Elementy geometrii wypukłej
ciąg dalszy

Przypomnienie: $A \subset \mathbb{R}^n$ jest wypukły, gdy $\forall_{x,y \in A}$ cały odcinek $[x,y]$ leży w A, tzn.

$$\forall_{x,y\in A} \ \forall_{\lambda \in [0,1]} \quad \lambda x + (1-\lambda) y \in A.$$

Wprost z definicji możemy sprawdzić, że

① część wspólna dowolnej rodziny zbiorów wypukłych jest zbiorem wypukłym

② jeżeli $T: \mathbb{R}^n \to \mathbb{R}^m$ jest liniowe, ~~to~~ a $A \subset \mathbb{R}^n$ jest wypukły, to $T(A) \subset \mathbb{R}^m$ też jest wypukły

Uwaga: Dla każdego $X \subset \mathbb{R}^n$ istnieje najmniejszy wypukły podzbiór $\mathbb{R}^n$ zawierający X; nazywamy go uwypukleniem lub otoczką wypukłą zbioru X i oznaczamy $\operatorname{conv}(X)$.

Dowód: $\operatorname{conv}(X) = \bigcap \{ Y \subset \mathbb{R}^n : Y \text{ wypukły i } X \subset Y \}$

Zadanko: $\operatorname{conv}(X) = \left\{ \sum_{i=1}^{k} \lambda_i x_i : k \in \mathbb{N}, \lambda_i \geq 0, \sum_{i=1}^{k} \lambda_i = 1, x_i \in X \right\}$

Niech A będzie zbiorem wypukłym.

Def: Punkt $x \in A$ nazywamy punktem ekstremalnym zbioru A, gdy x nie jest punktem wewnętrznym żadnego odcinka o końcach w A.

Zbiór wszystkich punktów ekstremalnych zbioru A oznaczamy $\operatorname{ex} A$

Przykłady: Gdy A jest wielościanem wypukłym, $\operatorname{ex} A$ jest zbiorem jego wierzchołków. Kula $B(0,1)$ nie ma punktów ekstremalnych, podobnie prosta $A = \{(x,0) : x \in \mathbb{R}\} \subset \mathbb{R}^2$.

Twierdzenie (Minkowski, Steiner)

Niech $A \subset \mathbb{R}^n$ będzie zwarty, niepusty i wypukły. Wówczas

a) $A = \operatorname{conv}(\operatorname{ex} A)$, w szczególności $\operatorname{ex} A \neq \emptyset$

b) jeżeli $f : A \to \mathbb{R}$ jest wypukła i ciągła, to f osiąga maksimum na A w jednym z punktów ekstremalnych A.

c) jeżeli $f : A \to \mathbb{R}$ jest wklęsła i ciągła, to f osiąga minimum na A w jednym z punktów ekstremalnych

d) jeżeli $f : A \to \mathbb{R}$ jest afiniczna, to zarówno minimum, jak i maksimum na A osiąga w punktach ekstremalnych A

Dowód: Zauważmy na początek, że c) wynika z b), wystarczy, dla wklęsłej $f : A \to \mathbb{R}$, zastosować b) do funkcji wypukłej $-f$. Dalej, d) wynika z b) i c), bo funkcja afiniczna jest równocześnie wklęsła i wypukła.

Wystarczy zatem wykazać a) i b).
Zrobimy to przez indukcję ze względu na wymiar n.

Dla $\boxed{n=1}$: jedyne zwarte, niepuste, wypukłe podzbiory $\mathbb{R}$ to odcinki domknięte (być może zdegenerowane do punktu), a więc $A=[a,b]$, $a\leq b$;
bez trudu sprawdzamy, że $\mathrm{ex}\,A=\{a,b\}$, $A=\mathrm{conv}(\mathrm{ex}\,A)$;
jeżeli f jest wypukła na $[a,b]$, to dla każdego $x\in[a,b]$
$x=\alpha a+(1-\alpha)b$ dla pewnego $\alpha\in[0,1]$
$f(x)\leq\alpha f(a)+(1-\alpha)f(b)\leq\max(f(a),f(b))$
skąd od razu wynika b).

Załóżmy zatem, że teza twierdzenia zachodzi dla wszystkich zwartych, niepustych, wypukłych podzbiorów $\mathbb{R}^{n-1}$.

Niech teraz $A\subset\mathbb{R}^n$ będzie wypukły, $f: A\to\mathbb{R}$ wypukła.

Ustalmy $x\in A$; wykażemy, że $x\in\mathrm{conv}(\mathrm{ex}\,A)$, a więc że x można zapisać jako kombinację wypukłą punktów ekstremalnych A.

Jeżeli $x\in\mathrm{ex}\,A$, to oczywiście $x\in\mathrm{conv}(\mathrm{ex}\,A)$;
załóżmy zatem, że $x\notin\mathrm{ex}\,A$. Wtedy x jest punktem wewnętrznym pewnego odcinka $[y',z']\subset A$, $y'\neq z'$.
Oznaczmy prostą przechodzącą przez y' i z' (i przez x) przez l. Zbiór $A\cap l$ jest zwartym, wypukłym, niepustym podzbiorem $l\simeq\mathbb{R}$, więc $A\cap l=[y,z]$

dla pewnych $y, z \in A$, $y \neq z$, przy czym punkty y i z leżą na brzegu A. Niech H_y i H_z oznaczają hiperpłaszczyzny podpierające A odpowiednio w punktach y i z, niech $A_y = H_y \cap A$, $A_z = H_z \cap A$.

Zbiory A_y i A_z są wypukłymi, zwartymi, niepustymi podzbiorami $(n-1)$-wymiarowych podprzestrzeni afinicznych H_y i H_z, możemy więc zastosować do nich założenie indukcyjne: $A_y = \mathrm{conv}(\mathrm{ex}\, A_y)$, $A_z = \mathrm{conv}(\mathrm{ex}\, A_z)$. Wykażemy, że $\mathrm{ex}\, A_y \subset \mathrm{ex}\, A$ i $\mathrm{ex}\, A_z \subset \mathrm{ex}\, A$. Niech bowiem $w \in \mathrm{ex}\, A_y$. Jeżeli w nie jest punktem ekstremalnym A, to jest punktem wewnętrznym pewnego odcinka ~~o jego końcach~~ $[v, z]$ o końcach $v, z \in A$. Oczywiście v, z nie mogą oba należeć do A_y, bo to przeczyłoby $w \in \mathrm{ex}\, A_y$, więc co najmniej jeden z nich leży poza A_y (czyli poza H_y). Aby $w \in H_y$ mogło być punktem wewnętrznym $[v, z]$, punkty v i z muszą leżeć po przeciwnych stronach H_y – ale H_y jest hiperpłaszczyzną podpierającą A w y, więc A leży tylko po jednej stronie H_y – sprzeczność. To, że $\mathrm{ex}\, A_z \subset \mathrm{ex}\, A$ dowodzimy dokładnie tak samo.

