Optymalizacja ciągła

3. Zbiory i funkcje wypukłe

Wojciech Kotłowski

Instytut Informatyki PP
http://www.cs.put.poznan.pl/wkotlowski/

14.03.2019

Kombinacja wypukła

Kombinacją wypukłą wektorów $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ nazywamy dowolny wektor:

$$\boldsymbol{x} = \lambda_1 \boldsymbol{x}_1 + \ldots + \lambda_k \boldsymbol{x}_k,$$

gdzie współczynniki λ_i są **nieujemne** i takie, że $\lambda_1 + \ldots + \lambda_k = 1$.

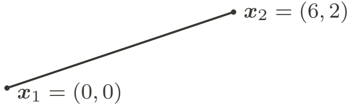
Kombinacja wypukła

Kombinacją wypukłą wektorów $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ nazywamy dowolny wektor:

$$\boldsymbol{x} = \lambda_1 \boldsymbol{x}_1 + \ldots + \lambda_k \boldsymbol{x}_k,$$

gdzie współczynniki λ_i są **nieujemne** i takie, że $\lambda_1 + \ldots + \lambda_k = 1$.

Przykład: zbiór wszystkich kombinacji wypukłych dwóch punktów $\boldsymbol{x}_1$ i $\boldsymbol{x}_2$ $\{\boldsymbol{x} \colon \boldsymbol{x} = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2, \lambda \in [0,1]\}$ to odcinek łączący $\boldsymbol{x}_1$ i $\boldsymbol{x}_2$:



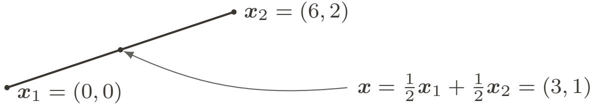
Kombinacja wypukła

Kombinacją wypukłą wektorów $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ nazywamy dowolny wektor:

$$\boldsymbol{x} = \lambda_1 \boldsymbol{x}_1 + \ldots + \lambda_k \boldsymbol{x}_k,$$

gdzie współczynniki λ_i są **nieujemne** i takie, że $\lambda_1 + \ldots + \lambda_k = 1$.

Przykład: zbiór wszystkich kombinacji wypukłych dwóch punktów $\boldsymbol{x}_1$ i $\boldsymbol{x}_2$ $\{\boldsymbol{x} \colon \boldsymbol{x} = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2, \lambda \in [0,1]\}$ to odcinek łączący $\boldsymbol{x}_1$ i $\boldsymbol{x}_2$:



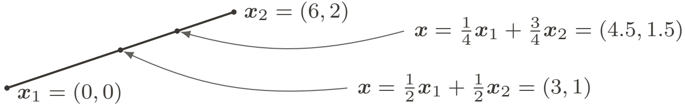
Kombinacja wypukła

Kombinacją wypukłą wektorów $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ nazywamy dowolny wektor:

$$\boldsymbol{x} = \lambda_1 \boldsymbol{x}_1 + \ldots + \lambda_k \boldsymbol{x}_k,$$

gdzie współczynniki λ_i są **nieujemne** i takie, że $\lambda_1 + \ldots + \lambda_k = 1$.

Przykład: zbiór wszystkich kombinacji wypukłych dwóch punktów $\boldsymbol{x}_1$ i $\boldsymbol{x}_2$ $\{\boldsymbol{x} \colon \boldsymbol{x} = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2, \lambda \in [0,1]\}$ to odcinek łączący $\boldsymbol{x}_1$ i $\boldsymbol{x}_2$:



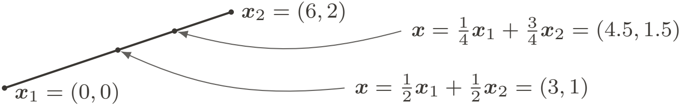
Kombinacja wypukła

Kombinacją wypukłą wektorów $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ nazywamy dowolny wektor:

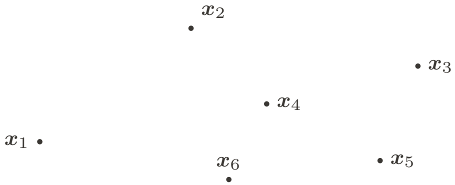
$$\boldsymbol{x} = \lambda_1 \boldsymbol{x}_1 + \ldots + \lambda_k \boldsymbol{x}_k,$$

gdzie współczynniki λ_i są **nieujemne** i takie, że $\lambda_1 + \ldots + \lambda_k = 1$.

Przykład: zbiór wszystkich kombinacji wypukłych dwóch punktów $\boldsymbol{x}_1$ i $\boldsymbol{x}_2$ $\{\boldsymbol{x} \colon \boldsymbol{x} = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2, \lambda \in [0,1]\}$ to odcinek łączący $\boldsymbol{x}_1$ i $\boldsymbol{x}_2$:



Przykład: Zbiór wszystkich kombinacji wypukłych k punktów?



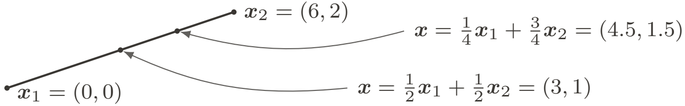
Kombinacja wypukła

Kombinacją wypukłą wektorów $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ nazywamy dowolny wektor:

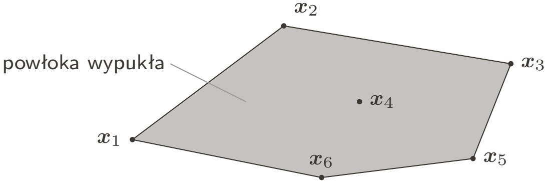
$$\boldsymbol{x} = \lambda_1 \boldsymbol{x}_1 + \ldots + \lambda_k \boldsymbol{x}_k,$$

gdzie współczynniki λ_i są **nieujemne** i takie, że $\lambda_1 + \ldots + \lambda_k = 1$.

Przykład: zbiór wszystkich kombinacji wypukłych dwóch punktów $\boldsymbol{x}_1$ i $\boldsymbol{x}_2$ $\left\{\boldsymbol{x} \colon \boldsymbol{x} = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2, \lambda \in [0,1]\right\}$ to odcinek łączący $\boldsymbol{x}_1$ i $\boldsymbol{x}_2$:



Przykład: Zbiór wszystkich kombinacji wypukłych k punktów?



Zbiór wypukły

Zbiór wypukły $\mathcal{X}$ to taki zbiór, że dla jakichkolwiek punktów $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in \mathcal{X}$, każda ich kombinacja wypukła należy do $\mathcal{X}$.

Zbiór wypukły

Zbiór wypukły $\mathcal{X}$ to taki zbiór, że dla jakichkolwiek punktów $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in \mathcal{X}$, każda ich kombinacja wypukła należy do $\mathcal{X}$.

Alternatywnie, zbiór wypukły $\mathcal{X}$ to taki zbiór, że dla jakichkolwiek dwóch punktów $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{X}$, odcinek je łączący w całości należy do $\mathcal{X}$.

$$\forall\, \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{X} \qquad \{\boldsymbol{x} \colon \boldsymbol{x} = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2, \lambda \in [0,1]\} \subseteq \mathcal{X}$$

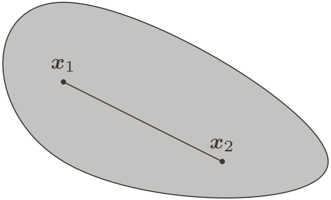
Zbiór wypukły

Zbiór wypukły $\mathcal{X}$ to taki zbiór, że dla jakichkolwiek punktów $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k \in \mathcal{X}$, każda ich kombinacja wypukła należy do $\mathcal{X}$.

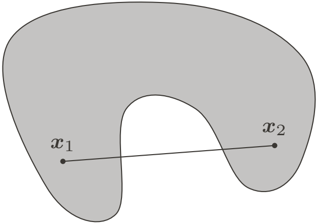
Alternatywnie, zbiór wypukły $\mathcal{X}$ to taki zbiór, że dla jakichkolwiek dwóch punktów $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{X}$, odcinek je łączący w całości należy do $\mathcal{X}$.

$$\forall\, \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{X} \qquad \{\boldsymbol{x} \colon \boldsymbol{x} = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2, \lambda \in [0,1]\} \subseteq \mathcal{X}$$

wypukły

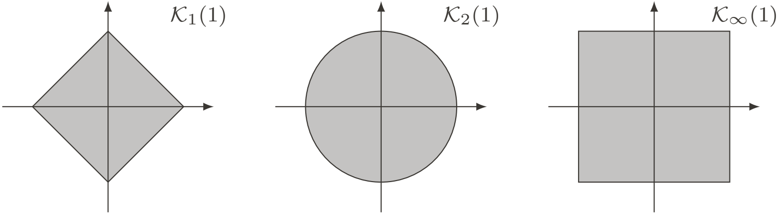


niewypukły



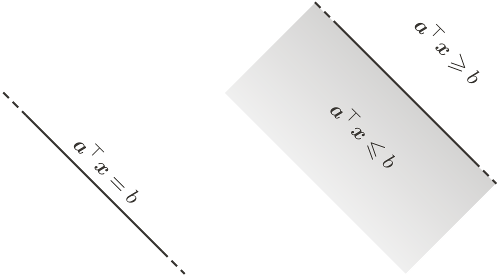
Przykłady zbiorów wypukłych

- Kula względem p-normy: $\mathcal{K}_p(r) = \{\boldsymbol{x}\colon |x_1|^p + \ldots + |x_n|^p \leqslant r^p\}$



Przykłady zbiorów wypukłych

- Hiperpłaszczyzna $\{\boldsymbol{x}\colon \boldsymbol{a}^\top \boldsymbol{x} = b\}$ lub półprzestrzeń $\{\boldsymbol{x}\colon \boldsymbol{a}^\top \boldsymbol{x} \leqslant b\}$



Przecięcie zbiorów wypukłych

Przecięcie zbiorów wypukłych jest zbiorem wypukłym

Przecięcie zbiorów wypukłych

Przecięcie zbiorów wypukłych jest zbiorem wypukłym

Dowód: Niech $\mathcal{C}_1, \ldots \mathcal{C}_m$ – zbiory wypukłe i niech $\mathcal{C} = \bigcap_{j=1}^{m} \mathcal{C}_j$

Przecięcie zbiorów wypukłych

Przecięcie zbiorów wypukłych jest zbiorem wypukłym

Dowód: Niech $\mathcal{C}_1, \ldots \mathcal{C}_m$ – zbiory wypukłe i niech $\mathcal{C} = \bigcap_{j=1}^m \mathcal{C}_j$

- Weźmy dowolne $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{C}$
  Z konstrukcji zbioru $\mathcal{C}$ mamy $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{C}_j$ dla każdego $j = 1, \ldots, m$

Przecięcie zbiorów wypukłych

Przecięcie zbiorów wypukłych jest zbiorem wypukłym

Dowód: Niech $\mathcal{C}_1, \ldots \mathcal{C}_m$ – zbiory wypukłe i niech $\mathcal{C} = \bigcap_{j=1}^m \mathcal{C}_j$

- Weźmy dowolne $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{C}$
  Z konstrukcji zbioru $\mathcal{C}$ mamy $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{C}_j$ dla każdego $j = 1, \ldots, m$
- Z wypukłości $\mathcal{C}_j$ mamy $\boldsymbol{x} = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2 \in \mathcal{C}_j$ dla każdego j

Przecięcie zbiorów wypukłych

Przecięcie zbiorów wypukłych jest zbiorem wypukłym

Dowód: Niech $\mathcal{C}_1, \ldots \mathcal{C}_m$ – zbiory wypukłe i niech $\mathcal{C} = \bigcap_{j=1}^{m} \mathcal{C}_j$

- Weźmy dowolne $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{C}$
  Z konstrukcji zbioru $\mathcal{C}$ mamy $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{C}_j$ dla każdego $j = 1, \ldots, m$
- Z wypukłości $\mathcal{C}_j$ mamy $\boldsymbol{x} = \lambda \boldsymbol{x}_1 + (1 - \lambda) \boldsymbol{x}_2 \in \mathcal{C}_j$ dla każdego j
- Stąd $\boldsymbol{x} \in \mathcal{C}$, czyli $\mathcal{C}$ jest zbiorem wypukłym

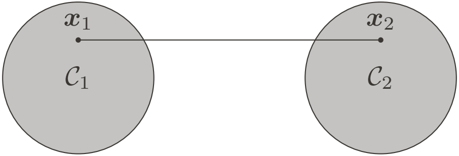
Przecięcie zbiorów wypukłych

Przecięcie zbiorów wypukłych jest zbiorem wypukłym

Dowód: Niech $\mathcal{C}_1, \ldots \mathcal{C}_m$ – zbiory wypukłe i niech $\mathcal{C} = \bigcap_{j=1}^m \mathcal{C}_j$

- Weźmy dowolne $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{C}$
  Z konstrukcji zbioru $\mathcal{C}$ mamy $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{C}_j$ dla każdego $j = 1, \ldots, m$
- Z wypukłości $\mathcal{C}_j$ mamy $\boldsymbol{x} = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2 \in \mathcal{C}_j$ dla każdego j
- Stąd $\boldsymbol{x} \in \mathcal{C}$, czyli $\mathcal{C}$ jest zbiorem wypukłym

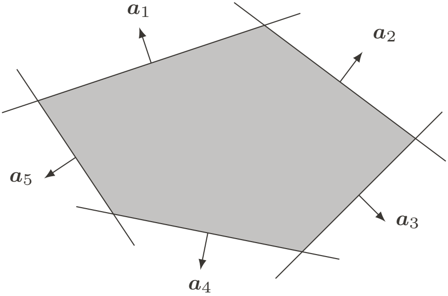
Uwaga: suma dwóch zbiorów wypukłych niekoniecznie jest zbiorem wypukłym (i zwykle nie jest)!



