Techniki Optymalizacji: Optymalizacja wypukła

Wojciech Kotłowski

Instytut Informatyki Politechniki Poznańskiej
email: imię.nazwisko@cs.put.poznan.pl

pok. 2 (CW) tel. (61)665-2936 konsultacje: piątek 15:10-16:40
Slajdy dostępne pod adresem: `http://www.cs.put.poznan.pl/wkotlowski/to/`

30.10.2018

Kombinacja wypukła

- Kombinacją wypukłą wektorów (punktów) $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$ nazywamy dowolny wektor:

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

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

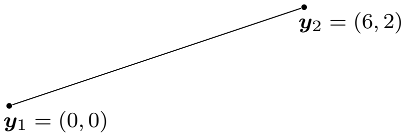
Kombinacja wypukła

- Kombinacją wypukłą wektorów (punktów) $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$ nazywamy dowolny wektor:

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

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

- Przykład: dla dwóch punktów $\boldsymbol{y}_1, \boldsymbol{y}_2$, ich wszystkie kombinacje wypukłe tworzą odcinek łączący $\boldsymbol{y}_1$ i $\boldsymbol{y}_2$.



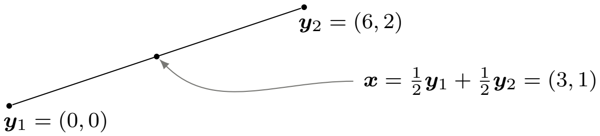
Kombinacja wypukła

- Kombinacją wypukłą wektorów (punktów) $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$ nazywamy dowolny wektor:

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

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

- Przykład: dla dwóch punktów $\boldsymbol{y}_1, \boldsymbol{y}_2$, ich wszystkie kombinacje wypukłe tworzą odcinek łączący $\boldsymbol{y}_1$ i $\boldsymbol{y}_2$.



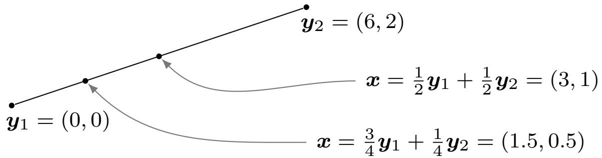
Kombinacja wypukła

- Kombinacją wypukłą wektorów (punktów) $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$ nazywamy dowolny wektor:

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

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

- Przykład: dla dwóch punktów $\boldsymbol{y}_1, \boldsymbol{y}_2$, ich wszystkie kombinacje wypukłe tworzą odcinek łączący $\boldsymbol{y}_1$ i $\boldsymbol{y}_2$.



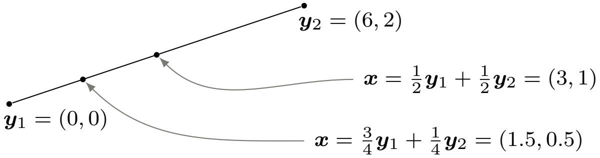
Kombinacja wypukła

- **Kombinacją wypukłą** wektorów (punktów) $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k$ nazywamy dowolny wektor:

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

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

- **Przykład**: dla dwóch punktów $\boldsymbol{y}_1, \boldsymbol{y}_2$, ich wszystkie kombinacje wypukłe tworzą odcinek łączący $\boldsymbol{y}_1$ i $\boldsymbol{y}_2$.



- **Pytanie**: jak wygląda kombinacja wypukła k punktów na płaszczyźnie?

Zbiór wypukły

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

Zbiór wypukły

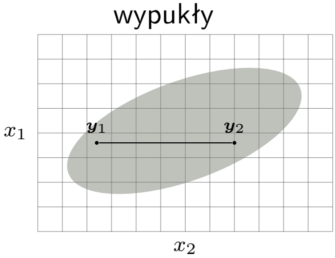
Zbiór wypukły $\mathcal{X}$ to taki zbiór, że dla jakichkolwiek punktów $\boldsymbol{y}_1, \ldots, \boldsymbol{y}_k \in \mathcal{X}$, każda ich kombinacja wypukła należy do $\mathcal{X}$.

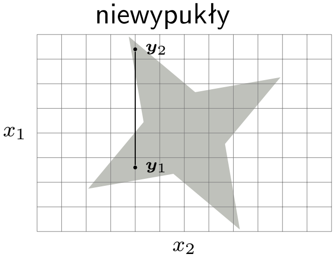
Alternatywnie, zbiór wypukły $\mathcal{X}$ to taki zbiór, że dla jakichkolwiek dwóch punktów $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathcal{X}$, odcinek je łączący w całości należy do $\mathcal{X}$.

