

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: wtorek 15:00-16:30
Slajdy dostępne pod adresem: <http://www.cs.put.poznan.pl/wkotlowski/>

02.12.2013

1 Łatwe i trudne w optymalizacji ciągłej

1 Łatwe i trudne w optymalizacji ciągłej

Kombinacja wypukła

Kombinacja wypukła

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

$$\mathbf{x} = \lambda_1 \mathbf{y}_1 + \dots + \lambda_k \mathbf{y}_k,$$

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

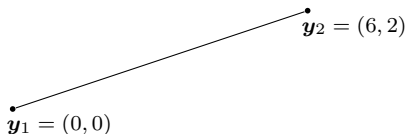
Kombinacja wypukła

- **Kombinacją wypukłą** wektorów (punktów) y_1, \dots, y_k nazywamy dowolny wektor:

$$x = \lambda_1 y_1 + \dots + \lambda_k y_k,$$

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

- **Przykład:** dla dwóch punktów y_1, y_2 , ich wszystkie kombinacje wypukłe tworzą odcinek łączący y_1 i y_2 .



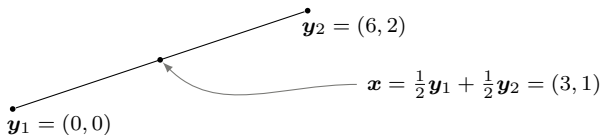
Kombinacja wypukła

- **Kombinacją wypukłą** wektorów (punktów) y_1, \dots, y_k nazywamy dowolny wektor:

$$x = \lambda_1 y_1 + \dots + \lambda_k y_k,$$

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

- **Przykład:** dla dwóch punktów y_1, y_2 , ich wszystkie kombinacje wypukłe tworzą odcinek łączący y_1 i y_2 .



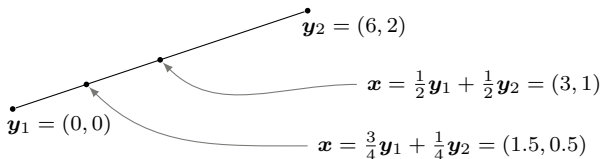
Kombinacja wypukła

- **Kombinacją wypukłą** wektorów (punktów) y_1, \dots, y_k nazywamy dowolny wektor:

$$x = \lambda_1 y_1 + \dots + \lambda_k y_k,$$

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

- **Przykład:** dla dwóch punktów y_1, y_2 , ich wszystkie kombinacje wypukłe tworzą odcinek łączący y_1 i y_2 .



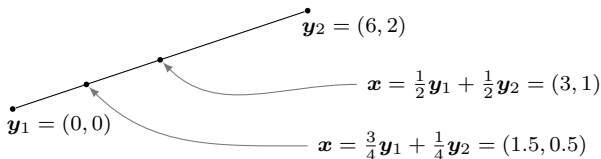
Kombinacja wypukła

- **Kombinacją wypukłą** wektorów (punktów) y_1, \dots, y_k nazywamy dowolny wektor:

$$x = \lambda_1 y_1 + \dots + \lambda_k y_k,$$

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

- **Przykład:** dla dwóch punktów y_1, y_2 , ich wszystkie kombinacje wypukłe tworzą odcinek łączący y_1 i y_2 .



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

Zbiór wypukły

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

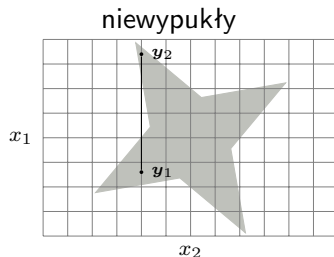
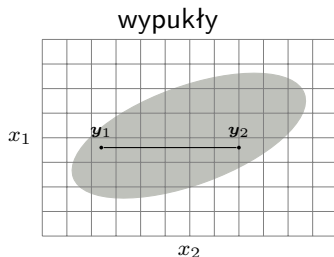
Zbiór wypukły \mathcal{X} to taki zbiór, że dla jakichkolwiek punktów $\mathbf{y}_1, \dots, \mathbf{y}_k \in \mathcal{X}$, każda ich kombinacja wypukła należy do \mathcal{X} .

