którym kierunku najszybciej maleje). Łatwo więc zrozumieć, że w maksimum (czy też minimum), gdzie nie ma czego wskazywać, gradient powinien być równy 0 (to oczywiście intuicja, nie precyzyjny dowód twierdzenia Fermata). Gdy jednak nasz ruch jest ograniczony do pewnej powierzchni, na przykład gdy szukamy maksimum funkcji $f(x, y, z) = x^2 + y^2 + z$ na sferze $S = \{(x, y, z) \in \mathbb{R}^3 : (x-1)^2 + y^2 + z^2 = 1\}$, metoda ta zawodzi — może się okazać, że, aby pójść w kierunku wskazywanym przez $\nabla f(x, y, z)$, musielibyśmy opuścić powierzchnię sfery. Okazuje się jednak (fakt ten podaję bez dowodu), że, o ile powierzchnia ta jest dostatecznie „przyzwoita", wciąż możemy wskazać kierunek najszybszego wzrostu f z punktu $p = (x_o, y_o, z_o) \in S$, spośród kierunków, w których można się udać nie opuszczając powierzchni sfery: wskazuje go rzut wektora $\nabla f(p)$ na przestrzeń styczną do S w punkcie p.

Łatwo teraz możemy wskazać, poprzez analogię z tw. Fermata, punkty S podejrzane o to, że f przyjmuje w nich ekstremum: będą to te punkty, w których rzut ∇f na przestrzeń styczną do S będzie równy 0, a więc punkty, w których ∇f jest prostopadły do S. Jak znaleźć takie punkty?

Powierzchnię S mamy zadaną przez warunek $g(x, y, z) = (x-1)^2 + y^2 + z^2 - 1 = 0$. Przywołajmy znów fakt z wykładu analizy: jeżeli $g(p) = 0$ i funkcja g jest w (p) różniczkowalna, to wektor $\nabla g(p)$ jest prostopadły do zbioru $\{(x, y, z) \in \mathbb{R}^3 : g(x, y, z) = 0\}$ (a więc do S) w punkcie p. Przestrzeń wektorów prostopadłych do sfery S w punkcie p jest jednowymiarowa, a więc jeżeli zarówno $\nabla g(p)$ i $\nabla f(p)$ do niej należą, to muszą być proporcjonalne. Innymi słowy, musi istnieć $\lambda \in \mathbb{R}$ takie, że $\nabla f(p) = \lambda \nabla g(p)$.

Gdy zbiór, na którym szukamy ekstremów, jest zadany większą liczbą warunków ($S = \{x \in \mathbb{R}^n : g_1(x) = 0, g_2(x) = 0, \ldots, g_k(x) = 0\}$), sprawa się komplikuje. Najczęściej przestrzeń wektorów prostopadłych do S w p (teraz już niekoniecznie jednowymiarowa) jest rozpięta przez wektory $\nabla g_1(p), \nabla g_2(p), \ldots, \nabla g_k(p)$ — a więc, by $\nabla f(p)$ do niej należał, potrzeba i wystarcza, by dla pewnych $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ zachodziło

$$\nabla f(p) = \lambda_1 \nabla g_1(p) + \lambda_2 \nabla g_2(p) + \cdots + \lambda_k \nabla g_k(p).$$

Może się jednak złośliwie zdarzyć tak, że wektory $\nabla g_1(p), \nabla g_2(p), \ldots, \nabla g_k(p)$ nie rozpinają całej przestrzeni wektorów prostopadłych do S w p, a nawet, że któraś z funkcji g_i nie jest w tym punkcie różniczkowalna. Rozmaite tego typu przykładu obejrzą Państwo na ćwiczeniach. Biorąc pod uwagę powyższą agitację (do dowodu jej trochę brakuje, ale dowód był już przecież w zeszłym semestrze) wraz z wszystkimi zastrzeżeniami możemy sformułować twierdzenie:

Twierdzenie 19. *Niech A będzie podzbiorem $\mathbb{R}^n$ zadanym warunkami $g_1(x) = 0, g_2(x) = 0, \ldots, g_k(x) = 0$. Załóżmy teraz, że funkcja f przyjmuje* *w $x_o \in A$ ekstremum. Jeżeli f oraz wszystkie funkcje g_i są różniczkowalne w x_o oraz wektory $\nabla g_1(x_o), \nabla g_2(x_o), \dots, \nabla g_k(x_o)$ są liniowo niezależne, to istnieją takie liczby rzeczywiste $\lambda_1, \lambda_2, \dots, \lambda_k$, że funkcja*

$$L(x) = f(x) - \lambda_1 g_1(x) - \lambda_2 g_2(x) - \cdots - \lambda_k g_k(x)$$

ma w punkcie x_o punkt krytyczny.

Joseph Louis Lagrange (1736–1813) — wybitny francuski matematyk i astronom włoskiego pochodzenia, twórca rachunku wariacyjnego i jeden z pionierów mechaniki teoretycznej

Funkcję L nazywamy funkcją Lagrange'a, a liczby λ_i — mnożnikami Lagrange'a.

Kilka uwag do twierdzenia:

(1) Warunek liniowej niezależności gradientów warunków (czyli wektorów $\nabla g_1(x_o), \nabla g_2(x_o), \dots, \nabla g_k(x_o)$) nazywać będziemy odtąd *warunkiem jakości więzów* w punkcie x_o. Ma on szczególnie prostą postać, gdy mamy do czynienia tylko z jedną funkcją g (a więc $k = 1$ i $A = \{x \in \mathbb{R}^n \ : \ g(x) = 0\}$) — oznacza on wtedy, że $\nabla g(x_o) \neq 0$. Warto też zauważyć, że warunek liniowej niezależności nie może być spełniony, gdy liczba warunków jest większa, niż wymiar przestrzeni ($k > n$).

(2) Punkty, w których funkcja f osiąga ekstrema są (przy spełnieniu podanych w twierdzeniu powyżej warunków) punktami krytycznymi funkcji Lagrange'a — nie oznacza to jednak, że zachowany jest charakter tych punktów krytycznych. Prosty przykład pokazuje, że w punkcie, w którym f ma maksimum, funkcja L wcale maksimum mieć nie musi. Rozważmy mianowicie funkcję $f(x,y) = x+y$ i poszukajmy jej maksimum na zbiorze $A = \{x \in \mathbb{R}^2 \ : \ x+y = 2\}$ — a więc $g(x,y) = x+y-2$. Oczywiście A jest prostą $y = 2-x$, na niej $f(x,y) = f(x, 2-x) = x(2-x)$. Nietrudno sprawdzić, że funkcja ta przyjmuje maksimum równe 1 w punkcie $x_o = 1$ (a więc $y_o = 2 - x_o = 1$), zatem szukanym maksimum jest 1 osiągane w $(x_o, y_o) = (1,1)$. Z drugiej strony $L(x,y) = xy - \lambda(x+y)$, więc punkty krytyczne L leżące w A spełniają układ równań

$$\nabla L(x,y) = (y-\lambda, x-\lambda) = (0,0)$$
$$x + y = 2,$$

co natychmiast daje nam jako rozwiązanie $(x,y) = (1,1)$ — zgodnie z oczekiwaniami — oraz $\lambda = 1$. Z drugiej strony, kładąc w L mnożnik $\lambda = 1$ otrzymujemy

$$D^2L(1,1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \text{— macierz nieokreślona,}$$

zatem $L(x,y)$ ma w $(1,1)$ nie maksimum, lecz punkt siodłowy.

8. Interpretacja mnożników Lagrange'a. Rozpatrzmy następujący problem:

Załóżmy, że funkcja $f(x,y)$ przyjmuje w (x_m, y_m) maksimum na zbiorze $g(x,y) = m$. Jak wartość tego maksimum zależy od m?

Oznaczmy przez $v(m)$ wartość tego maksimum: $v(m) = f(x_m, y_m)$ i załóżmy, że v jest różniczkowalną funkcją m (to założenie nie zawsze jest spełnione, ale to zupełnie inna historia). Zdefiniujmy teraz funkcję $H(x,y) = f(x,y) - v(g(x,y))$. Natychmiast widzimy, że

$$\begin{aligned} H(x_m, y_m) &= f(x_m, y_m) - v(g(x_m, y_m)) = f(x_m, y_m) - v(m) \\ &= f(x_m, y_m) - f(x_m, y_m) = 0. \end{aligned}$$

Z drugiej strony

$$\forall (x,y) \in \mathbb{R}^2 \quad H(x,y) = f(x,y) - \max_{g(\xi,\zeta)=g(x,y)} f(\xi,\zeta) \leq 0.$$

Oznacza to, że funkcja $H(x,y)$ ma w (x_m, y_m) maksimum lokalne, zatem

$$\begin{aligned} 0 = \frac{\partial H}{\partial x}(x_m, y_m) &= \frac{\partial f}{\partial x}(x_m, y_m) - v'(g(x_m, y_m))\frac{\partial g}{\partial x}(x_m, y_m) \\ &= \frac{\partial f}{\partial x}(x_m, y_m) - v'(m)\frac{\partial g}{\partial x}(x_m, y_m), \\ 0 = \frac{\partial H}{\partial y}(x_m, y_m) &= \frac{\partial f}{\partial y}(x_m, y_m) - v'(g(x_m, y_m))\frac{\partial g}{\partial y}(x_m, y_m) \\ &= \frac{\partial f}{\partial y}(x_m, y_m) - v'(m)\frac{\partial g}{\partial y}(x_m, y_m). \end{aligned}$$

Widać, że jeżeli położymy $\lambda = v'(m)$, otrzymamy warunki Lagrange'a dla naszego zagadnienia. Daje to nam — oprócz dowodu Twierdzenia Lagrange'a — interpretację mnożnika Lagrange'a: mówi on, jak szybko zmienia się wartość maksimum w miarę jak zmieniamy poziomicę funkcji g, na której szukamy maksimum związanego.

W podanym powyżej przykładzie założyliśmy, że funkcja f ma w (x_m, y_m) maksimum, ale równie dobrze moglibyśmy założyć, że ma w tym punkcie minimum — punkt ten byłby wówczas minimum funkcji $H(x,y)$. Nietrudno też uogólnić powyższy wynik na większą liczbę warunków i zmiennych.

9. Zastosowania ekonomiczne. Na początek bardzo prosty przykład: firma do produkcji potrzebuje dwóch składników x i y, częściowo wymiennych, kupowanych za ceny p i q. Wartość jej produkcji opisana jest przez funkcję produkcji $F(x,y)$. Typowym zagadnieniem jest: jak dobrać ilości składników x i y by zminimalizować koszty, utrzymując ten sam poziom produkcji? Oczywiście sprowadza się to do poszukiwania ekstremum związanego

$$\operatorname{argmin}_{(x,y)\in A}(px + qy), \qquad \text{gdzie } A = \{(x,y) \in \mathbb{R}^2 \ : \ F(x,y) = a\}.$$

Załóżmy, że funkcja $f(x,y) = px + qy$ przyjmuje (sensowne - np. nieujemne) minimum związane w punkcie (x_a, y_a). Funkcja Lagrange'a dla tego problemu to $L(x,y) = f(x,y) - \lambda(F(x,y) - a)$, a warunki Lagrange'a mają postać

$$\begin{aligned} \frac{\partial L}{\partial x}(x_a, y_a) &= \frac{\partial f}{\partial x}(x_a, y_a) - \lambda\frac{\partial F}{\partial x}(x_a, y_a) = p - \lambda\frac{\partial F}{\partial x}(x_a, y_a) = 0 \\ \frac{\partial L}{\partial y}(x_a, y_a) &= \frac{\partial f}{\partial y}(x_a, y_a) - \lambda\frac{\partial F}{\partial y}(x_a, y_a) = q - \lambda\frac{\partial F}{\partial y}(x_a, y_a) = 0 \end{aligned}$$

Jeżeli przez $K(a)$ oznaczymy minimalny koszt potrzebny do wyprodukowania towaru o wartości a, to, na mocy podanej dopiero co interpretacji mnożnika λ, wiemy, że λ jest równa kosztowi krańcowemu $K'(a)$. Warto też zauważyć, że $\lambda = \frac{\partial L}{\partial a}$.

Oczywiście ów minimalny koszt zależy nie tylko od wartości produkcji a, ale przede wszystkim od cen p i q, należałoby więc pisać $K(p,q,a)$ zamiast $K(a)$. Czy potrafimy powiedzieć, jak K zależy od tych cen?

By to ustalić, popatrzmy trochę inaczej na nasze zagadnienie optymalizacyjne: jak dotąd szukaliśmy minimum $f(x,y)$ przy ustalonych p, q i a, ale p i q trzymaliśmy rzeczywiście stałe, podczas gdy na zmienność a dawaliśmy ciche przyzwolenie, podkreślając np. że punkt, w którym osiągane jest minimum związane, zależy od a, gdyż oznaczaliśmy go przez (x_a, y_a) (w rzeczywistości wartości x_a i y_a zależą zarówno od a, jak i od p i q). Spróbujmy teraz dla odmiany ustalić na dobre q i a, a zgodzić się na zależność (x_a, y_a) od zmiennej p — od tej chwili nasze zagadnienie

$$\boxed{\text{znajdź argmin} f(x,y,p) = px + qy \text{ z warunkiem } F(x,y) = a}$$

ma rozwiązanie $(x(p), y(p))$ (jest to ten sam punkt, co poprzednio, tylko zwracamy szczególną uwagę na jego zależność od p — podobnie podkreślamy też zależność funkcji kosztów $f(x,y,p)$). Oczywiście

Funkcje x_a i y_a, zależne, jak podkreślaliśmy, w rzeczywistości od a, p i q, nazywane są *funkcjami warunkowego popytu* — warunkowego, bo uzależnionego od wielkości produkcji a.

$$\begin{aligned} K(p,q,a) &= f(x(p), y(p), p) \\ a &= F(x(p), y(p)), \end{aligned}$$

więc

$$\begin{aligned} \frac{\partial K}{\partial p}(p,q,a) &= \frac{\partial f}{\partial x}(x(p), y(p), p)x'(p) + \frac{\partial f}{\partial y}(x(p), y(p), p)y'(p) + \frac{\partial f}{\partial p}(x(p), y(p), p) \\ 0 &= \frac{\partial F}{\partial x}(x(p), y(p))x'(p) + \frac{\partial F}{\partial y}(x(p), y(p))y'(p). \end{aligned}$$

Mnożąc drugą równość stronami przez λ z warunków Lagrange'a i odejmując ją od pierwszej otrzymujemy

$$\begin{aligned} \frac{\partial K}{\partial p}(p,q,a) = &\left(\frac{\partial f}{\partial x}(x(p), y(p), p) - \lambda\frac{\partial F}{\partial x}(x(p), y(p))\right) x'(p) \\ &+ \left(\frac{\partial f}{\partial y}(x(p), y(p), p) - \lambda\frac{\partial F}{\partial y}(x(p), y(p))\right) y'(p) \\ &+ \frac{\partial f}{\partial p}(x(p), y(p), p). \end{aligned}$$

Zauważmy, że punkt $(x(p), y(p))$ spełnia warunki Lagrange'a, a więc wyrażenia w dużych nawiasach są równe 0. Ostatecznie

$$\frac{\partial K}{\partial p}(p,q,a) = \frac{\partial f}{\partial p}(x(p), y(p), p) = x(p) = x_a.$$

Analogicznie pokazujemy, że $\frac{\partial K}{\partial q}(p,q,a) = y_a$. Otrzymaliśmy więc interesujący wynik, znany jako lemat Shepharda:

Ronald William Shephard (1912-1982) — matematyk i ekonomista amerykański

TWIERDZENIE 20 (lemat Shepharda). *Pochodna funkcji kosztu względem ceny czynnika produkcji jest równa warunkowemu popytowi na ten czynnik.*

Wykład 5

10. Funkcje quasiwypukłe i quasiwklęsłe. Na początek rozważmy, jak poprzednio, funkcję wypukłą $f : A \subset \mathbb{R}^n \to \mathbb{R}$. Z definicji wypukłości mamy, że dla dowolnych $x, y \in A$ i $\alpha \in [0,1]$

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha)f(y).$$

Zauważmy, że prawa strona nigdy nie jest większa od $\max\{f(x), f(y)\}$. Załóżmy bowiem, że $f(x) \geq f(y)$. Wówczas $\alpha f(x) + (1-\alpha)f(y) \leq \alpha f(x) + (1-\alpha)f(x) = f(x)$. Analogicznie, jeżeli $f(y) \geq f(x)$, otrzymujemy $\alpha f(x) + (1-\alpha)f(y) \leq \alpha f(y) + (1-\alpha)f(y) = f(y)$. Tak więc dla funkcji wypukłych zachodzi nierówność

$$f(\alpha x + (1-\alpha)y) \leq \max\{f(x), f(y)\}. \tag{2}$$

Oczywiście własność tę mają nie tylko funkcje wypukłe. Mają ją na przykład wszystkie funkcje monotoniczne (dlaczego?), jak również funkcja

$$f : \mathbb{R} \to \mathbb{R} \qquad f(x) = \begin{cases} 1 & x < -1 \\ 0 & x \in [-1,1] \\ 1 & x > 1, \end{cases}$$

która nie jest ani wypukła, ani monotoniczna, ani nawet ciągła.

Definicja 10. Funkcję określoną na zbiorze wypukłym A, spełniającą dla dowolnych $x, y \in A$ i $\alpha \in [0,1]$ nierówność (2), nazywamy funkcją *quasiwypukłą*.

Wiemy zatem, że każda funkcja wypukła jest quasiwypukła, ale nie odwrotnie. Oczywiście w podobny sposób możemy zdefiniować funkcje quasiwklęsłe:

Definicja 11. Funkcję określoną na zbiorze wypukłym A, spełniającą dla dowolnych $x, y \in A$ i $\alpha \in [0,1]$ nierówność

$$f(\alpha x + (1-\alpha)y) \geq \min\{f(x), f(y)\}.$$

nazywamy funkcją *quasiwklęsłą*.

Łatwo zauważyć, że każda funkcja wklęsła, jak również każda funkcja monotoniczna jest quasiwklęsła.

Twierdzenie 21. *Funkcja $f : A \subset \mathbb{R}^n \to \mathbb{R}$ jest quasiwypukła wtedy i tylko wtedy, gdy dla dowolnego $a \in \mathbb{R}$ zbiór $X_a = \{x \in A \ : \ f(x) \leq a$ jest wypukły.*

Proof. Najpierw wykażemy, że dla funkcji quasiwypukłej zbiór X_a jest, dla dowolnego $a \in \mathbb{R}$, zbiorem wypukłym. Jest tak, jeżeli X_a jest zbiorem pustym (zbiór pusty jest, na mocy umowy, wypukły); możemy zatem w dalszym ciągu dowodu założyć, że $X_a \neq \varnothing$. Musimy wykazać, że jeżeli x i y należą do X_a, to również $\alpha x + (1-\alpha)y$ należy, dla dowolnego $\alpha \in [0,1]$, do X_a.

Skoro $x, y \in X_a$, to $f(x) \leq a$, $f(y) \leq a$. Z definicji quasiwypukłości wiemy natomiast, że

$$f(\alpha x + (1-\alpha)y) \leq \max\{f(x), f(y)\}$$

. Prawa strona jest oczywiście nie większa niż a, co dowodzi, że $f(\alpha x + (1-\alpha)y) \leq a$, czyli $\alpha x + (1-\alpha)y \in X_a$.

Pozostaje nam udowodnić wynikanie w drugą stronę: jeżeli dla każdego $a \in \mathbb{R}$ zbiór X_a jest wypukły, to funkcja f jest quasiwypukła (czyli spełniona jest nierówność (2)). Wybierzmy dowolne $x, y \in A$ oraz $\alpha \in [0,1]$. Oznaczmy przez a liczbę $\max\{f(x), f(y)\}$. Z założenia wiemy, że zbiór X_a jest wypukły. Co więcej, zarówno x, jak i y należą do X_a — a więc, z wypukłości X_a, również punkt $\alpha x + (1-\alpha)y$ należy do X_a. To jednak oznacza, że $f(\alpha x + (1-\alpha)y) \leq a = \max\{f(x), f(y)\}$, a więc funkcja f jest quasiwypukła.

□

11. Twierdzenia o oddzielaniu.

Na początku udowodnijmy bardzo użyteczny, choć prosty fakt:

TWIERDZENIE 22. *Funkcja ściśle wypukła f osiąga minimum na zbiorze wypukłym i domkniętym A w co najwyżej jednym punkcie.*

PROOF. Załóżmy, że $\min_{x\in A} f(x) = \alpha$ i że istnieją dwa różne punkty x i y takie, że $f(x) = f(y) = \alpha$. Z nierówności Jensena otrzymujemy wówczas

$$f\left(\frac{x+y}{2}\right) < \frac{f(x)+f(y)}{2} = a.$$

Ale z drugiej strony punkt $\frac{x+y}{2}$ należy do A, co przeczy temu, że $a = \min_A f$. □

Z ćwiczeń wiemy (jest to zresztą bardzo proste), że podprzestrzenie liniowe $\mathbb{R}^n$ są zbiorami wypukłymi — w szczególności wypukłe są hiperpłaszczyzny (proste w $\mathbb{R}^2$, płaszczyzny w $\mathbb{R}^3$, podprzestrzenie wymiaru 3 w $\mathbb{R}^4$ itd.). Łatwo też zauważyć, że wypukłe są półprzestrzenie, na które taka hiperpłaszczyzna dzieli całą przestrzeń.

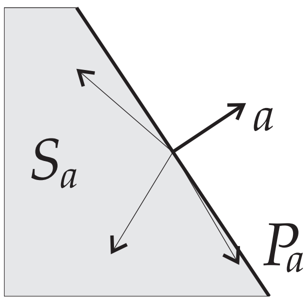
Zauważmy, że dla ustalonego $a \in \mathbb{R}^n$, $a \neq 0$, zbiór

$$P_a = \{x \in \mathbb{R}^n \;:\; \langle x\,,\,a\rangle = 0\}$$

opisuje hiperpłaszczyznę przechodzącą przez punkt $0 \in \mathbb{R}^n$. Jak łatwo zauważyć, jest to płaszczyzna prostopadła do wektora a — warunek na należenie do P_a to dokładnie warunek na to, by wektor łączący x z punktem 0 (czyli wektor wodzący punktu x) był prostopadły do wektora a.

A co możemy powiedzieć o zbiorze

$$S_a = \{x \in \mathbb{R}^n \;:\; \langle x\,,\,a\rangle \leq 0\}?$$



(Hiper)płaszczyzna P_a i półprzestrzeń S_a.