Zbiór wypukły

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

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





Funkcja wypukła

Funkcja $f(\boldsymbol{x})$ jest wypukła, jeśli dla dowolnych dwóch punktów $\boldsymbol{x}_1, \boldsymbol{x}_2$ i dowolnego $\lambda \in [0, 1]$,

$$f(\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) \leq \lambda f(\boldsymbol{x}_1) + (1-\lambda) f(\boldsymbol{x}_2)$$

Funkcja wypukła

Funkcja $f(\boldsymbol{x})$ jest wypukła, jeśli dla dowolnych dwóch punktów $\boldsymbol{x}_1, \boldsymbol{x}_2$ i dowolnego $\lambda \in [0, 1]$,

$$f(\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) \leq \lambda f(\boldsymbol{x}_1) + (1-\lambda) f(\boldsymbol{x}_2)$$

Innymi słowy, funkcja $f(\boldsymbol{x})$ jest wypukła, jeśli odcinek łączący dwa punkty na wykresie leży w całości powyżej lub na wykresie funkcji.

Innymi słowy, epigraf funkcji $f(\boldsymbol{x})$ (zbiór ograniczony od dołu wykresem) jest zbiorem wypukłym.

Funkcja wypukła

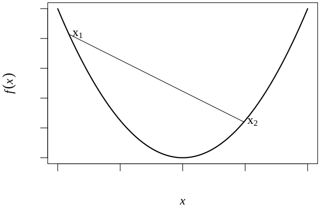
Funkcja $f(\boldsymbol{x})$ jest wypukła, jeśli dla dowolnych dwóch punktów $\boldsymbol{x}_1, \boldsymbol{x}_2$ i dowolnego $\lambda \in [0,1]$,

$$f(\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) \leq \lambda f(\boldsymbol{x}_1) + (1-\lambda) f(\boldsymbol{x}_2)$$

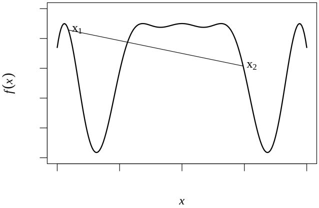
Innymi słowy, funkcja $f(\boldsymbol{x})$ jest wypukła, jeśli odcinek łączący dwa punkty na wykresie leży w całości powyżej lub na wykresie funkcji.

Innymi słowy, epigraf funkcji $f(\boldsymbol{x})$ (zbiór ograniczony od dołu wykresem) jest zbiorem wypukłym.

wypukła



niewypukła



Funkcja wypukła

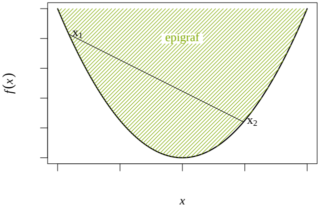
Funkcja $f(\boldsymbol{x})$ jest wypukła, jeśli dla dowolnych dwóch punktów $\boldsymbol{x}_1, \boldsymbol{x}_2$ i dowolnego $\lambda \in [0, 1]$,

$$f(\lambda \boldsymbol{x}_1 + (1 - \lambda)\boldsymbol{x}_2) \leq \lambda f(\boldsymbol{x}_1) + (1 - \lambda) f(\boldsymbol{x}_2)$$

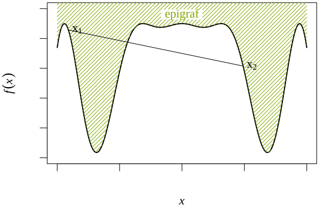
Innymi słowy, funkcja $f(\boldsymbol{x})$ jest wypukła, jeśli odcinek łączący dwa punkty na wykresie leży w całości powyżej lub na wykresie funkcji.