Przykłady zbiorów wypukłych

- Wielościan:

$$\{\boldsymbol{x}\colon \boldsymbol{A}\boldsymbol{x} \leqslant \boldsymbol{b}\} \;=\; \bigcap_{j=1}^{m} \{\boldsymbol{x}\colon \boldsymbol{a}_j^\top \boldsymbol{x} \leqslant b_j\}, \qquad \boldsymbol{A} = \begin{bmatrix} \boldsymbol{a}_1^\top \\ \boldsymbol{a}_2^\top \\ \cdots \\ \boldsymbol{a}_m^\top \end{bmatrix}$$



Funkcja wypukła

Funkcja $f(\boldsymbol{x})$ jest **wypukła**, jeśli dla dowolnych dwóch punktów $\boldsymbol{x}, \boldsymbol{y}$ i dowolnego $\lambda \in [0, 1]$,

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leqslant \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})$$

Innymi słowy, odcinek łączący dwa dowolne punkty na wykresie leży w całości powyżej lub na wykresie funkcji.

Jeśli nierówność jest ostra („$<$”), to funkcja jest **ściśle wypukła**

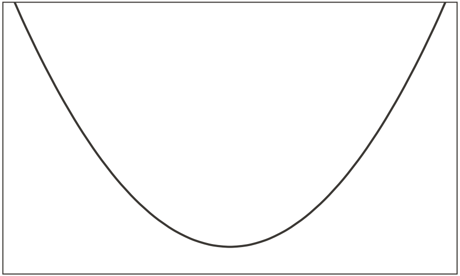
Funkcja wypukła

Funkcja $f(\boldsymbol{x})$ jest **wypukła**, jeśli dla dowolnych dwóch punktów $\boldsymbol{x}, \boldsymbol{y}$ i dowolnego $\lambda \in [0, 1]$,

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leqslant \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})$$

Innymi słowy, odcinek łączący dwa dowolne punkty na wykresie leży w całości powyżej lub na wykresie funkcji.

Jeśli nierówność jest ostra („<”), to funkcja jest **ściśle wypukła**



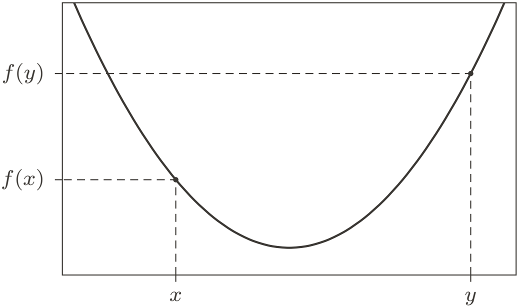
Funkcja wypukła

Funkcja $f(\boldsymbol{x})$ jest **wypukła**, jeśli dla dowolnych dwóch punktów $\boldsymbol{x}, \boldsymbol{y}$ i dowolnego $\lambda \in [0, 1]$,

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leqslant \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})$$

Innymi słowy, odcinek łączący dwa dowolne punkty na wykresie leży w całości powyżej lub na wykresie funkcji.

Jeśli nierówność jest ostra („$<$”), to funkcja jest **ściśle wypukła**



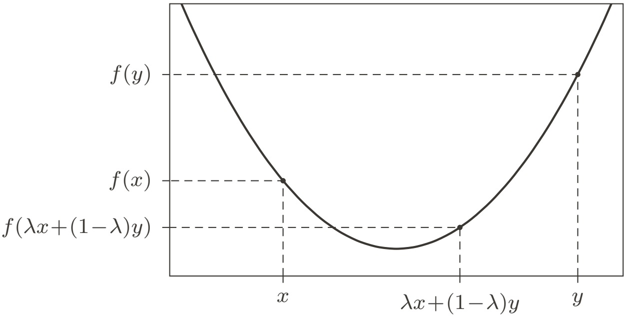
Funkcja wypukła

Funkcja $f(\boldsymbol{x})$ jest **wypukła**, jeśli dla dowolnych dwóch punktów $\boldsymbol{x}, \boldsymbol{y}$ i dowolnego $\lambda \in [0, 1]$,

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leqslant \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})$$

Innymi słowy, odcinek łączący dwa dowolne punkty na wykresie leży w całości powyżej lub na wykresie funkcji.

Jeśli nierówność jest ostra („$<$”), to funkcja jest **ściśle wypukła**



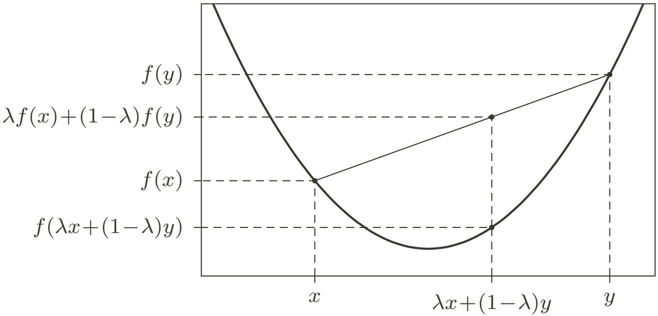
Funkcja wypukła

Funkcja $f(\boldsymbol{x})$ jest **wypukła**, jeśli dla dowolnych dwóch punktów $\boldsymbol{x}, \boldsymbol{y}$ i dowolnego $\lambda \in [0, 1]$,

$$f(\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y}) \leqslant \lambda f(\boldsymbol{x}) + (1 - \lambda) f(\boldsymbol{y})$$

Innymi słowy, odcinek łączący dwa dowolne punkty na wykresie leży w całości powyżej lub na wykresie funkcji.

Jeśli nierówność jest ostra („$<$”), to funkcja jest **ściśle wypukła**



Funkcja wypukła

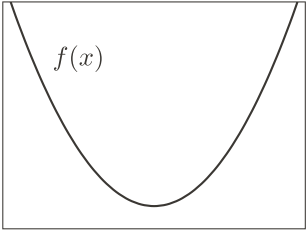
Funkcja $f(\boldsymbol{x})$ jest **wypukła**, jeśli dla dowolnych dwóch punktów $\boldsymbol{x}, \boldsymbol{y}$ i dowolnego $\lambda \in [0, 1]$,

$$f(\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y}) \leqslant \lambda f(\boldsymbol{x}) + (1 - \lambda) f(\boldsymbol{y})$$

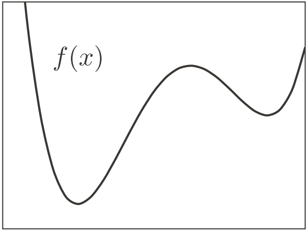
Innymi słowy, odcinek łączący dwa dowolne punkty na wykresie leży w całości powyżej lub na wykresie funkcji.

Jeśli nierówność jest ostra („$<$”), to funkcja jest **ściśle wypukła**

funkcja wypukła



funkcja niewypukła



Funkcja wypukła

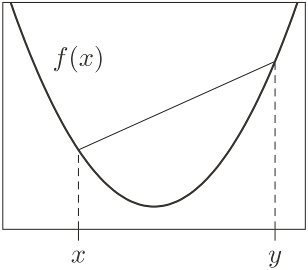
Funkcja $f(\boldsymbol{x})$ jest **wypukła**, jeśli dla dowolnych dwóch punktów $\boldsymbol{x}, \boldsymbol{y}$ i dowolnego $\lambda \in [0, 1]$,

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \leqslant \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})$$

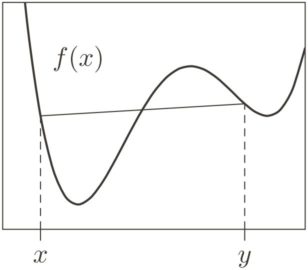
Innymi słowy, odcinek łączący dwa dowolne punkty na wykresie leży w całości powyżej lub na wykresie funkcji.

Jeśli nierówność jest ostra („$<$”), to funkcja jest **ściśle wypukła**

funkcja wypukła

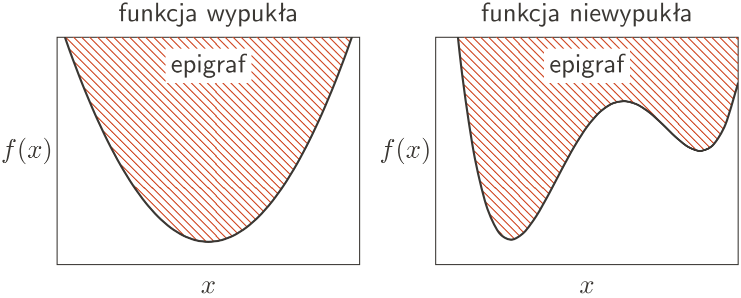


funkcja niewypukła



Funkcja wypukła

Funkcja jest wypukła wtedy i tylko wtedy gdy jej **epigraf** (zbiór ograniczony od dołu wykresem) jest **zbiorem wypukłym**



Funkcja wypukła

Jeśli funkcja $f(x)$ jest **wypukła**, to dla dowolnych $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k$ i dowolnych nieujemnych $\lambda_1, \ldots, \lambda_k$ takich, że $\lambda_1 + \ldots + \lambda_k = 1$,

$$f(\lambda_1 \boldsymbol{x}_1 + \ldots + \lambda_k \boldsymbol{x}_k) \leqslant \lambda_1 f(\boldsymbol{x}_1) + \ldots + \lambda_k f(\boldsymbol{x}_k)$$

Innymi słowy: wartość funkcji dla kombinacji wypukłej jest nie większa niż kombinacja wypukła wartości funkcji

Dowód przez prostą indukcję po k pomijamy

Minimalizacja funkcji wypukłych

Minimum lokalne: Funkcja $f(\boldsymbol{x})$ ma w punkcie $\boldsymbol{x}_0$ minimum lokalne, jeśli istnieje takie ϵ, że dla dowolnych $\boldsymbol{x}$ spełniających $\|\boldsymbol{x} - \boldsymbol{x}_0\| \leqslant \epsilon$, mamy $f(\boldsymbol{x}_0) \leqslant f(\boldsymbol{x})$

Minimalizacja funkcji wypukłych

Minimum lokalne: Funkcja $f(\boldsymbol{x})$ ma w punkcie $\boldsymbol{x}_0$ minimum lokalne, jeśli istnieje takie ϵ, że dla dowolnych $\boldsymbol{x}$ spełniających $\|\boldsymbol{x} - \boldsymbol{x}_0\| \leqslant \epsilon$, mamy $f(\boldsymbol{x}_0) \leqslant f(\boldsymbol{x})$

Twierdzenie: Rozważmy problem optymalizacji:

$$\begin{aligned} \text{min} \quad & f(\boldsymbol{x}) \\ \text{p.o.} \quad & \boldsymbol{x} \in \mathcal{X}, \end{aligned}$$

w którym zarówno funkcja $f(\boldsymbol{x})$, jak i zbiór rozwiązań dopuszczalnych $\mathcal{X}$ są **wypukłe**. Wtedy każde minimum lokalne funkcji $f(\boldsymbol{x})$ jest też jej minimum globalnym

Dowód

Dowód

- Niech $\boldsymbol{x}_0 \in \mathcal{X}$ – minimum lokalne. Załóżmy przeciwnie, że nie to jest minimum globalne, tj. istnieje $\boldsymbol{x}^* \in \mathcal{X}$ takie, że $f(\boldsymbol{x}^*) < f(\boldsymbol{x}_0)$