Zauważmy, że iloczyn skalarny $\langle x\,,\,a\rangle$ to iloczyn długości wektorów x i a oraz kosinusa kąta pomiędzy nimi. O ile tylko $a, x \neq 0$, to długości tych wektorów są zawsze dodatnie. Aby więc ich iloczyn skalarny mógł być niedodatni, niedodatni musi być kosinus — a to może mieć miejsce tylko wtedy, gdy kąt między a a x jest $\geq \pi/2 = 90^o$. Łatwo więc zrozumieć, że S_a to jedna z dwóch domkniętych półprzestrzeni, na jakie P_a dzieli całą przestrzeń — a dokładniej ta, do której nie należy a.

Jak możemy opisać hiperpłaszczyzny nieprzechodzące przez początek układu współrzędnych? Punkty hiperpłaszczyzny przechodzącej przez punkt z i prostopadłej do wektora a muszą spełniać warunek $\langle x - z\,,\,a\rangle = 0$, co można przepisać jako $\langle x\,,\,a\rangle = \langle z\,,\,a\rangle$. I odwrotnie, zbiór $\{x \in \mathbb{R}^n \;:\; \langle x\,,\,a\rangle = b\}$ to pewna hiperpłaszczyzna, nieprzechodząca (o ile $b \neq 0$) przez $0 \in \mathbb{R}^n$. Dzieli ona przestrzeń na dwie półprzestrzenie, opisane odpowiednio warunkiem $\langle x\,,\,a\rangle \geq b$ i $\langle x\,,\,a\rangle \leq b$.

Po co nam te wszystkie fakty dotyczące hiperpłaszczyzn? Otóż jedne z najważniejszych własności zbiorów wypukłych opisują tzw. twierdzenia o oddzielaniu: dwa rozłączne zbiory wypukłe można rozdzielić hiperpłaszczyzną tak, że jeden leży po jednej, a drugi – po drugiej jej stronie.

TWIERDZENIE 23 (Twierdzenie o oddzielaniu punktu od domkniętego zbioru wypukłego). *Załóżmy, że zbiór A jest domknięty i wypukły, a punkt y leży poza A. Istnieją wówczas liczba rzeczywista b i wektor $v \in \mathbb{R}^n$, $v \neq 0$, takie, że*

$$\forall x \in A \qquad \langle v\,,\,y\rangle < b \leq \langle v\,,\,x\rangle.$$

Nim przejdziemy do dowodu twierdzenia, zdradzę, skąd bierze się wektor v z jego tezy. Jest to mianowicie wektor $z - y$, gdzie z to punkt zbioru A leżący najbliżej punktu y. Wypadałoby wykazać, że punkt taki rzeczywiście istnieje.

Chciałoby się napisać, że funkcja $f(x) = \|y-x\|$ musi mieć na (domkniętym) zbiorze A minimum. Nie możemy jednak powołać się od razu na tw. Weierstrassa, gdyż zbiór A nie jest zbiorem zwartym. Radzimy sobie z tym używając prostej sztuczki: wybieramy dowolny punkt $v \in A$ i oznaczamy $\beta = \|y - v\|$. Poprawimy też nieco funkcję f — zamiast niej będziemy szukać minimum $g(x) = \|y - x\|^2$. Teraz zbiór $C = A \cap B(y, \beta)$ jest wypukły, domknięty i ograniczony — a więc zwarty. Funkcja g przyjmuje zatem na C minimum w pewnym punkcie z; oczywiście $\min_C g = \min_A g$, gdyż dla punktów z $A \setminus C$ kwadrat odległości od y jest większy niż β^2, a $g(z) = \|z - y\|^2 \leq \beta^2$.

Tw. Weierstrassa: funkcja ciągła na zbiorze zwartym przyjmuje zarówno minimum, jak i maksimum.

Przez $B(y, \beta)$ oznaczam domkniętą kulę o środku w y i promieniu β.

Funkcja g jest ściśle wypukła, toteż z twierdzonka z początku wykładu wynika, że istnieje dokładnie jeden punkt zbioru A najbliższy punktowi y.

Dowód Twierdzenia 23. Jak zapowiadałem, za punkt z bierzemy punkt zbioru A najbliższy punktowi y. Jest on, jak już zauważyliśmy, minimum funkcji $g(x) = \|x - y\|^2$ na zbiorze A. Oznacza to w szczególności, że funkcja g nie może maleć, gdy poruszamy się z z w kierunku wnętrza zbioru A — a zatem

$$\forall x \in A \quad g'_{x-z}(z) = \langle \nabla g(z) \, , \, x - z \rangle \geq 0$$

(pochodne kierunkowe w kierunkach do środka zbioru A muszą być nieujemne). Gradient funkcji g możemy łatwo policzyć: $\nabla g(z) = 2(z - y)$, toteż

$$\forall x \in A \quad \langle 2(z - y) \, , \, x - z \rangle \geq 0 \quad \Rightarrow \quad \langle z - y \, , \, x \rangle \geq \langle z - y \, , \, z \rangle$$

Połóżmy $v = z - y$, $b = \langle z - y \, , \, z \rangle = \langle v \, , \, z \rangle$. Mamy $\langle z - y \, , \, z - y \rangle = \|v\|^2 > 0$, więc $b = \langle z - y \, , \, z \rangle > \langle z - y \, , \, y \rangle$, co kończy dowód. □

Uwaga: w tezie Twierdzenia 23 możemy łatwo dostać obie nierówności ostre. Wystarczy wziąć $c = \frac{1}{2}(b + \langle v \, , \, y \rangle)$, by otrzymać

Twierdzenie 24 (Druga wersja Twierdzenia o oddzielaniu). *Załóżmy, że zbiór A jest domknięty i wypukły, a punkt y leży poza A. Istnieją wówczas liczba rzeczywista c i wektor $v \in \mathbb{R}^n$, $v \neq 0$, takie, że*

$$\forall x \in A \qquad \langle v \, , \, y \rangle < c < \langle v \, , \, x \rangle.$$

Twierdzenie 25 (Twierdzenie o płaszczyźnie podpierającej). *Niech A będzie domkniętym zbiorem wypukłym, zaś z — punktem leżącym na brzegu A. Istnieje hiperpłaszczyzna przechodząca przez z taka, że cały zbiór A leży po jednej jej stronie. Innymi słowy, istnieje $v \in \mathbb{R}^n$, $v \neq 0$, takie, że*

$$\forall x \in A \quad \langle v \, , \, z \rangle \leq \langle v \, , \, x \rangle$$

Tw. Bolzano-Weierstrassa: Z każdego ciągu elementów ograniczonego podzbioru $\mathbb{R}^n$ (innymi słowy: z każdego ograniczonego ciągu punktów z $\mathbb{R}^n$) można wybrać podciąg zbieżny.

Bernard Placidus Johann Nepomuk Bolzano (1781–1848) — czeski matematyk, filozof i teolog; ksiądz katolicki, bardzo oddany dobroczynności. Ma wielkie (choć nieco niedoceniane — bo zauważone długo po jego śmierci) zasługi w uporządkowaniu podstaw analizy matematycznej.

Proof. Skoro z leży na brzegu zbioru A, to istnieje ciąg punktów y_i leżących poza A taki, że $z = \lim_{i\to\infty} y_i$. Każdy z punktów y_i możemy oddzielić od zbioru A hiperpłaszczyzną $P_i = \{x \in \mathbb{R}^n \; : \; \langle v_i \, , \, x \rangle = \langle v_i \, , \, z_i \rangle\}$, gdzie punkt z_i jest punktem A leżącym najbliżej y_i, a $v_i = z_i - y_i$. Zauważmy, że warunek na P_i możemy przepisać jako

$$P_i = \{x \in \mathbb{R}^n \; : \; \langle \tilde{v}_i \, , \, x \rangle = \langle \tilde{v}_i \, , \, z_i \rangle\},$$

gdzie $\tilde{v}_i = \frac{v_i}{\|v_i\|}$.

Łatwo możemy wykazać, że punkty z_i dążą do z, mamy bowiem

$$\|z_i - z\| = \|(z_i - y_i) + (y_i - z)\| \leq \|z_i - y_i\| + \|y_i - z\| \leq 2\|y_i - z\|,$$

przy czym druga nierówność bierze się stąd, że z_i jest tym punktem w A, który leży najbliżej y_i, a z – pewnym innym punktem A, więc $\|z_i - y_i\| \leq \|z - y_i\|$. Wiemy jednak, że $\|y_i - z\| \xrightarrow{i\to\infty} 0$, więc i $\|z_i - z\| \xrightarrow{i\to\infty} 0$.

Punkty $\tilde{v}_i$ leżą na sferze jednostkowej, która jest zbiorem zwartym. Możemy powołać się zatem na Twierdzenie Bolzano–Weierstrassa , by wybrać z ciągu $\tilde{v}_i$ pewien podciąg zbieżny $\tilde{v}_{i_j} \xrightarrow{j\to\infty} v$. Zauważmy, że dla każdego j zachodzi nierówność

$$\forall x \in A \quad \langle \tilde{v}_{i_j}\, , \, x\rangle \geq \langle \tilde{v}_{i_j}\, , \, z_{i_j}\rangle.$$

Iloczyn skalarny jest ciągły, możemy więc przejść po obu stronach nierówności z j do ∞. Otrzymamy

$$\forall x \in A \quad \langle v\, , \, x\rangle \geq \langle v\, , \, z\rangle,$$

co oznacza, że płaszczyzna $P = \{x \in \mathbb{R}^n \ : \ \langle v\, , \, x\rangle = \langle v\, , \, z\rangle\}$ jest poszukiwaną płaszczyzną podpierającą w punkcie z. Jest jasne, że P przechodzi przez punkt z. □

Łatwo możemy zauważyć, że gdy zbiór A nie jest domknięty, to nie umiemy go w taki sposób, jaki podaje Twierdzenie 23, odseparować od niektórych punków z jego domknięcia. Niech, na przykład, zbiór $A \subset \mathbb{R}^2$ oznacza otwartą dolną półpłaszczyznę z dodanym punktem $(0, 0)$. To jest zbiór wypukły; jeżeli jednak spróbujemy oddzielić go prostą (hiperpłaszczyzną w $\mathbb{R}^2$) od punktu $(1, 0)$, który do A nie należy, okaże się, że nie ma dobrej kandydatki: wszystkie proste, poza tymi postaci $y = a$, $a > 0$, przecinają zbiór A; ale względem prostej $y = a$, przy $a > 0$, zarówno A, jak i $(1, 0)$ leżą po tej samej stronie. Możemy jednak uratować sytuację nieznacznie osłabiając tezę Twierdzenia 23:

Twierdzenie 26 (o oddzielaniu punktu od zbioru wypukłego — niekoniecznie domkniętego). *Niech U będzie wypukłym podzbiorem $\mathbb{R}^n$, a punkt x niech leży poza A. Istnieje wówczas wektor $v \in \mathbb{R}^n$ i stała $\alpha \in \mathbb{R}$ takie, że*

$$\forall y \in U \quad \langle x\, , \, v\rangle \leq \alpha \leq \langle y\, , \, v\rangle.$$

Proof. Mamy dwa przypadki:

- albo $x \notin \mathrm{cl}U$ — wówczas możemy zastosować znane nam twierdzenie o oddzielaniu — tym razem x od $\mathrm{cl}U$,
- albo $x \in \mathrm{cl}U$. Wówczas $\mathrm{cl}U$ ma w x hiperpłaszczyznę podpierającą, tj wektor v taki, że

$$\forall y \in \mathrm{cl}U \quad \langle x\, , \, v\rangle \leq \langle y\, , \, v\rangle.$$

  Oczywiście to, że nierówność powyższa zachodzi dla wszystkich $y \in \mathrm{cl}U$ jest mocniejszym warunkiem, niż to, że ma zachodzić dla wszystkich $y \in U$. Kładąc $\alpha = \langle x\, , \, v\rangle$ otrzymujemy tezę.

□

Twierdzenie o oddzielaniu można znacznie wzmocnić:

Twierdzenie 27 (Twierdzenie o oddzielaniu zbiorów wypukłych). *Załóżmy, że zbiory A i B są wypukłe, domknięte i rozłączne, i przynajmniej jeden z nich jest ograniczony. Istnieją wówczas liczba rzeczywista c i wektor $v \in \mathbb{R}^n$, $v \neq 0$, takie, że*

$$\forall x \in A,\ y \in B \qquad \langle v\, , \, y\rangle < c < \langle v\, , \, x\rangle.$$

Innymi słowy — istnieje hiperpłaszczyzna taka, że zbiór A leży po jednej, a B — po drugiej jej stronie.

Do dowodu tego twierdzenia potrzebujemy pewnego użytecznego narzędzia (wprowadzonego już na ćwiczeniach):

DEFINICJA 12. Niech A, $B \subset \mathbb{R}^n$. Zbiór
$$A + B = \{x + y \in \mathbb{R}^n \ : \ x \in A, \ y \in B\}$$
nazywamy *sumą Minkowskiego* zbiorów A i B. Analogicznie możemy zdefiniować *różnicę Minkowskiego*:
$$A - B = \{x - y \in \mathbb{R}^n \ : \ x \in A, \ y \in B\}.$$

Na ćwiczeniach sprawdziliśmy, że jeżeli zbiory A i B są wypukłe, to wypukłe są też $A + B$ oraz $A - B$. Łatwo można też sprawdzić, że gdy A i B są domknięte i przynajmniej jeden z nich jest zwarty (tj. domknięty i ograniczony), to domknięte są $A + B$ i $A - B$.

PROOF. Jak wspomniałem, zbiór $A - B$ jest wypukły i domknięty. Łatwo też można zauważyć, że $0 \notin A - B$. Gdyby bowiem $0 \in A - B$, to istniałyby punkty $x_o \in A$ i $y_o \in B$ takie, że $0 = x_o - y_o$, czyli $x_o = y_o$. Ale wówczas $x_o = y_o$ jest punktem wspólnym zbiorów A i B — a te miały być rozłączne.

Skoro punkt 0 leży poza zbiorem $A - B$, to istnieje, na mocy wersji Tw. o oddzielaniu z ostrymi nierównościami, płaszczyzna oddzielająca 0 od $A - B$, a więc wektor v i stała $\gamma \in \mathbb{R}$ taka, że
$$\forall z \in A - B \quad 0 = \langle v \, , \, 0 \rangle < \gamma < \langle v \, , \, z \rangle.$$
Zbiór $A - B$ to, z definicji, zbiór punktów postaci $x - y$, gdzie $x \in A$, $y \in B$. Wiemy zatem, że
$$\forall x \in A, \ y \in B \quad 0 = \langle v \, , \, 0 \rangle < \gamma < \langle v \, , \, x - y \rangle.$$
Oznacza to, że
$$\forall x \in A, \ y \in B \quad \langle v \, , \, x \rangle > \langle v \, , \, y \rangle + \gamma, \quad \gamma > 0.$$
W szczególności
$$\inf_{x \in A} \langle v \, , \, x \rangle \geq \sup_{y \in B} \langle v \, , \, y \rangle + \gamma.$$
Wystarczy więc za c wziąć liczbę $\sup_{y \in B} \langle v \, , \, y \rangle + \gamma/2$. □

Mamy też analogiczne twierdzenie w przypadku, gdy zbiory nie są domknięte, lub gdy oba są domknięte i żaden z nich nie jest ograniczony:

TWIERDZENIE 28 (o oddzielaniu dwóch — niekoniecznie domkniętych — zbiorów wypukłych). *Niech U, V będą wypukłymi podzbiorami $\mathbb{R}^n$ i niech $U \cap V = \varnothing$. Istnieje wówczas wektor $v \in \mathbb{R}^n$ i stała $\alpha \in \mathbb{R}$ takie, że*
$$\forall x \in U, \, y \in V \quad \langle x \, , \, v \rangle \leq \alpha \leq \langle y \, , \, v \rangle.$$

Dowód tego twierdzenia przebiega dokładnie tak samo, jak Twierdzenia 27, tylko zamiast powoływać się na Twierdzenie 23 należy powołać się na mające słabsze założenia Twierdzenie 26.

Na koniec przedstawimy ważny wniosek z Twierdzenia 25. Jak już wspominaliśmy, jeżeli funkcja wypukła ma punkt krytyczny, to automatycznie osiąga w nim minimum. A jak szukać maksimum funkcji wypukłej na zbiorze wypukłym A? Pomóc w tym może następujące twierdzenie:

TWIERDZENIE 29. *Niech $A \subset \mathbb{R}^n$ będzie zwartym, niepustym zbiorem wypukłym, a $f : A \to \mathbb{R}$ niech będzie funkcją wypukłą i ciągłą w A. Istnieje wówczas punkt $a \in \mathrm{ex}A$ taki, że $f(a) = \max_{x \in A} f(x)$.*

PROOF. Przeprowadzimy dowód indukcyjny, z indukcją ze względu na n – wymiar przestrzeni $\mathbb{R}^n$, w której leży zbiór A.

Na wstępie zauważmy, że zbiór A jest zwarty, więc na mocy Twierdzenia Weierstrassa funkcja f przyjmuje na A swoje kresy. Istnieje zatem punkt $x_o \in A$ taki, że $\max_{x \in A} f(x) = f(x_o)$. Musimy zatem jedynie (?) wykazać, że x_o możemy wybrać spośród punktów ekstremalnych A.

Załóżmy, że $n = 1$ – wówczas zbiór A musi być albo punktem $\{a\}$, albo odcinkiem domkniętym $[a, b]$. W pierwszym przypadku $\mathrm{ex}A = A = \{a\}$, nie ma więc nic do dowodu.

Niech zatem $A = [a, b]$ (wówczas $\mathrm{ex}A = \{a, b\}$). Wiemy, że $x_o = \alpha a + (1 - \alpha)b$ dla jakiegoś $\alpha \in [0, 1]$ (każdy punkt odcinka $[a, b]$ da się tak zapisać), zatem

$$\begin{aligned}\max_{x\in A} f(x) &= f(x_o) = f(\alpha a + (1-\alpha)b) \leq \alpha f(a) + (1-\alpha)f(b)\\ &\leq \alpha \max\{f(a), f(b)\} + (1-\alpha)\max\{f(a), f(b)\} = \max\{f(a), f(b)\} \leq \max_{x\in A} f(x)\end{aligned}$$

Jeżeli powyższy ciąg nierówności ma być prawdziwy, to wszystkie nierówności w środku muszą w rzeczywistości być równościami! W szczególności $f(x_o) = \max\{f(a), f(b)\}$, a zatem albo $f(b)$, albo $f(a)$ jest równe $f(x_o) = \max_{x\in A} f(x)$.

Rozstrzygnęliśmy w ten sposób przypadek $n = 1$. Załóżmy teraz, że teza twierdzenia zachodzi dla każdego zbioru A spełniającego założenia twierdzenia i leżącego w przestrzeni n wymiarowej. Wykażemy, że teza zachodzi również dla zbiorów z przestrzeni $n + 1$-wymiarowej.

Niech więc, jak poprzednio, $\max_{x\in A} f(x) = f(x_o)$ i niech $x_o \notin \mathrm{ex}A$. Wiemy wówczas, że x_o jest środkiem odcinka o końcach y_o i z_o, należących do A, przy czym $y_o \neq z_o$. Niech l oznacza prostą przechodzącą przez punkty y_o i z_o. Zbiór $l \cap A$ jest zwarty i wypukły, i zawarty w (jednowymiarowej!) prostej l, do tego zawiera odcinek $[y_o, z_o]$ – jest więc odcinkiem domkniętym. Oznaczmy jego końce przez y i z.

Z rozważań dla $n = 1$ wiemy, że albo $\max_{x\in[y,z]} f(x) = f(y)$, albo $\max_{x\in[y,z]} f(x) = f(z)$. Bez zmniejszenia ogólności możemy przyjąć, że zachodzi pierwsza możliwość; w szczególności otrzymujemy, że $f(x_o) \leq f(y)$, bo $x_o \in [y, z]$. Z drugiej strony $f(x_o)$ to maksimum f na A, więc $f(y) \leq f(x_o)$. Ostatecznie oznacza to, że $f(y) = f(x_o)$.

Punkt y leży na brzegu zbioru A, a więc A ma w y płaszczyznę podpierającą. Istnieje więc wektor v taki, że

$$\forall_{x\in A} g(x) := \langle x\,,\, v\rangle \leq \langle y\,,\, v\rangle.$$

Dalszy ciąg jest bardzo podobny do dowodu Twierdzenia Minkowskiego, jednak zamiast rzutowania na pierwszą współrzędną używamy funkcji g. Oznaczmy $c = \langle y\,,\, v\rangle$. Funkcja $g : A \to \mathbb{R}$ jest liniowa, zatem $g^{-1}(\{c\})$ to hiperpłaszczyzna wymiaru n, a zbiór $B = A \cap g^{-1}(\{c\})$ jest zwarty, wypukły, niepusty i zawarty w tej hiperpłaszczyźnie. Możemy więc do niego zastosować założenie indukcyjne: funkcja f przyjmuje na nim swoje maksimum w pewnym punkcie ekstremalnym:

$$\exists_{u\in \mathrm{ex}B} \quad f(u) = \max_{x\in B} f(x).$$

Z drugiej strony $y \in B$, a wiemy już, że $f(y$ to maksymalna wartość f w A, więc $f(u) = f(y) = \max_{x\in A} f(x)$. Pozostaje wykazać, że u jest punktem ekstremalnym nie tylko B, ale i A.

Załóżmy zatem nie wprost, że dla pewnych $v, w \in A$ mamy $u = \frac{1}{2}(v + w)$, $v \neq w$. Oczywiście oba punkty v, w nie mogą równocześnie leżeć w B, bo $u \in \mathrm{ex}B$ – załóżmy więc, że $v \notin B$. Oznacza to, że $g(v) < c$; o w i u wiemy natomiast, że $g(w) \leq c$, $g(u) = c$. To prowadzi do sprzeczności, gdyż

$$c = g(u) = g(\frac{1}{2}(v + w)) = \frac{1}{2}(g(v) + g(w)) < \frac{1}{2}(c + c) = c.$$

□

Wykład 6

12. Stożki. Spośród zbiorów wypukłych, szczególną rolę w optymalizacji odgrywają stożki

DEFINICJA 13. Zbiór $A \subset \mathbb{R}^n$ nazywamy *stożkiem*, jeżeli dla każdego $x \in A$ i $\lambda \geq 0$ punkt λx też należy do A.

Zauważmy, że punkt 0 – czubek stożka – należy do każdego stożka (wielu autorów rozróżnia *stożki zaostrzone* zawierające 0 i *tępe*, nie zawierające go – ułatwia to sformułowanie pewnych twierdzeń, komplikując inne).

