W trakcie wykładu omówiona została **metoda subgradientowa** szukania minimum funkcji wypukłej $f : \mathbb{R}^n \to \mathbb{R}$ na zbiorze wypukłym domkniętym C, zdefiniowana w następujący sposób:
Startujemy z dowolnego $x_0 \in C$.
Dla $k = 0, 1, 2, \ldots$, $x_{k+1} = [x_k - \alpha_k g_k]^+_C$, gdzie $g_k \in \mathbb{R}^n$ to dowolny subgradient funkcji f w punkcie x_k, a $\alpha_k \in \mathbb{R}$ – pozytywna długość kroku.

Lemat 1 *Niech (x_k) będzie ciągiem generowanym przez metodę subgradientową. Wtedy dla dowolnych $y \in C$ oraz $k \geqslant 0$:*

(a)

$$\|x_{k+1} - y\|^2 \leqslant \|x_k - y\|^2 - 2\alpha_k(f(x_k) - f(y)) + \alpha_k^2 \|g_k\|^2.$$

(b) Jeśli $f(y) < f(x_k)$, to

$$\|x_{k+1} - y\| < \|x_k - y\| \quad \text{o ile } \alpha_k \in \left(0, \frac{2(f(x_k) - f(y))}{\|g_k\|^2}\right).$$

Twierdzenie 1 *Załóżmy, że istnieje stała c taka że $c \geqslant \sup_k \|g_k\|$ dla $k = 0, 1, 2, \ldots$. Wtedy, jeśli ciąg (α_k) spełnia $\lim_{k\to\infty} \alpha_k = 0$, $\sum_{k=0}^{\infty} \alpha_k = \infty$, to $\liminf_{k\to\infty} f(x_k) = \inf_{x\in C} f(x)$.*

Wniosek 1 *Przy dodatkowym założeniu, że $\sum_{k=0}^{\infty} \alpha_k^2 < \infty$, ciąg (x_k) jest zbieżny do (jakiegoś) minimum funkcji f na zbiorze C, o ile takie istnieje.*

Dowód: Niech w punkcie x^* będzie minimum funkcji f na zbiorze C. Wtedy na mocy części (a) Lematu 1 mamy

$$\|x_{k+1} - x^*\|^2 \leqslant \|x_k - x^*\|^2 - 2\alpha_k(f(x_k) - f(x^*)) + \alpha_k^2 \|g_k\|^2.$$

Drugie wyrażenie po prawej stronie jest zawsze niedodatnie, zaś trzecie możemy ograniczyć z góry przez $\alpha_k^2 c^2$, a zatem

$$\|x_{k+1} - x^*\|^2 \leqslant \|x_k - x^*\|^2 + \alpha_k^2 c^2. \tag{1}$$

Iterując (1), dostajemy dla dowolnego k

$$\|x_{k+1} - x^*\|^2 \leqslant \|x_0 - x^*\|^2 + c^2 \sum_{i=0}^{k} \alpha_i^2 \leqslant \|x_0 - x^*\|^2 + c^2 \sum_{i=0}^{\infty} \alpha_i^2, \tag{2}$$

a zatem wszystkie wyrazy ciągu (x_k) leżą w kole domkniętym o środku w x^* i promieniu $(\|x_0 - x^*\|^2 + c^2 \sum_{i=0}^{\infty} \alpha_i^2)^{0.5} < \infty$, które jest zbiorem zwartym.

Zwartość oraz teza Twierdzenia 1 implikują, że ciąg (x_k) ma podciąg zbieżny x_{k_m} taki, że w jego granicy, x^{**}, jest minimum funkcji f na zbiorze C. Pokażemy teraz, że cały ciąg (x_k) zbiega do x^{**}.