Dowód

- Niech $\boldsymbol{x}_0 \in \mathcal{X}$ – minimum lokalne. Załóżmy przeciwnie, że nie to jest minimum globalne, tj. istnieje $\boldsymbol{x}^* \in \mathcal{X}$ takie, że $f(\boldsymbol{x}^*) < f(\boldsymbol{x}_0)$
- Ponieważ $\boldsymbol{x}_0$ jest minimum lokalnym, więc istnieje ϵ taki, że jeśli $\|\boldsymbol{x} - \boldsymbol{x}_0\| \leqslant \epsilon$, to $f(\boldsymbol{x}_0) \leqslant f(\boldsymbol{x})$

Dowód

- Niech $\boldsymbol{x}_0 \in \mathcal{X}$ – minimum lokalne. Załóżmy przeciwnie, że nie to jest minimum globalne, tj. istnieje $\boldsymbol{x}^* \in \mathcal{X}$ takie, że $f(\boldsymbol{x}^*) < f(\boldsymbol{x}_0)$
- Ponieważ $\boldsymbol{x}_0$ jest minimum lokalnym, więc istnieje ϵ taki, że jeśli $\|\boldsymbol{x} - \boldsymbol{x}_0\| \leqslant \epsilon$, to $f(\boldsymbol{x}_0) \leqslant f(\boldsymbol{x})$
- Weźmy kombinację wypukłą $\boldsymbol{y} = \lambda \boldsymbol{x}^* + (1 - \lambda)\boldsymbol{x}_0$. Zauważmy, że:

$$\|\boldsymbol{y} - \boldsymbol{x}_0\| \;=\; \|\lambda(\boldsymbol{x}^* - \boldsymbol{x}_0)\| \;=\; \lambda\|\boldsymbol{x}^* - \boldsymbol{x}_0\|$$

czyli dla odpowiednio małego λ, będziemy mieli $\|\boldsymbol{y} - \boldsymbol{x}_0\| \leqslant \epsilon$, a stąd $f(\boldsymbol{x}_0) \leqslant f(\boldsymbol{y})$

Dowód

- Niech $\boldsymbol{x}_0 \in \mathcal{X}$ – minimum lokalne. Załóżmy przeciwnie, że nie to jest minimum globalne, tj. istnieje $\boldsymbol{x}^* \in \mathcal{X}$ takie, że $f(\boldsymbol{x}^*) < f(\boldsymbol{x}_0)$
- Ponieważ $\boldsymbol{x}_0$ jest minimum lokalnym, więc istnieje ϵ taki, że jeśli $\|\boldsymbol{x} - \boldsymbol{x}_0\| \leqslant \epsilon$, to $f(\boldsymbol{x}_0) \leqslant f(\boldsymbol{x})$
- Weźmy kombinację wypukłą $\boldsymbol{y} = \lambda \boldsymbol{x}^* + (1-\lambda)\boldsymbol{x}_0$. Zauważmy, że:

$$\|\boldsymbol{y} - \boldsymbol{x}_0\| \;=\; \|\lambda(\boldsymbol{x}^* - \boldsymbol{x}_0)\| \;=\; \lambda\|\boldsymbol{x}^* - \boldsymbol{x}_0\|$$

  czyli dla odpowiednio małego λ, będziemy mieli $\|\boldsymbol{y} - \boldsymbol{x}_0\| \leqslant \epsilon$, a stąd $f(\boldsymbol{x}_0) \leqslant f(\boldsymbol{y})$
- Z drugiej strony, z wypukłości funkcji $f(\boldsymbol{x})$:

$$\begin{aligned} f(\boldsymbol{y}) &\leqslant \lambda f(\boldsymbol{x}^*) + (1-\lambda) f(\boldsymbol{x}_0) \\ &< \lambda f(\boldsymbol{x}_0) + (1-\lambda) f(\boldsymbol{x}_0) \;=\; f(\boldsymbol{x}_0) \end{aligned}$$

Dowód

- Niech $\boldsymbol{x}_0 \in \mathcal{X}$ – minimum lokalne. Załóżmy przeciwnie, że nie to jest minimum globalne, tj. istnieje $\boldsymbol{x}^* \in \mathcal{X}$ takie, że $f(\boldsymbol{x}^*) < f(\boldsymbol{x}_0)$
- Ponieważ $\boldsymbol{x}_0$ jest minimum lokalnym, więc istnieje ϵ taki, że jeśli $\|\boldsymbol{x} - \boldsymbol{x}_0\| \leqslant \epsilon$, to $f(\boldsymbol{x}_0) \leqslant f(\boldsymbol{x})$
- Weźmy kombinację wypukłą $\boldsymbol{y} = \lambda \boldsymbol{x}^* + (1-\lambda)\boldsymbol{x}_0$. Zauważmy, że:

$$\|\boldsymbol{y} - \boldsymbol{x}_0\| \;=\; \|\lambda(\boldsymbol{x}^* - \boldsymbol{x}_0)\| \;=\; \lambda\|\boldsymbol{x}^* - \boldsymbol{x}_0\|$$

  czyli dla odpowiednio małego λ, będziemy mieli $\|\boldsymbol{y} - \boldsymbol{x}_0\| \leqslant \epsilon$, a stąd $f(\boldsymbol{x}_0) \leqslant f(\boldsymbol{y})$
- Z drugiej strony, z wypukłości funkcji $f(\boldsymbol{x})$:

$$\begin{aligned} f(\boldsymbol{y}) &\leqslant \lambda f(\boldsymbol{x}^*) + (1-\lambda) f(\boldsymbol{x}_0) \\ &< \lambda f(\boldsymbol{x}_0) + (1-\lambda) f(\boldsymbol{x}_0) \;=\; f(\boldsymbol{x}_0) \end{aligned}$$

  sprzeczność!

Dygresja: funkcje wklęsłe i maksymalizacja

Funkcja $f(\boldsymbol{x})$ jest **wklęsła** gdy $-f(\boldsymbol{x})$ jest wypukła

Twierdzenie: Rozważmy problem optymalizacji:

$$\begin{aligned} \text{max} \quad & f(\boldsymbol{x}) \\ \text{p.o.} \quad & \boldsymbol{x} \in \mathcal{X}, \end{aligned}$$

w którym funkcja $f(\boldsymbol{x})$ jest **wklęsła**, a zbiór rozwiązań dopuszczalnych $\mathcal{X}$ jest **wypukły**. Wtedy każde maksimum lokalne funkcji $f(\boldsymbol{x})$ jest też jej maksimum globalnym.

Czy istnieje funkcja równocześnie wklęsła i wypukła?

Czy istnieje funkcja równocześnie wklęsła i wypukła?

Dla funkcji **wypukłej** zachodzi dla dowolnych $\boldsymbol{x}, \boldsymbol{y}, \lambda \in [0, 1]$:

$$f(\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y}) \ \leqslant \ \lambda f(\boldsymbol{x}) + (1 - \lambda) f(\boldsymbol{y})$$

Dla funkcji **wklęsłej** zachodzi dla dowolnych $\boldsymbol{x}, \boldsymbol{y}, \lambda \in [0, 1]$:

$$f(\lambda \boldsymbol{x} + (1 - \lambda)\boldsymbol{y}) \ \geqslant \ \lambda f(\boldsymbol{x}) + (1 - \lambda) f(\boldsymbol{y})$$

Czy istnieje funkcja równocześnie wklęsła i wypukła?

Dla funkcji **wypukłej** zachodzi dla dowolnych $\boldsymbol{x}, \boldsymbol{y}, \lambda \in [0,1]$:

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \ \leqslant \ \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})$$

Dla funkcji **wklęsłej** zachodzi dla dowolnych $\boldsymbol{x}, \boldsymbol{y}, \lambda \in [0,1]$:

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \ \geqslant \ \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})$$

Wniosek: Jeśli $f(\boldsymbol{x})$ jest równocześnie wklęsła i wypukła, to dla dowolnych $\boldsymbol{x}, \boldsymbol{y}, \lambda \in [0,1]$ zachodzi:

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \ = \ \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}),$$

czyli $f(\boldsymbol{x})$ jest funkcją **liniową**:

$$f(\boldsymbol{x}) \ = \ \boldsymbol{b}^\top \boldsymbol{x} + c$$

Różniczkowalne funkcje wypukłe

Jeśli funkcja $f(\boldsymbol{x})$ **różniczkowalna**, to jest **wypukła** wtedy i tylko wtedy gdy dla dowolnych $\boldsymbol{x}, \boldsymbol{x}_0$ zachodzi:

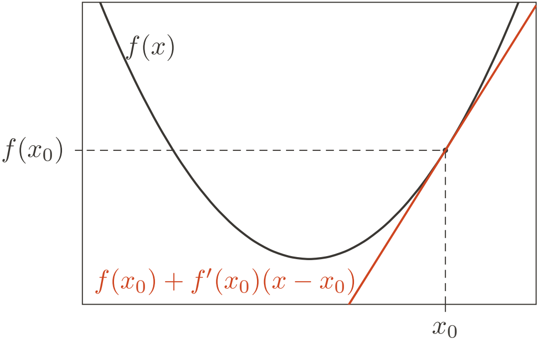
$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0)$$

Różniczkowalne funkcje wypukłe

Jeśli funkcja $f(\boldsymbol{x})$ **różniczkowalna**, to jest **wypukła** wtedy i tylko wtedy gdy dla dowolnych $\boldsymbol{x}, \boldsymbol{x}_0$ zachodzi:

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0)$$

Interpretacja: dla funkcji jednej zmiennej sprowadza się to do warunku $f(x) \geqslant f(x_0) + f'(x_0)(x - x_0)$, tzn. styczna do wykresu jest zawsze pod wykresem

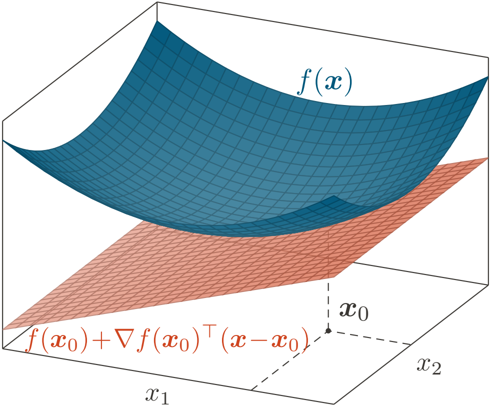


Różniczkowalne funkcje wypukłe

Jeśli funkcja $f(\boldsymbol{x})$ **różniczkowalna**, to jest **wypukła** wtedy i tylko wtedy gdy dla dowolnych $\boldsymbol{x}, \boldsymbol{x}_0$ zachodzi:

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0)$$

Interpretacja: w ogólności, hiperpłaszczyzna styczna pod wykresem



Różniczkowalne funkcje wypukłe

Jeśli funkcja $f(x)$ **różniczkowalna**, to jest **wypukła** wtedy i tylko wtedy gdy dla dowolnych $\boldsymbol{x}, \boldsymbol{x}_0$ zachodzi:

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0)$$

Dowód: w obie strony

Różniczkowalne funkcje wypukłe

Jeśli funkcja $f(x)$ **różniczkowalna**, to jest **wypukła** wtedy i tylko wtedy gdy dla dowolnych $\boldsymbol{x}, \boldsymbol{x}_0$ zachodzi:

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0)$$

Dowód: w obie strony

$\Rightarrow$ Zakładamy, że $f(\boldsymbol{x})$ jest wypukła, tzn. dla dowolnego $\lambda \in [0, 1]$:

$$\underbrace{f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{x}_0)}_{f(\boldsymbol{x}_0 + \lambda(\boldsymbol{x} - \boldsymbol{x}_0))} \leqslant \underbrace{\lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{x}_0)}_{f(\boldsymbol{x}_0) + \lambda(f(\boldsymbol{x}) - f(\boldsymbol{x}_0))}$$

Różniczkowalne funkcje wypukłe

Jeśli funkcja $f(x)$ **różniczkowalna**, to jest **wypukła** wtedy i tylko wtedy gdy dla dowolnych $\boldsymbol{x}, \boldsymbol{x}_0$ zachodzi:

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0)$$

Dowód: w obie strony

$\Rightarrow$ Zakładamy, że $f(\boldsymbol{x})$ jest wypukła, tzn. dla dowolnego $\lambda \in [0, 1]$:

$$\underbrace{f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{x}_0)}_{f(\boldsymbol{x}_0 + \lambda(\boldsymbol{x} - \boldsymbol{x}_0))} \leqslant \underbrace{\lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{x}_0)}_{f(\boldsymbol{x}_0) + \lambda(f(\boldsymbol{x}) - f(\boldsymbol{x}_0))}$$

Stąd:

$$\frac{f(\boldsymbol{x}_0 + \lambda(\boldsymbol{x} - \boldsymbol{x}_0)) - f(\boldsymbol{x}_0)}{\lambda} \leqslant f(\boldsymbol{x}) - f(\boldsymbol{x}_0)$$