Stożek, który jest równocześnie zbiorem wypukłym, nazywamy, jak łatwo się domyślić, *stożkiem wypukłym.*

TWIERDZENIE 30. *Zbiór* $A \subset \mathbb{R}^n$ *jest stożkiem wypukłym wtedy i tylko wtedy, gdy*

$$\forall x, y \in A,\ \forall \lambda, \mu \geq 0 \quad \lambda x + \mu y \in A. \tag{3}$$

PROOF. Dowód jest bardzo prosty. Załóżmy najpierw, że A jest stożkiem wypukłym. Niech $x, y \in A$, $\lambda, \mu \geq 0$. Jeżeli $\lambda = \mu = 0$, to punkt $\lambda x + \mu y = 0$ należy do A, gdyż A jest stożkiem, a punkt $0 \in \mathbb{R}^n$ należy do każdego stożka. Możemy więc założyć, że λ i μ nie są równocześnie równe 0. Wówczas, z wypukłości A, punkt $z = \frac{\lambda}{\lambda+\mu}x + \frac{\mu}{\lambda+\mu}y$ należy do A (mamy $\frac{\lambda}{\lambda+\mu} + \frac{\mu}{\lambda+\mu} = 1$, $\frac{\lambda}{\lambda+\mu}, \frac{\mu}{\lambda+\mu} \in [0,1]$). Następnie, skoro A jest stożkiem, punkt $(\lambda+\mu)z = \lambda x + \mu y$ należy do A.

Musimy jeszcze udowodnić wynikanie w drugą stronę. Załóżmy, że zbiór A ma własność (3). Mamy udowodnić, że A jest wypukły oraz że jest stożkiem. Jest to jednak oczywiste: jeżeli weźmiemy w (3) λ i μ takie, że $\lambda + \mu = 1$, otrzymamy warunek na wypukłość, kładąc natomiast $\mu = 0$ — warunek na to, że A jest stożkiem. □

DEFINICJA 14. Stożkiem wypukłym *generowanym przez zbiór wypukły* A nazywamy zbiór

$$\mathrm{cone}(A) = \{x \in \mathbb{R}^n \ :\ x = \lambda y \text{ dla pewnych } \lambda \geq 0 \text{ i } y \in A\}.$$

Łatwo można sprawdzić (ćwiczenia), że zbiór $\mathrm{cone}(A)$ jest najmniejszym stożkiem wypukłym zawierającym zbiór A.

DEFINICJA 15. Niech, jak zwykle, A będzie zbiorem wypukłym, $a \in A$. Stożek $\mathrm{cone}(A - \{a\})$ (dalej będziemy pisać po prostu $A - a$) nazywamy *stożkiem kierunków osiągalnych* lub *dopuszczalnych* zbioru A w punkcie a i oznaczamy $F_A(a)$. Wektory należące do $F_A(a)$ nazywamy *wektorami osiągalnymi* dla A w a.

TWIERDZENIE 31. *Następujące warunki są równoważne:*

a) $v \in F_A(a)$ *i* $v \neq 0$,
b) *istnieje* $\tau_o > 0$ *takie, że dla wszystkich* $\tau \in [0, \tau_o)$ *punkt* $a + \tau v$ *leży w* A.

Innymi słowy – wektory osiągalne w a to te, w kierunku których możemy się poruszać przez pewien (być może bardzo krótki) czas, nie opuszczając zbioru A.

PROOF. Udowodnimy najpierw, że z a) wynika b).

Załóżmy, że $v \in F_A(a) = \mathrm{cone}(A - a)$, $v \neq 0$. Oznacza to, że istnieje $y \in (A - a)$ i $\lambda > 0$ taka, że $v = \lambda y$ (*a priori* λ mogłaby być 0, ale wtedy v byłoby równe 0). Skoro $y \in (A - a)$, to istnieje $z \in A$ takie, że $y = z - a$, a zatem $v = \lambda(z - a)$. Oczywiście $z \neq a$, w przeciwnym przypadku bowiem v byłoby równe 0.

Mamy więc następujące pytanie: czy istnieje τ_o takie, że dla $\tau \in [0, \tau_o)$ punkt $a + \tau v = a + \tau\lambda(z - a)$ leży w A? Odpowiedź jest twierdząca:

$$a + \tau v = a + \tau\lambda(z - a) = (1 - \tau\lambda)a + \tau\lambda z$$

jest kombinacją liniową punktów a i z, a gdy $0 \leq \tau\lambda \leq 1$, jest kombinacją wypukłą – więc z wypukłości A badany punkt leży w A, wystarczy więc wziąć $\tau_o \leq 1/\lambda$.

Teraz wykażemy, że z b) wynika a).

Skoro zachodzi b), to istnieje $\tau > 0$ takie, że $a + \tau v \in A$. Oznacza to, że $z = \tau v \in A - a$, a więc biorąc $\lambda = 1/\tau$ widzimy, że $v = \lambda z$, gdzie $\lambda > 0$, $z \in A - a$ – a to z definicji oznacza, że $v \in F_A(a) = \mathrm{cone}(A - a)$. □

Twierdzenie 32. *Część wspólna dowolnej rodziny stożków wypukłych jest stożkiem wypukłym.*

Proof. Dowód jest bardzo prosty. Wiemy już, że część wspólna dowolnej rodziny zbiorów wypukłych jest zbiorem wypukłym. Pozostaje sprawdzić, czy, podobnie, część wspólna dowolnej rodziny stożków jest stożkiem. Niech więc $\{C_i\}_{i\in I}$ będzie badaną rodziną stożków. Aby stwierdzić, że $C = \bigcap_{i\in I} C_i$ jest stożkiem, musimy udowodnić, że dla każdego $x \in C$ i $\lambda \geq 0$ punkt λx też leży w C. Jest to jednak oczywiste: skoro $x \in C$, to dla każdego $i \in I$ mamy $x \in C_i$, a że C_i jest stożkiem, to $\lambda x \in C_i$. Zatem $\lambda x \in C$. □

Twierdzenie 33 (o tym, jak obliczać stożki osiągalne dla części wspólnej). *Niech A i B będą niepustymi zbiorami wypukłymi, $x \in A\cap B$. Wówczas $F_{A\cap B}(x) = F_A(x) \cap F_B(x)$.*

Proof. Na początek zauważmy, że $A \cap B - x = (A - x) \cap (B - x)$. Jeżeli bowiem $y \in A \cap B - x$, to znaczy, że istnieje $z \in A \cap B$ takie, że $y = z - x$; innymi słowy: $y + x \in A\cap B$. To jest równoważne stwierdzeniu, że $y + x \in A$ i $y + x \in B$, czyli $y \in A - x$ i $y \in B - x$, to zaś jest równoważne temu, że $y \in (A - x) \cap (B - x)$.

Teraz, z definicji, $F_{A\cap B}(x) = \mathrm{cone}(A\cap B - x) = \mathrm{cone}((A - x)\cap(B - x))$. Twierdzę, że ten ostatni stożek jest równy $\mathrm{cone}(A - x)\cap \mathrm{cone}(B - x) = F_A(x)\cap F_B(x)$. Łatwo możemy zauważyć, że

$$\mathrm{cone}((A - x) \cap (B - x)) \subset \mathrm{cone}(A - x) \cap \mathrm{cone}(B - x),$$

gdyż

(1) z zadania z ćwiczeń wiemy, że stożek $\mathrm{cone}((A - x)\cap(B - x))$ jest najmniejszym stożkiem wypukłym zawierającym zbiór $(A - x) \cap (B - x)$,
(2) a zbiór $\mathrm{cone}(A - x)\cap\mathrm{cone}(B - x)$ jest stożkiem wypukłym (z Tw. 32); $A - x \subset \mathrm{cone}(A-x)$, $B-x \subset \mathrm{cone}(B-x)$, więc $(A-x)\cap(B-x) \subset \mathrm{cone}(A-x)\cap\mathrm{cone}(B-x)$ – jest więc jakimś innym stożkiem wypukłym zawierającym $(A-x)\cap(B-x)$.

Wykażemy teraz zawieranie w przeciwną stronę. Niech $z \in \mathrm{cone}(A - x) \cap \mathrm{cone}(B - x)$. Możemy założyć, że $z \neq 0$, bo $z = 0$ z definicji należy do dowolnego stożka, w szczególności do $\mathrm{cone}((A - x) \cap (B - x))$.

$$0 \neq z \in \mathrm{cone}(A-x)\cap\mathrm{cone}(B-x) \;\Leftrightarrow\; \exists\lambda, \mu > 0,\ a \in A-x,\ b \in B-x \;:\; z = \lambda a \wedge z = \mu b.$$

Możemy założyć, że $\lambda, \mu \neq 0$, bo gdy któraś z nich jest zero, to i $z = 0$. Jeżeli teraz $\lambda = \mu$, to $a = z/\lambda = b \in (A - x) \cap (B - x)$, a zatem $z \in \mathrm{cone}((A - x) \cap (B - x))$. Co jednak, gdy $\lambda \neq \mu$?

Załóżmy bez zmniejszenia ogólności, że $\lambda > \mu$ i zauważmy, że zbiór $B-x$ jest wypukły i zawiera 0 (dlaczego?). Zawiera też $b = z/\mu$, a zatem $a = \mu/\lambda b + (1 - \mu/\lambda)0 \in B - x$. Stąd $z/\lambda = a \in (A - x) \cap (B - x)$, czyli $z \in \mathrm{cone}((A - x) \cap (B - x))$. □

Definicja 16. Niech A będzie stożkiem wypukłym. Zbiór

$$A^o = \{y \in \mathbb{R}^n \;:\; \langle x\,,\, y\rangle \leq 0 \text{ dla wszystkich } x \in A\}$$

nazywamy *stożkiem polarnym* stożka A.

Na ćwiczeniach sprawdzą Państwo, że A^o jest rzeczywiście stożkiem, co więcej – stożkiem wypukłym. A^o jest zawsze zbiorem domkniętym, niezależnie od tego, czy stożek A jest domknięty, czy nie.

Nas będą szczególnie interesować stożki polarne do stożków kierunków osiągalnych:

Definicja 17. Niech $A \subset \mathbb{R}^n$ będzie zbiorem wypukłym. Stożek $(F_A(x))^o$, polarny do stożka kierunków osiągalnych dla A w x, nazywamy *stożkiem normalnym* zbioru A w punkcie x i oznaczamy $N_A(x)$.

Do kompletu ważnych stożków związanych ze zbiorami wypukłymi brak nam jeszcze jednego.

Definicja 18. Niech, jak poprzednio, $A \subset \mathbb{R}^n$ będzie wypukły. Domknięcie stożka $F_A(x)$ nazywamy *stożkiem stycznym* do A w x i oznaczamy $T_A(x)$.

Stożki styczne definiuje się również dla zbiorów innych, niż wypukłe – i wówczas definicja ta jest bardziej złożona (mogą się na nią Państwo natknąć w literaturze). Nam, na szczęście, nie będzie ona potrzebna.

Twierdzenie 34. *Niech $A \subset \mathbb{R}^n$ będzie wypukły. Wówczas $(T_A(x))^o = N_A(x)$, zaś $(N_A(x))^o = T_A(x)$.*

Proof. Najpierw wykażemy pierwszą równość. Z definicji stożka polarnego wiemy, że

$$(T_A(x))^o = \{y \in \mathbb{R}^n \ : \ \langle z\,,\, y\rangle \le 0 \text{ dla wszystkich } z \in T_A(x)\},$$
$$N_A(x) = (F_A(x))^o = \{y \in \mathbb{R}^n \ : \ \langle z\,,\, y\rangle \le 0 \text{ dla wszystkich } z \in F_A(x)\}.$$

Oczywiście $F_A(x) \subset T_A(x)$, więc warunek „dla wszystkich $z \in T_A(x)$” jest mocniejszy (stawia większe wymagania) niż warunek „dla wszystkich $z \in F_A(x)$”. Oznacza to automatycznie, że $(T_A(x))^o \subset N_A(x)$. Załóżmy zatem, że istnieje $z_o \in N_A(x)$ nie należące do $(T_A(x))^o$. To, że $z_o \notin (T_A(x))^o$ oznacza, że dla pewnego $u \in T_A(x)$ zachodzi $\langle z_o\,,\, u\rangle > 0$. Z drugiej strony u leży w domknięciu $F_A(x)$, więc istnieje ciąg (u_k) punktów $F_A(x)$ zbieżny do u. Dla każdego u_k mamy $\langle z_o\,,\, u_k\rangle \le 0$, gdyż z_o in $N_A(x)$, $u_k \in F_A(x)$ i w nierówności tej możemy przejść z k do granicy. Otrzymujemy $\langle u\,,\, z_o\rangle \le 0$, wbrew założeniu.

Teraz druga równość. Mamy

$$(N_A(x))^o = \{y \in \mathbb{R}^n \ : \ \langle z\,,\, y\rangle \le 0 \text{ dla wszystkich } z \in N_A(x)\}$$

Zauważmy najpierw, że jeżeli $y \in F_A(x)$, to dla wszystkich $z \in N_A(x)$ mamy $\langle z\,,\, y\rangle \le 0$. Oznacza to, że $F_A(x) \subset (N_A(x))^o$. Z drugiej strony stożek $(N_A(x))^o$ jest, jako stożek polarny, zbiorem domkniętym (ćwiczenia), jeżeli więc zawiera $F_A(x)$, to zawiera i jego domknięcie $T_A(x)$ (bo domknięcie X jest najmniejszym zbiorem domkniętym zawierającym X – i każdy inny zbiór domknięty zawierający X musi cl$X$ zawierać). Mamy zatem $T_A(x) \subset (N_A(x))^o$.

Załóżmy teraz, że w $(N_A(x))^o$ jest punkt y nie należący do $T_A(x)$. Stożek $T_A(x)$ domknięty, możemy więc oddzielić go od y hiperpłaszczyzną – istnieje $v \in \mathbb{R}^n$ takie, że

$$\langle z\,,\, v\rangle < \langle y\,,\, v\rangle \text{ dla wszystkich } z \in T_A(x). \tag{4}$$

Wykażemy teraz, że dla wszystkich $z \in T_A(x)$ zachodzi $\langle z\,,\, v\rangle \le 0$. Załóżmy bowiem, że dla pewnego z_o jest $\langle z_o\,,\, v\rangle > 0$. Wówczas, dla dowolnego dodatniego λ, punkt λz_o też leży w $T_A(x)$, więc

$$\lambda\langle z_o\,,\, v\rangle = \langle \lambda z_o\,,\, v\rangle < \langle y\,,\, v\rangle,$$

a to nie może być prawda, bo biorąc λ duże możemy lewą stronę nierówności uczynić dowolnie dużą, a prawa od λ nie zależy.

Wiemy zatem, że dla wszystkich $z \in T_A(x)$ zachodzi $\langle z\,,\, v\rangle \le 0$. Oznacza to, że $v \in (T_A(x))^o = N_A(x)$. Z drugiej strony, kładąc w (4) $z = 0$ widzimy, że $\langle y\,,\, v\rangle > 0$. Oznacza to, że, wbrew założeniu, y nie może należeć do $(N_A(x))^o$, bo wektory z tego stożka mają niedodatnie iloczyny skalarne z wektorami z $N_A(x)$. □

Druga część udowodnionego przez nas twierdzenia jest szczególnym przypadkiem twierdzenia mówiącego, że dla dowolnego stożka wypukłego A zachodzi $(A^o)^o = \mathrm{cl}A$. Dowód wygląda w zasadzie tak samo, można zresztą uniknąć powtarzania go: jeżeli A jest stożkiem wypukłym, to $F_A(0) = A$ (natychmiast z definicji), zatem z powyższego twierdzenia mamy

$$\mathrm{cl}A = \mathrm{cl}F_A(0) = T_A(0) = (N_A(0))^o = ((F_A(0))^o)^o = (A^o)^o.$$

Zbiory wypukłe, z jakimi się stykamy w zagadnieniach optymalizacji, mają najczęściej postać

$$\{x \in \mathbb{R}^n \ : \ f_1(x) \le 0,\ f_2(x) \le 0,\ \ldots, f_k(x) \le 0\}.$$

Czasem, oprócz nierówności, pojawiają się również warunki typu $g(x) = 0$. Musimy nauczyć się, jak oblicza się stożki $F_A(x_o)$, $T_A(x_o)$ i $N_A(x_o)$ dla zbiorów tej postaci. Na początek zajmiemy się najprostszym przypadkiem, zbiorem opisanym jedną tylko nierównością $\{f(x) \leq 0\}$, przy założeniu, że f jest różniczkowalna w z i, aby zapewnić sobie wypukłość badanego zbioru, że jest quasiwypukła.

Twierdzenie 35. *Niech $A = \{x \in \mathbb{R}^n \ : \ f(x) \leq 0\}$, przy czym f jest funkcją quasiwypukłą i różniczkowalną w $x_o \in A$. Wówczas*

a) *jeżeli $f(x_o) < 0$, to $N_A(x_o) = \{0\}$,*

b) *jeżeli $f(x_o) = 0$, ale $\nabla f(x_o) \neq 0$, to $N_A(x_o) = \{\lambda \nabla f(x_o) \ : \ \lambda \geq 0\}$ (jest to stożek generowany przez wektor $\nabla f(x_o)$).*

Niestety, jeżeli zachodzi przypadek b) i $\nabla f(x_o) = 0$, nie umiemy tak łatwo (bez np. narysowania zbioru A) wyznaczyć stożka normalnego do A. Warto jednak pamiętać, że jeżeli funkcja f jest ściśle wypukła, to warunek $\nabla f(x_o)$ automatycznie oznacza, że f ma w x_o minimum globalne, a zbiór A składa się wyłącznie z punktu x_o. Wówczas $N_A(x_o) = \mathbb{R}^n$.

Dowód tego twierdzenia jest dość pracochłonny, zostawimy go na następny wykład. Znając stożek normalny możemy łatwo, korzystając z Twierdzenia 34, wyznaczyć stożek styczny – wystarczy wskazać stożek polarny do znalezionego $N_A(x_o)$.

Twierdzenie 36. *Niech $A = \{x \in \mathbb{R}^n \ : \ f(x) \leq 0\}$, przy czym f jest funkcją quasiwypukłą i różniczkowalną w $x_o \in A$. Wówczas*

a) *jeżeli $f(x_o) < 0$, to $T_A(x_o) = \mathbb{R}^n$,*

b) *jeżeli $f(x_o) = 0$, ale $\nabla f(x_o) \neq 0$, to $T_A(x_o) = \{v \in \mathbb{R}^n \ : \ \langle v \, , \, \nabla f(x_o) \rangle \leq 0\}$.*

Najtrudniej wskazać stożek kierunków osiągalnych – tu bez rysunku (albo ostatecznie dobrej wyobraźni) się nie obędzie, bo trzeba ustalić, które z wektorów stożka stycznego „wskazują” poza zbiór A (i je odrzucić).

Twierdzenie 37 (o postaci $F_A(x_o)$). *Niech $A = \{x \in \mathbb{R}^n \ : \ f(x) \leq 0\}$, $x_o \in \partial A$ i f niech będzie funkcją quasiwypukłą, dwukrotnie różniczkowalną w x_o. Jeżeli $\nabla f(x_o) \neq 0$, a $D^2 f(x_o)$ jest macierzą dodatnio określoną, to $F_A(x_o) = \{x \in \mathbb{R}^n \ : \ \langle x \, , \, \nabla f(x_o) \rangle < 0\} \cup \{0\}$.*

Dowód pominiemy, nie jest bardzo trudny, ale wymaga użycia rozwinięć Taylora – przypomina dowód warunków IIgo rzędu na wypukłość funkcji.

Wykład 7

Zaczniemy od obiecanego dowodu Twierdzenia 35.

Proof. Podpunkt a) jest dość łatwy. Jeżeli funkcja f jest różniczkowalna w x_o, to jest też w x_o ciągła, zatem warunek $f(x) < 0$ jest spełniony nie tylko w x_o, ale i we wszystkich punktach dostatecznie bliskich x_o – a więc wszystkie punkty dostatecznie bliskie x_o leżą w A (warunek $f(x) < 0$ opisuje zbiór otwarty, choć niekoniecznie równy wnętrzu zbioru A!). Oznacza to, że z x_o w dowolną stronę możemy wyprowadzić (być może bardzo króciutki) odcinek leżący w całości w A – a więc każdy wektor z $\mathbb{R}^n$ jest kierunkiem dopuszczalnym: $F_A(x_o) = \mathbb{R}^n$. Pozostaje tylko prościutkie ćwiczenie – sprawdzenie, że $(\mathbb{R}^n)^o = \{0\}$. To bardzo łatwe, proszę się nad tym zastanowić.

Przejdźmy do punktu b). Niech v będzie dowolnym wektorem dopuszczalnym w x_o ($v \in F_A(x_o)$). Wówczas, dla t dostatecznie małych (ale nieujemnych), $x_o + vt \in A$. W szczególności $f(x_o + vt) \le 0$, zatem

$$f(x_o + vt) - f(x_o) \le 0 \;\Rightarrow\; \frac{f(x_o + vt) - f(x_o)}{t} \le 0.$$

Przechodząc w ułamku z t do zera otrzymujemy, że $f'_v(x_o) \le 0$, ale $f'_v(x_o) = \langle \nabla f(x_o)\,,\, v\rangle \le 0$. Ta nierówność zachodzi dla dowolnego $v \in F_A(x_o)$, co oznacza, że $\nabla f(x_o) \in F_A(x_o)^o = N_A(x_o)$. Oczywiście wówczas dla każdego $\lambda \ge 0$ również i $\lambda\nabla f(x_o) \in N_A(x_o)$, bo $N_A(x_o)$ jest stożkiem. Pozostaje wykazać, że w $N_A(x_o)$ nic poza tym nie ma.

To wymaga trochę pracy. Założymy, że $w \in N_A(x_o)$ i wykażemy, że $w = \lambda\nabla f(x_o)$ dla jakiegoś $\lambda \ge 0$.

Oznaczmy $a = \frac{\langle w\,,\,\nabla f(x_o)\rangle}{\|\nabla f(x_o)\|^2}$. Jeżeli $\alpha > a$, to

$$\begin{aligned}\langle \nabla f(x_o)\,,\, w - \alpha\nabla f(x_o)\rangle &= \langle w\,,\,\nabla f(x_o)\rangle - \alpha\|\nabla f(x_o)\|^2 \\ &< \langle w\,,\,\nabla f(x_o)\rangle - \frac{\langle w\,,\,\nabla f(x_o)\rangle}{\|\nabla f(x_o)\|^2}\|\nabla f(x_o)\|^2 = 0,\end{aligned}$$

Oznacza to, że, jeżeli oznaczymy $v_\alpha = w - \alpha\nabla f(x_o)$, to $f'_{v_\alpha}(x_o) < 0$. Wiemy więc, że w kierunku w_α funkcja maleje, zatem $v_\alpha \in F_A(x_o)$.