Stąd wynika już, że $y \in A_y = \mathrm{conv}(\mathrm{ex}\, A_y) \subset \mathrm{conv}(\mathrm{ex}\, A)$ i $z \in A_z = \mathrm{conv}(\mathrm{ex}\, A_z) \subset \mathrm{conv}(\mathrm{ex}\, A)$;

zbiór $\mathrm{conv}(\mathrm{ex}\,A)$ jest wypukły i zawiera y i z, więc $x \in [y,z] \subset \mathrm{conv}(\mathrm{ex}\,A)$. Z dowolności wyboru x mamy zatem $A \subset \mathrm{conv}(\mathrm{ex}\,A)$; odwrotne zawieranie jest oczywiste.

Załóżmy teraz dodatkowo (przy tych samych oznaczeniach), że f przyjmuje w x maksimum na A. Z wypukłości f na $[y,z]$ dostajemy od razu, że

$$\underset{\max\limits_A f}{\overset{\|}{f(x)}} \leq \max(f(y), f(z)) \leq \max_A f, \text{ skąd}$$

$$f(y) = \max_A f \text{ lub } f(z) = \max_A f.$$

Załóżmy bez straty ogólności, że $f(y) = \max\limits_A f$.
Z założenia indukcyjnego zastosowanego do A_y f przyjmuje na A_y maksimum w pewnym punkcie ekstremalnym A_y, więc $f(y) \leq \max\limits_{\mathrm{ex}\,A_y} f$, ale

$$\max_A f = f(y) \leq \max_{\mathrm{ex}\,A_y} f \leq \max_A f, \text{ więc}$$

$$\max_{\mathrm{ex}\,A_y} f = \max_A f.$$ Wiemy już jednak, że punkty ekstremalne A_y są punktami ekstremalnymi A.

To kończy dowód b).

$\square$.

Uwaga: Wystarczy prześledzić dowód, by zauważyć, że
- twierdzenie, a dokładniej punkty b) i c), zachodzi dla funkcji odpowiednio quasiwypukłych i quasiwklęsłych
- założenie ciągłości f jest potrzebne tylko po to, by zagwarantować to, że f osiąga na A maksimum (odp. minimum). Jeżeli f jest quasiwypukła i osiąga na A maksimum, to osiąga je w pewnym punkcie ekstremalnym A, bez względu na to, czy jest ciągła.

Def: Zbiór $A \subset \mathbb{R}^n$ nazywamy stożkiem, gdy $\forall_{x \in A} \forall_{\lambda \geq 0} \lambda x \in A$. Stożek, który jest zbiorem wypukłym, nazywamy stożkiem wypukłym.

Zadanko: $A \subset \mathbb{R}^n$ jest stożkiem wypukłym wtedy i tylko wtedy, gdy $\forall_{x,y \in A} \forall_{\lambda, \mu \geq 0} \lambda x + \mu y \in A$.

Def: Dla dowolnego $A \subset \mathbb{R}^n$ definiujemy stożek generowany przez A, $\operatorname{cone}(A) = \{\lambda x : \lambda \geq 0, x \in A\}$.

Zadanie: • Jeżeli $A \subset \mathbb{R}^n$ jest wypukły, to cone(A) jest stożkiem wypukłym
• cone(A) jest najmniejszym stożkiem zawierającym A, tzn. gdy $A \subset C$ i C jest stożkiem, to $\mathrm{cone}(A) \subset C$.

Def: Niech $A \subset \mathbb{R}^n$ będzie wypukły i niech $a \in A$. Stożek $F_A(a) = \mathrm{cone}\,(A - a)$ nazywamy stożkiem kierunków osiągalnych zbioru A w a.

{ Powinienem tu napisać $A - \{a\}$ ← to różnica Minkowskiego zbiorów, ale dla prostszego zapisu pomijam nawiasy { }

Twierdzenie: Następujące warunki są równoważne:

a) $v \in F_A(a)$

b) istnieje $t_0 > 0$ takie, że $\forall_{t \in [0, t_0]}\ a + tv \in A$.

Dowód: a) ⇒ b)

Gdy $v = 0$, b) jest oczywiste. Załóżmy zatem, że $v \neq 0$.

Skoro $v \in F_A(a) = \mathrm{cone}\,(A - a)$, to istnieje $\lambda \geq 0$ i $x \in A$ takie, że $v = \lambda(x - a)$, a więc $x = a + \frac{1}{\lambda} v$.

Skoro A jest wypukły, $a \in A$, $x \in A$, to $[a, x] \subset A$.

Z drugiej strony $[a, x] = \{a + tv \ : \ t \in [0, \frac{1}{\lambda}]\}$.

To dowodzi b).

b) ⇒ a) Jak poprzednio, gdy $v = 0$, a) zachodzi, możemy więc założyć, że $v \neq 0$.

Istnieje $t>0$ takie, że $a+tv\in A$, zatem

$tv\in A-a \Rightarrow v=\frac{1}{t}(tv)\in \text{cone}(A-a)=F_A(a)$,

bo $\frac{1}{t}\geq 0$, $tv\in A-a$; to dowodzi a). □.

~~Ćwiczenie~~: ~~Co~~

<u>Zadanko</u>: Część wspólna dowolnej rodziny stożków jest stożkiem.

<u>Wniosek</u>: Część wspólna dowolnej rodziny stożków wypukłych jest stożkiem wypukłym.

<u>Twierdzonko</u>: Jeżeli $A,B\subset\mathbb{R}^n$ są wypukłe i niepuste; $x\in A\cap B$, to $F_{A\cap B}(x)=F_A(x)\cap F_B(x)$

<u>Dowód</u>: Zauważmy, że $A\cap B-x=(A-x)\cap(B-x)$; $\text{cone}(A-x)\cap\text{cone}(B-x)$ jest stożkiem wypukłym zawierającym $A-x$ oraz $B-x$, a więc zawierającym $A\cap B-x$; stożek $\text{cone}(A\cap B-x)$ jest najmniejszym stożkiem wypukłym zawierającym $A\cap B-x$, zatem

$$F_{A\cap B}(x)=\text{cone}(A\cap B-x)\subset\text{cone}(A-x)\cap\text{cone}(B-x)$$
$$=F_A(x)\cap F_B(x)$$

W drugą stronę też prosto: jeżeli $v\in F_A(x)\cap F_B(x)$, to istnieją: $t_0>0$ t.ż. $\forall_{t\in[0,t_0]}\ x+tv\in A$

$t_1>0$ t.ż. $\forall_{t\in[0,t_1]}\ x+tv\in B$

więc dla $t\in[0,\min(t_0,t_1)]$ $x+tv\in A\cap B$, czyli $v\in F_{A\cap B}(x)$.