Różniczkowalne funkcje wypukłe

Jeśli funkcja $f(x)$ **różniczkowalna**, to jest **wypukła** wtedy i tylko wtedy gdy dla dowolnych $\boldsymbol{x}, \boldsymbol{x}_0$ zachodzi:

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0)$$

Dowód: w obie strony

$\Rightarrow$ Zakładamy, że $f(\boldsymbol{x})$ jest wypukła, tzn. dla dowolnego $\lambda \in [0, 1]$:

$$\underbrace{f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{x}_0)}_{f(\boldsymbol{x}_0 + \lambda(\boldsymbol{x} - \boldsymbol{x}_0))} \leqslant \underbrace{\lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{x}_0)}_{f(\boldsymbol{x}_0) + \lambda(f(\boldsymbol{x}) - f(\boldsymbol{x}_0))}$$

Stąd:

$$\lim_{\lambda \to 0} \frac{f(\boldsymbol{x}_0 + \lambda(\boldsymbol{x} - \boldsymbol{x}_0)) - f(\boldsymbol{x}_0)}{\lambda} \leqslant f(\boldsymbol{x}) - f(\boldsymbol{x}_0)$$

Różniczkowalne funkcje wypukłe

Jeśli funkcja $f(x)$ **różniczkowalna**, to jest **wypukła** wtedy i tylko wtedy gdy dla dowolnych $\boldsymbol{x}, \boldsymbol{x}_0$ zachodzi:

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0)$$

Dowód: w obie strony

$\Rightarrow$ Zakładamy, że $f(\boldsymbol{x})$ jest wypukła, tzn. dla dowolnego $\lambda \in [0, 1]$:

$$\underbrace{f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{x}_0)}_{f(\boldsymbol{x}_0 + \lambda(\boldsymbol{x} - \boldsymbol{x}_0))} \leqslant \underbrace{\lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{x}_0)}_{f(\boldsymbol{x}_0) + \lambda(f(\boldsymbol{x}) - f(\boldsymbol{x}_0))}$$

Stąd:

$$\underbrace{\lim_{\lambda \to 0} \frac{f(\boldsymbol{x}_0 + \lambda(\boldsymbol{x} - \boldsymbol{x}_0)) - f(\boldsymbol{x}_0)}{\lambda}}_{\text{pochodna kierunkowa wzdłuż wektora } \boldsymbol{x} - \boldsymbol{x}_0} \leqslant f(\boldsymbol{x}) - f(\boldsymbol{x}_0)$$

Różniczkowalne funkcje wypukłe

Jeśli funkcja $f(x)$ **różniczkowalna**, to jest **wypukła** wtedy i tylko wtedy gdy dla dowolnych $\boldsymbol{x}, \boldsymbol{x}_0$ zachodzi:

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0)$$

Dowód: w obie strony

$\Rightarrow$ Zakładamy, że $f(\boldsymbol{x})$ jest wypukła, tzn. dla dowolnego $\lambda \in [0, 1]$:

$$\underbrace{f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{x}_0)}_{f(\boldsymbol{x}_0 + \lambda(\boldsymbol{x} - \boldsymbol{x}_0))} \leqslant \underbrace{\lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{x}_0)}_{f(\boldsymbol{x}_0) + \lambda(f(\boldsymbol{x}) - f(\boldsymbol{x}_0))}$$

Stąd:

$$\nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0) \leqslant f(\boldsymbol{x}) - f(\boldsymbol{x}_0)$$

co dowodzi nierówności w ramce

Różniczkowalne funkcje wypukłe

Jeśli funkcja $f(\boldsymbol{x})$ **różniczkowalna**, to jest **wypukła** wtedy i tylko wtedy gdy dla dowolnych $\boldsymbol{x}, \boldsymbol{x}_0$ zachodzi:

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0)$$

Dowód: w obie strony

$\Leftarrow$ Zakładamy, że nierówność w ramce jest spełniona i dowodzimy wypukłości $f(\boldsymbol{x})$

Weźmy dowolne $\boldsymbol{x}_1, \boldsymbol{x}_2$, $\lambda \in [0,1]$, niech $\boldsymbol{x}_0 = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2$

Mamy: $\boldsymbol{x}_1 - \boldsymbol{x}_0 = (1-\lambda)(\boldsymbol{x}_1 - \boldsymbol{x}_2)$, $\quad \boldsymbol{x}_2 - \boldsymbol{x}_0 = \lambda(\boldsymbol{x}_2 - \boldsymbol{x}_1)$

Różniczkowalne funkcje wypukłe

Jeśli funkcja $f(\boldsymbol{x})$ **różniczkowalna**, to jest **wypukła** wtedy i tylko wtedy gdy dla dowolnych $\boldsymbol{x}, \boldsymbol{x}_0$ zachodzi:

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0)$$

Dowód: w obie strony

$\Leftarrow$ Zakładamy, że nierówność w ramce jest spełniona i dowodzimy wypukłości $f(\boldsymbol{x})$

Weźmy dowolne $\boldsymbol{x}_1, \boldsymbol{x}_2$, $\lambda \in [0,1]$, niech $\boldsymbol{x}_0 = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2$

Mamy: $\quad \boldsymbol{x}_1 - \boldsymbol{x}_0 = (1-\lambda)(\boldsymbol{x}_1 - \boldsymbol{x}_2), \quad \boldsymbol{x}_2 - \boldsymbol{x}_0 = \lambda(\boldsymbol{x}_2 - \boldsymbol{x}_1)$

Używając nierówności w ramce **dwukrotnie** dla $\boldsymbol{x} = \boldsymbol{x}_1$ i $\boldsymbol{x} = \boldsymbol{x}_2$:

$$\begin{aligned} f(\boldsymbol{x}_1) &\geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x}_1 - \boldsymbol{x}_0) \\ f(\boldsymbol{x}_2) &\geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x}_2 - \boldsymbol{x}_0) \end{aligned}$$

Różniczkowalne funkcje wypukłe

Jeśli funkcja $f(\boldsymbol{x})$ **różniczkowalna**, to jest **wypukła** wtedy i tylko wtedy gdy dla dowolnych $\boldsymbol{x}, \boldsymbol{x}_0$ zachodzi:

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0)$$

Dowód: w obie strony

$\Leftarrow$ Zakładamy, że nierówność w ramce jest spełniona i dowodzimy wypukłości $f(\boldsymbol{x})$

Weźmy dowolne $\boldsymbol{x}_1, \boldsymbol{x}_2$, $\lambda \in [0,1]$, niech $\boldsymbol{x}_0 = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2$

Mamy: $\boldsymbol{x}_1 - \boldsymbol{x}_0 = (1-\lambda)(\boldsymbol{x}_1 - \boldsymbol{x}_2)$, $\quad \boldsymbol{x}_2 - \boldsymbol{x}_0 = \lambda(\boldsymbol{x}_2 - \boldsymbol{x}_1)$

Używając nierówności w ramce **dwukrotnie** dla $\boldsymbol{x} = \boldsymbol{x}_1$ i $\boldsymbol{x} = \boldsymbol{x}_2$:

$$\begin{aligned} f(\boldsymbol{x}_1) &\geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (1-\lambda)(\boldsymbol{x}_1 - \boldsymbol{x}_2) \\ f(\boldsymbol{x}_2) &\geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top \lambda(\boldsymbol{x}_2 - \boldsymbol{x}_1) \end{aligned}$$

Różniczkowalne funkcje wypukłe

Jeśli funkcja $f(\boldsymbol{x})$ **różniczkowalna**, to jest **wypukła** wtedy i tylko wtedy gdy dla dowolnych $\boldsymbol{x}, \boldsymbol{x}_0$ zachodzi:

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0)$$

Dowód: w obie strony

$\Leftarrow$ Zakładamy, że nierówność w ramce jest spełniona i dowodzimy wypukłości $f(\boldsymbol{x})$

Weźmy dowolne $\boldsymbol{x}_1, \boldsymbol{x}_2$, $\lambda \in [0,1]$, niech $\boldsymbol{x}_0 = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2$

Mamy: $\quad \boldsymbol{x}_1 - \boldsymbol{x}_0 = (1-\lambda)(\boldsymbol{x}_1 - \boldsymbol{x}_2), \quad \boldsymbol{x}_2 - \boldsymbol{x}_0 = \lambda(\boldsymbol{x}_2 - \boldsymbol{x}_1)$

Używając nierówności w ramce **dwukrotnie** dla $\boldsymbol{x} = \boldsymbol{x}_1$ i $\boldsymbol{x} = \boldsymbol{x}_2$:

$$\begin{aligned} \lambda f(\boldsymbol{x}_1) &\geqslant \lambda f(\boldsymbol{x}_0) + \lambda \nabla f(\boldsymbol{x}_0)^\top (1-\lambda)(\boldsymbol{x}_1 - \boldsymbol{x}_2) \\ (1-\lambda) f(\boldsymbol{x}_2) &\geqslant (1-\lambda) f(\boldsymbol{x}_0) + (1-\lambda) \nabla f(\boldsymbol{x}_0)^\top \lambda(\boldsymbol{x}_2 - \boldsymbol{x}_1) \end{aligned}$$

Różniczkowalne funkcje wypukłe

Jeśli funkcja $f(\boldsymbol{x})$ **różniczkowalna**, to jest **wypukła** wtedy i tylko wtedy gdy dla dowolnych $\boldsymbol{x}, \boldsymbol{x}_0$ zachodzi:

$$f(\boldsymbol{x}) \geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0)$$

Dowód: w obie strony

$\Leftarrow$ Zakładamy, że nierówność w ramce jest spełniona i dowodzimy wypukłości $f(\boldsymbol{x})$

Weźmy dowolne $\boldsymbol{x}_1, \boldsymbol{x}_2$, $\lambda \in [0,1]$, niech $\boldsymbol{x}_0 = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2$

Mamy: $\boldsymbol{x}_1 - \boldsymbol{x}_0 = (1-\lambda)(\boldsymbol{x}_1 - \boldsymbol{x}_2)$, $\quad \boldsymbol{x}_2 - \boldsymbol{x}_0 = \lambda(\boldsymbol{x}_2 - \boldsymbol{x}_1)$

Używając nierówności w ramce **dwukrotnie** dla $\boldsymbol{x} = \boldsymbol{x}_1$ i $\boldsymbol{x} = \boldsymbol{x}_2$:

$$\begin{aligned} \lambda f(\boldsymbol{x}_1) &\geqslant \lambda f(\boldsymbol{x}_0) + \lambda \nabla f(\boldsymbol{x}_0)^\top (1-\lambda)(\boldsymbol{x}_1 - \boldsymbol{x}_2) \\ (1-\lambda) f(\boldsymbol{x}_2) &\geqslant (1-\lambda) f(\boldsymbol{x}_0) + (1-\lambda) \nabla f(\boldsymbol{x}_0)^\top \lambda(\boldsymbol{x}_2 - \boldsymbol{x}_1) \end{aligned}$$

Sumujemy obie nierówności, człon z gradientem znika, co daje:

$$\lambda f(\boldsymbol{x}_1) + (1-\lambda) f(\boldsymbol{x}_2) \geqslant f(\boldsymbol{x}_0),$$

czyli $f(\boldsymbol{x})$ jest wypukła

Różniczkowalne funkcje wypukłe

Niech $f(\boldsymbol{x})$ będzie dwukrotnie różniczkowalna z ciągłymi pochodnymi. Funkcja $f(\boldsymbol{x})$ jest wypukła wtedy i tylko wtedy gdy hesjan $\nabla^2 f(\boldsymbol{x})$ jest **dodatnio półokreślony** dla dowolnego $\boldsymbol{x}$

- Dodatnia półokreśloność oznacza, że dla dowolnego wektora $\boldsymbol{v}$,

$$\boldsymbol{v}^\top \nabla^2 f(\boldsymbol{x}) \boldsymbol{v} \ \geqslant \ 0$$

- Dla funkcji jednej zmiennej sprowadza się to do znanego warunku:

$$f''(x) \geqslant 0$$

Różniczkowalne funkcje wypukłe

Niech $f(\boldsymbol{x})$ będzie dwukrotnie różniczkowalna z ciągłymi pochodnymi. Funkcja $f(\boldsymbol{x})$ jest wypukła wtedy i tylko wtedy gdy hesjan $\nabla^2 f(\boldsymbol{x})$ jest **dodatnio półokreślony** dla dowolnego $\boldsymbol{x}$