Podobnie, jeżeli $\beta < a$, to

$$\begin{aligned}\langle \nabla f(x_o)\,,\, -w + \beta\nabla f(x_o)\rangle &= -\langle w\,,\,\nabla f(x_o)\rangle + \beta\|\nabla f(x_o)\|^2 \\ &< -\langle w\,,\,\nabla f(x_o)\rangle + \frac{\langle w\,,\,\nabla f(x_o)\rangle}{\|\nabla f(x_o)\|^2}\|\nabla f(x_o)\|^2 = 0,\end{aligned}$$

skąd $u_\beta = -w + \beta\nabla f(x_o) \in F_A(x_o)$. Wybierzmy teraz dwa ciągi: $\alpha_k \searrow a$, $\beta_k \nearrow a$. Dla każdego k mamy $v_{\alpha_k}, u_{\beta_k} \in F_A(x_o)$, więc

$$\langle w\,,\, v_{\alpha_k}\rangle \le 0 \qquad \langle w\,,\, u_{\beta_k}\rangle \le 0,$$

bo $w \in N_A(x_o)$. Rozpisując mamy

$$\begin{aligned}0 \ge &\langle w\,,\, v_{\alpha_k}\rangle = \|w\|^2 - \alpha_k\langle w\,,\,\nabla f(x_o)\rangle \\ 0 \ge &\langle w\,,\, u_{\beta_k}\rangle = -\|w\|^2 + \beta_k\langle w\,,\,\nabla f(x_o)\rangle\end{aligned}$$

Nierówności te będą prawdziwe, gdy przejdziemy z k do nieskończoności. Otrzymamy

$$\begin{aligned}0 \ge &\|w\|^2 - a\langle w\,,\,\nabla f(x_o)\rangle = \|w\|^2 - \frac{\langle w\,,\,\nabla f(x_o)\rangle^2}{\|\nabla f(x_o)\|^2}, \\ 0 \ge &-\|w\|^2 + a\langle w\,,\,\nabla f(x_o)\rangle = -\|w\|^2 + \frac{\langle w\,,\,\nabla f(x_o)\rangle^2}{\|\nabla f(x_o)\|^2}\end{aligned}$$

Prawe strony różnią się tylko znakiem - jeżeli więc obie mają być niedodatnie, to obie muszą być równe zero:

$$\|w\|^2\|\nabla f(x_o)\|^2 = \langle w\,,\,\nabla f(x_o)\rangle^2$$

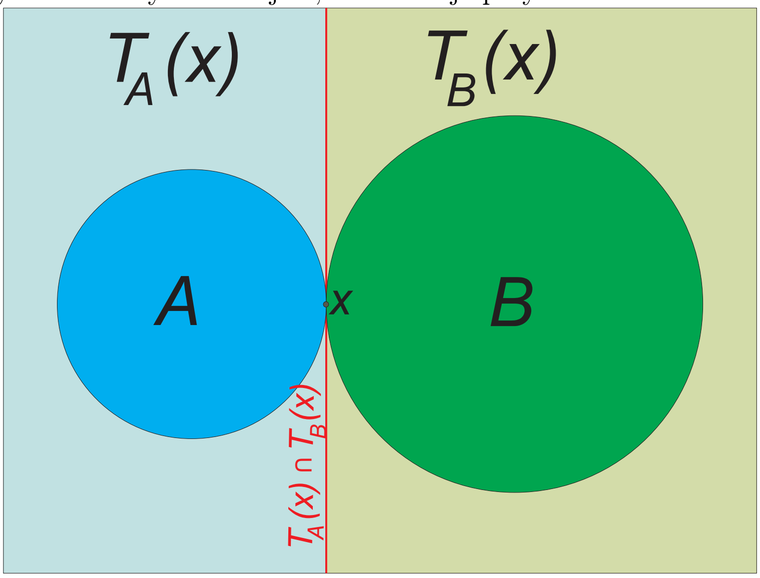
Z drugiej strony $\langle w\,,\,\nabla f(x_o)\rangle^2 = \|w\|^2\|\nabla f(x_o)\|^2\cos^2\measuredangle(w, \nabla f(x_o))$, zatem $\measuredangle(w, \nabla f(x_o)) = 0^o$ lub 180^o. Gdyby teraz $\measuredangle(w, \nabla f(x_o)) = 180^o$, to $w = -\lambda\nabla f(x_o)$ dla pewnego $\lambda > 0$. Wykazaliśmy wcześniej, że istnieją wektory $v \in F_A(x_o)$ takie, że $\langle \nabla f(x_o)\,,\, v\rangle < 0$ (np. wektor v_α zdefiniowany wcześniej). Wówczas jednak $\langle w\,,\, v\rangle = \langle -\lambda\nabla f(x_o)\,,\, v\rangle > 0$, co przeczy temu, że $w \in N_A(x_o)$. Stąd $\measuredangle(w, \nabla f(x_o)) = 0^o$, czyli $w = \lambda\nabla f(x_o)$ dla pewnego $\lambda \ge 0$. □

Zastanówmy się teraz przez chwilę, co możemy powiedzieć o stożkach osiągalnym, stycznym i normalnym do zbioru $A \cap B$, gdy znamy je dla zbiorów A i B?

Najprościej jest z $F_{A\cap B}$: nietrudno zauważyć, że kierunek osiągalny dla $A \cap B$ w x jest osiągalny zarówno dla A, jak i dla B. I odwrotnie - jeżeli w pewnym kierunku v możemy się poruszać z x tak, że nie opuścimy (przez pewien czas) ani A, ani B (czyli $v \in F_A(x) \cap F_B(x)$, to nie opuścimy w szczególności $A \cap B$, a więc jest to kierunek dopuszczalny w x dla $A \cap B$ ($v \in F_{A\cap B}$). Wykazaliśmy tym samym, że

$$F_{A\cap B}(x) = F_A(x) \cap F_B(x).$$

Wydawałoby się, że taki sam wzór powinien zachodzić dla stożków stycznych. Niestety tak nie jest, co ilustruje przykład.



W przykładzie $A\cap B = \{x\}$, więc oczywiście $T_{A\cap B}(x) = \{0\}$ (w x nie ma żadnych niezerowych kierunków osiągalnych dla $A \cap B$). Widzimy jednak, że $T_A(x) \cap T_B(x)$ jest całą prostą zawierającą x.

Okazuje się, że jeżeli tylko zbiory wypukłe A i B przecinają się "dostatecznie tłusto" — a więc jeżeli $A \cap intB \neq \varnothing$ lub $B \cap intA \neq \varnothing$, to

$$T_{A\cap B}(x) = T_A(x) \cap T_B(x).$$

Szczególny przypadek tego wzoru — gdy nasz zbiór jest opisany nierównościami na funkcjach quasiwypukłych, spełniających w x warunek jakości więzów, wykażemy za chwilę.

A jak to jest ze stożkami normalnymi? Zauważmy na początek, że jeżeli $v \in \mathbb{N}_A(x)$, $w \in N_B(x)$, to $v + w \in N_{A\cap B}(x)$ — wprost z definicji. Wiemy bowiem, że dla każdego $z \in F_{A\cap B}(x) = F_A(x) \cap F_B(x)$ zachodzi

$$\langle z\,,\,v\rangle \le 0, \qquad \langle z\,,\,w\rangle \le 0.$$

Gdy dodamy te dwie nierówności stronami, otrzymamy $\langle z\,,\,v + w\rangle \le 0$. Skoro nierówność ta zachodzi dla dowolnego $z \in F_{A\cap B}(x)$, to dowodzi ona, że $v + w \in F_{A\cap B}(x)^o = N_{A\cap B}(x)$.

Wykazaliśmy powyżej, że $N_A(x) + N_B(x) \subset N_{A\cap B}(x)$. Okazuje się, że

Twierdzenie 38. *Jeżeli $T_{A\cap B}(x) = T_A(x) \cap T_B(x)$ (np. gdy $A \cap intB \neq \varnothing$ lub $B \cap intA \neq \varnothing$), to*

$$N_{A\cap B}(x) = N_A(x) + N_B(x)$$

.

Proof. Dowód (z dokładnością do nieudowodnionego faktu o stożkach stycznych) jest dość prosty. Wiemy, że $N_A(x) + N_B(x) \subset N_{A\cap B}(x)$; załóżmy, że istnieje $y \neq 0$,

$y \in N_{A\cap B}(x) \setminus (N_A(x) + N_B(x))$. Istnieje wówczas hiperpłaszczyzna rozdzielająca y od $N_A(x) + N_B(x)$, to jest $w \in \mathbb{R}^n$ takie, że

$$\forall v \in N_A(x) + N_B(x) \quad \langle w\,,\,v\rangle \leq 0, \qquad \langle w\,,\,y\rangle > 0.$$

Wiemy też, że $N_A(x) \subset N_A(x) + N_B(x)$ oraz $N_B(x) \subset N_A(x) + N_B(x)$ (dlaczego?), a więc $\langle w\,,\,v\rangle \leq 0$ dla wszystkich $v \in N_A(x)$ oraz $N_B(x)$. Stąd $w \in N_A(x)^o = T_A(x)$ i, analogicznie, $w \in T_B(x)$. Dostajemy zatem $w \in T_A(x) \cap T_B(x) = T_{A\cap B}(x)$. To jednak jest sprzeczne z $\langle w\,,\,y\rangle > 0$, gdyż iloczyn skalarny wektora z $N_{A\cap B}(x)$ z wektorem z $T_{A\cap B}(x)$ jest niedodatni. □

Do rozważań o postaci stożków stycznych i normalnych do zbiorów opisanych nierównościami i równaniami wrócimy za kilka wykładów. Na razie przytoczę jeszcze dwa ważne, choć dość techniczne twierdzenia.

Twierdzenie 39 (Twierdzenie o alternatywie). *Rozważmy układ n równań liniowych $Ax = b$, gdzie A jest macierzą m na n, $b \in \mathbb{R}^m$. Zachodzi wówczas jedna i tylko jedna z poniższych możliwości.*

a) *$Ax = b$ ma rozwiązanie $x \in \mathbb{R}^n$ o wszystkich współrzędnych nieujemnych (w skrócie — nieujemne),*
b) *istnieje rozwiązanie $y \in \mathbb{R}^m$ układu nierówności $y^T A \geq 0$ [2], $\langle y\,,\,b\rangle < 0$.*

Proof. Załóżmy, że zachodzi a), a więc, że $Ax = b$ ma nieujemne rozwiązanie. Wówczas

$$\langle y\,,\,b\rangle = \langle y\,,\,Ax\rangle = y^T Ax = \langle y^T A\,,\,x\rangle$$

Gdyby teraz zachodziło b), to wektor $y^T A$ miałby wszystkie współrzędne nieujemne, podobnie jak x — a więc ich iloczyn skalarny byłby również nieujemny, co przeczy warunkowi $\langle y\,,\,b\rangle < 0$. wykazaliśmy zatem, że jeżeli zachodzi a), to nie zachodzi b).

Wykażemy teraz, że jeżeli nie zachodzi a), to musi zajść b).

Oznaczmy kolumny macierzy A przez $\mathrm{a}_1, \mathrm{a}_2, \ldots, \mathrm{a}_n \in \mathbb{R}^m$ i rozważmy stożek wypukły rozpięty przez te wektory.

$$X = \{x \in \mathbb{R}^n \ : \ x = \sum_{i=1}^{m} \lambda_i \mathrm{a}_i, \ \lambda_1, \ldots, \lambda_m \geq 0\}.$$

Fakt, że równanie $\mathrm{a}_1 x_1 + \mathrm{a}_2 x_2 + \cdots + \mathrm{a}_n x_n = Ax = b$ nie ma rozwiązań nieujemnych oznacza dokładnie tyle, że b leży poza stożkiem X (bo X składa się z wszystkich nieujemnych kombinacji wektorów $\mathrm{a}_1, \ldots, \mathrm{a}_n$). Możemy zatem skorzystać z Twierdzenia o oddzielaniu: istnieje wektor $y \in \mathbb{R}^m$ i stała $c \in \mathbb{R}$ taka, że

$$\forall x \in X \qquad \langle y\,,\,b\rangle < c \leq \langle y\,,\,x\rangle. \tag{5}$$

Gdy skorzystamy z faktu, że $0 \in X$ i wstawimy 0 w miejsce x, otrzymamy $\langle y\,,\,b\rangle < 0$. Pozostało wykazać, że $y^T A$ ma wszystkie współrzędne nieujemne. Wykażemy to przez zaprzeczenie.

Współrzędne $y^T A$ są postaci $\langle y\,,\,\mathrm{a}_i\rangle$. Załóżmy, że i-ta współrzędna $y^T A$ jest ujemna, a więc że $\langle y\,,\,\mathrm{a}_i\rangle < 0$. Wówczas wielkość $\langle y\,,\,\lambda \mathrm{a}_i\rangle$ jest, przy dostatecznie dużym λ, dowolnie mała, w szczególności możemy dobrać λ na tyle duże, by $\langle y\,,\,\lambda \mathrm{a}_i\rangle < \langle y\,,\,b\rangle$. Z drugiej strony a_i należy do stożka X, podobnie więc i $\lambda \mathrm{a}_i$ należy do X. Przeczy to jednak nierówności (5), która miała zachodzić dla wszystkich $x \in X$, w szczególności dla $\lambda \mathrm{a}_i$. □

Warto zapamiętać interpretację geometryczną tego twierdzenia: warunek a) jest spełniony wtedy i tylko wtedy, gdy wektor b leży w stożku wypukłym rozpiętym przez kolumny macierzy A, a więc przez wektory $\mathrm{a}_1, \mathrm{a}_2, \ldots, \mathrm{a}_n$. Warunek b) spełniony jest wtedy i tylko wtedy, gdy istnieje wektor y, który tworzy wraz z każdym z a_i kąt $\leq 90^o$, podczas gdy z b tworzy kąt rozwarty.

[2] po lewej stronie nierówności jest wektor; to, że ma być on ≥ 0 rozumiemy tak, że każda z jego współrzędnych jest nieujemna.

Wniosek 39.1. *Niech B będzie macierzą $m\times n$, a $K = \{x \in \mathbb{R}^n \ : \ Bx \leq 0$ (łatwo sprawdzić, że K jest stożkiem wypukłym). Wówczas*

$$K^o = \{y \in \mathbb{R}^n \ : \ \textit{istnieje nieujemne rozwiązanie równania } B^T x = y\}$$

Proof. Zastosujemy twierdzenie o alternatywie dla $A = B^T$ i $y = -z$: warunek b) ma postać:

istnieje rozwiązanie z układu nierówności $z^T B^T \leq 0$, $\langle z\, ,\, b\rangle > 0$.

Zauważmy jeszcze, że wektor $z^T B^T$ to transpozycja wektora Bz, więc warunek $z^T B^T \leq 0$ jest równoważny $Bz \leq 0$.

Jak wygląda zaprzeczenie warunku b)? Oczywiście ńie istnieje rozwiązanie...", czyli gdy dla każdego z spełniającego $Bz \leq 0$ mamy $sazb \leq 0$. Jest to jednak równoważne temu, że $b \in K^o$. Z drugiej strony, na mocy Twierdzenia o alternatywie, zaprzeczenie b) jest równoważne warunkowi a). W naszej sytuacji ma on postać

istnieje nieujemne rozwiązanie x układu równań $B^T x = b$.

Otrzymujemy zatem, że powyższa linijka jest warunkiem koniecznym i dostatecznym na to, by b leżało w K^o. □

Gyula vel Julius Farkas (1847–1930), węgierski matematyk i fizyk pochodzenia żydowskiego.

Innym wnioskiem z Twierdzenia o alternatywie jest ważne

Twierdzenie 40 (Lemat Farkasa). *Jak poprzednio, rozważmy układ n równań liniowych $Ax = b$, gdzie A jest macierzą m na n, $b \in \mathbb{R}^m$. Następujące warunki są równoważne:*

- *istnieje nieujemne (tj. o wszystkich współrzędnych nieujemnych) rozwiązanie równania $Ax = b$*
- *$\langle y\, ,\, b\rangle \geq 0$ dla wszystkich $y \in \mathbb{R}^m$ takich, że yA ma wszystkie współrzędnie nieujemne.*

Interpretacja geometryczna: pierwszy z warunków oznacza, że b leży w stożku X, a drugi — że jeżeli tylko y leży pod kątem $\leq 90^o$ do każdego z wektorów a_i ($\langle y\, ,\, \mathrm{a}_i\rangle \geq 0$), to również kąt miedzy y a b jest co najwyżej 90^o.

Proof. Jeżeli $Ax = b$ ma nieujemne rozwiązanie, to na mocy Twierdzenia o alternatywie układ nierówności $y^T A \geq 0$, $\langle y\, ,\, b\rangle < 0$ nie ma rozwiązań — a więc jeżeli tylko $y^T A \geq 0$, to $\langle y\, ,\, b\rangle \geq 0$. I odwrotnie - jeżeli ów układ nierówności nie ma rozwiązania (a więc, równoważnie, $(y^T A \geq 0) \Rightarrow (\langle y\, ,\, b\rangle \geq 0)$), to $Ax = b$ dla pewnego $x \geq 0$. □

Za kilka wykładów potrzebny nam będzie jeszcze jeden wniosek z Twierdzenia o alternatywie, dla kompletności podam go teraz:

Paul Albert Gordan (1837–1912) — urodzony (i wykształcony) we Wrocławiu wybitny algebraik niemiecki. Pamiętany dzięki swoim wynikom w teorii niezmienników oraz jako promotor doktoratu jednej z pierwszych wielkich matematyczek w historii - Emmy Noether

Twierdzenie 41 (Twierdzenie Gordana). *Niech M będzie macierzą $m\times n$. Zachodzi wówczas jedna i tylko jedna z poniższych możliwości:*

a) *istnieje rozwiązanie układu nierówności $Mx < 0$,*
b) *istnieje niezerowe rozwiązanie $y \in \mathbb{R}^m$ układu nierówności*

$$\begin{cases} y \geq 0 \ \textit{(tj. jako wektor ma wszystkie współrzędne nieujemne)} \\ y^T M = 0. \end{cases}$$

Proof. Na początek zauważmy, że układ $Mx < 0$ ma rozwiązanie wtedy i tylko wtedy, gdy rozwiązanie ma układ $Mx + s\vec{e}^T \leq 0$, $s > 0$, gdzie $\vec{e} = (1, 1, \dots, 1)$ (trywialne, ćwiczenia ?). Warunek a) możemy zatem zapisać następująco

ą) istnieje rozwiązanie układu nierówności

$$(M, \vec{e}) \begin{pmatrix} x \\ s \end{pmatrix} \leq 0, \quad (0, 0, \dots, 1) \begin{pmatrix} x \\ s \end{pmatrix} > 0.$$

Jeżeli przyjmiemy $M' = (M, \vec{e})^T$, $z = \begin{pmatrix} -x \\ -s \end{pmatrix}$, $b = (0, 0, \dots, 1)$, otrzymamy å) istnieje rozwiązanie układu nierówności

$$zM' \geq 0, \quad \langle z\,,\, b\rangle < 0.$$

Ten układ ma taką samą postać, jak układ b) Twierdzenia o alternatywie — wiemy zatem, że albo on ma rozwiązanie, albo ma rozwiązanie układ a) z Tw. o alternatywie:

ß) $M'y = b$ ma rozwiązanie $y \in \mathbb{R}^m$ o wszystkich współrzędnych nieujemnych

Ten ostatni warunek tłumaczy się natychmiast na

$\flat$) $y^T M = 0$, $y^T \vec{e} = 1$ i y ma wszystkie współrzędne nieujemne.

Jeżeli ten ostatni układ ma rozwiązanie, to oczywiście y jest niezerowe, bo $y^T \vec{e} = 1$, więc rozwiązanie ma układ b). Pozostaje nam tylko pokazać, że jest i odwrotnie - jeżeli rozwiązanie ma b), to ma je również $\flat$).

Skoro $y \neq 0$ i $y \geq 0$, to $\langle y\,,\, \vec{e}\rangle > 0$. Jeżeli teraz położymy $\tilde{y} = \frac{y}{\langle y\,,\, \vec{e}\rangle}$, będziemy mieli $\tilde{y}^T M = 0$, $\tilde{y}^T \vec{e} = 1$ i $\tilde{y} \geq 0$, a więc $\tilde{y}$ jest rozwiązaniem $\flat$). □

Wykłady 8 i 9

13. Subróżniczki i subgradienty funkcji wypukłych.

Definicja 19. Mówimy, że $a \in \mathbb{R}^n$ jest *subgradientem* f w punkcie $x_o \in A$, jeżeli dla każdego $x \in A$ zachodzi

$$f(x) \geq f(x_o) + \langle a\,,\, x - x_o\rangle$$

Przykład: niech $f(x) = |x|$, $A = \mathbb{R}$. Wówczas $a = 0$ jest subgradientem f w $x = 0$ (oczywiste). Subgradientem jest też $a = 1/2$, gdyż

$$\forall x \in \mathbb{R} \quad |x| \geq \frac{1}{2}x.$$

Przykład pokazuje, że funkcja może mieć w punkcie więcej, niż jeden subgradient.

Definicja 20. Zbiór wszystkich subgradientów funkcji f w punkcie $x_o \in A$ nazywamy *subróżniczką* f w x_o i oznaczamy $\partial f(x_o)$. Subróżniczki możemy definiować w ten sposób nie żądając, by funkcja f była wypukła, choć dla funkcji niewypukłych nie maja one zbyt wielkiego zastosowania.

Twierdzenie 42. *Subróżniczka funkcji wypukłej $f : A \to \mathbb{R}$ w $x_o \in A$ jest zbiorem wypukłym.*

Proof. Musimy wykazać, że jeżeli a i b są subgradientami f w x_o, czyli

$$f(x_o) + \langle a\,,\, x - x_o\rangle \leq f(x)$$
$$f(x_o) + \langle b\,,\, x - x_o\rangle \leq f(x),$$

zaś $p \in [0, 1]$, to $pa + (1 - p)b \in \partial f(x_o)$. Mamy

$$f(x_o)+\langle pa+(1-p)b\,,\, x-x_o\rangle = p(f(x_o)+\langle a\,,\, x-x_o\rangle)+(1-p)\langle b\,,\, x-x_o\rangle \leq pf(x)+(1-p)f(x) = f(x).$$

□

Twierdzenie 43. *Jeżeli funkcja f jest wypukła, to dla każdego $x_o \in intA$ istnieje subgradient f w x_o.*

Proof. Jak wiemy, nadwykres f, oznaczany epi$f$, jest wypukłym podzbiorem $\mathbb{R}^{n+1} = R^n \times \mathbb{R}$, a punkty wykresu f leżą na jego brzegu. Istnieje zatem hiperpłaszczyzna podpierająca epi$f$ w punkcie $(x_o, f(x_o))$, a więc taki niezerowy wektor $v = (\alpha, \beta) \in \mathbb{R}^n \times \mathbb{R}$, że

$$\forall (x, y) \in \text{epi} f \quad \langle (x, y)\,,\, v\rangle \leq \langle (x_o, f(x_o))\,,\, v\rangle,$$

czyli

$$\langle \alpha\,,\, (x - x_o)\rangle + \beta(y - f(x_o)) \leq 0 \tag{6}$$

Liczba β nie może być dodatnia, gdyż w przeciwnym wypadku wyrażenie $\beta(y - f(x_o))$ można by uczynić dowolnie dużym, biorąc dostatecznie duże y (a w epi$f$ y nie jest ograniczone), nierówność nie mogłaby więc zachodzić.