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

□.

Niech A będzie wypukły.

<u>Twierdzenie</u>: Domknięcie stożka kierunków osiągalnych $F_A(x)$ to stożek styczny $T_x A$, tzn. $\operatorname{cl} F_A(x) = T_x(A)$.

<u>Uwaga</u>: Będę przez jakiś czas używał nieco innego oznaczenia na $T_x A$; $T_x A \equiv T_A(x)$.

<u>Lemat</u>: Dla każdego $A \subset \mathbb{R}^n$ i $x \in A$ stożek $T_A(x)$ jest domknięty.

<u>Dowód</u>: Jeżeli $x \notin \operatorname{Acc} A$, to $T_A(x) = \{0\}$ i teza zachodzi.

Załóżmy zatem, że $x \in \operatorname{Acc} A$ i niech $0 \neq v \in \operatorname{cl} T_A(x)$, a więc $0 \neq v$ jest granicą wektorów z $T_A(x)$, $v = \lim_{i \to \infty} v_i$, $v_i \in T_A(x)$ dla $i = 1, 2, \ldots$ Możemy założyć, że $\forall_i\ v_i \neq 0$ (jest tak dla dost. dużych i; pozostałe wyrazy ciągu możemy odrzucić).

Skoro $\overset{0}{\underset{\#}{}}\!\!v_i \in T_A(x)$, to istnieje ciąg $(y_{i,j})_{j \in \mathbb{N}}$ taki, że $y_{i,j} \in A \setminus \{x\}$, $\frac{y_{i,j} - x}{\|y_{i,j} - x\|} \xrightarrow{j \to \infty} \frac{v_i}{\|v_i\|}$, a więc w szczególności istnieje $j_i \in \mathbb{N}$ takie, że $\left\| \frac{y_{i,j_i} - x}{\|y_{i,j_i} - x\|} - \frac{v_i}{\|v_i\|} \right\| < \frac{1}{i}$. Rozważmy zatem ciąg $(y_{i,j_i})_{i \in \mathbb{N}}$:

$$\left\| \frac{y_{i,j_i} - x}{\|y_{i,j_i} - x\|} - \frac{v}{\|v\|} \right\| \leq \left\| \frac{y_{i,j_i} - x}{\|y_{i,j_i} - x\|} - \frac{v_i}{\|v_i\|} \right\| + \left\| \frac{v_i}{\|v_i\|} - \frac{v}{\|v\|} \right\| <$$

$$< \frac{1}{i} + \frac{v_i \|v\| - v \|v_i\|}{\|v_i\| \cdot \|v\|} \xrightarrow[i \to \infty]{} 0 \quad \Big\} \begin{array}{l} \text{tu ważne jest,} \\ \text{że } v_i \to v \neq 0. \end{array}$$

Ten rachunek dowodzi, że $v \in T_A(x)$, a zatem $\mathrm{cl}\, T_A(x) \setminus \{0\} \subset T_A(x)$. Ale oczywiście $0 \in T_A(x)$, więc $\mathrm{cl}\, T_A(x) \cancel{=} \subset T_A(x)$, skąd $T_A(x) = \mathrm{cl}\, T_A(x)$.

□.

<u>Dowód twierdzenia</u>

Jeżeli $0 \neq v \in F_A(x)$, to dla dużych k $\overbrace{x + \frac{1}{k} v}^{y_k \,=} \in A$, więc $\frac{y_k - x}{\|y_k - x\|} = \frac{v}{\|v\|} \longrightarrow \frac{v}{\|v\|}$, skąd $F_A(x) \setminus \{0\} \subset T_A(x)$; oczywiście $0 \in T_A(x) \Rightarrow F_A(x) \subset T_A(x)$. Stąd i z domkniętości $T_A(x)$ wynika $\mathrm{cl} F_A(x) \subset T_A(x)$.

Niech teraz $v \in T_A(x)$, $v \neq 0$. Istnieją wówczas $y_k \in A \setminus \{x\}$ takie, że $\frac{y_k - x}{\|y_k - x\|} \longrightarrow \frac{v}{\|v\|}$.

Zauważmy, że dla $t \in [0, \|y_k - x\|]$ punkty $x + t \frac{y_k - x}{\|y_k - x\|} \in [x, y_k] \subset A$

bo $x, y_k \in A$ i A jest wypukły.

Stąd $\frac{y_k - x}{\|y_k - x\|} \in F_A(x)$, a zatem $\frac{v}{\|v\|} \in \mathrm{cl}\, F_A(x)$, a że domknięcie stożka jest stożkiem, $v \in \mathrm{cl}\, F_A(x)$.

Dowiedliśmy zatem, że $T_A(x) \setminus \{0\} \subset \mathrm{cl}\, F_A(x)$; oczywiście $0 \in \mathrm{cl}\, F_A(x)$, więc $T_A(x) \subset \mathrm{cl}\, F_A(x)$.

□.

Def. Niech $A \subset \mathbb{R}^n$ będzie stożkiem.
Zbiór $A^\circ = \{ y \in \mathbb{R}^n : \langle x, y \rangle \leq 0 \quad \forall_{x \in A} \}$ nazywamy **stożkiem polarnym** stożka A.

Uwaga: A° jest domkniętym stożkiem wypukłym.

Dowód: • A° jest stożkiem, bo jeżeli $y \in A^\circ$, $\lambda \geq 0$,
to $\forall_{x \in A}$ $\langle \cancel{y}, x, \lambda y \rangle = \underset{\geq 0}{\lambda} \underset{\leq 0}{\langle x, y \rangle} \leq 0$, więc $\lambda y \in A^\circ$.

• A° jest wypukły, bo jeżeli $y_1, y_2 \in A^\circ$, $\lambda, \mu \geq 0$,
to $\forall_{x \in A}$ $\langle x, \lambda y_1 + \mu y_2 \rangle = \underset{\geq 0}{\lambda} \underset{\leq 0}{\langle x, y_1 \rangle} + \underset{\geq 0}{\mu} \underset{\leq 0}{\langle x, y_2 \rangle} \leq 0$,
czyli $\lambda y_1 + \mu y_2 \in A^\circ$.

• A° jest domknięty. Jeżeli bowiem $y_k \in A^\circ$, $y = \lim_{k \to \infty} y_k$,
to $\forall_{x \in A}$ $\langle y_k, x \rangle \leq 0$, więc $\langle y, x \rangle \leq 0 \Rightarrow y \in A^\circ$.

$$\langle y_k, x \rangle \xrightarrow{k \to \infty} \langle y, x \rangle$$

$\square$.

Def: Niech $A \subset \mathbb{R}^n$ będzie wypukły, $x \in A$.
Stożek $F_A(x)^\circ$ nazywamy **stożkiem normalnym** zbioru A w x i oznaczamy $N_A(x)$.