Dowód: w obie strony

Różniczkowalne funkcje wypukłe

Niech $f(\boldsymbol{x})$ będzie dwukrotnie różniczkowalna z ciągłymi pochodnymi. Funkcja $f(\boldsymbol{x})$ jest wypukła wtedy i tylko wtedy gdy hesjan $\nabla^2 f(\boldsymbol{x})$ jest **dodatnio półokreślony** dla dowolnego $\boldsymbol{x}$

Dowód: w obie strony

$\Leftarrow$ Zakładamy, hesjan jest dodatnio półokreślony. Ze wzoru Taylora:

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0) + \frac{1}{2} \underbrace{(\boldsymbol{x} - \boldsymbol{x}_0)^\top \nabla^2 f(\boldsymbol{\xi})(\boldsymbol{x} - \boldsymbol{x}_0)}_{\geqslant 0 \text{ (dodatnio półokreśl.)}}$$

Różniczkowalne funkcje wypukłe

Niech $f(\boldsymbol{x})$ będzie dwukrotnie różniczkowalna z ciągłymi pochodnymi. Funkcja $f(\boldsymbol{x})$ jest wypukła wtedy i tylko wtedy gdy hesjan $\nabla^2 f(\boldsymbol{x})$ jest **dodatnio półokreślony** dla dowolnego $\boldsymbol{x}$

Dowód: w obie strony

$\Leftarrow$ Zakładamy, hesjan jest dodatnio półokreślony. Ze wzoru Taylora:

$$\begin{aligned} f(\boldsymbol{x}) &= f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0) + \frac{1}{2} \underbrace{(\boldsymbol{x} - \boldsymbol{x}_0)^\top \nabla^2 f(\boldsymbol{\xi})(\boldsymbol{x} - \boldsymbol{x}_0)}_{\geqslant 0 \text{ (dodatnio półokreśl.)}} \\ &\geqslant f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0), \end{aligned}$$

stąd funkcja jest wypukła z poprzedniego twierdzenia

Różniczkowalne funkcje wypukłe

Niech $f(\boldsymbol{x})$ będzie dwukrotnie różniczkowalna z ciągłymi pochodnymi. Funkcja $f(\boldsymbol{x})$ jest wypukła wtedy i tylko wtedy gdy hesjan $\nabla^2 f(\boldsymbol{x})$ jest **dodatnio półokreślony** dla dowolnego $\boldsymbol{x}$

Dowód: w obie strony

$\Rightarrow$ Niech $f(\boldsymbol{x})$ będzie wypukła. Załóżmy przeciwnie, że istnieje $\boldsymbol{x}_0$ i wektor $\boldsymbol{v}$ taki, że:

$$\boldsymbol{v}^\top \nabla^2 f(\boldsymbol{x}_0) \boldsymbol{v} < 0$$

Różniczkowalne funkcje wypukłe

Niech $f(\boldsymbol{x})$ będzie dwukrotnie różniczkowalna z ciągłymi pochodnymi. Funkcja $f(\boldsymbol{x})$ jest wypukła wtedy i tylko wtedy gdy hesjan $\nabla^2 f(\boldsymbol{x})$ jest **dodatnio półokreślony** dla dowolnego $\boldsymbol{x}$

Dowód: w obie strony

$\Rightarrow$ Niech $f(\boldsymbol{x})$ będzie wypukła. Załóżmy przeciwnie, że istnieje $\boldsymbol{x}_0$ i wektor $\boldsymbol{v}$ taki, że:

$$\boldsymbol{v}^\top \nabla^2 f(\boldsymbol{x}_0) \boldsymbol{v} < 0$$

Z ciągłości hesjanu zachodzi to również w małym otoczeniu $\mathcal{S}$ punktu $\boldsymbol{x}_0$. Przyjmijmy długość wektora $\boldsymbol{v}$ na tyle małą, aby $\boldsymbol{x}_0 + \boldsymbol{v} \in \mathcal{S}$

Różniczkowalne funkcje wypukłe

Niech $f(\boldsymbol{x})$ będzie dwukrotnie różniczkowalna z ciągłymi pochodnymi. Funkcja $f(\boldsymbol{x})$ jest wypukła wtedy i tylko wtedy gdy hesjan $\nabla^2 f(\boldsymbol{x})$ jest **dodatnio półokreślony** dla dowolnego $\boldsymbol{x}$

Dowód: w obie strony

$\Rightarrow$ Niech $f(\boldsymbol{x})$ będzie wypukła. Załóżmy przeciwnie, że istnieje $\boldsymbol{x}_0$ i wektor $\boldsymbol{v}$ taki, że:

$$\boldsymbol{v}^\top \nabla^2 f(\boldsymbol{x}_0)\boldsymbol{v} < 0$$

Z ciągłości hesjanu zachodzi to również w małym otoczeniu $\mathcal{S}$ punktu $\boldsymbol{x}_0$. Przyjmijmy długość wektora $\boldsymbol{v}$ na tyle małą, aby $\boldsymbol{x}_0 + \boldsymbol{v} \in \mathcal{S}$

Ze wzoru Taylora dla $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{v}$:

$$\begin{aligned} f(\boldsymbol{x}) &= f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0) + \frac{1}{2} \overbrace{\underbrace{(\boldsymbol{x} - \boldsymbol{x}_0)}_{\boldsymbol{v}}{}^\top \nabla^2 f(\boldsymbol{\xi}) \underbrace{(\boldsymbol{x} - \boldsymbol{x}_0)}_{\boldsymbol{v}}}^{<0} \\ &< f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top (\boldsymbol{x} - \boldsymbol{x}_0), \end{aligned}$$

ponieważ $\boldsymbol{\xi}$ leży między $\boldsymbol{x}_0$ a $\boldsymbol{x}_0 + \boldsymbol{v}$, stąd $\boldsymbol{\xi} \in \mathcal{S}$.

Różniczkowalne funkcje wypukłe

Niech $f(\boldsymbol{x})$ będzie dwukrotnie różniczkowalna z ciągłymi pochodnymi. Funkcja $f(\boldsymbol{x})$ jest wypukła wtedy i tylko wtedy gdy hesjan $\nabla^2 f(\boldsymbol{x})$ jest **dodatnio półokreślony** dla dowolnego $\boldsymbol{x}$

Dowód: w obie strony

$\Rightarrow$ Niech $f(\boldsymbol{x})$ będzie wypukła. Załóżmy przeciwnie, że istnieje $\boldsymbol{x}_0$ i wektor $\boldsymbol{v}$ taki, że:

$$\boldsymbol{v}^\top \nabla^2 f(\boldsymbol{x}_0)\boldsymbol{v} < 0$$

Z ciągłości hesjanu zachodzi to również w małym otoczeniu $\mathcal{S}$ punktu $\boldsymbol{x}_0$. Przyjmijmy długość wektora $\boldsymbol{v}$ na tyle małą, aby $\boldsymbol{x}_0 + \boldsymbol{v} \in \mathcal{S}$

Ze wzoru Taylora dla $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{v}$:

$$\begin{aligned} f(\boldsymbol{x}) &= f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top(\boldsymbol{x} - \boldsymbol{x}_0) + \frac{1}{2}\overbrace{\underbrace{(\boldsymbol{x} - \boldsymbol{x}_0)}_{\boldsymbol{v}}{}^\top \nabla^2 f(\boldsymbol{\xi})\underbrace{(\boldsymbol{x} - \boldsymbol{x}_0)}_{\boldsymbol{v}}}^{<0} \\ &< f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)^\top(\boldsymbol{x} - \boldsymbol{x}_0), \end{aligned}$$

ponieważ $\boldsymbol{\xi}$ leży między $\boldsymbol{x}_0$ a $\boldsymbol{x}_0 + \boldsymbol{v}$, stąd $\boldsymbol{\xi} \in \mathcal{S}$. Tym samym funkcja nie spełnia poprzedniego twierdzenia, więc nie jest wypukła

Różniczkowalne funkcje wypukłe

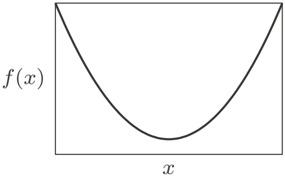
Niech $f(\boldsymbol{x})$ będzie dwukrotnie różniczkowalna z ciągłymi pochodnymi. Funkcja $f(\boldsymbol{x})$ jest wypukła wtedy i tylko wtedy gdy hesjan $\nabla^2 f(\boldsymbol{x})$ jest **dodatnio półokreślony** dla dowolnego $\boldsymbol{x}$

Niech $f(\boldsymbol{x})$ będzie dwukrotnie różniczkowalna z ciągłymi pochodnymi. Jeśli hesjan $\nabla^2 f(\boldsymbol{x})$ jest **dodatnio określony** dla dowolnego $\boldsymbol{x}$, to funkcja jest **ściśle wypukła**. Stwierdzenie w odwrotną stronę nie zachodzi.

Przykłady funkcji wypukłych (jedna zmienna)

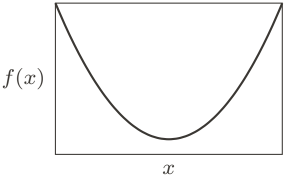
Funkcja kwadratowa
$f(x) = ax^2 + bx + c$ dla $a \geqslant 0$

$$f'(x) = 2ax + b, \quad f''(x) = 2a \geqslant 0$$



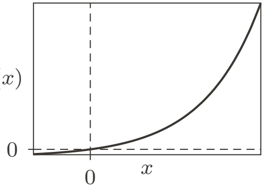
Przykłady funkcji wypukłych (jedna zmienna)

Funkcja kwadratowa
$f(x) = ax^2 + bx + c$ dla $a \geqslant 0$



$$f'(x) = 2ax + b, \quad f''(x) = 2a \geqslant 0$$

Funkcja wykładnicza $f(x) = e^x$

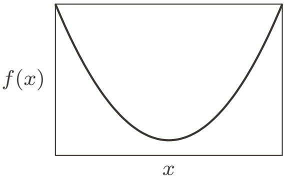


$$f'(x) = e^x, \quad f''(x) = e^x > 0$$

Przykłady funkcji wypukłych (jedna zmienna)

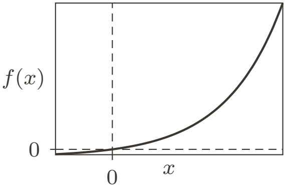
Funkcja kwadratowa
$f(x) = ax^2 + bx + c$ dla $a \geqslant 0$

$$f'(x) = 2ax + b, \quad f''(x) = 2a \geqslant 0$$



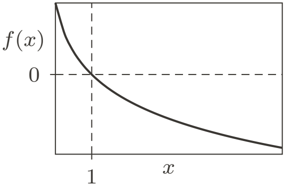
Funkcja wykładnicza $f(x) = e^x$

$$f'(x) = e^x, \quad f''(x) = e^x > 0$$



Ujemny logarytm $f(x) = -\ln(x)$
(dla $x > 0$)

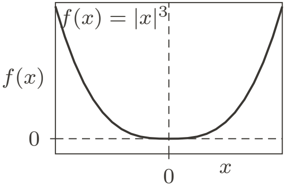
$$f'(x) = -\frac{1}{x}, \quad f''(x) = \frac{1}{x^2} > 0$$



Przykłady funkcji wypukłych (jedna zmienna)

Funkcja $f(x) = |x|^\alpha$ dla $\alpha \geqslant 1$
(w szczególności $f(x) = |x|$)

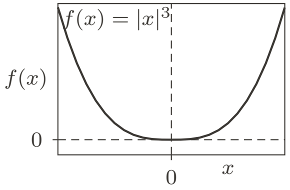
Nie jest różniczkowalna



Przykłady funkcji wypukłych (jedna zmienna)

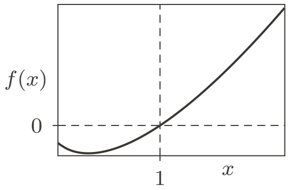
Funkcja $f(x) = |x|^\alpha$ dla $\alpha \geqslant 1$
(w szczególności $f(x) = |x|$)

Nie jest różniczkowalna



Funkcja $f(x) = x \ln x$ dla $x > 0$

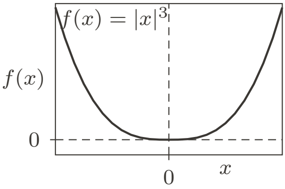
$$f'(x) = \ln x + 1, \quad f''(x) = \frac{1}{x} > 0$$



Przykłady funkcji wypukłych (jedna zmienna)

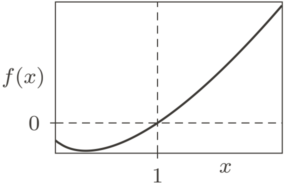
Funkcja $f(x) = |x|^\alpha$ dla $\alpha \geqslant 1$
(w szczególności $f(x) = |x|$)