Innymi słowy, epigraf funkcji $f(\boldsymbol{x})$ (zbiór ograniczony od dołu wykresem) jest zbiorem wypukłym.

wypukła



niewypukła



Minimum lokalne

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

Minimum lokalne

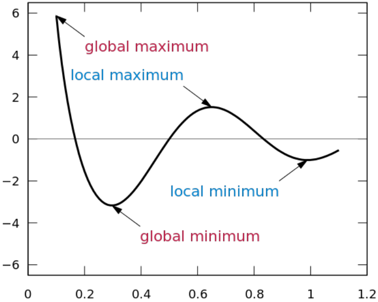
Funkcja $f(\boldsymbol{x})$ ma w punkcie $\boldsymbol{x}_0$ minimum lokalne, jeśli istnieje takie ϵ, że dla dowolnych $\boldsymbol{x}$ spełniających $\|\boldsymbol{x} - \boldsymbol{x}_0\| \leq \epsilon$, mamy $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x})$.

Funkcja $f(\boldsymbol{x})$ ma w punkcie $\boldsymbol{x}_0$ minimum globalne, jeśli dla dowolnych $\boldsymbol{x}$, mamy $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x})$.

Minimum lokalne

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

Funkcja $f(\boldsymbol{x})$ ma w punkcie $\boldsymbol{x}_0$ minimum globalne, jeśli dla dowolnych $\boldsymbol{x}$, mamy $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x})$.



(źródło: wikipedia)

Optymalizacja wypukła

Fakt

Rozważmy minimalizację funkcji wypukłej $f(\boldsymbol{x})$ na wypukłym zbiorze rozwiązań dopuszczalnych $\mathcal{X}$. 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 jest minimum globalne, tj. istnieje $\boldsymbol{x}_1 \in \mathcal{X}$ takie, że $f(\boldsymbol{x}_1) < f(\boldsymbol{x}_0)$.

Dowód:

- Niech $\boldsymbol{x}_0 \in \mathcal{X}$ – minimum lokalne. Załóżmy przeciwnie, że nie jest minimum globalne, tj. istnieje $\boldsymbol{x}_1 \in \mathcal{X}$ takie, że $f(\boldsymbol{x}_1) < f(\boldsymbol{x}_0)$.
- Ponieważ $\boldsymbol{x}_0$ jest minimum lokalnym, więc istnieje ϵ taki, że jeśli $\|\boldsymbol{x} - \boldsymbol{x}_0\| \leq \epsilon$, to $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x})$.

Dowód:

- Niech $\boldsymbol{x}_0 \in \mathcal{X}$ – minimum lokalne. Załóżmy przeciwnie, że nie jest minimum globalne, tj. istnieje $\boldsymbol{x}_1 \in \mathcal{X}$ takie, że $f(\boldsymbol{x}_1) < f(\boldsymbol{x}_0)$.
- Ponieważ $\boldsymbol{x}_0$ jest minimum lokalnym, więc istnieje ϵ taki, że jeśli $\|\boldsymbol{x} - \boldsymbol{x}_0\| \leq \epsilon$, to $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x})$.
- Weźmy kombinację wypukłą $\boldsymbol{x}_2 = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_0$. Zauważmy, że:
$$\|\boldsymbol{x}_2 - \boldsymbol{x}_0\| = \|\lambda(\boldsymbol{x}_1 - \boldsymbol{x}_0)\| = \lambda \|\boldsymbol{x}_1 - \boldsymbol{x}_0\|$$
czyli dla odpowiednio małego λ, będziemy mieli $\|\boldsymbol{x}_2 - \boldsymbol{x}_0\| \leq \epsilon$, a stąd $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x}_2)$.

Dowód:

- Niech $\boldsymbol{x}_0 \in \mathcal{X}$ – minimum lokalne. Załóżmy przeciwnie, że nie jest minimum globalne, tj. istnieje $\boldsymbol{x}_1 \in \mathcal{X}$ takie, że $f(\boldsymbol{x}_1) < f(\boldsymbol{x}_0)$.
- Ponieważ $\boldsymbol{x}_0$ jest minimum lokalnym, więc istnieje ϵ taki, że jeśli $\|\boldsymbol{x} - \boldsymbol{x}_0\| \leq \epsilon$, to $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x})$.
- Weźmy kombinację wypukłą $\boldsymbol{x}_2 = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_0$. Zauważmy, że:
$$\|\boldsymbol{x}_2 - \boldsymbol{x}_0\| = \|\lambda(\boldsymbol{x}_1 - \boldsymbol{x}_0)\| = \lambda\|\boldsymbol{x}_1 - \boldsymbol{x}_0\|$$
czyli dla odpowiednio małego λ, będziemy mieli $\|\boldsymbol{x}_2 - \boldsymbol{x}_0\| \leq \epsilon$, a stąd $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x}_2)$.
- Z drugiej strony, z wypukłości funkcji $f(\boldsymbol{x})$:
$$f(\boldsymbol{x}_2) \leq \lambda f(\boldsymbol{x}_1) + (1-\lambda) f(\boldsymbol{x}_0) < f(\boldsymbol{x}_0)$$