Weźmy dowolny $\varepsilon > 0$ i niech $k^* \in (k_m)_{m=0}^{\infty}$ będzie taką liczbą naturalną, że $\sum_{i=k^*}^{\infty} \alpha_i^2 < \frac{\varepsilon}{2c^2}$ oraz $\|x_{k^*} - x^{**}\| < \frac{\varepsilon}{2}$. Wtedy na mocy części (a) Lematu 1 dla dowolnego $k > k^*$ mamy (na podobnej zasadzie jak w przypadku dowodu nierówności (1–2))

$$\|x_k - x^{**}\|^2 \leqslant \|x_{k-1} - x^{**}\|^2 + \alpha_{k-1}^2 c^2 \leqslant \ldots \leqslant \|x_{k^*} - x^{**}\| + c^2 \sum_{i=k^*}^{k-1} \alpha_i^2 < \frac{\varepsilon}{2} + c^2 \frac{\varepsilon}{2c^2} = \varepsilon.$$

Wobec dowolności ε oznacza to, że $x_k \to_{k\to\infty} x^{**}$. $\square$

Twierdzenie 2 *Jeśli $\alpha_k = \frac{f(x_k) - f_{min}}{\|g_k\|^2}$, $k = 0, 1, 2, \ldots$, gdzie $f_{min} = \min_{x\in C} f(x)$, to ciąg przybliżeń metody subgradientowej jest zbieżny do (jakiegoś) minimum funkcji f na zbiorze C, o ile takie istnieje.*

Dowód: Niech w x^* będzie minimum funkcji f na zbiorze C. Na mocy części (a) Lematu 1 mamy:

$$\begin{aligned} \|x_{k+1} - x^*\| &\leqslant \|x_k - x^*\| - 2\alpha_k(f(x_k) - f_{min}) + \alpha_k^2\|g_k\|^2 \\ &= \|x_k - x^*\| - \frac{2(f(x_k) - f_{min})^2}{\|g_k\|^2} + \frac{(f(x_k) - f_{min})^2}{\|g_k\|^2} \\ &= \|x_k - x^*\| - \frac{(f(x_k) - f_{min})^2}{\|g_k\|^2} \leqslant \|x_k - x^*\|^2. \end{aligned} \tag{3}$$

A zatem ciąg (x_k) jest ograniczony. Na mocy faktu udowodnionego na ćwiczeniach ciąg (g_k) jest również ograniczony. Oznacza to, że oba ciągi są zawarte w pewnych kulach domkniętych, które są zbiorami zwartymi.

Załóżmy teraz, że $f(x_k) \not\to_{k\to\infty} f_{min}$. Istnieje wtedy podciąg (x_{k_m}) ciągu (x_k), taki że $(f(x_{k_m}))$ i (g_{k_m}) są zbieżne, przy czym granicą tego pierwszego nie jest f_{min}. To oczywiście implikuje, że granica drugiego ciągu musi być niezerowa (w przeciwnym razie ciąg $(f(x_{k_m}))$ musiałby zbiegać do minimum funkcji). Oznacza to, że dla pewnego $\varepsilon > 0$,

$$\lim_{m\to\infty} \frac{(f(x_{k_m}) - f_{min})^2}{\|g_k\|^2} > 2\varepsilon.$$

To na mocy (3) oznacza, że dla nieskończenie wielu k,

$$\|x_{k+1} - x^*\|^2 < \|x_k - x^*\|^2 - \varepsilon.$$

Ponieważ jednak ciąg $(\|x_k - x^*\|)$ jest nierosnący, to oznaczałoby, że dla pewnego k $\|x_{k+1} - x^*\|^2 < 0$, co jest niemożliwe.

Na koniec załóżmy, że istnieją dwa podciągi ciągu (x_k) zbieżne do różnych granic $\widehat{x}$ i $\overline{x}$. Wtedy nierówność (3) daje w granicy równość $\|\widehat{x} - x^*\|^2 = \|\overline{x} - x^*\|^2$ dla dowolnego x^* takiego, że $f(x^*) = f_{min}$. W szczególności, jeśli za x^* podstawimy $\overline{x}$, dostajemy $\|\widehat{x} - \overline{x}\|^2 = 0$, a zatem $\overline{x} = \widehat{x}$. Sprzeczność. □

Uwaga Oczywiście na ogół nie znamy dokładnej wartości f_{min} przed znalezieniem minimum. Łatwo jednak zauważyć, że twierdzenie pozostanie w mocy, gdy za f_{min} we wzorze na α_k podstawimy dostatecznie dobry estymator tej wartości. W przypadku wielu problemów optymalizacyjnych jest to możliwe.