Nie jest różniczkowalna



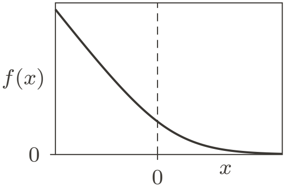
Funkcja $f(x) = x \ln x$ dla $x > 0$

$$f'(x) = \ln x + 1, \quad f''(x) = \frac{1}{x} > 0$$



Funkcja $f(x) = \ln(1 + e^{-x})$

$$f'(x) = -\frac{1}{1+e^x}, f''(x) = \frac{e^x}{(1+e^x)^2} > 0$$



Średnia geometryczna a arytmetyczna

Z wypukłości ujemnego logarytmu mamy dla dowolnych $x_1, \ldots, x_n$ i $\lambda_1 = \ldots = \lambda_n = \frac{1}{n}$:

$$-\ln\left(\frac{1}{n}x_1 + \ldots + \frac{1}{n}x_n\right) \leqslant \frac{1}{n}\left(-\ln(x_1)\right) + \ldots + \frac{1}{n}\left(-\ln(x_n)\right)$$

Średnia geometryczna a arytmetyczna

Z wypukłości ujemnego logarytmu mamy dla dowolnych $x_1, \ldots, x_n$ i $\lambda_1 = \ldots = \lambda_n = \frac{1}{n}$:

$$-\ln\left(\frac{1}{n}x_1 + \ldots + \frac{1}{n}x_n\right) \leqslant \frac{1}{n}\left(-\ln(x_1)\right) + \ldots + \frac{1}{n}\left(-\ln(x_n)\right)$$

Po złożeniu logarytmów po prawej stronie w jeden:

$$-\ln\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \leqslant -\ln\left(\left(\prod_{i=1}^{n} x_i\right)^{1/n}\right)$$

Średnia geometryczna a arytmetyczna

Z wypukłości ujemnego logarytmu mamy dla dowolnych $x_1, \ldots, x_n$ i $\lambda_1 = \ldots = \lambda_n = \frac{1}{n}$:

$$-\ln\left(\frac{1}{n}x_1 + \ldots + \frac{1}{n}x_n\right) \leqslant \frac{1}{n}\left(-\ln(x_1)\right) + \ldots + \frac{1}{n}\left(-\ln(x_n)\right)$$

Po złożeniu logarytmów po prawej stronie w jeden:

$$-\ln\left(\frac{1}{n}\sum_{i=1}^{n} x_i\right) \leqslant -\ln\left(\left(\prod_{i=1}^{n} x_i\right)^{1/n}\right)$$

Po przemnożeniu przez -1 i opuszczeniu logarytmów:

$$\frac{1}{n}\sum_{i=1}^{n} x_i \geqslant \left(\prod_{i=1}^{n} x_i\right)^{1/n}$$

Otrzymujemy znane twierdzenie: **średnia arytmetyczna** jest **nie mniejsza** od **średniej geometrycznej**

Przykłady funkcji wypukłych (wiele zmiennych)

Dowolna **norma** $f(x) = \|x\|$ (np. euklidesowa) jest funkcją wypukłą

Przykłady funkcji wypukłych (wiele zmiennych)

Dowolna **norma** $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$ (np. euklidesowa) jest funkcją wypukłą

Dowód: Z nierówności trójkąta $\|\boldsymbol{u}+\boldsymbol{v}\| \leqslant \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$ i skalowania $\|\alpha\boldsymbol{v}\| = \alpha\|\boldsymbol{v}\|$ (dla $\alpha \geqslant 0$):

$$f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \quad = \quad \|\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\|$$

Przykłady funkcji wypukłych (wiele zmiennych)

Dowolna **norma** $f(x) = \|x\|$ (np. euklidesowa) jest funkcją wypukłą

Dowód: Z nierówności trójkąta $\|\boldsymbol{u} + \boldsymbol{v}\| \leqslant \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$ i skalowania $\|\alpha \boldsymbol{v}\| = \alpha \|\boldsymbol{v}\|$ (dla $\alpha \geqslant 0$):

$$\begin{aligned} f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= \|\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}\| \\ &\leqslant \|\lambda \boldsymbol{x}\| + \|(1-\lambda)\boldsymbol{y}\| \end{aligned}$$

Przykłady funkcji wypukłych (wiele zmiennych)

Dowolna **norma** $f(x) = \|x\|$ (np. euklidesowa) jest funkcją wypukłą

Dowód: Z nierówności trójkąta $\|\boldsymbol{u} + \boldsymbol{v}\| \leqslant \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$ i skalowania $\|\alpha \boldsymbol{v}\| = \alpha \|\boldsymbol{v}\|$ (dla $\alpha \geqslant 0$):

$$
\begin{aligned}
f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= \|\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}\| \\
&\leqslant \|\lambda \boldsymbol{x}\| + \|(1-\lambda)\boldsymbol{y}\| \\
&= \lambda \|\boldsymbol{x}\| + (1-\lambda)\|\boldsymbol{y}\|
\end{aligned}
$$

Przykłady funkcji wypukłych (wiele zmiennych)

Dowolna **norma** $f(x) = \|x\|$ (np. euklidesowa) jest funkcją wypukłą

Dowód: Z nierówności trójkąta $\|\boldsymbol{u} + \boldsymbol{v}\| \leqslant \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$ i skalowania $\|\alpha \boldsymbol{v}\| = \alpha \|\boldsymbol{v}\|$ (dla $\alpha \geqslant 0$):

$$
\begin{aligned}
f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= \|\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}\| \\
&\leqslant \|\lambda \boldsymbol{x}\| + \|(1-\lambda)\boldsymbol{y}\| \\
&= \lambda \|\boldsymbol{x}\| + (1-\lambda)\|\boldsymbol{y}\| \\
&= \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})
\end{aligned}
$$

Przykłady funkcji wypukłych (wiele zmiennych)

Maksimum $f(\boldsymbol{x}) = \max_{i=1,\ldots,n}\{x_i\}$ jest funkcją wypukłą

Przykłady funkcji wypukłych (wiele zmiennych)

Maksimum $f(\boldsymbol{x}) = \max_{i=1,\ldots,n}\{x_i\}$ jest funkcją wypukłą

Dowód:

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \quad = \quad \max_i \{\lambda x_i + (1-\lambda) y_i\}$$

Przykłady funkcji wypukłych (wiele zmiennych)

Maksimum $f(\boldsymbol{x}) = \max_{i=1,\ldots,n}\{x_i\}$ jest funkcją wypukłą

Dowód:

$$\begin{aligned} f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= \max_i\{\lambda x_i + (1-\lambda)y_i\} \\ &\leqslant \max_i\{\lambda x_i\} + \max_i\{(1-\lambda)y_i\} \end{aligned}$$

Przykłady funkcji wypukłych (wiele zmiennych)

Maksimum $f(\boldsymbol{x}) = \max_{i=1,\ldots,n}\{x_i\}$ jest funkcją wypukłą

Dowód:

$$\begin{aligned} f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= \max_i\{\lambda x_i + (1-\lambda)y_i\} \\ &\leqslant \max_i\{\lambda x_i\} + \max_i\{(1-\lambda)y_i\} \\ &= \lambda \max_i\{x_i\} + (1-\lambda)\max_i\{y_i\} \end{aligned}$$

Przykłady funkcji wypukłych (wiele zmiennych)

Maksimum $f(\boldsymbol{x}) = \max_{i=1,\ldots,n}\{x_i\}$ jest funkcją wypukłą

Dowód:

$$\begin{aligned} f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= \max_i \{\lambda x_i + (1-\lambda) y_i\} \\ &\leqslant \max_i\{\lambda x_i\} + \max_i\{(1-\lambda)y_i\} \\ &= \lambda \max_i\{x_i\} + (1-\lambda)\max_i\{y_i\} \\ &= \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}) \end{aligned}$$

Przykłady funkcji wypukłych (wiele zmiennych)

Funkcja kwadratowa $f(x) = \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{x} + c$ dla **dodatnio półokreślonej** macierzy $\boldsymbol{A}$ jest wypukła

Przykłady funkcji wypukłych (wiele zmiennych)

Funkcja kwadratowa $f(\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{x} + c$ dla **dodatnio półokreślonej** macierzy $\boldsymbol{A}$ jest wypukła

Dowód: Wynika natychmiast z $\nabla^2 f(\boldsymbol{x}) = 2\boldsymbol{A}$ (już to policzyliśmy!)

Przykłady funkcji wypukłych (wiele zmiennych)

Funkcja kwadratowa $f(\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{x} + c$ dla **dodatnio półokreślonej** macierzy $\boldsymbol{A}$ jest wypukła

Dowód: Wynika natychmiast z $\nabla^2 f(\boldsymbol{x}) \;=\; 2\boldsymbol{A}$ (już to policzyliśmy!)

Przykład 1: Funkcja liniowa $f(\boldsymbol{x}) = \boldsymbol{b}^\top \boldsymbol{x} + c$ jest wypukła

Przykłady funkcji wypukłych (wiele zmiennych)

Funkcja kwadratowa $f(\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{x} + c$ dla **dodatnio półokreślonej** macierzy $\boldsymbol{A}$ jest wypukła

Dowód: Wynika natychmiast z $\nabla^2 f(\boldsymbol{x}) = 2\boldsymbol{A}$ (już to policzyliśmy!)

Przykład 1: Funkcja liniowa $f(\boldsymbol{x}) = \boldsymbol{b}^\top \boldsymbol{x} + c$ jest wypukła

Mamy $\boldsymbol{A} = \boldsymbol{0}$, co jest macierzą dodatnio półokreśloną, ponieważ dla dowolnego wektora $\boldsymbol{v}$:

$$\boldsymbol{v}^\top \boldsymbol{A}\boldsymbol{v} = 0 \geqslant 0$$

Przykłady funkcji wypukłych (wiele zmiennych)

Funkcja kwadratowa $f(\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{x} + c$ dla **dodatnio półokreślonej** macierzy $\boldsymbol{A}$ jest wypukła

Dowód: Wynika natychmiast z $\nabla^2 f(\boldsymbol{x}) \; = \; 2\boldsymbol{A}$ (już to policzyliśmy!)

Przykład 2: Funkcja $f(\boldsymbol{x}) = \|\boldsymbol{x}\|^2$ jest wypukła

Przykłady funkcji wypukłych (wiele zmiennych)

Funkcja kwadratowa $f(\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{x} + c$ dla **dodatnio półokreślonej** macierzy $\boldsymbol{A}$ jest wypukła

Dowód: Wynika natychmiast z $\nabla^2 f(\boldsymbol{x}) \ = \ 2\boldsymbol{A}$ (już to policzyliśmy!)

Przykład 2: Funkcja $f(\boldsymbol{x}) = \|\boldsymbol{x}\|^2$ jest wypukła

Ponieważ $f(\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{x} = \boldsymbol{x}^\top \boldsymbol{I}\boldsymbol{x}$, mamy $\boldsymbol{A} = \boldsymbol{I}$
$\boldsymbol{I}$ jest dodatnio określona: dla dowolnego niezerowego wektora $\boldsymbol{v}$

$$\boldsymbol{v}^\top \boldsymbol{I}\boldsymbol{v} \ = \ \|\boldsymbol{v}\|^2 \ > \ 0$$

Przykłady funkcji wypukłych (wiele zmiennych)

Funkcja kwadratowa $f(\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{x} + c$ dla **dodatnio półokreślonej** macierzy $\boldsymbol{A}$ jest wypukła

Dowód: Wynika natychmiast z $\nabla^2 f(\boldsymbol{x}) = 2\boldsymbol{A}$ (już to policzyliśmy!)

Przykład 3: Funkcja $f(\boldsymbol{x}) = (y - \boldsymbol{a}^\top \boldsymbol{x})^2$, $y \in \mathbb{R}$, $\boldsymbol{a}, \boldsymbol{x} \in \mathbb{R}^n$ jest wypukła

Przykłady funkcji wypukłych (wiele zmiennych)

Funkcja kwadratowa $f(\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{x} + c$ dla **dodatnio półokreślonej** macierzy $\boldsymbol{A}$ jest wypukła

Dowód: Wynika natychmiast z $\nabla^2 f(\boldsymbol{x}) = 2\boldsymbol{A}$ (już to policzyliśmy!)