Twierdzenie: Niech $A \subset \mathbb{R}^n$ będzie zbiorem wypukłym. Wówczas $T_A(x)^\circ = N_A(x)$ i $N_A(x)^\circ = T_A(x)$.

Nim przejdziemy do dowodu, zauważmy następujący

Wniosek: Jeżeli $A \subset \mathbb{R}^n$ jest stożkiem wypukłym, to $(A^\circ)^\circ = \mathrm{cl}\, A$.

Dowód wniosku: $F_A(0) = A$ (trywialne ćwiczenie), więc $N_A(0) = F_A(0)^\circ = A^\circ$, $N_A(0)^\circ = (A^\circ)^\circ$

$$\underset{T_A(0) = \mathrm{cl} F_A(0) = \mathrm{cl}\, A.}{\overset{N_A(0)^\circ}{\|}} \quad \square$$

Dowód twierdzenia

Zauważmy na początek, że $T_A(x)^\circ \subset N_A(x)$.

Mamy bowiem

$$T_A(x)^\circ = \{y \in \mathbb{R}^n : \langle y, z\rangle \leq 0 \ \forall_{z \in T_A(x)}\} \overset{\text{bo } F_A(x) \subset T_A(x)}{\subset} \{y \in \mathbb{R}^n : \langle y, z\rangle \leq 0 \ \forall_{z \in F_A(x)}\} = F_A(x)^\circ = N_A(x).$$

Załóżmy, że $T_A(x)^\circ \neq N_A(x)$, a więc że istnieje $y \in N_A(x) \setminus T_A(x)^\circ$. Skoro $y \notin T_A(x)^\circ$, to dla pewnego $z \in T_A(x)$ $\langle y, z\rangle > 0$. Wiemy jednak, że $T_A(x) = \mathrm{cl}\, F_A(x)$, a więc istnieje ciąg $z_k \in F_A(x)$ t.ż $z_k \to z$; $\underset{0}{\overset{\langle z_k, y\rangle}{\wedge|}} \xrightarrow[k\to\infty]{} \langle z, y\rangle > 0$ ↯ sprzeczność.

bo $z_k \in F_A(x)$,
$y \in N_A(x) = F_A(x)^\circ$

Stąd $T_A(x)^\circ = N_A(x)$.

Skoro $N_A(x) = F_A(x)^\circ$, to $\forall_{y \in F_A(x)} \forall_{z \in N_A(x)} \langle y, z \rangle \leq 0$

To oznacza, że $F_A(x) \subset N_A(x)^\circ$. Wiemy też, że $N_A(x)^\circ$ jest domknięty (każdy stożek polarny jest domknięty) $\Longrightarrow T_A(x) = \mathrm{cl}\, F_A(x) \subset N_A(x)^\circ$.

Załóżmy, że istnieje $y \in N_A(x)^\circ \setminus T_A(x)$.
Możemy wówczas oddzielić y od $T_A(x)$ hiperpłaszczyzną: istnieje $v \in \mathbb{R}^n$ takie, że $\forall_{z \in T_A(x)}$

$$\langle z, v \rangle < \langle y, v \rangle \qquad (*)$$

Wykażemy, że $\forall_{z \in T_A(x)} \langle z, v \rangle \leq 0$. Gdyby bowiem dla pewnego $z_0 \in T_A(x)$ było $\langle z_0, v \rangle > 0$, to biorąc w $(*)$ $z = \lambda z_0$ i przyjmując duże $\lambda > 0$ moglibyśmy lewą stronę nierówności $(*)$ uczynić dowolnie dużą (a jest ona ograniczona z góry przez $\langle y, v \rangle$).
Stąd $\forall_{z \in T_A(x)} \langle z, v \rangle \leq 0 \Rightarrow v \in T_A(x)^\circ = N_A(x)$.

Mamy jednak wtedy, biorąc w $(*)$ $z = 0$, że $\langle y, v \rangle > 0$, choć $v \in N_A(x)$, $y \in N_A(x)^\circ$. ↯

To dowodzi, że $N_A(x)^\circ = T_A(x)$.

$\square$.

Ważnym motorem rozwoju geometrii wypukłej było badanie układów nierówności liniowych.

<u>Twierdzenie o alternatywie</u> (lemat Farkasa):

Gyula Farkas
1847 – 1930
matematyk i fizyk węgierski

Niech A będzie macierzą $m \times n$ i niech $b \in \mathbb{R}^m$.
Zachodzi wówczas dokładnie jedna z dwóch możliwości:

albo a) równanie $Ax = b$ ma rozwiązanie x o wszystkich współrzędnych nieujemnych (będziemy to oznaczać $x \geq 0$)

albo b) istnieje rozwiązanie układu nierówności

$$y^T A \geq 0 \; , \quad \langle y, b \rangle < 0$$

tu znów umawiamy się, że ta nierówność zachodzi po współrzędnych, tzn. $(y^T A)_i \geq 0$ dla wszystkich i.

<u>Dowód</u>: Załóżmy, że zachodzi a) i równocześnie zachodzi b). Wówczas

$$0 > \langle y, b \rangle = \langle y, Ax \rangle = y^T A x = \langle \underline{y^T A}, \underline{x} \rangle$$

ten wektor ma wszystkie współrzędne nieujemne

i ten też, więc ich iloczyn skalarny nie może być ujemny. ↯.

Stąd a) $\Rightarrow \sim$ b).

Załóżmy, że nie zachodzi a). Oznaczmy kolumny macierzy A przez $A_1, \dots, A_n$ i niech $X = \{ \sum_{i=1}^{n} \lambda_i A_i : \lambda_i \geq 0 \}$ będzie stożkiem wypukłym rozpiętym przez wektory $A_1, \dots, A_n$.

To, że nie zachodzi a), a więc że $\sum_{i=1}^{n} x_i A_i = Ax = b$ nie ma rozwiązania $x \geq 0$ oznacza, że $b \notin X$.

Możemy zatem oddzielić b od X hiperpłaszczyzną: istnieją $y \in \mathbb{R}^m$ i $c \in \mathbb{R}$ takie, że

$$\forall_{x \in X} \quad \langle y, b\rangle < c \leq \langle y, x\rangle \qquad (*)$$

Biorąc $x = 0$ widzimy, że $\langle y, b\rangle < 0$. Pozostaje sprawdzić, czy $y^T A \geq 0$.

Zauważmy, że $y^T A = (\langle y, A_i\rangle)_{i=1,\dots,n}$. Gdyby dla pewnego $i \in \{1,\dots,n\}$ $\langle y, A_i\rangle < 0$, to biorąc w $(*)$ $x = \lambda A_i$ (oczywiście $\lambda A_i \in X$ $\forall_{\lambda \geq 0}$) otrzymalibyśmy

$$(**) \quad c \leq \underset{\substack{\vee\!\!|\\ 0}}{\lambda} \underbrace{\langle y, A_i\rangle}_{<0}$$

i prawą stronę tej nierówności możemy, biorąc duże $\lambda > 0$, uczynić dowolnie małą, więc $(**)$ nie może zachodzić.