Dowód:

- Niech $\boldsymbol{x}_0 \in \mathcal{X}$ – minimum lokalne. Załóżmy przeciwnie, że nie jest minimum globalne, tj. istnieje $\boldsymbol{x}_1 \in \mathcal{X}$ takie, że $f(\boldsymbol{x}_1) < f(\boldsymbol{x}_0)$.
- Ponieważ $\boldsymbol{x}_0$ jest minimum lokalnym, więc istnieje ϵ taki, że jeśli $\|\boldsymbol{x} - \boldsymbol{x}_0\| \leq \epsilon$, to $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x})$.
- Weźmy kombinację wypukłą $\boldsymbol{x}_2 = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_0$. Zauważmy, że:
$$\|\boldsymbol{x}_2 - \boldsymbol{x}_0\| = \|\lambda(\boldsymbol{x}_1 - \boldsymbol{x}_0)\| = \lambda\|\boldsymbol{x}_1 - \boldsymbol{x}_0\|$$
czyli dla odpowiednio małego λ, będziemy mieli $\|\boldsymbol{x}_2 - \boldsymbol{x}_0\| \leq \epsilon$, a stąd $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x}_2)$.
- Z drugiej strony, z wypukłości funkcji $f(\boldsymbol{x})$:
$$f(\boldsymbol{x}_2) \leq \lambda f(\boldsymbol{x}_1) + (1-\lambda) f(\boldsymbol{x}_0) < f(\boldsymbol{x}_0)$$

sprzeczność!

Podsumowanie

- Problemy wypukłe mają tylko globalne minima
  $\Longrightarrow$ każda metoda szukająca lokalnych minimów, znajdzie również globalne minimum!

Podsumowanie

- Problemy wypukłe mają tylko globalne minima
  $\Longrightarrow$ każda metoda szukająca lokalnych minimów, znajdzie również globalne minimum!
- Problemy niewypukłe są często dużo trudniejsze:
  $\Longrightarrow$ Wiele (często: wykładniczo wiele) lokalnych minimów.

Podsumowanie

- Problemy wypukłe mają tylko globalne minima
$\Longrightarrow$ każda metoda szukająca lokalnych minimów, znajdzie również globalne minimum!
- Problemy niewypukłe są często dużo trudniejsze:
$\Longrightarrow$ Wiele (często: wykładniczo wiele) lokalnych minimów.
- Jeśli problem jest niewypukły, możesz starać się rozwiązać „najbliższe" wypukłe przybliżenie problemu (tzw. relaksacja).
Bardzo popularna metoda w uczeniu maszynowym!

Podsumowanie

- Problemy wypukłe mają tylko globalne minima
  $\Longrightarrow$ każda metoda szukająca lokalnych minimów, znajdzie również globalne minimum!
- Problemy niewypukłe są często dużo trudniejsze:
  $\Longrightarrow$ Wiele (często: wykładniczo wiele) lokalnych minimów.
- Jeśli problem jest niewypukły, możesz starać się rozwiązać „najbliższe" wypukłe przybliżenie problemu (tzw. relaksacja).
  Bardzo popularna metoda w uczeniu maszynowym!
- W mojej części przedmiotu będziemy rozwiązywać wyłącznie problemy wypukłe.

Podsumowanie

- Problemy wypukłe mają tylko globalne minima
  $\Longrightarrow$ każda metoda szukająca lokalnych minimów, znajdzie również globalne minimum!
- Problemy niewypukłe są często dużo trudniejsze:
  $\Longrightarrow$ Wiele (często: wykładniczo wiele) lokalnych minimów.
- Jeśli problem jest niewypukły, możesz starać się rozwiązać „najbliższe" wypukłe przybliżenie problemu (tzw. relaksacja). Bardzo popularna metoda w uczeniu maszynowym!
- W mojej części przedmiotu będziemy rozwiązywać wyłącznie problemy wypukłe.
- W części prof. Jaszkiewicza będziecie rozwiązywać problemy wyłącznie niewypukłe.