Gdyby teraz $\beta = 0$, to mielibyśmy $\langle \alpha\,, (x - x_o)\rangle \le 0$ dla wszystkich $x \in A$. Z drugiej strony, skoro $x_o \in intA$, to dla dostatecznie małego $\varepsilon > 0$ punkt $\bar{x} = x_o + \varepsilon\alpha$ też leży w A, więc

$$\langle \alpha\,, (\bar{x} - x_o)\rangle = \varepsilon\|\alpha\|^2 \le 0,$$

to zaś może mieć miejsce tylko wtedy, gdy $\alpha = 0$. To jednak przeczy temu, że v miał być wektorem niezerowym. Wiemy zatem, że $\beta < 0$. Dzieląc obie strony (6) przez $|\beta|$ otrzymujemy

$$\langle \alpha/|\beta|\,, (x - x_o)\rangle - y + f(x_o) \le 0,$$

czyli

$$y \ge f(x_o) + \langle \alpha/|\beta|\,, (x - x_o)\rangle \quad \text{dla wszystkich } (x, y) \in \text{epi}f,$$

Dla dowolnego $x \in A$ punkt $(x, f(x))$ leży na wykresie f, a więc $(x, f(x)) \in \text{epi}f$. Możemy więc w powyższej nierówności położyć $y = f(x)$, otrzymując, że $\alpha/|\beta|$ jest subgradientem f w x_o. □

Zachodzi też odwrotne twierdzenie, ale tylko wtedy, gdy A jest zbiorem otwartym:

Twierdzenie 44. *Niech $A \subset \mathbb{R}^n$ będzie otwarty i wypukły. Jeżeli funkcja f ma w każdym punkcie A subgradient, to jest wypukła.*

Proof. Dowód powtarza w zasadzie ideę dowodu warunków pierwszego rzędu dla wypukłości. Niech $\lambda \in [0, 1]$. Skoro zbiór A jest wypukły, to dla dowolnych $x, y \in A$ punkt $z = \lambda x + (1 - \lambda)y$ też należy do A. Z założenia w punkcie z funkcja ma subgradient, czyli wektor a taki, że $f(x) \ge f(z) + \langle a\,, x - z\rangle$ oraz $f(y) \ge f(z) + \langle a\,, y - z\rangle$. Mnożąc pierwszą z tych równości przez λ, drugą przez $1 - \lambda$, a następnie dodając stronami, otrzymujemy (dla dowolnego $\lambda \in [0, 1]$, $x, y \in A$)

$$\begin{aligned} \lambda f(x) + (1 - \lambda)f(y) &\ge \lambda f(z) + (1 - \lambda)f(z) + \langle a\,, \lambda(x - z) + (1 - \lambda)(y - z)\rangle \\ &= f(z) = f(\lambda x + (1 - \lambda)y), \end{aligned}$$

a więc definicję wypukłości f. □

Okazuje się, że jeżeli funkcja f jest wypukła i różniczkowalna w $x_o \in A$, to $\partial f(x_o)$ składa się tylko z jednego punktu: $\partial f(x_o) = \{\nabla f(x_o)\}$.

To, że gradient funkcji f w punkcie x_o jest jej subgradientem, jest jasne, jeżeli f jest różniczkowalna w całym A — mówią o tym warunki pierwszego rzędu na wypukłość. Wspomniany powyżej fakt jest znacznie mocniejszy, gdyż wymaga jedynie różniczkowalności w jednym punkcie.

Twierdzenie 45. *Niech A będzie wypukły i otwarty, zaś $f : A \to \mathbb{R}$ niech będzie wypukła. Jeżeli f jest różniczkowalna w $x_o \in A$, to $\partial f(x_o) = \{\nabla f(x_o)\}$.*

Proof. Ustalmy $v \in \mathbb{R}^n$ i przyjrzyjmy się funkcji $\phi(t) = f(x_o + tv)$. Funkcja ta jest różniczkowalna w $t = o$ (jako złożenie funkcji różniczkowanych), a jej rozwinięcie Taylora stopnia 1 ma postać

$$f(x_o + tv) = \phi(t) = \phi(0) + t\phi'(0) + o(t) = f(x_o) + t\langle \nabla f(x_o)\,, v\rangle + o(t)$$

Wiemy też, że w x_o istnieje jakiś subgradient f — oznaczmy go a.

$$f(x_o + tv) \ge f(x_o) + \langle a\,, x_o + tv - x_o\rangle = f(x_o) + t\langle a\,, v\rangle$$

Mamy zatem

$$f(x_o) + t\langle \nabla f(x_o)\,, v\rangle + o(t) \ge f(x_o) + t\langle a\,, v\rangle,$$

czyli, dla $t > 0$, $\langle \nabla f(x_o)\,, v\rangle + o(t)/t \ge \langle a\,, v\rangle$. Pamiętając, że $o(t)/t \xrightarrow{t\to 0} 0$ możemy przejść z t do 0, otrzymując $\langle \nabla f(x_o)\,, v\rangle \ge \langle a\,, v\rangle$, czyli $\langle a - \nabla f(x_o)\,, v\rangle \le 0$. Nierówność ta zachodzi dla dowolnego wektora $v \in \mathbb{R}^n$, możemy więc przyjąć $v = a - \nabla f(x_o)$, co da $\|a - \nabla f(x_o)\| \leq 0$, a więc $a = \nabla f(x_o)$. □

Możemy teraz sformułować i udowodnić wersję twierdzenia Fermata dla funkcji wypukłych, nieróżniczkowalnych:

TWIERDZENIE 46 (Warunek pierwszego rzędu na minimum bez ograniczeń). *Niech $f : \mathbb{R}^n \to \mathbb{R}$ będzie funkcją wypukłą. Następujące warunki są równoważne:*

a) $0 \in \partial f(x_o)$,
b) *f przyjmuje w x_o minimum globalne.*

PROOF. Jeżeli $0 \notin \partial f(x_o)$, to

$$\text{nieprawda, że } (\forall x \in \mathbb{R}^n \quad f(x) \geq f(x_o) + \langle 0\,,\, x - x + o\rangle = f(x_o))\,,$$

czyli

$$\exists x \in \mathbb{R}^n \quad f(x) < f(x_o).$$

Tym samym warunek a) jest warunkiem koniecznym na to, by w x_o było minimum f.

Jeżeli zaś w $0 \in \partial f(x_o)$, to z definicji

$$\forall x \in \mathbb{R}^n \quad f(x) \geq f(x_o) + \langle 0\,,\, x - x_o\rangle = f(x_o),$$

a więc f przyjmuje w x_o minimum. □

Wiemy oczywiście, że nie każda funkcja wypukła jest różniczkowalna – znamy mnóstwo prostych przykładów (np $f(x) = |x|$). Okazuje się jednak, że funkcje wypukłe nie mogą być zbyt paskudne:

- funkcja wypukła spełnia warunek Lipschitza,
- funkcja wypukła jest różniczkowalna „prawie wszędzie" (zbiór punktów, w których jest nieróżniczkowalna nie może być za duży, cokolwiek miało by to znaczyć),
- funkcja wypukła jest dwukrotnie różniczkowalna „prawie wszędzie" (j.w.),

Dowody tych faktów (a przynajmniej pierwszych dwóch) nie są trudne, ale nie będą nam potrzebne. Dla nas ważniejsze będzie inne twierdzenie tego typu:

TWIERDZENIE 47. *Funkcja wypukła $f : A \subset \mathbb{R}^n \to \mathbb{R}$ ma, dla każdego punktu $x_o \in A$ i wektora $v \in F_A(x_o)$, pochodną kierunkową $f'_v(x_o)$. W szczególności, gdy x_o leży we wnętrzu A, funkcja f ma pochodną kierunkową w x_o dla dowolnego wektora v.*

SZKIC. Trzeba wykazać, że dla każdego $v \in F_A(x_o)$ istnieje granica

$$f'_v(x_o) = \lim_{\tau \to 0^+} \frac{f(x_o + \tau v) - f(x_o)}{\tau}.$$

Oznaczmy $I(\tau) = \frac{f(x_o+\tau v)-f(x_o)}{\tau}$ (ma to sens dla τ dostatecznie małych). Weźmy $0 < \sigma < \tau$. Możemy teraz napisać

$$x_o + \sigma v = (1 - \frac{\sigma}{\tau})x_o + \frac{\sigma}{\tau}(x_o + \tau v)$$

(zapisaliśmy $x_o + \sigma v$ jako kombinację wypukłą x_o i $(x_o + \tau v)$); z wypukłości f mamy więc

$$f(x_o + \sigma v) \leq (1 - \frac{\sigma}{\tau})f(x_o) + \frac{\sigma}{\tau}f(x_o + \tau v).$$

Chwila przekształceń prowadzi do nierówności $I(\sigma) \leq I(\tau)$ – a zatem funkcja $I(\cdot)$ jest rosnąca (a więc maleje, gdy maleje jej argument). Chwili zachodu wymaga jeszcze wykazanie, że jest dla $\tau > 0$ ograniczona z dołu (przez jeszcze inny iloraz różnicowy) – ten fragment pominiemy – i tak dowodzimy, że istnieje granica. □

Z tego szkicu dowodu wynika wart zapamiętania wniosek – wersja warunku Igo rzędu na wypukłość:

Wniosek 47.1. *Dla dowolnych $x, x_o \in A$ zachodzi nierówność*

$$f(x) \geq f(x_o) + f'_{x-x_o}(x_o).$$

Oczywiście, gdy f jest w x_o różniczkowalna, otrzymujemy znany nam warunek Igo rzędu na wypukłość f. Gdy jednak f nie jest w x_o różniczkowalna, definiujemy $I(\tau) = \frac{f(x_o+\tau(x-x_o))-f(x_o)}{\tau}$ i tak samo jak wyżej dowodzimy, że jest to rosnąca funkcja zmiennej τ. W szczególności $I(1) \geq \lim_{\tau\to 0^+} I(\tau) = f'_{x-x_o}(x_o)$. Po trywialnych przekształceniach z nierówności tej dostajemy tezę.

Pochodne kierunkowe funkcji wypukłej mają bliski związek z jej subróżniczką:

Twierdzenie 48. *Niech $f : A \subset \mathbb{R}^n \to \mathbb{R}$ będzie funkcją wypukłą, A otwarty, $x_o \in A$. Wektor g jest subgradientem f w x_o wtedy i tylko wtedy, gdy*

$$\forall v \in \mathbb{R}^n \quad f'_v(x_o) \geq \langle g \, , \, v\rangle. \tag{7}$$

Proof. Niech wektor g spełnia (7). Na mocy Wniosku 47.1

$$f(x) \geq f(x_o) + f'_{x-x_o}(x_o) \geq f(x_o) + \langle g \, , \, x - x_o\rangle,$$

co dowodzi, że g jest subgradientem f w x_o.

Jeżeli g jest subgradientem f w x_o, to z definicji, dla każdego $v \in \mathbb{R}^n$ i $\tau > 0$,

$$f(x_o + \tau v) \geq f(x_o) + \langle g \, , \, \tau v\rangle,$$

czyli

$$\frac{f(x) - f(x_o)}{\tau v} \geq \frac{\langle g \, , \, \tau v\rangle}{\tau} = \langle g \, , \, v\rangle.$$

Biorąc $\tau \to 0^+$ dostajemy tezę. □

Wniosek 48.1. *Niech f, A, x_o będą jak w poprzednim twierdzeniu. Wówczas dla każdego wektora $v \in \mathbb{R}^n$*

$$\sup_{g\in\partial f(x_o)} \langle g \, , \, v\rangle = f'_v(x_o).$$

Wniosek 48.2. *Niech f, A, x_o będą jak w poprzednim twierdzeniu. Wówczas $\partial f(x_o)$ jest zbiorem domkniętym, wypukłym i ograniczonym.*

Proof. Wypukłość subróżniczki dowiedliśmy już wcześniej. Jej domkniętość wynika stąd, że warunek na subgradient jest zadany przez nieostrą nierówność, jeżeli więc mamy zbieżny do g ciąg subgradientów g_i w x_o, to stąd, że dla każdego $x \in A$

$$f(x) - f(x_o) \geq \langle g_i \, , \, x - x_o\rangle$$

możemy przechodząc z i do ∞ wywnioskować, że taka sama nierówność zachodzi dla g w miejsce g_i – a więc że g też jest subgradientem. Pozostaje ograniczoność $\partial f(x_o)$.

Na początek zauważmy, że w punkcie x_o wyrażenie $M = \sup_{|v|=1} f'_v(x_o)$ jest skończone. Wynika to stąd, że

- z otwartości A istnieje $\varepsilon > 0$ takie, że punkty x spełniające $|x - x_o| = \varepsilon$ (a więc punkty postaci $x_o + \varepsilon v$, gdzie $|v| = 1$) leżą w A;
- istnieje zatem (dzięki zwartości sfery $\{v \; : \; |v| = 1\}$) skończone supremum ilorazów różnicowych

$$M = \sup_{|v|=1} I_v(\varepsilon), \text{ gdzie } I_v(\varepsilon) = \frac{f(x_o + \varepsilon v) - f(x_o)}{\varepsilon},$$

- z dowodu Twierdzenia 47 wiemy, że dla każdego v zachodzi $f'_v(x_o) \leq I_v(\varepsilon)$.

Wiemy jednak z Twierdzenia 48, że dla każdego $g \in \partial f(x_o)$ i $v \in \mathbb{R}^n$ zachodzi $\langle g\,,\, v\rangle \leq f'_v(x_o)$. W szczególności, kładąc $v = g/|g|$ (oczywiście dla $g \neq 0$), mamy

$$|g| = \langle g\,,\, \frac{g}{|g|}\rangle \leq f'_{g/|g|}(x_o) \leq M.$$

□

Twierdzenie 49 (O postaci subróżniczki dla maksimum dwóch funkcji wypukłych). *Niech funkcja f określona na zbiorze otwartym i wypukłym A będzie zadana wzorem*

$$f(x) = \max\{g(x), h(x)\}.$$

Wówczas

$$\partial f(x) = \begin{cases} \partial g(x) & \textit{gdy } g(x) > h(x) \\ \partial h(x) & \textit{gdy } h(x) > g(x) \\ \mathrm{conv}(\partial h(x) \cup \partial g(x)) & \textit{gdy } g(x) = h(x). \end{cases}$$

Przez $\mathrm{conv}(A)$ oznaczam najmniejszy zbiór wypukły zawierający zbiór A – nazywam go **uwypukleniem** lub **powłoką wypukłą** zbioru A. Jest on równy

$$\mathrm{conv}A = \{\lambda_1 x_1 + \cdots + \lambda_k x_k \;:\; k \in \mathbb{N},\, \lambda_i \geq 0,\, \sum \lambda_i = 1,\, x_i \in A\},$$

choć taka definicja jest mało użyteczna w praktyce; z reguły łatwo zgadnąć, jak taki zbiór wygląda. Dla przykładu – uwypuklenie zbioru złożonego z 3 niewspółliniowych punktów na płaszczyźnie to trójkąt, którego punkty te są wierzchołkami. Można wykazać, że uwypuklenie zwartego podzbioru $\mathbb{R}^n$ jest zbiorem zwartym.

Proof. Funkcje f, g i h, jako funkcje wypukłe na zbiorze otwartym, są ciągłe. Dlatego też zbiór $\{x \;:\; g(x) > h(x)\}$, podobnie jak $\{x \;:\; h(x) > g(x)\}$, jest zbiorem otwartym; na pierwszym z nich $f \equiv g$, na drugim – $f \equiv h$, jest więc oczywiste, że pokrywają się również ich subróżniczki. Pozostaje nam zbadać postać ∂f na zbiorze $K = \{x \;:\; g(x) = h(x)\}$.

Oczywiście dla dowolnych $x, x_o \in A$ zachodzą nierówności $f(x) - f(x_o) \geq g(x) - f(x_o)$ oraz $f(x) - f(x_o) \geq h(x) - f(x_o)$. Jeżeli teraz $x_o \in K$, mamy

$$f(x) - f(x_o) \geq g(x) - f(x_o) = g(x) - g(x_o) \geq \langle v\,,\, x - x_o\rangle \quad \forall v \in \partial g(x_o)$$
$$f(x) - f(x_o) \geq h(x) - f(x_o) = h(x) - h(x_o) \geq \langle w\,,\, x - x_o\rangle \quad \forall v \in \partial h(x_o).$$

Oznacza to, że subgradienty zarówno g, jak i h w x_o są subgradientami funkcji f w x_o, a więc $\partial g(x_o) \cup \partial h(x_o) \subset \partial f(x_o)$.

Subróżniczka f w x_o jest zbiorem wypukłym, a zatem zawierając $\partial g(x_o) \cup \partial h(x_o)$ musi zawierać też uwypuklenie tego zbioru.

Załóżmy teraz, że w $\partial f(x_o)$ istnieje subgradient u nienależący do $B = \mathrm{conv}(\partial g(x_o) \cup \partial h(x_o))$. Możemy zatem oddzielić u od zbioru wypukłego B (jest to zbiór zwarty, gdyż $\partial g(x_o) \cup \partial h(x_o)$ jest zwarty): istnieje stała c i wektor v taki, że dla dowolnego p z B mamy $\langle u\,,\, v\rangle > c > \langle p\,,\, v\rangle$. Wiemy zatem, że dla ε dostatecznie małych, by $x_o + \varepsilon v$ leżało w A,

$$\begin{aligned} \frac{f(x_o + \varepsilon v) - f(x_o)}{\varepsilon} &\geq \frac{1}{\varepsilon}\langle u\,,\, \varepsilon v\rangle = \langle u\,,\, v\rangle > c > \max_{p \in B}\langle p\,,\, v\rangle \\ &\geq \max\{\sup_{p \in \partial g(x_o)}\langle p\,,\, v\rangle,\, \sup_{p \in \partial h(x_o)}\langle p\,,\, v\rangle\} = \max\{g'_v(x_o), h'_v(x_o)\} \end{aligned} \tag{8}$$

(maksimum w pierwszej linijce istnieje, gdyż zbiór B jest zwarty). Z drugiej strony, z definicji f,

$$\max\{\frac{g(x_o + \varepsilon v) - g(x_o)}{\varepsilon}, \frac{h(x_o + \varepsilon v) - h(x_o)}{\varepsilon}\}\frac{f(x_o + \varepsilon v) - f(x_o)}{\varepsilon}.$$

Przechodząc z ε do zera dostajemy $\max\{g'_v(x_o), h'_v(x_o)\} \geq f'_v(x_o)$, podczas gdy to samo przejście w (8) daje $f'_v(x_o) \geq c > \max\{g'_v(x_o), h'_v(x_o)\}$. Dostaliśmy w ten sposób sprzeczność, co dowodzi, że nie ma u leżącego w $\partial f(x_o)$ i poza B. □

Dotąd rozważaliśmy jedynie „prawdziwe" funkcje wypukłe – przyjmujące jedynie wartości liczbowe. Z różnych względów w analizie wypukłej rozważa się również funkcje wypukłe przyjmujące oprócz wartości skończonych również wartość $+\infty$. Najważniejszymi przykładami takich funkcji są tzw. funkcje indykatorowe: dla domkniętego zbioru wypukłego $A \subset \mathbb{R}^n$ definiuję

$$I_A(x) = \begin{cases} 0 & \text{gdy } x \in A \\ +\infty & \text{gdy } x \notin A. \end{cases}$$

Funkcję tę nazywamy *funkcją indykatorową* zbioru A. Łatwo możemy sprawdzić, że tak zdefiniowana funkcja jest wypukła na $\mathbb{R}^n$: by tak było, ma zachodzić

$$\forall x, y \in \mathbb{R}^n,\ \forall \alpha \in [0,1] \qquad f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y).$$

Jeżeli którykolwiek z punktów x, y leży poza A, to prawa strona nierówności jest nieskończona (i nierówność jest prawdziwa, niezależnie od wartości lewej strony). Jeżeli natomiast x i y leżą w A, to również $\alpha x + (1-\alpha)y$ leży w A i obie strony mają wartość 0.

Jaka jest subróżniczka I_A w punkcie $x_o \in A$? Aby g było subgradientem I_A w x_o, musi zachodzić nierówność

$$\forall y \in \mathbb{R}^n \quad I_A(y) = I_A(y) - I_A(x_o) \geq \langle g\,,\, y - x_o\rangle.$$

Jeżeli $y \notin A$, to lewa strona nierówności jest nieskończona, nierówność jest więc spełniona dla dowolnego g. Jakieś warunki na g możemy zatem dostać tylko biorąc $y \in A$. Wtedy nierówność przyjmuje postać

$$\forall y \in A \ 0 \geq \langle g\,,\, y - x_o\rangle,$$

co jest równoważne

$$\forall z \in A - x_o \ 0 \geq \langle g\,,\, z\rangle.$$

Jeżeli powyższa nierówność zachodzi dla pewnego z, to jest prawdziwa również dla λz w miejsce z dla dowolnego $\lambda \geq 0$. Oznacza to, że powyższy warunek jest równoważny

$$\forall z \in \operatorname{cone}(A - x_o) = F_A(x_o) \ 0 \geq \langle g\,,\, z\rangle,$$

a więc że $g \in (F_A(x_o))^o = N_A(x_o)$. Tak więc subróżniczka $\partial I_A(x_o)$ jest równa stożkowi normalnemu zbioru A w x_o – w szczególności na brzegu A może być nieograniczona. Zauważmy, że wcześniej dowiedliśmy, że subróżniczka funkcji wypukłej określonej na zbiorze *otwartym* jest ograniczona.

Jean Jacques Moreau (1923-) – wybitny francuski mechanik teoretyczny i matematyk, specjalista od mechaniki niegładkiej (nieróżniczkowalnej).

R.(alph) Tyrrell Rockafellar (1935-) – matematyk amerykański, jeden z twórców współczesnej optymalizacji wypukłej

Na koniec podam, bez dowodu, jeszcze jedno ważne twierdzenie, ułatwiające obliczanie subróżniczek funkcji wypukłych:

TWIERDZENIE 50 (Moreau-Rockafellara). *Niech funkcje f i g będą funkcjami wypukłymi, przyjmującymi wartości skończone, określonymi na niepustym zbiorze otwartym wypukłym A. Wówczas dla każdego $x \in A$ zachodzi równość $\partial(f+g)(x) = \partial f(x) + \partial g(x)$, przy czym suma po prawej stronie jest sumą Minkowskiego zbiorów.*

Twierdzenie to jest prawdziwe również dla niekoniecznie otwartych zbiorów A, i dla funkcji przyjmujących wartości nieskończone, jednak musimy wówczas zmienić nieco założenia: x jest punktem, w którym obie funkcje przyjmują wartości skończone, a funkcja g ma przynajmniej jeden punkt, w którym jest skończona i ciągła (nie ma tej własności np. funkcja indykatorowa prostej w $\mathbb{R}^2$).