Stąd $y^T A \geq 0$ i b) zachodzi, a więc $\sim a) \Rightarrow b)$

□.

<u>Wniosek</u> (lemat Gordana)

Paul Gordan 1837 – 1912
niemiecki matematyk pochodzenia żydowskiego urodzony we Wrocławiu, uczeń Jacobiego, promotor Emmy Noether.

Niech M będzie macierzą $m \times n$.

Albo a) istnieje rozwiązanie x układu nierówności $Mx < 0$,

albo b) istnieje <u>niezerowe</u> rozwiązanie y układu nierówności $y \geq 0$, $y^T M = 0$.

<u>Dowód</u>: Oznaczmy $\mathbb{1} = (1,1,\dots,1) \in \mathbb{R}^n$.

Zauważmy, że układ nierówności $(\cdot)$ $Mx > 0$ ma rozwiązanie wtedy i tylko wtedy, gdy rozwiązanie ma układ nierówności

$(\cdot\cdot)$ $Mx + s\cdot\mathbb{1} \leq 0,\ s > 0 \qquad (s \in \mathbb{R})$,

a ten z kolei można zapisać jako

$(\because)$ $(M, \mathbb{1})\binom{x}{s} \leq 0, \quad (0,\dots,0,1)\binom{x}{s} > 0$

Wprowadźmy nowe oznaczenia: niech $M' = (M,\mathbb{1})^T$, $y = -\binom{x}{s}$, $b = (0,\dots,0,1)$. Transponując pierwszą nierówność w $(\because)$ stronami możemy $(\because)$ zapisać jako

$(::)$ $y^T M' \geq 0, \quad \langle y, b\rangle < 0$

i układ $(\cdot)$ ma rozwiązanie x wtedy i tylko wtedy, gdy $(::)$ ma rozwiązanie y. Z twierdzenia o alternatywie $(::)$ ma rozwiązanie wtedy i tylko wtedy, gdy układ $(:\cdot:)$ $M'x = b,\ x \geq 0$ nie ma rozwiązania. Wracając do poprzednich oznaczeń, z $(:\cdot:)$ otrzymujemy

$(:::)$ $x^T M = 0, \quad x^T\cdot\mathbb{1} = 1, \quad x \geq 0$

$\left\{ \begin{array}{l} \text{a) } (\cdot) \text{ ma rozwiązanie} \\ \Updownarrow \\ (:::) \text{ nie ma rozwiązania.} \end{array} \right.$

Jeżeli $(:::)$ ma rozwiązanie, to $x \neq 0$, bo $x^T\mathbb{1} = 1$, więc układ b) ma rozwiązanie niezerowe. Jeżeli zaś b) ma rozwiązanie niezerowe $y \geq 0$, to $y^T\mathbb{1} > 0$ i biorąc $x = \frac{y}{\langle y, \mathbb{1}\rangle}$ dostajemy rozwiązanie $(:::)$. Stąd b) zachodzi $\Leftrightarrow$ $(:::)$ ma rozwiązanie $\Leftrightarrow$ $(\cdot)$ nie ma rozw.

Naszym celem jest dowód następującego twierdzenia, udowodnionego w latach 30' XX wieku przez Williama Karusha (1917-1997) (w jego pracy magisterskiej), a ponownie odkrytego w 1951 roku przez Harolda Kuhna (1925-2014) i Alberta Tuckera (1905-1995).

Twierdzenie (warunki konieczne KKT)

Niech $\mathcal{U} \subset \mathbb{R}^n$ będzie otwarty,

$f, g_i, h_j : \mathcal{U} \to \mathbb{R}$, $i = 1, \dots, k$, $j = 1, \dots, l$ niech będą różniczkowalne, klasy C^1,

niech $S = \{x \in \mathcal{U} : g_i(x) \leq 0,\ h_j(x) = 0,\ i = 1, \dots, k,\ j = 1, \dots, l\}$

i załóżmy, że dla pewnego $x_0 \in S$

- f przyjmuje w x_0 maksimum (odp. minimum) lokalne na S

warunki jakości więzów

- $\nabla h_1(x_0), \dots, \nabla h_l(x_0)$ są liniowo niezależne w x_0
- spełniony jest warunek Mangasariana-Fromovitza:

układ równań i nierówności

$$\begin{cases} \langle v, \nabla h_j(x_0) \rangle = 0 & j = 1, \dots, l \\ \langle v, \nabla g_i(x_0) \rangle < 0 & \text{dla wszystkich } i \in \{1, \dots, k\} \text{ takich, że } g_i(x_0) = 0. \end{cases}$$

ma rozwiązanie

Wówczas istnieją mnożniki Lagrange'a $\lambda_1, \dots, \lambda_k \geq 0$ (odp. $\lambda_1, \dots, \lambda_k \leq 0$) oraz $\mu_1, \dots, \mu_l \in \mathbb{R}$ takie, że spełniony jest warunek Lagrange'a

$$\nabla f(x_0) = \sum_{i=1}^{k} \lambda_i \nabla g_i(x_0) + \sum_{j=1}^{l} \mu_j \nabla h_j(x_0)$$

oraz $\lambda_i g_i(x_0) = 0$ dla $i = 1, \dots, k$.

Na początek trochę terminologii: mówimy, że więz g_i jest aktywny w $x_0 \in S$, gdy $g_i(x_0)=0$. Oznaczmy przez $I(x_0)$ zbiór tych wszystkich indeksów $i \in \{1,\dots,n\}$, dla których więz g_i jest aktywny w x_0.
Ostatni warunek ($\lambda_i g_i(x_0) = 0$ dla $i=1,\dots,n$) oznacza, że mnożniki odpowiadające nieaktywnym więzom w x_0 są równe 0; oznacza to, że warunek Lagrange'a możemy zapisać jako

$$\nabla f(x_0) = \sum_{i\in I(x_0)} \lambda_i \nabla g_i(x_0) + \sum_{j=1}^{\ell} \mu_j \nabla h_j(x_0),$$

a warunek Mangasariana-Fromovitza przeformułować do

$$\begin{cases} \langle v, \nabla h_j(x_0)\rangle = 0 & j=1,\dots,\ell \\ \langle v, \nabla g_i(x_0)\rangle < 0 & i\in I(x_0) \end{cases}$$ ma rozwiązanie.

Nim przejdziemy do dowodu warunków KKT, sformułujmy, w terminach stożków, warunek konieczny na to, by $f: U \to \mathbb{R}$, różniczkowalna w punkcie $x_0 \in A \subset U$, przyjmowała w tym punkcie ekstremum lokalne na A.