Przykład 3: Funkcja $f(\boldsymbol{x}) = (y - \boldsymbol{a}^\top \boldsymbol{x})^2$, $y \in \mathbb{R}$, $\boldsymbol{a}, \boldsymbol{x} \in \mathbb{R}^n$ jest wypukła

$$f(\boldsymbol{x}) = y^2 - 2y\boldsymbol{a}^\top \boldsymbol{x} + (\boldsymbol{a}^\top \boldsymbol{x})^2 = \underbrace{y^2}_{c} + \underbrace{(-2y\boldsymbol{a})}_{\boldsymbol{b}}{}^\top \boldsymbol{x} + \boldsymbol{x}^\top \underbrace{\boldsymbol{a}\boldsymbol{a}^\top}_{\boldsymbol{A}} \boldsymbol{x}$$

Przykłady funkcji wypukłych (wiele zmiennych)

Funkcja kwadratowa $f(\boldsymbol{x}) = \boldsymbol{x}^\top \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}^\top \boldsymbol{x} + c$ dla **dodatnio półokreślonej** macierzy $\boldsymbol{A}$ jest wypukła

Dowód: Wynika natychmiast z $\nabla^2 f(\boldsymbol{x}) \ = \ 2\boldsymbol{A}$ (już to policzyliśmy!)

Przykład 3: Funkcja $f(\boldsymbol{x}) = (y - \boldsymbol{a}^\top \boldsymbol{x})^2$, $y \in \mathbb{R}$, $\boldsymbol{a}, \boldsymbol{x} \in \mathbb{R}^n$ jest wypukła

$$f(\boldsymbol{x}) \ = \ y^2 - 2y\boldsymbol{a}^\top \boldsymbol{x} + (\boldsymbol{a}^\top \boldsymbol{x})^2 \ = \ \underbrace{y^2}_{c} + \underbrace{(-2y\boldsymbol{a})}_{\boldsymbol{b}}{}^\top \boldsymbol{x} + \boldsymbol{x}^\top \underbrace{\boldsymbol{a}\boldsymbol{a}^\top}_{\boldsymbol{A}} \boldsymbol{x}$$

$\boldsymbol{A} = \boldsymbol{a}\boldsymbol{a}^\top \in \mathbb{R}^{n \times n}$ jest dodatnio półokreślona, gdyż dla dowolnego $\boldsymbol{v}$:

$$\boldsymbol{v}^\top \boldsymbol{A}\boldsymbol{v} \ = \ \boldsymbol{v}^\top \boldsymbol{a}\boldsymbol{a}^\top \boldsymbol{v} \ = \ (\boldsymbol{v}^\top \boldsymbol{a})^2 \geqslant 0$$

Operacje zachowujące wypukłość

Maksimum funkcji **wypukłych** jest funkcją wypukłą

$$f(\boldsymbol{x}) \;=\; \max_{j=1,\ldots,m} \{f_j(\boldsymbol{x})\}, \qquad (f_1, \ldots, f_m \text{ wypukłe})$$

Operacje zachowujące wypukłość

Maksimum funkcji **wypukłych** jest funkcją wypukłą

$$f(\boldsymbol{x}) \;=\; \max_{j=1,\ldots,m} \{f_j(\boldsymbol{x})\}, \qquad (f_1, \ldots, f_m \text{ wypukłe})$$

Dowód:

$$f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \quad = \quad \max_j \big\{ f_j(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \big\}$$

Operacje zachowujące wypukłość

Maksimum funkcji **wypukłych** jest funkcją wypukłą

$$f(\boldsymbol{x}) = \max_{j=1,\ldots,m} \{f_j(\boldsymbol{x})\}, \qquad (f_1, \ldots, f_m \text{ wypukłe})$$

Dowód:

$$\begin{aligned} f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= \max_j \{f_j(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y})\} \\ &\overset{(*)}{\leqslant} \max_j \{\lambda f_j(\boldsymbol{x}) + (1-\lambda) f_j(\boldsymbol{y})\} \end{aligned}$$

Operacje zachowujące wypukłość

Maksimum funkcji **wypukłych** jest funkcją wypukłą

$$f(\boldsymbol{x}) = \max_{j=1,\ldots,m} \{f_j(\boldsymbol{x})\}, \qquad (f_1, \ldots, f_m \text{ wypukłe})$$

Dowód:

$$\begin{aligned} f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= \max_j \{f_j(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y})\} \\ &\overset{(*)}{\leqslant} \max_j \{\lambda f_j(\boldsymbol{x}) + (1-\lambda) f_j(\boldsymbol{y})\} \end{aligned}$$

gdzie w $(*)$ wykorzystaliśmy wypukłość $f_1, \ldots, f_m$ oraz fakt, że jeśli $a_j \leqslant b_j$ dla $j = 1, \ldots, m$ to $\max_{j=1,\ldots,m}\{a_j\} \leqslant \max_{j=1,\ldots,m}\{b_j\}$

Operacje zachowujące wypukłość

Maksimum funkcji **wypukłych** jest funkcją wypukłą

$$f(\boldsymbol{x}) \;=\; \max_{j=1,\ldots,m} \{f_j(\boldsymbol{x})\}, \qquad (f_1, \ldots, f_m \text{ wypukłe})$$

Dowód:

$$\begin{aligned}
f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= \max_j \{f_j(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y})\} \\
&\overset{(*)}{\leqslant} \max_j \{\lambda f_j(\boldsymbol{x}) + (1-\lambda) f_j(\boldsymbol{y})\} \\
&\leqslant \max_j \{\lambda f_j(\boldsymbol{x})\} + \max_j \{(1-\lambda) f_j(\boldsymbol{y})\}
\end{aligned}$$

gdzie w $(*)$ wykorzystaliśmy wypukłość $f_1, \ldots, f_m$ oraz fakt, że jeśli $a_j \leqslant b_j$ dla $j = 1, \ldots, m$ to $\max_{j=1,\ldots,m}\{a_j\} \leqslant \max_{j=1,\ldots,m}\{b_j\}$

Operacje zachowujące wypukłość

Maksimum funkcji **wypukłych** jest funkcją wypukłą

$$f(\boldsymbol{x}) = \max_{j=1,\ldots,m} \{f_j(\boldsymbol{x})\}, \qquad (f_1, \ldots, f_m \text{ wypukłe})$$

Dowód:

$$\begin{aligned} f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= \max_j \{f_j(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y})\} \\ &\overset{(*)}{\leqslant} \max_j \{\lambda f_j(\boldsymbol{x}) + (1-\lambda) f_j(\boldsymbol{y})\} \\ &\leqslant \max_j \{\lambda f_j(\boldsymbol{x})\} + \max_j \{(1-\lambda) f_j(\boldsymbol{y})\} \\ &= \lambda \max_j \{f_j(\boldsymbol{x})\} + (1-\lambda) \max_j \{f_j(\boldsymbol{y})\} \end{aligned}$$

gdzie w $(*)$ wykorzystaliśmy wypukłość $f_1, \ldots, f_m$ oraz fakt, że jeśli $a_j \leqslant b_j$ dla $j = 1, \ldots, m$ to $\max_{j=1,\ldots,m}\{a_j\} \leqslant \max_{j=1,\ldots,m}\{b_j\}$

Operacje zachowujące wypukłość

Maksimum funkcji **wypukłych** jest funkcją wypukłą

$$f(\boldsymbol{x}) \;=\; \max_{j=1,\ldots,m} \{f_j(\boldsymbol{x})\}, \qquad (f_1,\ldots,f_m \text{ wypukłe})$$

Dowód:

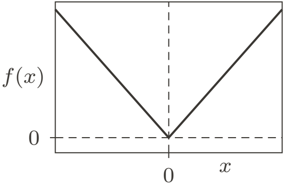
$$\begin{aligned}
f(\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}) &= \max_j \{f_j(\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y})\} \\
&\overset{(*)}{\leqslant} \max_j \{\lambda f_j(\boldsymbol{x})+(1-\lambda)f_j(\boldsymbol{y})\} \\
&\leqslant \max_j \{\lambda f_j(\boldsymbol{x})\} + \max_j \{(1-\lambda)f_j(\boldsymbol{y})\} \\
&= \lambda \max_j \{f_j(\boldsymbol{x})\} + (1-\lambda)\max_j \{f_j(\boldsymbol{y})\} \\
&= \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}),
\end{aligned}$$

gdzie w $(*)$ wykorzystaliśmy wypukłość $f_1,\ldots,f_m$ oraz fakt, że jeśli $a_j \leqslant b_j$ dla $j=1,\ldots,m$ to $\max_{j=1,\ldots,m}\{a_j\} \leqslant \max_{j=1,\ldots,m}\{b_j\}$

Przykłady funkcji wypukłych

- Wartość bezwzględna $f(x) = |x|$

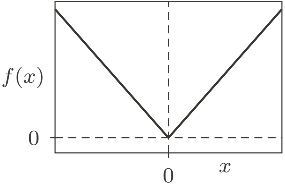
$$f(x) = |x| = \max\{x, -x\}$$



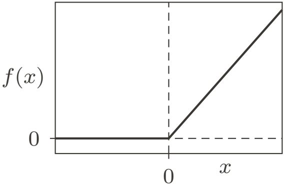
Przykłady funkcji wypukłych

- Wartość bezwzględna $f(x) = |x|$

$$f(x) = |x| = \max\{x, -x\}$$



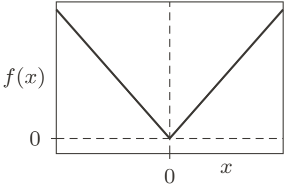
- Funkcja $f(x) = \max\{0, x\}$



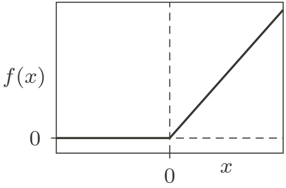
Przykłady funkcji wypukłych

- Wartość bezwzględna $f(x) = |x|$

$$f(x) = |x| = \max\{x, -x\}$$



- Funkcja $f(x) = \max\{0, x\}$



- Maksimum z m funkcji liniowych:

$$f(\boldsymbol{x}) \;=\; \max\left\{\boldsymbol{a}_1^\top \boldsymbol{x} + b_1, \boldsymbol{a}_2^\top \boldsymbol{x} + b_2, \ldots, \boldsymbol{a}_m^\top \boldsymbol{x} + b_m\right\}$$

Operacje zachowujące wypukłość

Złożenie funkcji **wypukłej** z funkcją **liniową** jest funkcją wypukłą

$$f(\boldsymbol{x}) \;=\; g(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}),$$

gdzie $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, $g\colon \mathbb{R}^m \to \mathbb{R}$ jest wypukła

Operacje zachowujące wypukłość

Złożenie funkcji **wypukłej** z funkcją **liniową** jest funkcją wypukłą

$$f(\boldsymbol{x}) \;=\; g(\boldsymbol{A}\boldsymbol{x}+\boldsymbol{b}),$$

gdzie $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, $g\colon \mathbb{R}^m \to \mathbb{R}$ jest wypukła

Dowód:

$$f(\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y}) \;=\; g\Big(\boldsymbol{A}(\lambda\boldsymbol{x}+(1-\lambda)\boldsymbol{y})+\boldsymbol{b}\Big)$$

Operacje zachowujące wypukłość

Złożenie funkcji **wypukłej** z funkcją **liniową** jest funkcją wypukłą

$$f(\boldsymbol{x}) \;=\; g(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}),$$

gdzie $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, $g \colon \mathbb{R}^m \to \mathbb{R}$ jest wypukła

Dowód:

$$\begin{aligned} f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \;&=\; g\Big(\boldsymbol{A}(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) + \boldsymbol{b}\Big) \\ &=\; g\Big(\lambda(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}) + (1-\lambda)(\boldsymbol{A}\boldsymbol{y} + \boldsymbol{b})\Big) \end{aligned}$$

Operacje zachowujące wypukłość

Złożenie funkcji **wypukłej** z funkcją **liniową** jest funkcją wypukłą

$$f(\boldsymbol{x}) \;=\; g(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}),$$

gdzie $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, $g \colon \mathbb{R}^m \to \mathbb{R}$ jest wypukła

Dowód:

$$\begin{aligned} f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \quad &= \quad g\Big(\boldsymbol{A}(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) + \boldsymbol{b}\Big) \\ &= \quad g\Big(\lambda(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}) + (1-\lambda)(\boldsymbol{A}\boldsymbol{y} + \boldsymbol{b})\Big) \\ \text{(wypukłość } g\text{)} \quad &\leqslant \quad \lambda g(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}) + (1-\lambda) g(\boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}) \end{aligned}$$

Operacje zachowujące wypukłość

Złożenie funkcji **wypukłej** z funkcją **liniową** jest funkcją wypukłą

$$f(\boldsymbol{x}) \;=\; g(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}),$$

gdzie $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{A} \in \mathbb{R}^{m \times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, $g \colon \mathbb{R}^m \to \mathbb{R}$ jest wypukła