1. Warunki Karusha-Kuhna-Tackera

1. Optymalizacja z warunkami postaci $g(x) \leq 0$.

Nim zbadamy, jak szukać ekstremum funkcji $f(x)$ na zbiorze A zadanym przy pomocy nierówności (więzu) $g(x) \leq 0$, przypomnijmy sobie argumentację do twierdzenia o mnożnikach Lagrange'a. Dla uproszczenia załóżmy, że mamy do czynienia z tylko jednym warunkiem $g(x) = 0$. Wspomniane twierdzenie mówi, że w punkcie ekstremalnym x_o, spełniającym $g(x_o) = 0$, mamy $\nabla f(x) = \lambda \nabla g(x_o)$. Przypomnijmy sobie teraz, że $\nabla g(x_o)$ wskazuje kierunek, w którym funkcja g najszybciej rośnie – a więc na zewnątrz zbioru $\{g(x) \leq 0\}$. Jeżeli teraz $\lambda < 0$, to $\nabla f(x_o)$ jest skierowany do wewnątrz zbioru $A = \{g(x_o) \leq 0\}$ i $f(x)$ rośnie, gdy z x_o przesuwamy się w głąb zbioru A. Nawet więc jeżeli na zbiorze $\{g(x) = 0\}$ w x_o funkcja f osiąga maksimum, to jeżeli mamy do dyspozycji cały zbiór A, funkcja f może przyjmować większe wartości niż w x_o. Inaczej jest, gdy $\lambda > 0$ - wówczas $\nabla f(x_o)$ jest skierowany na zewnątrz zbioru A i funkcja f maleje, gdy z x_o ruszamy wgłąb A.

Jeżeli maksimum $f(x)$ jest osiągane w punkcie $x_o \in A$ i x_o leży na brzegu A, a więc $g(x_o) = 0$, mówimy, że wiąz ten jest w x_o *aktywny*. Z powyższej argumentacji widzimy, że w tym przypadku musi zachodzić

$$\nabla f(x_o) - \lambda \nabla g(x_o) = 0, \ g(x_o) = 0 \text{ oraz } \lambda > 0$$

Odwrotnie jest, gdy poszukujemy minimum funkcji $f(x)$ na A — na brzegu A, czyli na zbiorze $\{g(x) = 0\}$, szukamy tych punktów x_o, dla których spełnione jest twierdzenie Lagrange'a *oraz* $\lambda < 0$ — tylko w nich wartości funkcji nie będą większe od tych osiąganych punktach bliskich x_o, ale leżących we wnętrzu A.

Jeżeli natomiast punkt x_o leży we wnętrzu zbioru A i funkcja $f(x)$ przyjmuje w nim maksimum (czy też minimum), to x_o jest po prostu ekstremum lokalnym funkcji f, a więc $\nabla f(x_o) = 0$. Możemy to oczywiście zapisać następująco:

$$\nabla f(x_o) - \lambda \nabla g(x_o) = 0, \ g(x_o) \leq 0 \text{ oraz } \lambda = 0.$$

Ogólnie więc możemy przedstawić następujące warunki konieczne na to, by punkt x_o był rozwiązaniem zagadnienia

$\boxed{\text{znajdź argmax} f(x) \text{ dla } x \in A = \{g(x) \leq 0,\}}$

oczywiście przy nieustającym założeniu, że funkcje f i g są różniczkowalne w sposób ciągły:

$$\nabla f(x_o) - \lambda \nabla g(x_o) = 0, \ g(x_o) \leq 0, \ \lambda \geq 0 \text{ oraz } \lambda g(x_o) = 0.$$

Uwaga: na brzegu A korzystamy z twierdzenia o mnożnikach Lagrange'a — a więc dotyczą nas wszystkie zastrzeżenia związane z tym twierdzeniem.

W szczególności jeżeli $g(x_o) = 0$, to podany powyżej warunek jest w mocy, o ile $\nabla g(x_o) \neq 0$.

Jak powyższe „twierdzenie” uogólnia się do większej liczby warunków? Otóż zachodzi

William Karush (1917–1997) udowodnił to twierdzenie w swojej pracy magisterskiej w latach 30tych. Ponownie odkryli je i opublikowali w 1951 Harold Kuhn (ur. 1925) i Albert Tucker (1905–1995), wybitni amerykańscy specjaliści od teorii gier i programowania nieliniowego.

TWIERDZENIE 51 (warunki konieczne Karusha-Kuhna-Tuckera (KKT) dla maksimum przy ograniczeniach nierównościowych). *Niech funkcje f oraz g_i, $i = 1, \dots, k$, będą różniczkowalne, i niech punkt $x_o \in \mathbb{R}^n$ będzie rozwiązaniem zagadnienia*

$$\boxed{\textit{znajdź } \operatorname{argmax} f \textit{ na zbiorze } A = \{x \in \mathbb{R}^n \ : \ g_i(x) \leq 0, \ i = 1, \dots, n\}.}$$

Jeżeli gradienty tych więzów, które są aktywne *w x_o (a więc tych g_i, dla których $g_i(x_o) = 0$), są w x_o liniowo niezależne, to istnieją mnożniki Lagrange'a $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ takie, że spełnione są warunki*

- $\nabla f(x_o) - \sum_{i=1}^{k} \lambda_i \nabla g_i(x_o) = 0$
- $\lambda_i \geq 0$, $g_i(x_o) \leq 0$ *oraz* $\lambda_i g_i(x_o) = 0$ *dla wszystkich* $i = 1, \dots, k$.

Pierwszy z warunków można wyrazić tworząc funkcję Lagrange'a, analogicznie jak w twierdzeniu o mnożnikach Lagrange'a:

$$L(x) = f(x) - \sum_{i=1}^{k} \lambda_i g_i(x),$$

i żądając, by $\nabla L(x_o) = 0$.

Jak wyglądają warunki KKT, gdy poszukujemy nie $\operatorname{argmax} f$, ale $\operatorname{argmin} f$? Wystarczy pamiętać, że jeżeli f osiąga minimum w punkcie x_o, to $-f$ ma w tym samym punkcie maksimum. Okazuje się jednak, że zmiana znaku funkcji f ma dokładnie ten sam skutek, co zostawienie f bez zmian, a zastąpienie wszystkich λ_i przez $-\lambda_i$ (dlaczego?). Dostajemy wówczas analogiczne twierdzenie

TWIERDZENIE 52 (warunki konieczne KKT dla minimum przy ograniczeniach nierównościowych). *Niech funkcje f oraz g_i, $i = 1, \dots, k$, będą różniczkowalne, i niech punkt $x_o \in \mathbb{R}^n$ będzie rozwiązaniem zagadnienia*

$$\boxed{\textit{znajdź } \operatorname{argmin} f \textit{ na zbiorze } A = \{x \in \mathbb{R}^n \ : \ g_i(x) \leq 0, \ i = 1, \dots, n\}.}$$

Jeżeli gradienty tych więzów, które są aktywne *w x_o (a więc tych g_i, dla których $g_i(x_o) = 0$), są w x_o liniowo niezależne, to istnieją mnożniki Lagrange'a $\lambda_1, \dots, \lambda_k \in \mathbb{R}$ takie, że spełnione są warunki*

- $\nabla f(x_o) - \sum_{i=1}^{k} \lambda_i \nabla g_i(x_o) = 0$
- $\lambda_i \leq 0$, $g_i(x_o) \leq 0$ *oraz* $\lambda_i g_i(x_o) = 0$ *dla wszystkich* $i = 1, \dots, k$.

To nie jest jeszcze pełna postać warunków koniecznych KKT. Widać jednak, że twierdzenie to nie wykracza daleko poza to, czego nauczyli się Państwo w pierwszym semestrze — mnożniki Lagrange'a, punkty krytyczne i troszkę argumentów geometrycznych. Warunki te są ważne i znane dlatego, że przy pewnych założeniach na funkcje f i g_i są one warunkami dostatecznymi istnienia optimum.

2. Ogólna postać warunków koniecznych KKT. Na poprzednim wykładzie omówiliśmy warunki konieczne pierwszego rzędu, jakie musi spełniać punkt, w którym osiągane jest lokalne ekstremum związane funkcji na zbiorze opisanym skończoną liczbą nierówności $g_i(x) \leq 0$, $i = 1, \ldots, k$. Wcześniej przypomnieliśmy sobie metodę mnożników Lagrange'a, dającą warunki konieczne istnienia ekstremum lokalnego, gdy zbiór opisany jest pewną liczbą równości $h_j(x) = 0$, $j = 1, \ldots, l$. Najwyższy czas połączyć te warunki i przedstawić ogólną postać warunków Karusha-Kuhna-Tuckera, jakie muszą być spełnione w punkcie, w którym funkcja $f(x)$ przyjmuje lokalne ekstremum związane na zbiorze opisanym zarówno równościami, jak i nierównościami.

TWIERDZENIE 53 (warunki konieczne Karusha-Kuhna-Tuckera (KKT) dla maksimum). *Niech funkcje f, g_i, $i = 1, \ldots, k$ oraz h_j, $j = 1, \ldots, l$ będą różniczkowalne, i niech punkt $x_o \in \mathbb{R}^n$ będzie rozwiązaniem zagadnienia*

$$\boxed{\begin{array}{c} \textit{znajdź } \mathrm{argmax} f \textit{ na zbiorze} \\ A = \{x \in \mathbb{R}^n \ : \ g_i(x) \leq 0, \ h_j(x) = 0, \ i = 1, \ldots, k, \ j = 1, \ldots, l\} \\ . \end{array}}$$

Jeżeli gradienty tych więzów, które są aktywne *w x_o (a więc wszystkich h_j i tych g_i, dla których $g_i(x_o) = 0$), są w x_o liniowo niezależne, to istnieją mnożniki Lagrange'a $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ i $\mu_1, \ldots, \mu_l$ takie, że spełnione są warunki*

- $\nabla f(x_o) - \sum_{i=1}^{k} \lambda_i \nabla g_i(x_o) - \sum_{j=1}^{l} \mu_j \nabla h_j(x_o) = 0,$

(a więc x_o jest punktem krytycznym funkcji Lagrange'a

$$L(x) = f(x) - \sum_{i=1}^{k} \lambda_i g_i(x) - \sum_{j=1}^{l} \mu_j h_j(x) \ )$$

- $h_j(x_o) = 0$ *dla wszystkich* $j = 1 \ldots, l$,
- $\lambda_i \geq 0$, $g_i(x_o) \leq 0$ *oraz* $\lambda_i g_i(x_o) = 0$ *dla wszystkich* $i = 1, \ldots, k$.

Gdy szukamy $\mathrm{argmin} f$, zmienia się funkcja Lagrange'a:

$$L(x) = f(x) + \sum_{i=1}^{k} \lambda_i g_i(x) - \sum_{j=1}^{l} \mu_j h_j(x),$$

zatem pierwszy z warunków przyjmuje postać

- $\nabla f(x_o) + \sum_{i=1}^{k} \lambda_i \nabla g_i(x_o) - \sum_{j=1}^{l} \mu_j \nabla h_j(x_o) = 0,$

a pozostałe nie zmienią się (to, czy przed drugą sumą damy plus, czy minus, nie ma znaczenia).

Warto zauważyć (i zapamiętać), że na mnożniki λ_i, związane z warunkami nierównościowymi, nakładamy warunek $\lambda_i \geq 0$, podczas gdy na μ_j nie nakładamy żadnych ograniczeń.

3. Warunki dostateczne KKT. Poniższe twierdzenie pokazuje, że jeżeli funkcje f, g_i i h_j występujące w twierdzeniu KKT mają pewne szczególne własności, to warunki konieczne podane w tym twierdzeniu są również warunkami dostatecznymi.

TWIERDZENIE 54 (Warunki dostateczne Karusha-Kuhna-Tuckera dla maksimum).
Załóżmy, że w punkcie $x_o \in \mathbb{R}^n$ spełnione są warunki konieczne KKT:

(1) *Funkcje $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, k$ i $h_j : \mathbb{R}^n \to \mathbb{R}$, $j = 1 \dots, l$ są różniczkowalne w x_o,*
(2) *Istnieją mnożniki Lagrange'a $\lambda_1, \dots, \lambda_k$ i $u_1, \dots, u_l$ takie, że*

- $\nabla f(x_o) - \sum_{i=1}^{k} \lambda_i \nabla g_i(x_o) - \sum_{j=1}^{l} u_j \nabla h_j(x_o) = 0,$
- $h_j(x_o) = 0$ *dla wszystkich* $j = 1 \dots, l,$
- $\lambda_i \geq 0,\ g_i(x_o) \leq 0$ *oraz* $\lambda_i g_i(x_o) = 0$ *dla wszystkich* $i = 1, \dots, k$.

Załóżmy co więcej, że
(3) *funkcje $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, k$, są quasiwypukłe,*
(4) *funkcje $h_j : \mathbb{R}^n \to \mathbb{R}$ są afiniczne,*
(5) *funkcja $f : \mathbb{R}^n \to \mathbb{R}$ jest quasiwklęsła i spełnia co najmniej jeden z warunków:*
   - $\nabla f(x_o) \neq 0$
   - *lub funkcja f jest złożeniem funkcji ściśle rosnącej H i funkcji wklęsłej $\tilde{f}$ (czyli $f(x) = H(\tilde{f}(x))$).*

Funkcje afiniczne na $\mathbb{R}^n$ to funkcje postaci $\alpha + b^T \cdot x$, gdzie $\alpha \in \mathbb{R}$, $b \in \mathbb{R}^n$

Wówczas $x_o = \operatorname{argmax}_{x \in A} f(x)$, gdzie

$$A = \{x \in \mathbb{R}^n \ : \ g_i(x) \leq 0,\ h_j(x) = 0,\ i = 1, \dots, k,\ j = 1, \dots, l\}.$$

Analogiczne twierdzenie zachodzi dla minimum - potrzebne jest tylko kilka drobnych modyfikacji:

TWIERDZENIE 55 (Warunki dostateczne Karusha-Kuhna-Tuckera dla minimum).
Załóżmy, że w punkcie $x_o \in \mathbb{R}^n$ spełnione są warunki konieczne KKT:

(1) *Funkcje $f : \mathbb{R}^n \to \mathbb{R}$, $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, k$ i $h_j : \mathbb{R}^n \to \mathbb{R}$, $j = 1 \dots, l$ są różniczkowalne w x_o,*
(2) *Istnieją mnożniki Lagrange'a $\lambda_1, \dots, \lambda_k$ i $u_1, \dots, u_l$ takie, że*

- $\nabla f(x_o) + \sum_{i=1}^{k} \lambda_i \nabla g_i(x_o) - \sum_{j=1}^{l} u_j \nabla h_j(x_o) = 0,$
- $h_j(x_o) = 0$ *dla wszystkich* $j = 1 \dots, l,$
- $\lambda_i \geq 0,\ g_i(x_o) \leq 0$ *oraz* $\lambda_i g_i(x_o) = 0$ *dla wszystkich* $i = 1, \dots, k$.

Załóżmy co więcej, że
(3) *funkcje $g_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, k$, są quasiwypukłe,*
(4) *funkcje $h_j : \mathbb{R}^n \to \mathbb{R}$ są afiniczne,*

(5) *funkcja* $f : \mathbb{R}^n \to \mathbb{R}$ *jest quasiwypukła i spełnia co najmniej jeden z warunków:*
- $\nabla f(x_o) \neq 0$
- *lub funkcja* f *jest złożeniem funkcji ściśle rosnącej* H *i funkcji wypukłej* $\tilde{f}$ *(czyli* $f(x) = H(\tilde{f}(x))$*).*

Wówczas $x_o = \mathrm{argmin}_{x\in A} f(x)$, *gdzie*

$$A = \{x \in \mathbb{R}^n \ : \ g_i(x) \leq 0, \ h_j(x) = 0, \ i = 1, \ldots, k, \ j = 1, \ldots, l\}.$$

Przedstawię tu dowód Twierdzenia 54 przy nieco uproszczonym warunku *(5)*:
założymy, że funkcja $f : \mathbb{R}^n \to \mathbb{R}$ jest wklęsła.

Proof. Załóżmy, że $x_o \neq \mathrm{argmax}_A f(x)$, a więc że istnieje punkt $y_o \in A$ taki, że $f(y_o) > f(x_o)$. Oznaczmy $v = y_o - x_o$. Zauważmy, że zbiór A jest wypukły, więc wszystkie punkty postaci $x_o + tv = ty_o + (1-t)x_o$ dla $t \in [0, 1]$ należą do A. Mamy

$$\begin{aligned}\nabla f(x_o) \cdot v = f'_v(x_o) &= \lim_{t\to 0} \frac{f(x_o + tv) - f(x_o)}{t} \\ &= \lim_{t\to 0} \frac{f(ty_o + (1-t)x_o) - f(x_o)}{t} \\ &\geq \lim_{t\to 0} \frac{tf(y_o) + (1-t)f(x_o) - f(x_o)}{t} \\ &= f(y_o) - f(x_o) > 0.\end{aligned}$$

Zauważmy z drugiej strony, że jeżeli wiąz g_i jest aktywny w x_o (czyli $g_i(x_o) = 0$), to

$$\nabla g_i(x_o) \cdot v = \lim_{t\to 0} \frac{g_i(x_o + tv) - g_i(x_o)}{t} \leq 0,$$

a że $\lambda_i \geq 0$, mamy $\lambda_i \nabla g_i(x_o) \cdot v \leq 0$. Jeżeli wiąz g_i nie jest aktywny w x_o, to nierówność ta też zachodzi, bo w takim przypadku $\lambda_i = 0$.

Mamy też

$$\nabla h_j(x_o) \cdot v = \lim_{t\to 0} \frac{h_j(x_o + tv) - h_j(x_o)}{t} = 0,$$

bo punkty $x_o + tv$ oraz x_o należą do A (a w punktach A wszystkie funkcje h_j są równe 0). Jeżeli teraz wypiszemy warunek Lagrange'a w działaniu na wektor v:

$$\begin{aligned}0 &= \left(\nabla f(x_o) - \sum_{i=1}^{k} \lambda_i \nabla g_i(x_o) - \sum_{j=1}^{l} u_j \nabla h_j(x_o)\right) \cdot v \\ &= \nabla f(x_o) \cdot v - \sum_{i=1}^{k} \lambda_i \nabla g_i(x_o) \cdot v - \sum_{j=1}^{l} u_j \nabla h_j(x_o) \cdot v \\ &= \nabla f(x_o) \cdot v - \sum_{i=1}^{k} \lambda_i \nabla g_i(x_o) \cdot v,\end{aligned}$$

więc

$$\nabla f(x_o) \cdot v = \sum_{i=1}^{k} \lambda_i \nabla g_i(x_o) \cdot v$$

Otrzymaliśmy sprzeczność, gdyż lewa strona równości jest dodatnia, a prawa jest sumą liczb niedodatnich. □

4. Geometryczna interpretacja warunków KKT. Przypomnijmy warunki konieczne Karusha-Kuhna-Tuckera na to, by funkcja f miała maksimum w punkcie $x_o \in A$,

$$A = \{x \in \mathbb{R}^n \ : \ g_i(x) \leq 0, \, h_j(x) = 0 \, , i = 1, \ldots, k \, j = 1, \ldots, l\},$$

(1) w punkcie x_o zawodzi warunek jakości więzów, czyli wektory $\nabla h_j(x_o)$, $j = 1, \ldots, l$ wraz z wektorami $\nabla g_i(x_o)$ dla tych i, dla których $g_i(x_o) = 0$ (a zatem gradienty wszystkich więzów, które są w x_o aktywne) nie są liniowo niezależne

lub

(2) Istnieją λ_i, $i = 1, \ldots, k$ oraz μ_j, $j = 1, \ldots, l$ takie, że spełniony jest następujący układ równań i nierówności:
- $\nabla f(x_o) = \sum_{i=1}^k \lambda_i \nabla g_i(x_o) + \sum_{j=1}^l \mu_j \nabla h_j(x_o)$
- $\lambda_i \geq 0$ dla $i = 1, \ldots, k$
- $\lambda_i g_i(x_o) = 0$ dla $i = 1, \ldots, k$

Załóżmy teraz dodatkowo, że zbiór A jest wypukły (jest tak, gdy spełnione są warunki dostateczne KKT, a więc gdy funkcje g_i są quasiwypukłe, a h_j — afiniczne) i zastanówmy się raz jeszcze, czego należy oczekiwać w punkcie x_o, w którym funkcja f przyjmuje maksimum.

Jest jasne, że jeżeli v jest kierunkiem dopuszczalnym w punkcie x_o, to funkcja f nie może w kierunku v rosnąć, a zatem pochodna kierunkowa w kierunku v musi być niedodatnia. Przypomnijmy, że $f'_v(x_o) = \langle \nabla f(x_o) \, , \, v \rangle$, zatem

$$\forall v \in F_A(x_o) \quad \langle \nabla f(x_o) \, , \, v \rangle \leq 0$$

Oznacza to jednak, z definicji stożka polarnego, że

$$\nabla f(x_o) \in F_A(x_o)^o = N_A(x_o) \tag{9}$$

I to jest poszukiwany przez nas warunek konieczny na to, by w punkcie x_o było maksimum. Nim przeformułujemy go do użytecznej postaci, zastanówmy się, co powinno spełniać minimum? Oczywiście pochodna kierunkowa w kierunkach dopuszczalnych nie może być dodatnia, bo wówczas funkcja rosła by w tych kierunkach. Stąd $\forall v \in F_A(x_o)$ $\langle \nabla f(x_o) \, , \, v \rangle \geq 0$, a zatem $\langle -\nabla f(x_o) \, , \, v \rangle \leq 0$, czyli

$$-\nabla f(x_o) \in T_A(x_o)^o = N_A(x_o).$$

Jak z geometrycznego warunku (9) wywnioskować warunki KKT?

Będziemy musieli nauczyć się wyznaczać stożki normalne do zbiorów zadanych, jak A, równościami i nierównościami na funkcjach różniczkowalnych. Umiemy to już zrobić w sytuacji, gdy A jest opisany jedną nierównością (i spełniony jest warunek jakości więzów):

Twierdzenie. *Niech $A = \{x \in \mathbb{R}^n \ : \ g(x) \leq 0\}$, przy czym g jest funkcją quasiwypukłą i różniczkowalną w $x_o \in A$. Wówczas*

a) *jeżeli $g(x_o) < 0$, to $T_A(x_o) = \mathbb{R}^n$, a $N_A(x_o) = \{0\}$,*

b) *jeżeli $g(x_o) = 0$, ale $\nabla g(x_o) \neq 0$, to $T_A(x_o) = \{\langle v \, , \, \nabla g(x_o) \rangle \leq 0\}$, a $N_A(x_o) = \{\lambda \nabla f(x_o) \ : \ \lambda \geq 0\}$.*

Wiemy też, że stożek styczny do części wspólnej dwóch zbiorów jest częścią wspólną stożków stycznych do tych zbiorów, pod warunkiem, że oba te zbiory przecinają się „dostatecznie tłusto" (cokolwiek miałoby to znaczyć). Udowodniliśmy też, że jeżeli $T_{A \cap B}(x_o) = T_A(x_o) \cap T_B(x_o)$, to (niezależnie od tego, jak przecinają się zbiory A i B) zachodzi $N_{A \cap B}(x_o) = N_A(x_o) + N_B(x_o)$. Pozostaje nam udowodnić twierdzenie, które tłumaczy znaczenie warunku jakości więzów.