<u>Lemat</u>: Niech $U \subset \mathbb{R}^n$ będzie otwarty, $A \subset U$ i załóżmy, że funkcja $f: U \to \mathbb{R}$ przyjmuje w $x_0 \in A$ maksimum lokalne (odp. minimum lokalne) na A. Jeżeli f jest różniczkowalna w x_0, to

$$\nabla f(x_0) \in N_A(x_0) = T_A(x_0)^\circ \quad (\text{odp. } -\nabla f(x_0) \in N_A(x_0)).$$

Dowód: Załóżmy, że f przyjmuje w x_0 maksimum lokalne na A, jest w x_0 różniczkowalna, ale $\nabla f(x_0) \notin N_A(x_0)$. To oznacza, że dla pewnego wektora $v \in T_A(x_0)$ zachodzi $\langle \nabla f(x_0), v \rangle > 0$. Z nierówności tej wynika, że $v \neq 0$, więc istnieją $y_k \in A \setminus \{x_0\}$, $y_k \to x_0$, takie, że $\frac{y_k - x_0}{\|y_k - x_0\|} \to \frac{v}{\|v\|}$.

Wtedy $\underset{\text{dla dost. dużych } k}{f(y_k) - f(x_0) \leq 0}$ (bo f ma w x_0 maks. lokalne), więc

$$0 \geq \frac{f(y_k) - f(x_0)}{\|y_k - x_0\|} = \frac{Df(x_0)(y_k - x_0)}{\|y_k - x_0\|} + \frac{R(y_k - x_0)}{\|y_k - x_0\|}$$

$$= \langle \nabla f(x_0), \frac{y_k - x_0}{\|y_k - x_0\|} \rangle \text{ + reszta} \xrightarrow{k \to \infty} \frac{\langle \nabla f(x_0), v \rangle}{\|v\|} > 0 \ \lightning.$$

Stąd $\nabla f(x_0) \in N_A(x_0)$.

Dla minimum lokalnego dowód jest taki sam.

Przyjrzyjmy się teraz zbiorowi S z twierdzenia KKT.

Mamy $S = \{x \in \mathcal{U} : g_i(x) \leq 0,\ h_j(x) = 0,\ \begin{smallmatrix} i = 1, \dots, k \\ j = 1, \dots, \ell \end{smallmatrix} \} =$

$$= \bigcap_{i=1}^{k} G_i \cap H, \qquad \text{gdzie } G_i = \{x \in \mathcal{U} : g_i(x) \leq 0\}$$
$$H = \{x \in \mathcal{U} : h_j(x) = 0,\ j = 1, \dots, \ell\}.$$

Jeżeli spełniony jest warunek jakości więzów, dotyczący więzów równościowych h_i, tzn. że gradienty $Dh_j(x_0)$, $j = 1, \dots, \ell$, są liniowo niezależne, to zbiór H, jako poziomica funkcji $h=(h_1,\dots,h_l)$, jest w otoczeniu punktu x_0 rozmaitością, $T_{x_0}H = T_H(x_0) = \ker Dh$ jest $n-l$ wymiarową przestrzenią liniową, prostopadłą do $\nabla h_1(x_0),\dots,\nabla h_l(x_0)$, a stożek normalny do H w x_0 też jest podprzestrzenią liniową, tym razem l wymiarową, rozpiętą przez $\nabla h_1(x_0),\dots,\nabla h_l(x_0)$.

<u>Zadanko:</u> Jeżeli $A\subset\mathbb{R}^n$ jest podprzestrzenią liniową, to $A^\circ = A^\perp$.

Trochę trudniej ze stożkami: stycznym i normalnym do G_i w x_0.

Jeżeli $g_i(x_0)<0$ (tzn. więz g_i jest nieaktywny w x_0), to $x_0\in \operatorname{int} G_i$, $T_{G_i}(x_0)=\mathbb{R}^n$, $N_{G_i}(x_0)=\{0\}$.

Co, jeżeli $g_i(x_0)=0$? Zauważmy, że nie musi to oznaczać, że $x_0\in\partial G_i$, jeżeli np. $g_i(x) = -\sum_{j=1}^n x_j^2$, to $g_i(0)=0$, ale $G_i=\mathbb{R}^n$. Jeżeli jednak $x_0\in \operatorname{int}G_i$, to wiemy już, że $N_{G_i}(x_0)=\{0\}$; wiemy też, że wówczas g_i ma w x_0 maksimum lokalne (na $\mathbb{R}^n$), więc $\nabla g_i(x_0)=0$. W tym przypadku zatem $N_{G_i}(x_0)=\{\lambda\nabla g_i(x_0):\lambda\ge 0\}$.

Jeżeli zaś $\nabla g_i(x_0)\neq 0$, to $x_0\in\partial G_i$, bo g_i rośnie w kierunku $v=\nabla g_i(x_0)$ i $g_i(x_0+tv)>g_i(x_0)$ dla dostatecznie małych $t>0$.

Wykażemy, że gdy $\nabla g_i(x_0) \neq 0$, to i w tym przypadku $N_{G_i}(x_0) = \{\lambda \nabla g_i(x_0) : \lambda \geq 0\}$. Oznaczmy bowiem $V = \{\lambda \nabla g_i(x_0) : \lambda \geq 0\}$. V jest stożkiem wypukłym, domkniętym. Niech teraz $v \in V^o$, tzn. $\langle v, \nabla g_i(x_0)\rangle \leq 0$. Wtedy dla każdego $k \in \mathbb{N}$ $\langle v - \frac{1}{k}\nabla g_i(x_0), \nabla g_i(x_0)\rangle < 0$, jeżeli więc oznaczymy $v_k = v - \frac{1}{k}\nabla g_i(x_0)$, to

$$g_i(x_0 + \tfrac{1}{l} v_k) = g_i(x_0) + Dg_i(x_0)(\tfrac{1}{l} v_k) + R(\tfrac{1}{l} v_k) =$$

$$= \cancel{0} \tfrac{1}{l}\Big(\underbrace{\langle \nabla g_i(x_0), v_k\rangle}_{<0} + \underbrace{\frac{R(\frac{1}{l} v_k)}{\|\frac{1}{l} v_k\|}}_{\to 0} \|v_k\|\Big)$$

i dla dużych l znak nawiasu jest ujemny. Stąd $x_0 + \frac{1}{l} v_k \in G_i$ dla dużych l, i od razu dostajemy, że $v_k \in T_{G_i}(x_0)$.

Z drugiej strony $T_{G_i}(x_0)$ jest domknięty, $v_k \xrightarrow{k\to\infty} v$, więc $v \in T_{G_i}(x_0)$. Wykazaliśmy zatem, że $V^o \subset T_{G_i}(x_0)$.