Dowód:

$$\begin{aligned}
f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= g\Big(\boldsymbol{A}(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) + \boldsymbol{b}\Big) \\
&= g\Big(\lambda(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}) + (1-\lambda)(\boldsymbol{A}\boldsymbol{y} + \boldsymbol{b})\Big) \\
\text{(wypukłość } g\text{)} \quad &\leqslant \lambda g(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}) + (1-\lambda) g(\boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}) \\
&= \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})
\end{aligned}$$

Operacje zachowujące wypukłość

Złożenie funkcji **wypukłej** z funkcją **liniową** jest funkcją wypukłą

$$f(\boldsymbol{x}) \;=\; g(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}),$$

gdzie $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{A} \in \mathbb{R}^{m\times n}$, $\boldsymbol{b} \in \mathbb{R}^m$, $g\colon \mathbb{R}^m \to \mathbb{R}$ jest wypukła

Dowód:

$$\begin{aligned}
f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \quad &= \quad g\Big(\boldsymbol{A}(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) + \boldsymbol{b}\Big) \\
&= \quad g\Big(\lambda(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}) + (1-\lambda)(\boldsymbol{A}\boldsymbol{y} + \boldsymbol{b})\Big) \\
\text{(wypukłość } g\text{)} \quad &\leqslant \quad \lambda g(\boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}) + (1-\lambda) g(\boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}) \\
&= \quad \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})
\end{aligned}$$

Uwaga: Szczególny (i bardzo użyteczny) przypadek gdy $m = 1$:

$$f(\boldsymbol{x}) \;=\; g(\boldsymbol{a}^\top \boldsymbol{x} + b)$$

Przykłady funkcji wypukłych

Następujące funkcje są wypukłe jako złożenie funkcji wypukłej g z funkcją liniową:

- $f(\boldsymbol{x}) = (y - \boldsymbol{a}^\top \boldsymbol{x})^2$ (dla $g(x) = x^2$)
- $f(\boldsymbol{x}) = |y - \boldsymbol{a}^\top \boldsymbol{x}|$ (dla $g(x) = |x|$)
- $f(\boldsymbol{x}) = \max\{0, 1 - \boldsymbol{a}^\top \boldsymbol{x}\}$ (dla $g(x) = \max\{0, x\}$)
- $f(\boldsymbol{x}) = \ln(1 + e^{-\boldsymbol{a}^\top \boldsymbol{x}})$ (dla $g(x) = \ln(1 + e^{-x})$)
- $f(\boldsymbol{x}) = \|\boldsymbol{A}\boldsymbol{x} - \boldsymbol{b}\|^2$ (dla $g(\boldsymbol{y}) = \|\boldsymbol{y}\|^2$)

Operacje zachowujące wypukłość

Suma dowolnych funkcji **wypukłych** z **nieujemnymi** współczynnikami jest funkcją wypukłą

$$f(\boldsymbol{x}) \;=\; \alpha_1 f_1(\boldsymbol{x}) + \ldots \alpha_m f_m(\boldsymbol{x}), \qquad \alpha_1, \ldots, \alpha_m \geqslant 0$$

Operacje zachowujące wypukłość

Suma dowolnych funkcji **wypukłych** z **nieujemnymi** współczynnikami jest funkcją wypukłą

$$f(\boldsymbol{x}) \;=\; \alpha_1 f_1(\boldsymbol{x}) + \ldots \alpha_m f_m(\boldsymbol{x}), \qquad \alpha_1, \ldots, \alpha_m \geqslant 0$$

Dowód:

$$f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \;=\; \sum_{j=1}^{m} \alpha_j f_j(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y})$$

Operacje zachowujące wypukłość

Suma dowolnych funkcji **wypukłych** z **nieujemnymi** współczynnikami jest funkcją wypukłą

$$f(\boldsymbol{x}) \;=\; \alpha_1 f_1(\boldsymbol{x}) + \ldots \alpha_m f_m(\boldsymbol{x}), \qquad \alpha_1, \ldots, \alpha_m \geqslant 0$$

Dowód:

$$\begin{aligned} f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \quad &= \quad \sum_{j=1}^{m} \alpha_j f_j(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \\ \text{(wypukłość } f_j\text{)} \quad &\leqslant \quad \sum_{j=1}^{m} \alpha_j \left(\lambda f_j(\boldsymbol{x}) + (1-\lambda) f_j(\boldsymbol{y})\right) \end{aligned}$$

Operacje zachowujące wypukłość

Suma dowolnych funkcji **wypukłych** z **nieujemnymi** współczynnikami jest funkcją wypukłą

$$f(\boldsymbol{x}) \;=\; \alpha_1 f_1(\boldsymbol{x}) + \ldots \alpha_m f_m(\boldsymbol{x}), \qquad \alpha_1, \ldots, \alpha_m \geqslant 0$$

Dowód:

$$\begin{aligned} f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= \sum_{j=1}^{m} \alpha_j f_j(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \\ \text{(wypukłość } f_j\text{)} &\leqslant \sum_{j=1}^{m} \alpha_j \left(\lambda f_j(\boldsymbol{x}) + (1-\lambda) f_j(\boldsymbol{y})\right) \\ &= \lambda \Bigg(\sum_{j=1}^{m} \alpha_j f_j(\boldsymbol{x}) \Bigg) + (1-\lambda) \Bigg(\sum_{j=1}^{m} \alpha_j f_j(\boldsymbol{y}) \Bigg) \end{aligned}$$

Operacje zachowujące wypukłość

Suma dowolnych funkcji **wypukłych** z **nieujemnymi** współczynnikami jest funkcją wypukłą

$$f(\boldsymbol{x}) \;=\; \alpha_1 f_1(\boldsymbol{x}) + \ldots \alpha_m f_m(\boldsymbol{x}), \qquad \alpha_1, \ldots, \alpha_m \geqslant 0$$

Dowód:

$$\begin{aligned}
f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= \sum_{j=1}^{m} \alpha_j f_j(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \\
\text{(wypukłość } f_j\text{)} &\leqslant \sum_{j=1}^{m} \alpha_j \left(\lambda f_j(\boldsymbol{x}) + (1-\lambda) f_j(\boldsymbol{y})\right) \\
&= \lambda\left(\sum_{j=1}^{m} \alpha_j f_j(\boldsymbol{x})\right) + (1-\lambda)\left(\sum_{j=1}^{m} \alpha_j f_j(\boldsymbol{y})\right) \\
&= \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y})
\end{aligned}$$

Przykłady funkcji wypukłych

Wniosek: następujące funkcje są wypukłe (dla $\boldsymbol{x} \in \mathbb{R}^n$)

$$
\begin{aligned}
f(\boldsymbol{x}) &= e^{x_1} + \ldots e^{x_n} \\
f(\boldsymbol{x}) &= x_1 \ln x_1 + \ldots x_n \ln x_n \\
f(\boldsymbol{x}) &= (y_1 - \boldsymbol{a}_1^\top \boldsymbol{x}_1)^2 + \ldots + (y_m - \boldsymbol{a}_m^\top \boldsymbol{x}_m)^2 \\
f(\boldsymbol{x}) &= \ln(1 + e^{-\boldsymbol{a}_1^\top \boldsymbol{x}}) + \ldots + \ln(1 + e^{-\boldsymbol{a}_m^\top \boldsymbol{x}})
\end{aligned}
$$

Operacje zachowujące wypukłość

Złożenie funkcji $f(\boldsymbol{x}) = h(g(\boldsymbol{x}))$, gdzie $g\colon \mathbb{R}^n \to \mathbb{R}$ jest **wypukła**, a $h\colon \mathbb{R} \to \mathbb{R}$ jest **wypukła i niemalejąca**, jest funkcją wypukłą

Operacje zachowujące wypukłość

Złożenie funkcji $f(\boldsymbol{x}) = h(g(\boldsymbol{x}))$, gdzie $g\colon \mathbb{R}^n \to \mathbb{R}$ jest **wypukła**, a $h\colon \mathbb{R} \to \mathbb{R}$ jest **wypukła i niemalejąca**, jest funkcją wypukłą

Dowód:

$$f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) \quad = \quad h\Big(g\big(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\big)\Big)$$

Operacje zachowujące wypukłość

Złożenie funkcji $f(\boldsymbol{x}) = h(g(\boldsymbol{x}))$, gdzie $g\colon \mathbb{R}^n \to \mathbb{R}$ jest **wypukła**, a $h\colon \mathbb{R} \to \mathbb{R}$ jest **wypukła i niemalejąca**, jest funkcją wypukłą

Dowód:

$$\begin{aligned} f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= h\Big(g\big(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\big)\Big) \\ &\overset{(*)}{\leqslant} h\Big(\lambda g(\boldsymbol{x}) + (1-\lambda)g(\boldsymbol{y})\Big) \end{aligned}$$

Operacje zachowujące wypukłość

Złożenie funkcji $f(\boldsymbol{x}) = h(g(\boldsymbol{x}))$, gdzie $g\colon \mathbb{R}^n \to \mathbb{R}$ jest **wypukła**, a $h\colon \mathbb{R} \to \mathbb{R}$ jest **wypukła i niemalejąca**, jest funkcją wypukłą

Dowód:

$$\begin{aligned} f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) &= h\Big(g\big(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}\big)\Big) \\ &\overset{(*)}{\leqslant} h\Big(\lambda g(\boldsymbol{x}) + (1-\lambda) g(\boldsymbol{y})\Big) \end{aligned}$$

gdzie w $(*)$ wykorzystaliśmy wypukłość g i monotoniczność h

Operacje zachowujące wypukłość

Złożenie funkcji $f(\boldsymbol{x}) = h(g(\boldsymbol{x}))$, gdzie $g\colon \mathbb{R}^n \to \mathbb{R}$ jest **wypukła**, a $h\colon \mathbb{R} \to \mathbb{R}$ jest **wypukła i niemalejąca**, jest funkcją wypukłą

Dowód:

$$
\begin{array}{rcl}
f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) & = & h\Big(g\big(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}\big)\Big) \\
 & \overset{(*)}{\leqslant} & h\Big(\lambda g(\boldsymbol{x}) + (1-\lambda) g(\boldsymbol{y})\Big) \\
\text{(wypukłość } h\text{)} & \leqslant & \lambda h(g(\boldsymbol{x})) + (1-\lambda) h(g(\boldsymbol{y}))
\end{array}
$$

gdzie w $(*)$ wykorzystaliśmy wypukłość g i monotoniczność h

Operacje zachowujące wypukłość

Złożenie funkcji $f(\boldsymbol{x}) = h(g(\boldsymbol{x}))$, gdzie $g\colon \mathbb{R}^n \to \mathbb{R}$ jest **wypukła**, a $h\colon \mathbb{R} \to \mathbb{R}$ jest **wypukła i niemalejąca**, jest funkcją wypukłą

Dowód:

$$\begin{array}{rcl}
f(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y}) & = & h\Big(g(\lambda \boldsymbol{x} + (1-\lambda)\boldsymbol{y})\Big) \\
 & \overset{(*)}{\leqslant} & h\Big(\lambda g(\boldsymbol{x}) + (1-\lambda) g(\boldsymbol{y})\Big) \\
\text{(wypukłość } h\text{)} & \leqslant & \lambda h(g(\boldsymbol{x})) + (1-\lambda) h(g(\boldsymbol{y})) \\
 & = & \lambda f(\boldsymbol{x}) + (1-\lambda) f(\boldsymbol{y}),
\end{array}$$

gdzie w $(*)$ wykorzystaliśmy wypukłość g i monotoniczność h

Operacje zachowujące wypukłość

Złożenie funkcji $f(\boldsymbol{x}) = h(g(\boldsymbol{x}))$, gdzie $g\colon \mathbb{R}^n \to \mathbb{R}$ jest **wypukła**, a $h\colon \mathbb{R} \to \mathbb{R}$ jest **wypukła i niemalejąca**, jest funkcją wypukłą

Dowód:

$$\begin{array}{rcl} f(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}) & = & h\Big(g\big(\lambda\boldsymbol{x} + (1-\lambda)\boldsymbol{y}\big)\Big) \\ & \overset{(*)}{\leqslant} & h\Big(\lambda g(\boldsymbol{x}) + (1-\lambda)g(\boldsymbol{y})\Big) \\ \text{(wypukłość } h\text{)} & \leqslant & \lambda h(g(\boldsymbol{x})) + (1-\lambda)h(g(\boldsymbol{y})) \\ & = & \lambda f(\boldsymbol{x}) + (1-\lambda)f(\boldsymbol{y}), \end{array}$$

gdzie w $(*)$ wykorzystaliśmy wypukłość g i monotoniczność h

Przykład: funkcja $f(\boldsymbol{x}) = e^{g(\boldsymbol{x})}$ jest wypukła jeśli $g(\boldsymbol{x})$ jest wypukła