Niech

$$A = \{x \in \mathbb{R}^n \ : \ g_i(x) \le 0,\, i = 1, \dots, k,\}$$

przy dodatkowym założeniu, że w punkcie x_o gradienty wszystkich aktywnych więzów są liniowo niezależne. Zakładamy dodatkowo, że funkcje g_i są quasiwypukłe (dzięki temu zbiór A jest wypukły). Udowodnimy, że zachodzi

Przypadek, gdy zbiór A jest opisany zarówno nierównościami, jak i równościami, jest nieco bardziej złożony, ale metoda dowodu jest w zasadzie ta sama.

TWIERDZENIE 56. *Jeżeli $x_o \in A$, $g_i(x_o) = 0$ dla $i = 1, \dots, m$, $g_i(x_o) < 0$ dla $i = m+1, \dots, k$ oraz $\{\nabla g_i(x_o)\}_{i=1,\dots,m}$ są liniowo niezależne, to*

$$T_A(x_o) = \{x \in \mathbb{R}^n \ : \ \nabla g_i(x_o) \le 0,\, i = 1, \dots, m\}$$

Zauważmy, że jeżeli zachodzi powyższe twierdzenie, to $T_A(x_o) = \bigcap_i T_{A_i}(x_o)$, gdzie

$$A_i = \{x \in \mathbb{R}^n \ : \ g_i(x) \le 0\}.$$

Skoro tak, to wiemy już, że

Ta suma ($\sum$) to oczywiście suma Minkowskiego zbiorów $N_{A_i}(x_o)$.

$$N_A(x_o) = \sum_{i=1}^{k} N_{A_i}(x_o)$$

No, ale każdy z $N_{A_i}(x_o)$ jest bądź $\{0\}$ (gdy wiąz g_i jest nieaktywny), bądź $N_{A_i}(x_o) = \{\lambda_i \nabla g_i(x_o) \ : \ \lambda_i \ge 0\}$. Stąd też

$$N_A(x_o) = \{\sum_{i=1}^{m} \lambda_i \nabla g_i(x_o) \ : \ \lambda_i \ge 0, i = 1, \dots, m\}$$

Do sumy tej możemy bezkarnie dopisać gradienty nieaktywnych więzów ∇g_i ($i = m+1, \dots, k$), zastrzegając, że λ_i występujące przy nich będą równe 0 (co możemy zagwarantować sobie żądając, by $g_i(x_o)\lambda_i = 0$). W ten sposób znany nam już warunek geometryczny

$$\nabla f(x_o) \in N_A(x_o)$$

przyjmuje postać znaną z warunków KKT:

$$\begin{cases} \nabla f(x_o) = \sum_{i=1}^{k} \lambda_i \nabla g_i(x_o) & \\ \lambda_i \ge 0 & i = 1, \dots, k \\ g_i(x_o)\lambda_i = 0 & i = 1, \dots, k \end{cases} \tag{10}$$

Pozostaje udowodnić prawdziwość wzoru na stożek styczny do A. Będą nam do tego potrzebne dwa narzędzia: udowodnione już Twierdzenie Gordana:

TWIERDZENIE (Twierdzenie Gordana). *Niech M będzie macierzą $m \times n$. Zachodzi wówczas jedna i tylko jedna z poniższych możliwości:*

a) *istnieje rozwiązanie układu nierówności $Mx < 0$,*
b) *istnieje niezerowe rozwiązanie $y \in \mathbb{R}^m$ układu nierówności*

$$\begin{cases} y \ge 0 \ \text{(tj. jako wektor ma wszystkie współrzędne nieujemne)} \\ y^T M = 0. \end{cases}$$

oraz następujące

TWIERDZENIE 57 (Twierdzonko). *Niech stożek C będzie opisany układem nierówności liniowych:*

$$C = \{x \in \mathbb{R}^n \ : \ Mx \le 0\}$$

(M jest pewną macierzą $n \times m$). Oznaczmy $C_o = \{x \in \mathbb{R}^n \ : \ Mx < 0\}$. Równość $C = cl\, C_o$ zachodzi wtedy i tylko wtedy, gdy $C_o \neq \varnothing$.

Proof. Jeżeli $C_o = \varnothing$, to oczywiście równość nie może zachodzić, bo $cl\, \varnothing = \varnothing$, podczas gdy $0 \in C$, więc C nie jest zbiorem pustym. Załóżmy więc, że $C_o \neq \varnothing$. Oczywiście $C_o \subset C$, wystarczy więc wykazać, że każdy punkt stożka C jest granicą ciągu punktów z C_o.

Skoro $C_o \neq \varnothing$, to istnieje $y \in C_o$. Wybierzmy teraz dowolne $x \in C$ i przyjmijmy $z_k = x + \frac{1}{k}y$. Mamy $Mz_k = Mx + \frac{1}{k}My < 0$, więc $z_k \in C_o$, a w oczywisty sposób $\lim_{k\to\infty} z_k = x$. □

Dowód Twierdzenia 56. Oznaczmy

$$C = \{x \in \mathbb{R}^n \ : \ \langle x\,,\, \nabla g_i(x_o)\rangle \leq 0,\ i = 1, \dots, m\}$$
$$C_o = \{x \in \mathbb{R}^n \ : \ \langle x\,,\, \nabla g_i(x_o)\rangle < 0,\ i = 1, \dots, m\}$$

Najpierw wykażemy, że $T_A(x_o) \subset C$. Jest to łatwe, bo każdy wektor styczny $v \in T_A(x_o)$ jest granicą ciągu v_k kierunków osiągalnych dla A w x_o. Każdy z nich jest osiągalny dla każdego ze zbiorów A_i, $i = 1, \dots, m$, zachodzi więc $\langle v_k\,,\, \nabla g_i(x_o)\rangle \leq 0$ dla $i = 1, \dots, m$. Wystarczy w tej ostatniej nierówności przejść z k do nieskończoności, by otrzymać, że $\langle v\,,\, \nabla g_i(x_o)\rangle \leq 0$, czyli $v \in C$.

Zauważmy też, że każdy wektor z C_o jest wektorem osiągalnym. Jeżeli bowiem $v \in C_o$, to $(g_i)'_v(x_o) = \langle v\,,\, \nabla g_i(x_o)\rangle < 0$, więc g_i maleje w kierunku v i poruszając się w jego kierunku nie opuścimy przez jakiś czas $A_i = \{x \in \mathbb{R}^n \ : \ g_i(x) \leq 0\}$. Skoro nierówność ta zachodzi dla każdego $i = 1, \dots, m$, to nie opuścimy żadnego z A_i — a więc nie opuścimy i A.

Liniowa niezależność $\nabla g_i(x_o)$, $i = 1, \dots, m$, oznacza, że równanie

$$\sum y_i \nabla g_i(x_o) = 0$$

nie ma niezerowych rozwiązań. Jeżeli oznaczymy przez M macierz, której kolumnami są $\nabla g_i(x_o)$, równanie to przyjmie postać $y^T M = 0$. Układ b) z Twierdzenia Gordana nie ma zatem, dla takiego M, rozwiązania — a więc ma go układ a): istnieje x takie, że Mx ma wszystkie współrzędne ujemne. No, ale te współrzędne to $\langle x\,,\, \nabla g_i(x_o)\rangle$ dla $\nabla g_i(x_o)$, czyli $x \in C_o$. Wiemy w szczególności, że $C_o \neq \varnothing$, a zatem, na mocy Twierdzonka, $cl\, C_o = C$

Domknięcie clC_o jest najmniejszym zbiorem domkniętym zawierającym C_o, stożek $T_A(x_o)$ jest domknięty, więc $C = clC_o \subset T_A(x_o)$. Skoro $C \subset T_A(x_o)$ i $T_A(x_o) \subset C$, to $T_A(x_o) = C$. □

5. Ekstrema związane funkcji nieróżniczkowalnych. Cała teoria subróżniczek i subgradientów nie miałaby za wiele sensu, gdyby dawało się ją zastosować jedynie do ekstremów globalnych, bez żadnych ograniczeń. Znany nam z poprzedniej części warunek to, by f miała minimum w $x_o \in A$: $-\nabla f(x_o) \in N_A(x_o)$ daje się również wypowiedzieć w języku subróżniczek:

Twierdzenie 58 (Warunek pierwszego rzędu na minimum z ograniczeniami). *Niech $A \subset \mathbb{R}^n$ będzie zbiorem wypukłym, a $f : \mathbb{R}^n \to \mathbb{R}$ — funkcją wypukłą. Warunkiem koniecznym i dostatecznym na to, by funkcja f przyjmowała w x_o minimalną wartość na zbiorze A jest to, by*

$$-\partial f(x_o) \cap N_A(x_o) \neq \varnothing. \tag{11}$$

Proof. Załóżmy, że warunek (11) jest spełniony. Istnieje zatem $a \in \partial f(x_o)$ takie, że $-a \in N_A(x_o)$, a więc że

$$\forall x \in A \quad \langle x - x_o\,,\, -a\rangle \leq 0,$$

czyli $\langle x - x_o\,,\, a\rangle \geq 0$. Stąd

$$\forall x \in A \quad f(x) \overset{\text{bo } a\in\partial f(x_o)}{\geq} f(x_o) + \langle a\,,\, x - x_o\rangle \geq f(x_o),$$

co dowodzi, że f ma w x_o minimum.

W drugą stronę jest trudniej.

Rozpatrzmy dwa podzbiory $\mathbb{R}^{n+1}$:

$$B_1 = \{(x - x_o, y) \in \mathbb{R}^n \times \mathbb{R} \ : \ x \in \mathbb{R}^n, \, y > f(x) - f(x_o)\}$$
$$B_2 = \{(x - x_o, y) \in \mathbb{R}^n \times \mathbb{R} \ : \ x \in A, \, y \leq 0\}.$$

Łatwo można sprawdzić (ćwiczenia), że oba te zbiory są wypukłe. Co więcej, $B_1 \cap B_2 = \varnothing$, gdyż gdyby istniał punkt $(x - x_o, y)$ należący równocześnie do obu zbiorów, to istniałby $x \in A$ taki, że $0 \geq y > f(x) - f(x_o)$, co przeczy założeniu, że f ma w x_o minimum. Istnieje zatem hiperpłaszczyzna rozdzielająca B_1 i B_2, a więc taki niezerowy wektor $(\xi, \mu) \in \mathbb{R}^n \times \mathbb{R}$ i liczba $\alpha \in \mathbb{R}$, że

$$\forall x \in \mathbb{R}, \, y > f(x) - f(x_o) \quad \langle \xi \, , \, x - x_o \rangle + \mu y \leq \alpha \tag{12}$$
$$\forall x \in A, \, y \leq 0 \quad \langle xi \, , \, x - x_o \rangle + \mu y \geq \alpha \tag{13}$$

Kładąc w drugim z równań $x = x_o$, $y = 0$ otrzymujemy, że $\alpha \leq 0$. Kładąc w pierwszym z równań $x = x_o$ dostajemy

$$\forall y > 0 \quad \mu y \leq \alpha.$$

Ta ostatnia nierówność może zachodzić dla wszystkich $y > 0$ tylko wtedy, gdy $\alpha = 0$ i $\mu \leq 0$.

Gdyby $\mu = 0$, to z pierwszego z równań mamy

$$\forall x \in \mathbb{R}^n \quad \langle \xi \, , \, x - x_o \rangle = 0,$$

w szczególności gdy weźmiemy $x = x_o + \xi$, otrzymamy że $\xi = 0$ — sprzeczność, bo (ξ, μ) miał być wektorem niezerowym. Wiemy zatem, że $\mu < 0$. Dzieląc (12) i (13) przez $-\mu$ otrzymujemy

$$\forall x \in \mathbb{R}^n, \, y > f(x) - f(x_o) \qquad y \geq \langle -\frac{1}{\mu}\xi \, , \, x - x_o \rangle \tag{14}$$
$$\forall x \in A, \, y \leq 0 \qquad \langle -\frac{1}{\mu}\xi \, , \, x - x_o \rangle - y \geq 0 \tag{15}$$

Kładąc w drugim z równań $y = 0$ i oznaczając $\zeta = -\frac{1}{\mu}\xi$ otrzymujemy, że

$$\forall x \in A, \quad \langle \zeta \, , \, x - x_o \rangle \geq 0,$$

zaś z pierwszego, że $f(x) - f(x_o) \geq \langle \zeta \, , \, x - x_o \rangle$ — a więc że $\zeta \in \partial f(x_o)$. □

Oczywiście zadanie maksymalizacji funkcji wklęsłej f na zbiorze A możemy natychmiast zamienić na minimalizację funkcji wypukłej $-f$ na A — należy jedynie pamiętać, by, uzyskawszy wartość minimalną, zmienić jej znak — w ten sposób otrzymamy poszukiwane maksimum f.

Trudniej jest niestety z maksymalizacją funkcji wypukłej na zbiorze wypukłym A. W tym przypadku możemy łatwo podać warunek konieczny:

TWIERDZENIE 59. *Jeżeli funkcja f przyjmuje w x_o maksimum lokalne, to $\partial f(x_o) \subset N_A(x_o)$.*

PROOF. Skoro w x_o funkcja f przyjmuje maksimum lokalne, to dla punktu $y \in A$ dostatecznie bliskiego x_o zachodzi $f(y) \leq f(x_o)$.

Weźmy dowolne $x \in A$ i zauważmy, że biorąc $\lambda > 0$ dostatecznie małe możemy sprawić, by punkt $y = x_o + \lambda(x - x_o)$ był dowolnie blisko punktu x_o. Mamy zatem dla takiego λ

$$f(x_o) \geq f(y) = f(x_o + \lambda(x - x_o)) \quad \Rightarrow \quad f(y) - f(x_o) \leq 0.$$

Niech teraz $\xi \in \partial f(x_o)$. Z definicji subgradientu wiemy, że

$$f(y) - f(x_o) \geq \langle \xi \, , \, y - x_o \rangle = \lambda \langle \xi \, , \, x - x_o \rangle$$

Stąd dla dowolnego $x \in A$ mamy $\langle \xi \, , \, x - x_o \rangle$, a to dowodzi, że $\xi \in N_A(x_o)$. □

To nie jest bardzo dobre narzędzie do poszukiwania maksimum funkcji — nie ma szans, by był to warunek dostateczny. Zauważmy, że na przykład jeżeli funkcja wypukła f ma w punkcie x_o gradient równy 0, to z wcześniejszych twierdzeń wiadomo, że ma ona w x_o minimum. Z drugiej strony $\partial f(x_o) = \{\nabla f(x_o)\} = \{0\}$, a więc oczywiście $\partial f(x_o) \subset N_A(x_o)$.

Istotną pomocą przy poszukiwaniu maksimów jest (udowodnione dawno temu) twierdzenie, że funkcja wypukła przyjmuje maksimum w punkcie ekstremalnym zbioru A. W przypadku, gdy zbiór A jest wielościanem, punktów tych (wierzchołków wielościanu) jest skończenie wiele, co więcej są algorytmy (np. warianty metody Simplex) pozwalające w sposób systematyczny przeszukiwać te wierzchołki.

Problem znajdowania minimum funkcji wklęsłej, jak poprzednio, możemy zamienić na problem znajdowania maksimum funkcji wypukłej, rozważając funkcję $-f$ w miejsce f.

Wykład 10

6. Zadanie dualne Lagrange'a. Wróćmy do typowego zadania optymalizacji z ograniczeniami równościowymi i nierównościowymi:

Szukamy minimum funkcji $f(x)$ na zbiorze A, gdzie

$$A = \{x \in \mathbb{R}^n \; : \; g_i(x) \leq 0,\, h_j(x) = 0\,, i = 1, \ldots, k\, j = 1, \ldots, l\},$$

Przy rozwiązywaniu tak sformułowanego zadania wprowadziliśmy funkcję Lagrange'a

$$L(x) = f(x) + \sum_{i=1}^{k} \lambda_i g_i(x) - \sum_{j=1}^{l} \mu_j h_j(x),$$

w rzeczywistości jednak nie korzystaliśmy z niej, formułując warunki KKT używając bezpośrednio funkcji f, g i h. Przyjrzymy się jej teraz staranniej.

Zauważmy po pierwsze, że funkcję L możemy traktować nie tylko jako funkcję zmiennej x, ale również i mnożników λ_i i μ_j. Od tych nowych zmiennych funkcja L zależy liniowo, możemy więc łatwo stwierdzić, że

$$\frac{\partial L(x, \lambda, \mu)}{\partial \lambda_i}(x, \lambda, \mu) = g_i(x) \qquad \frac{\partial L(x, \lambda, \mu)}{\partial \mu_j}(x, \lambda, \mu) = h_j(x).$$

Przyjmijmy następującą, wygodną konwencję: przez ∇_x oznaczać będziemy zwykły gradient względem zmiennych x (więc $\nabla_x F = (\frac{\partial F}{\partial x_1}, \frac{\partial F}{\partial x_2}, \ldots, \frac{\partial F}{\partial x_n})$), przez ∇_λ — gradient względem zmiennych λ ($\nabla_\lambda F = (\frac{\partial F}{\partial \lambda_1}, \frac{\partial F}{\partial \lambda_2}, \ldots, \frac{\partial F}{\partial \lambda_k})$), analogicznie wprowadzamy ∇_μ. Pozwala to w szczególnie prosty sposób zapisać warunki konieczne KKT:

$$\begin{aligned} \nabla_x L(x, \lambda, \mu) &= 0, \\ \nabla_\lambda L(x, \lambda, \mu) &\leq 0, \\ \nabla_\mu L(x, \lambda, \mu) &= 0, \\ \sum_{i=1}^{k} \lambda_i \frac{\partial L}{\partial \lambda_i}(x, \lambda, \mu) &= 0, \quad \lambda_i \geq 0. \end{aligned}$$

Zdefiniujmy teraz dwie nowe funkcje, związane z funkcją Lagrange'a:

$$L_P(x) = \sup_{\lambda \geq 0, \mu \in \mathbb{R}^l} L(x, \lambda, \mu)$$
$$L_D(\lambda, \mu) = \inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu).$$

W definicji funkcji L_D często zamiast brać infimum po wszystkich $x \in \mathbb{R}^n$, wydzielamy część warunków na zbiór A, tak by $A = A_1 \cap A_2$, funkcję Lagrange'a L budujemy korzystając jedynie z warunków zawartych w A_1, a infimum w L_D bierzemy po $x \in A_2$.

TWIERDZENIE 60. *Zadania minimalizacji $f(x)$ na zbiorze A i $L_P(x)$ na $\mathbb{R}^n$ są równoważne, czyli*

$$\mathrm{argmin}_{x \in A} f(x) = \mathrm{argmin}_{x \in \mathbb{R}^n} L_P(x) \quad \textit{oraz} \quad \min_{x \in A} f(x) = \min_{x \in \mathbb{R}^n} L_P(x)$$

PROOF. Okazuje się, że

$$L_P(x) = \begin{cases} f(x) & x \in A, \\ \infty & x \notin A, \end{cases}$$

Rzeczywiście, jeżeli $x \in A$, to $g_i(x) \leq 0$, zaś $h_j(x) = 0$. Stąd też $\sum \mu_j h_j(x) = 0$, zaś $\sum \lambda_i g_i(x) \leq 0$. To ostatnie wyrażenie przyjmuje wartość 0, gdy wszystkie λ_i są równe 0, a więc $\sup_{\lambda \geq 0} \sum \lambda_i g_i(x) = 0$.

Jeżeli natomiast $x \notin A$, to albo dla pewnego j zachodzi $h_j(x) \neq 0$, albo . Biorąc μ_j tego samego znaku co $h_j(x)$ i zwiększając $|\mu_j|$ możemy dowolnie powiększyć $\mu_j h_j(x)$. Podobnie, gdy dla pewnego i zachodzi $g_i(x) > 0$, to biorąc duże λ_i możemy uczynić wyrażenie $\lambda_i g_i(x)$ dowolnie dużym. Dlatego też w tym przypadku $L_P(x) = \sup_{\lambda \geq 0, \mu \in \mathbb{R}^l} L(x, \lambda, \mu) = \infty$.

Oczywiście funkcja L_P może przyjmować minimum tylko tam, gdzie jest skończona — a więc na zbiorze A, tam, gdzie pokrywa się z funkcją f. □

Innymi słowy — zamiast szukać $\min_{x \in A} f(x)$, możemy szukać

$$\min_{x \in \mathbb{R}^n} \sup_{\lambda \geq 0, \mu \in \mathbb{R}^l} L(x, \lambda, \mu). \tag{16}$$

Oczywiście jest wszystko jedno, czy minimum szukamy po $x \in \mathbb{R}^n$, czy po $x \in A$ — będą one równe.

Z zadaniem naszym związane jest tak zwane *zadanie dualne*:

DEFINICJA 21. *Zadaniem dualnym* do zadania minimalizacji funkcji f na zbiorze A nazywamy zadanie maksymalizacji funkcji $L_D(\lambda, \mu)$ na zbiorze $\{(\lambda, \mu) \in \mathbb{R}^k \times \mathbb{R}^l \ : \ \lambda \geq 0\}$.

Zadanie dualne polega, jak widać, na znalezieniu

$$\sup_{\lambda \geq 0, \mu \in \mathbb{R}^l} \inf_{x \in A} L(x, \lambda, \mu)$$

— a więc na rozwiązaniu zadania powstałego przez zamienienie w (16) minimum (a w zasadzie infimum) z supremum.

Dla odróżnienia, zadanie minimalizacji funkcji L_P na $\mathbb{R}^n$ (równoważne, jak już wiemy, minimalizacji f na A) nazywać będziemy *zadaniem prymarnym.* Dla uproszczenia sformułowań twierdzeń, rozwiązaniem zadania – czy to prymarnego, czy dualnego — będziemy nazywać zarówno wartość (minimalną lub maksymalną) optymalizowanej funkcji, jak i punkt, w którym ta optymalna wartość będzie osiągana. Tak więc "$x_o \in A$ jest rozwiązaniem zadania prymarnego" oznacza, że $x_o = \mathrm{argmin}_A f(x)$.