W drugą stronę: wiemy, że g_i ma w x_0 maksimum lokalne na G_i, więc $\nabla g_i(x_0) \in N_{G_i}(x_0) = T_{G_i}(x_0)^o$. Tym samym $\forall_{v \in T_{G_i}(x_0)}\ \forall_{w \in V}\ \langle w, v\rangle \leq 0 \Rightarrow T_{G_i}(x_0) \subset V^o$.

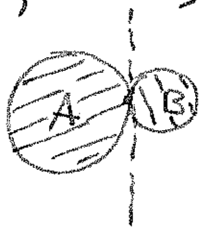
Stąd $T_{G_i}(x_0) = V^o$, ale V jest stożkiem wypukłym domkniętym, więc $N_{G_i}(x_0) = T_{G_i}(x_0)^o = (V^o)^o = V$.

Ostatecznie więc: jeżeli $g_i(x_0) < 0$, to $N_{G_i}(x_0) = \{0\}$,
jeżeli $g_i(x_0) = 0$, to $N_{G_i}(x_0) = \{\lambda \nabla g_i(x_0) : \lambda \geq 0\}$

i $T_{G_i}(x_0) = V^o = N_{G_i}(x_0)^o$.

Naszym celem jest ustalenie, jak wygląda stożek normalny do $S = H \cap \bigcap_{i=1}^{k} G_i$. Znamy już postać stożka normalnego do każdej ze składowych...

Wiemy już, że gdy $x \in A \cap B$ i zbiory A, B są wypukłe, to $F_A(x) \cap F_B(x) = F_{A\cap B}(x)$. Niestety dla stożków stycznych tak już być nie musi: jeżeli A i B są parą zewnętrznie stycznych kół na płaszczyźnie,



$x_0 \in A \cap B$
$(\{x_0\} = A \cap B)$,

to $T_{A\cap B}(x_0) = \{0\}$, podczas gdy $T_A(x_0) \cap T_B(x_0)$ jest prostą styczną do obu kół w x_0.

A co ze stożkami normalnymi?

<u>Twierdzenie</u>: Niech $C_1, C_2 \subset \mathbb{R}^n$ będą stożkami wypukłymi domkniętymi. Wówczas $(C_1 \cap C_2)^\circ = C_1^\circ + C_2^\circ$

suma Minkowskiego

<u>Dowód</u>: Załóżmy, że $v \in C_1^\circ + C_2^\circ$.

Wówczas $v = u + w$, gdzie $u \in C_1^\circ$, $w \in C_2^\circ$, a więc

$$\forall_{x \in C_1 \cap C_2} \quad \langle u, x \rangle \leq 0, \text{ bo } u \in C_1^\circ,\ x \in C_1$$

$$+ \quad \langle w, x \rangle \leq 0, \text{ bo } w \in C_2^\circ,\ x \in C_2$$

$$\forall_{x \in C_1 \cap C_2} \quad \langle v, x \rangle \leq 0, \text{ a więc } v \in (C_1 \cap C_2)^\circ.$$

Tak więc $C_1^\circ + C_2^\circ \subset (C_1 \cap C_2)^\circ$.

Załóżmy teraz, że $C_1^\circ + C_2^\circ \subsetneq (C_1 \cap C_2)^\circ$, a więc że istnieje $y \in (C_1 \cap C_2)^\circ \setminus C_1^\circ + C_2^\circ$.

Możemy wtedy oddzielić y od $C_1^\circ + C_2^\circ$ hiperpłaszczyzną: istnieje $w \in \mathbb{R}^n$ takie, że

$$\underline{\forall_{v \in C_1^\circ + C_2^\circ} \ \langle w, v \rangle \le 0}, \qquad \underline{\langle w, y \rangle > 0}$$

$$\Downarrow$$

$$\forall_{v \in C_1^\circ} \ \langle w, v \rangle \le 0 \Rightarrow w \in (C_1^\circ)^\circ = \operatorname{cl} C_1 = C_1$$

oraz

$$\forall_{v \in C_2^\circ} \ \langle w, v \rangle \le 0 \Rightarrow w \in (C_2^\circ)^\circ = \operatorname{cl} C_2 = C_2$$

Stąd $w \in C_1 \cap C_2$. Z drugiej strony $y \in (C_1 \cap C_2)^\circ$, ale $\langle w, y \rangle > 0$ – sprzeczność.

To dowodzi, że $C_1^\circ + C_2^\circ = (C_1 \cap C_2)^\circ$. $\square$

~~Wniosek: Jeżeli $T_{A \cap B}(x) = T_A(x) \cap T_B(x)$, to $N_{A \cap B}(x) = T_{A \cap B}$~~

W ogólności stożki styczne do zbiorów, które nie są wypukłe, nie muszą być stożkami wypukłymi. W naszej jednak sytuacji, choć zbiory G_i nie muszą być wypukłe, to, jak to już ustaliliśmy, $T_{G_i}(x_0)$ i $N_{G_i}(x_0)$ są stożkami wypukłymi i spełniają $T_{G_i}(x_0)^\circ = N_{G_i}(x_0)$ i $N_{G_i}(x_0)^\circ = T_{G_i}(x_0)$.

Jeżeli więc dla punktu $x_0 \in G_i \cap G_j$ skądinąd wiemy, że $T_{G_i \cap G_j}(x_0) = T_{G_i}(x_0) \cap T_{G_j}(x_0)$, to z Lematu

$$N_{G_i \cap G_j}(x_0) = T_{G_i \cap G_j}(x_0)^\circ = \left(T_{G_i}(x_0) \cap T_{G_j}(x_0)\right)^\circ =$$

$$= T_{G_i}(x_0)^\circ + T_{G_j}(x_0)^\circ = N_{G_i}(x_0) + N_{G_j}(x_0)$$

i ogólniej, jeżeli $T_{\bigcap_{i=1}^{k} G_i}(x_0) = \bigcap_{i=1}^{k} T_{G_i}(x_0)$,

to $N_{\bigcap_{i=1}^{k} G_i}(x_0) = \sum_{i=1}^{k} N_{G_i}(x_0)$. $\Big\{$ dowód przez oczywistą indukcję.

$\uparrow$ suma Minkowskiego.

<u>Zadanie</u>: Podaj przykład $A, B \subset \mathbb{R}^n$ wypukłych takich, że $A \cap B \neq \emptyset_*$ i dla pewnego $x_0 \in A \cap B$ $N_{A \cap B}(x_0) \neq N_A(x_0) + N_B(x_0)$.

Lemat: Załóżmy, że $g_1,\dots,g_k$ i $h_1,\dots,h_l$ spełniają w $x_0\in S$ warunki jakości więzów, tzn.