Często zdarza się, że nasze zadanie optymalizacji jest sformułowane nieco inaczej: mamy znaleźć minimum funkcji określonej na pewnym zbiorze wypukłym A_1, ale minimum wśród wszystkich punktów spełniających pewne, opisane równaniami i nierównościami, warunki. Innymi słowy — poszukujemy minimum $f(x)$ na zbiorze $A_1 \cap A_2$, przy czym zbiór A_1 jest wypukły (i nie jest dla nas bardzo ważne, jak jest określony), a zbiór A_2 to zbiór opisany równaniami i nierównościami — jak zbiór A w pierwotnym sformułowaniu problemu. Możemy wówczas sformułować funkcję Lagrange'a $L(x, \lambda, \mu)$ używając wyłącznie warunków opisujących zbiór A_2, po czym skonstruować funkcje $L_P(x)$ i $L_D(\lambda, \mu)$. W definicji tej pierwszej nie wprowadzamy żadnych zmian, natomiast drugą definiujemy jako $\inf_{x \in A_1} L(x, \lambda, \mu)$. W dalszym ciągu zadanie minimalizacji L_P (tylko tym razem na zbiorze A_1) będzie równoważne pierwotnemu zadaniu minimalizacji (dowód nie wymaga w zasadzie żadnych zmian).

Nawet, gdy zadanie mamy sformułowane w standardowy sposób, możemy je w powyższy sposób przeformułować, wykorzystując część warunków tak, by opisywały zbiór A_1, a resztę wykorzystując do skonstruowania funkcji Lagrange'a i zadania dualnego. Po co to robić, pokaże twierdzenie o silnej dualności, które poznają Państwo pod koniec tego wykładu.

Odtąd, dla prostszego zapisu, zbiór $\{(\lambda, \mu) \in \mathbb{R}^k \times \mathbb{R}^l \ : \ \lambda \geq 0\}$ oznaczać będziemy przez Λ.

Twierdzenie 61 (O słabej dualności). *Dla dowolnego $x \in \mathbb{R}^n$ ($x \in A_1$ w przypadku podziału zbioru warunków) i $(\lambda, \mu) \in \Lambda$ zachodzi*

$$f(x) \geq L_D(\lambda, \mu)$$

Proof. Skoro $x \in A$, to $h_j(x) = 0$ i $g_i(x) \leq 0$ dla $i = 1, \ldots, k$, $j = 1, \ldots, l$. Jeżeli więc $(\lambda, \mu) \in \Lambda$, to $\lambda_i g_i(x) \leq 0$. Mamy zatem

$$\begin{aligned}
L_D(\lambda, \mu) &= \inf_{z \in \mathbb{R}^n} (f(z) + \sum_i \lambda_i g_i(z) + \sum_j \mu_j h_j(z)) \\
&\leq f(x) + \sum_i \lambda_i g_i(x) + \sum_j \mu_j h_j(x) \\
&\leq f(x).
\end{aligned}$$

Gdy f jest określona jedynie na zbiorze A_1, należy w powyższym rachunku zamienić $\mathbb{R}^n$ na A_1. □

Wnioski:

(1) Na mocy Twierdzenia 60 zachodzi również

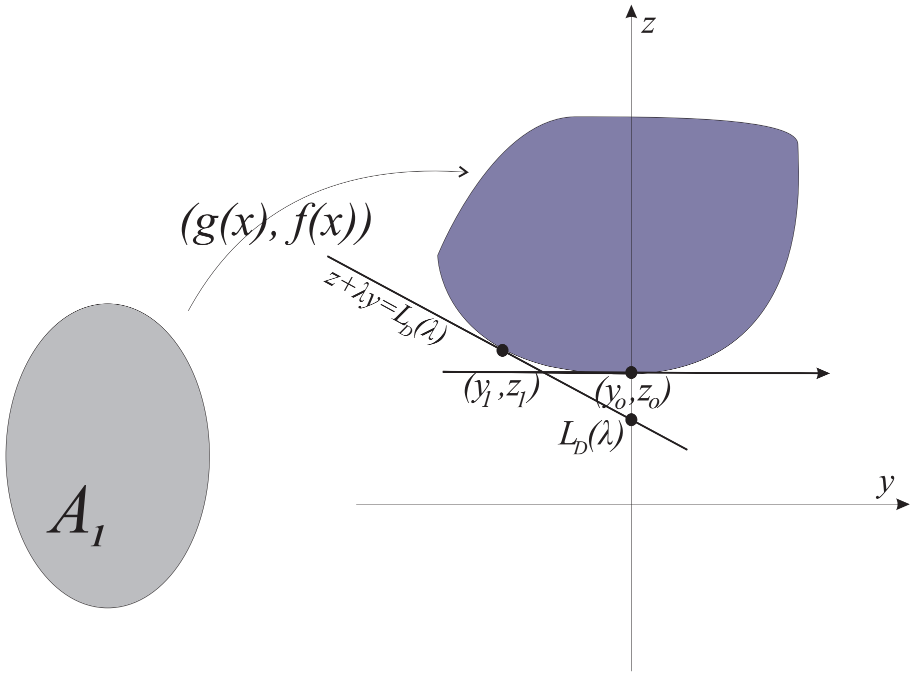
$$L_D(\lambda, \mu) \leq L_P(\lambda, \mu)$$

(2) Rozwiązanie zadania prymarnego jest nie mniejsze ($\leq$) od rozwiązania zadania dualnego.

(3) Jeżeli istnieją punkty $x_o \in A$ i $(\lambda_o, \mu_o) \in \Lambda$ takie, że $f(x_o) = L_D(\lambda_o, \mu_o)$, to x_o jest rozwiązaniem zadania prymarnego, a (λ_o, μ_o) — rozwiązaniem zadania dualnego.

7. Interpretacja geometryczna zadania dualnego. Rozważmy dla uproszczenia zadanie minimalizacji funkcji wypukłej f na zbiorze

$$A = \{x \in \mathbb{R}^n \ : \ g(x) \leq 0\}.$$



Przyjrzyjmy się przekształceniu $(g(x), f(x)) : A_1 \to \mathbb{R}^2$ (gdy $A_1 = \mathbb{R}^n$, otrzymujemy przypadek standardowego zadania, bez podziału warunków między A_1 a A_2). Jego obrazem jest, na rysunku, zbiór oznaczony kolorem niebieskim — będę go, dla uproszczenia, nazywał N. Rozwiązaniem zadania prymarnego jest współrzędna z takiego punktu spośród leżących w zbiorze N i mających współrzędną y niedodatnią (bo $g(x) \leq 0$), którego współrzędna z jest najmniejsza. Jest to oczywiście punkt (y_o, z_o), a więc rozwiązaniem jest z_o.

Aby znaleźć wartość funkcji $L_D(\lambda)$ dla jakiegoś $\lambda \geq 0$, musimy znaleźć minimum wyrażenia $z+\lambda y$ dla $(y, z) \in N$, $y \leq 0$. Wzór $z+\lambda y = c$ dla różnych stałych c opisuje rodzinę prostych równoległych — by więc znaleźć poszukiwane minimum trzeba wziąć taką prostą i zmniejszać stałą c (co odpowiada opuszczaniu tej prostej w dół). W pewnym momencie (dla pewnego c) prosta ta będzie się stykać ze zbiorem N (będzie go podpierać), a przy dalszym zmniejszaniu c nie będzie już miała z N żadnych punktów wspólnych. Ta wartość c jest minimalna - a więc jest równa $L_D(\lambda)$. Wartość c możemy łatwo odczytać z wykresu — jest to współrzędna z punktu przecięcia znalezionej prostej z osią z.

Z rysunku widać, że największa możliwa wartość $L_D(\lambda)$ jest przyjmowana dla prostej przechodzącej w (y_o, z_o) i jest ona równa z_o.

W tym przypadku zatem rozwiązania zadań prymarnego i dualnego są sobie równe. Nie zawsze tak jest, co pokazuje następujący przykład

Przykład: Rozważmy zagadnienie minimalizacji funkcji $f(x) = \sqrt[3]{x}$ na półprostej $x \geq 0$. Funkcja Lagrange'a jest równa $L(x, \lambda) = \sqrt[3]{x} - \lambda x$, więc $L_D(\lambda) = \inf_{x \in \mathbb{R}} L(x, \lambda) = -\infty$ dla wszystkich λ, podczas, gdy rozwiązaniem zadania prymarnego jest oczywiście 0.

Gdybyśmy jednak chcieli powtórzyć ten rachunek dla $f(x) = x$, otrzymalibyśmy

$$L_D(x) = \begin{cases} -\infty & \lambda \neq 1 \\ 0 & \lambda = 1, \end{cases}$$

więc

$$\max_{\lambda \geq 0} L_D(\lambda) = 0 = \min_{x \geq 0} f(x)$$

.

DEFINICJA 22. Różnicę między rozwiązaniem zadania prymarnego i dualnego nazywamy *odstępem dualności.*

Słabe twierdzenie o dualności mówi, że odstęp ten jest zawsze nieujemny. Powstaje ważne pytanie - kiedy jest on równy zeru (a więc kiedy możemy rozwiązać zadanie dualne *zamiast* zadania prymarnego i twierdzić, że otrzymamy w ten sposób dobry wynik. Jest sporo różnych sformułowań warunków gwarantujących zerowy odstęp dualności, jeden z najważniejszych podaje poniższe

TWIERDZENIE 62 (o silnej dualności). *Jeżeli*

- *zbiór A_1 jest niepusty i wypukły,*
- *funkcja f jest wypukła,*
- $A_2 = \{x \in \mathbb{R}^n \ : \ g_i(x) \leq 0,\, h_j(x) = 0,\, i = 1, \dots, k,\, j = 1, \dots, l\}$, *przy czym funkcje g_i są wypukłe, a h_j — afiniczne,*
- $\inf_{x \in A_1 \cap A_2} f(x) < \infty$ *(chcemy wykluczyć trywialny przypadek funkcji, która na całym zbiorze $A_1 \cap A_2$ jest równa $+\infty$),*
- *istnieje $x_o \in \mathrm{ri} A_1$ takie, że $g_i(x_o) < 0$, $h_j(x_o) = 0$ dla wszystkich $i = 1, \dots, k$, $j = 1, \dots, l$ (warunek regularności),*

to rozwiązania zadania prymarnego i dualnego są równe (nie ma odstępu dualności).

W sformułowaniu twierdzenia pojawia się niezdefiniowane jak dotąd pojęcie: $\mathrm{ri}A$, czyli *relatywne wnętrze* (ang. relative interior) zbioru A_1. Jest to uogólnienie pojęcia wnętrza. Zanim podamy precyzyjną definicję, zauważmy, że gdy rozpatrujemy koło $\{x^2 + y^2 \leq 1\}$ jako podzbiór płaszczyzny (x, y), to jego wnętrzem jest otwarte koło $x^2 + y^2 < 1$. Jeżeli jednak to samo koło umieścimy w przestrzeni trójwymiarowej (x, y, z), jako $\{x^2 + y^2 \leq 1, z = 0\}$, będzie ono zbiorem o pustym wnętrzu. Z „punktu widzenia" koła nic się nie zmieniło, zmiana pojęcia wnętrza przyszła „z zewnątrz", ze zmianą przestrzeni, w której owemu kołu się przyglądamy. Aby zaradzić takim sytuacjom, wprowadzamy pojęcie *relatywnego (względnego) wnętrza* — a więc wnętrza obliczonego w najmniejszej podprzestrzeni afinicznej, jaka zawiera nasz zbiór. Dla naszego przykładowego koła będzie to płaszczyzna $z = 0$.

Sformalizowana definicja wygląda tak:

DEFINICJA 23. *Relatywne wnętrze* zbioru $X \subset \mathbb{R}^n$ to

$$\mathrm{ri}X = \{x \in X \ : \ \exists_{\varepsilon > 0}\, B(x, \varepsilon) \cap Aff(X) \subset X\},$$

gdzie przez $B(x, \varepsilon)$ oznaczam kulę o środku w x i promieniu ε, a przez $Aff(X)$ — część wspólną wszystkich podprzestrzeni afinicznych $\mathbb{R}^n$ zawierających X (a więc najmniejszą podprzestrzeń afiniczną zawierającą X, bo część wspólna podprzestrzeni afinicznych jest podprzestrzenią afiniczną).

Oczywiście wnętrze X jest zawarte w ri$X$, pojęcie to zaczyna być jednak naprawdę użyteczne dopiero wtedy, gdy int$X = \varnothing$.

Proof. Przy dowodzie twierdzenia możemy ograniczyć się do przypadku, w którym:

- zbiór A_1 ma niepuste wnętrze,
- wektory gradientów ograniczeń równościowych h_j są liniowo niezależne.

Rzeczywiście, wiemy, że zbiór A_1 ma niepuste relatywne wnętrze — jeżeli więc jego wnętrze jest puste, to możemy przeformułować całe nasze zadanie tak, by działo się nie w $\mathbb{R}^n$ (które jest „za duże"), ale w przestrzeni $Aff(A_1)$, w której zbiór A_1 ma już wnętrze niepuste.

Jeżeli natomiast gradienty ograniczeń równościowych są liniowo zależne, to oznacza jedynie, że niektóre z nich opisują hiperpłaszczyzny nie wnoszące nic nowego do opisu zbioru. Możemy zignorować zatem te warunki (kładąc odpowiednie μ_j równe 0).

Zdefiniujmy zbiór

$$X = \{(a_1, a_2, \ldots, a_k, b_1, b_2, \ldots, b_l, c) \in R^{k+l+1} \ :$$
$$\text{istnieje } x \in A_1 \text{ t.ż. } g_i(x) \le a_i,\ h_j(x) = b_j,\ f(x) \le c,\ i = 1, \ldots, k,\ j = 1, \ldots, l\}$$

Jest to zbiór wypukły. Można to łatwo sprawdzić bezpośrednio, można też zauważyć, że wypukły jest (jako część wspólna zbiorów wypukłych) zbiór

$$Y = \{(a_1, a_2, \ldots, a_k, b_1, b_2, \ldots, b_l, c, x_1, \ldots, x_n) \in R^{k+l+1+n} \ :$$
$$x \in A_1,\ g_i(x) \le a_i,\ h_j(x) = b_j,\ f(x) \le c,\ i = 1, \ldots, k,\ j = 1, \ldots, l\},$$

a zbiór X to rzut zbioru Y na pierwsze $(k + l + 1)$ zmiennych. Rzut jest przekształceniem liniowym — a więc wypukłym — zatem rzut zbioru wypukłego jest zbiorem wypukłym.

Niech teraz y będzie rozwiązaniem zadania prymarnego – a więc minimum funkcji f na zbiorze $A_1 \cap A_2$. Punkt $(0, \ldots, 0, y)$ należy do zbioru X, gdyż punkt x_o, w którym to minimum jest osiągnięte, spełnia wszystkie równania i nierówności definiujące zbiór A_2. Widać łatwo, że nie może on leżeć we wnętrzu zbioru X, bo wówczas do X należałyby punkty postaci $(0, 0, y - \varepsilon)$ dla ε dostatecznie małych, co przeczyło by minimalności y.

Skoro punkt $(0, 0, \ldots, 0, y)$ leży na brzegu X, to istnieje hiperpłaszczyzna podpierająca zbiór X w punkcie $(0, \ldots, 0, y)$, a więc niezerowy wektor $v = (\lambda, \mu, \beta)$ taki, że $\langle v\,,\,(0, \ldots, 0, y)\rangle \le \langle v\,,\,(a, b, c)\rangle$ dla każdego punktu $(a, b, c) \in X$. Równoważnie,

$$\beta y \le \langle \lambda\,,\,a\rangle + \langle \mu\,,\,b\rangle + \beta c \qquad \forall (a, b, c) \in X. \tag{17}$$

Zauważmy następnie, że w zbiorze X zmienne a i c nie są ograniczone z góry - jeżeli $(a, b, c) \in X$, to dla dowolnych $\alpha \in \mathbb{R}$, $\tilde{a} \in \mathbb{R}^k$, α, $a_i \ge 0$, mamy $(a + \tilde{a}, b, c) \in X$ oraz $(a, b, c + \alpha) \in X$. Gdyby teraz któraś z liczb λ_i czy β była ujemna, to powiększając odpowiednią współrzędną a_i lub c w (17) moglibyśmy prawą stronę nierówności dowolnie pomniejszać, co przeczyłoby nierówności. wiemy zatem, że λ_i, $i = 1, \ldots, k$ oraz β są nieujemne.

Teraz wykażemy, że $\beta > 0$. Zauważmy, że jeżeli $x \in A_1$, to $(g(x), h(x), f(x)) \in X$. Dlatego też, gdyby $\beta = 0$, to dla każdego $x \in A_1$ zachodziłoby

$$0 \le \langle \lambda\,,\,g(x)\rangle + \langle \mu\,,\,h(x)\rangle. \tag{18}$$

Kładąc w tej nierówności $x = x_o$, gdzie x_o jest punktem spełniającym warunek regularności z treści twierdzenia, widzimy, że

$$0 \le \langle \lambda\,,\,g(x_o)\rangle \quad \text{oraz } g_i(x_o) < 0 \quad i = 1, \ldots, k.$$

Widać natychmiast, że wszystkie współczynniki λ_i muszą być równe 0. Musimy teraz skorzystać z faktu, że ograniczenia równościowe h_j są afiniczne - a więc że istnieją wektory $\chi_j \in \mathbb{R}^l$ i liczby $\xi \in \mathbb{R}$ takie, że $h_j(x) = \langle \chi\,,\,x\rangle + \xi$. Wektory χ_j to gradienty ograniczeń równościowych, na mocy więc założenia są liniowo niezależne.

Wróćmy do (18) i zapiszmy je korzystając z χ i ξ oraz z wiedzy, że $\lambda = 0$.

$$0 \le \sum_j \mu_j \langle \chi_j\,,\,x\rangle + \xi_j. \tag{19}$$

Wiemy, że $h_j(x_o) = \langle \chi_j\,,\,x_o\rangle + \xi_j = 0$, możemy więc odjąć tę wielkość od (19), otrzymując

$$0 \le \sum_j \mu_j \langle \chi_j\,,\,x - x_o\rangle = \langle \sum_j \mu_j \chi_j\,,\,x - x_o\rangle \quad \forall x \in A_1.$$

Oznacza to, że wektor $-\sum_j \mu_j \chi_j$ należy do $N_{A_1}(x_o)$. Z drugiej strony wiemy, że x_o jest punktem wewnętrznym zbioru A_2, więc $N_{A_1}(x_o) = \{0\}$, zatem $\sum_j \mu_j \chi_j = 0$. Co więcej, wektory χ_j są liniowo niezależne, co oznacza, że $\mu_j = 0$ dla $j = 1, \ldots, l$. Otrzymaliśmy sprzeczność — wektor (λ, μ, β) jest równy 0, a miał być niezerowy. To dowodzi, że $\beta > 0$.

Wróćmy do wzoru (17). Możemy skorzystać raz jeszcze z faktu, że jeżeli $x \in A_1$, to $(g(x), h(x), f(x)) \in X$, zatem

$$\beta y \leq \langle \lambda\,,\, g(x)\rangle + \langle \mu\,,\, h(x)\rangle + \beta f(x). \tag{20}$$

Dzieląc obie strony (20) prze β dostajemy, że

$$y = inf_{x \in A_1 \cap A_2} f(x) \leq \langle \lambda/\beta\,,\, g(x)\rangle + \langle \mu/\beta\,,\, h(x)\rangle + f(x) = L(x, \lambda/\beta, \mu/\beta)$$

dla pewnych $(\lambda/\beta, \mu/\beta) \in \Lambda$ i dowolnego $x \in A_1$. Oznacza to, że

$$L_D(\lambda/\beta, \mu/\beta) = \inf_{x \in A_1} L(x, \lambda/\beta, \mu/\beta) \geq y = inf_{x \in A_1 \cap A_2} f(x). \tag{21}$$

Ze słabego twierdzenia o dualności wynika już, że

- w (21) w rzeczywistości jest równość, nie nierówność,
- $L_D(\lambda/\beta, \mu/\beta)$ jest rozwiązaniem zadania dualnego,
- nie ma zatem odstępu dualności,
- wektory $\lambda/\beta, \mu/\beta$ są wektorami mnożników Lagrange'a.

□

Aby obliczyć dualną funkcję Lagrange'a L_D musimy obliczyć infimum wyrażenia $f(x) + \langle(\lambda, \mu)\,,\,(g(x), h(x)\rangle$ po $x \in A_1$. Czasem jest to nietrywialne, okazuje się jednak, że otrzymana w ten sposób funkcja ma pewne szczególne własności, przy bardzo słabych założeniach na funkcje g i h.

Twierdzenie 63. *Niech A_1 będzie zwartym podzbiorem $\mathbb{R}^n$, a funkcje $f : \mathbb{R}^n \to \mathbb{R}$ i $w : \mathbb{R}^n \to \mathbb{R}^{k+l}$ niech będą ciągłe. Funkcja $\phi : \mathbb{R}^{k+l} \to \mathbb{R}$ zdefiniowana jako*

$$\phi(\xi) = \inf_{x \in A_1} \{f(x) + \langle \xi\,,\, w\rangle\}$$

jest funkcją wklęsłą.

Założenie zwartości A_1, jak i ciągłości funkcji f i w, służą jedynie temu, by funkcja ϕ przyjmowała wartości skończone — jeżeli dopuścimy funkcje wklęsłe przyjmujące wartości nieskończone, to teza twierdzenia będzie prawdziwa bez tych założeń.

Oczywiście w naszym zagadnieniu wektor ξ to wektor mnożników Lagrange'a (λ, μ), a $w(x)$ to zapisana wektorowo funkcja warunków $(g(x), h(x))$.

Twierdzenie to pokazuje, że zadanie maksymalizacji funkcji dualnej Lagrange'a (gdy już damy radę ją wyznaczyć) jest prostsze, niż można by się spodziewać, gdyż maksymalizacja funkcji wklęsłych (podobnie jak minimalizacja wypukłych) jest stosunkowo prosta. Wiemy na przykład, że jeżeli ϕ jest różniczkowalna, to jej (dowolny) punkt krytyczny jest jej globalnym maksimum.

Proof. Niech $\alpha \in [0, 1]$.

$$\begin{aligned}\phi(\alpha\xi + (1-\alpha)\zeta) &= \inf_{x \in A_1} \{f(x) + \langle \alpha\xi + (1-\alpha)\zeta, w\,,\,\rangle\} \\ &= \inf_{x \in A_1} \{f(x) + \alpha\langle \xi, w\,,\,+\rangle(1-\alpha)\langle \zeta, w\,,\,\rangle\} \\ &\leq \alpha \inf_{x \in A_1} \{f(x) + \langle \xi, w\,,\,\rangle\} + (1-\alpha) \inf_{x \in A_1} \{f(x) + \langle \zeta, w\,,\,\rangle\} \\ &= \alpha\phi(\xi) + (1-\alpha)\phi(\zeta).\end{aligned} \tag{22}$$

□