- $\nabla h_1(x_0),\dots,\nabla h_l(x_0)$ są liniowo niezależne
- istnieje rozwiązanie v układu

$$(*)\ \begin{cases} \langle v, \nabla h_j(x_0)\rangle = 0 & j=1,\dots,l \\ \langle v, \nabla g_i(x_0)\rangle \leq 0 & i\in I(x_0) \end{cases}$$

$\leftarrow$ te i, dla których $g_i(x_0)=0$.

bo $T_{G_i}(x_0)=\mathbb{R}^n$ gdy $i\notin I(x_0)$

Wówczas $T_S(x_0)=\bigcap_{i=1}^{k} T_{G_i}(x_0)\cap T_H(x_0) = \bigcap_{i\in I(x_0)} T_{G_i}(x_0)\cap T_H(x_0)$

Wniosek: Jeżeli spełnione są założenia Lematu, to $N_S(x_0)=\sum_{i=1}^{n} N_{G_i}(x_0)+N_H(x_0) = \sum_{i\in I(x_0)} N_{G_i}(x_0)+N_H(x_0)$

$N_{G_i}(x_0)=\{0\}$ gdy $i\notin I(x_0)$.

Dowód Lematu

Oznaczmy przez K zbiór wszystkich $v\in\mathbb{R}^n$ spełniających $(*)$.

Z założenia $K\neq\emptyset$.

Niech $\overline{K}=\bigcap_{i=1}^{k} T_{G_i}(x_0)\cap T_H(x_0)=$

$$=\left\{ v\in\mathbb{R}^n : \begin{array}{ll} \langle v,\nabla g_i(x_0)\rangle \leq 0 & \text{dla } i\in I(x_0), \\ \langle v,\nabla h_j(x_0)\rangle = 0 & \text{dla } j=1,\dots,l \end{array} \right\}$$

Chcemy ustalić związki między K, $\overline{K}$ i $T_S(x_0)$.

Oczywiście jeżeli $v\in T_S(x_0)$, to w szczególności $v\in T_{G_i}(x_0)$ dla $i=1,\dots,k$ i $v\in T_H(x_0)$, więc $T_S(x_0)\subset\overline{K}$.

Teraz wykażemy, że $K\subset T_S(x_0)$. Niech $v\in K$.

Wtedy $v\in T_H(x_0)=\{w\in\mathbb{R}^n : \langle w,\nabla h_j(x_0)\rangle = 0 \text{ dla } j=1,\dots,l\}$, i $v\neq 0$, więc istnieje ciąg punktów $y_m\in H\setminus\{x_0\}$ takich, że $\frac{y_m-x_0}{\|y_m-x_0\|}\to\frac{v}{\|v\|}$. Warunek $\langle v,\nabla g_i(x_0)\rangle<0$, $i\in I(x_0)$,

oznacza, że $D_v g_i(x_0)<0$, skąd łatwo możemy wywnioskować (szczegóły Państwu zostawiam jako ćwiczenie), że dla dużych m $g_i(y_m)<0$ (dla $i\in I(x_0)$).
Dla $i\notin I(x_0)$ wiemy, że $g_i(x_0)<0$, więc również $g_i(y_m)<0$ dla dużych m, bo $y_m\to x_0$.
Oznacza to, że $y_m\in S=H\cap\bigcap_{i=1}^{k}G_i$, a więc $v\in T_S(x_0)$.

Na koniec wykażemy, że $cl\,K=\overline{K}$.
Niech bowiem $y\in\overline{K}$ i niech $w\in K$ (tu korzystamy z założenia, że $K\neq\emptyset$)
Wtedy $y_m=y+\frac{1}{m}w\in K$, bo $\langle y_m,\nabla g_i(x_0)\rangle=$

$$=\underset{\overset{\wedge\!\shortparallel}{0}}{\langle y,\nabla g_i(x_0)\rangle}+\frac{1}{m}\underset{\overset{\wedge}{0}}{\langle w,\nabla g_i(x_0)\rangle}<0 \text{ dla } i\in I(x_0),$$

$y_m\to y$, więc $\overline{K}\subset cl\,K$, więc $K\subset\overline{K}\subset cl\,K$, ale $\overline{K}$ jest domknięty $\Rightarrow$ $\overline{K}=cl\,K$.
Ostatecznie $K\subset T_S(x_0)\subset\overline{K}=cl\,K$, ale stożek $T_S(x_0)$ jest domknięty (wykazaliśmy, że każdy stożek styczny jest domknięty), więc $T_S(x_0)=\overline{K}$. □.

No i w zasadzie udowodniliśmy twierdzenie KKT: jeżeli f przyjmuje w x_0 maksimum lokalne na S, to $\nabla f(x_0)\in N_S(x_0)\underset{\text{na mocy Wniosku}}{=}\sum_{i=1}^{n}N_{G_i}(x_0)+N_H(x_0)=$

$$=\left\{v\in\mathbb{R}^n : v=\sum_{i\in I(x_0)}\lambda_i\nabla g_i(x_0)+\sum_{j=1}^{l}\mu_j\nabla h_j(x_0)\right\},$$

a to jest właśnie warunek Lagrange'a. □.

Na koniec przyjrzyjmy się warunkowi Mangasariana-Fromovitza. Wygląda na dość trudny do sprawdzenia, ale poniższy lemat pokazuje, że jest spełniony w naturalnych sytuacjach:

<u>Lemat:</u> Jeżeli w punkcie $x_0 \in S$ gradienty $\nabla h_j(x_0)$, $j = 1, \dots, l$ oraz $\nabla g_i(x_0)$, $i \in I(x_0)$ są liniowo niezależne, to w x_0 spełniony jest warunek M.-F.

<u>Dowód</u>

Użyjemy lematu Gordana na przestrzeni liniowej $T_H(x_0)$: niech $M = (\nabla g_{i_1}(x_0), \dots, \nabla g_{i_m}(x_0))$,

$\{i_1, \dots, i_m\} = I(x_0)$

na przestrzeni $T_H(x_0)$ równanie $y^T M = 0$ nie ma rozwiązań niezerowych (bo kolumny M są liniowo niezależne), więc istnieje na $T_H(x_0)$ rozwiązanie układu nierówności

$$Mv < 0$$

$$= (\langle \nabla g_{i_1}(x_0), v \rangle, \langle \nabla g_{i_2}(x_0), v \rangle, \dots, \langle \nabla g_{i_m}(x_0), v \rangle),$$

tzn. istnieje $v \in \mathbb{R}^n$ spełniające

- $\langle v, \nabla h_j(x_0) \rangle = 0$ dla $j = 1, \dots, l$ (bo $v \in T_H(x_0)$)
- $\langle v, \nabla g_i(x_0) \rangle < 0$ dla $i \in I(x_0)$

a zatem zbiór K z lematu jest niepusty, i warunek M.-F. jest spełniony. □