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

Zbiór wypukły

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

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



Funkcja wypukła

Funkcja $f(x)$ jest **wypukła**, jeśli dla dowolnych dwóch punktów x_1, x_2 i dowolnego $\lambda \in [0, 1]$,

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

Funkcja wypukła

Funkcja $f(x)$ jest **wypukła**, jeśli dla dowolnych dwóch punktów x_1, x_2 i dowolnego $\lambda \in [0, 1]$,

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

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

Innymi słowy, **epigraf** funkcji $f(x)$ (zbiór ograniczony od dołu wykresem) jest **zbiorem wypukłym**.

Funkcja wypukła

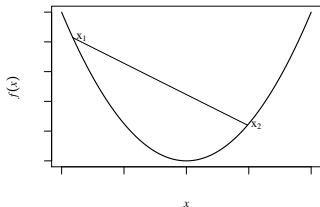
Funkcja $f(x)$ jest **wypukła**, jeśli dla dowolnych dwóch punktów x_1, x_2 i dowolnego $\lambda \in [0, 1]$,

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

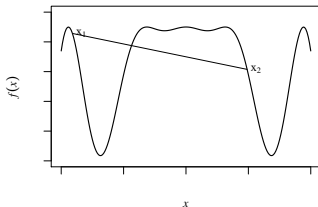
Innymi słowy, funkcja $f(x)$ jest **wypukła**, jeśli odcinek łączący dwa punkty na wykresie leży w całości powyżej lub na wykresie funkcji.

Innymi słowy, **epigraf** funkcji $f(x)$ (zbiór ograniczony od dołu wykresem) jest **zbiorem wypukłym**.

wypukła



niewypukła



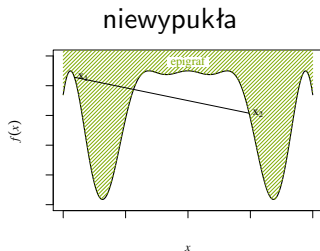
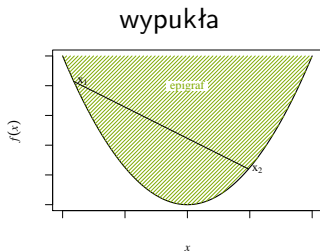
Funkcja wypukła

Funkcja $f(x)$ jest **wypukła**, jeśli dla dowolnych dwóch punktów x_1, x_2 i dowolnego $\lambda \in [0, 1]$,

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)$$

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

Innymi słowy, **epigraf** funkcji $f(x)$ (zbiór ograniczony od dołu wykresem) jest **zbiorem wypukłym**.



Minimum lokalne

Minimum lokalne

Funkcja $f(x)$ ma w punkcie x_0 **minimum lokalne**, jeśli istnieje takie ϵ , że dla dowolnych x spełniających $\|x - x_0\| \leq \epsilon$, mamy $f(x_0) \leq f(x)$.

Minimum lokalne

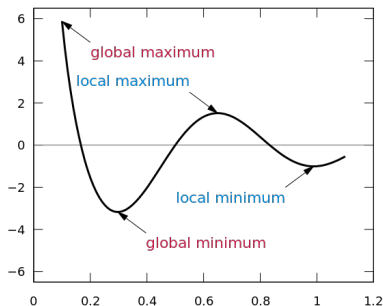
Funkcja $f(x)$ ma w punkcie x_0 **minimum lokalne**, jeśli istnieje takie ϵ , że dla dowolnych x spełniających $\|x - x_0\| \leq \epsilon$, mamy $f(x_0) \leq f(x)$.

Funkcja $f(x)$ ma w punkcie x_0 **minimum globalne**, jeśli dla dowolnych x , mamy $f(x_0) \leq f(x)$.

Minimum lokalne

Funkcja $f(x)$ ma w punkcie x_0 **minimum lokalne**, jeśli istnieje takie ϵ , że dla dowolnych x spełniających $\|x - x_0\| \leq \epsilon$, mamy $f(x_0) \leq f(x)$.

Funkcja $f(x)$ ma w punkcie x_0 **minimum globalne**, jeśli dla dowolnych x , mamy $f(x_0) \leq f(x)$.



(źródło: wikipedia)

Fakt

Rozważmy minimalizację funkcji wypukłej $f(x)$ na wypukłym zbiorze rozwiązań dopuszczalnych \mathcal{X} . Wtedy każde minimum lokalne funkcji $f(x)$ jest też jej minimum globalnym.

Dowód:

Dowód:

- Niech $x_0 \in \mathcal{X}$ – minimum lokalne. Załóżmy przeciwnie, że nie jest minimum globalne, tj. istnieje $x_1 \in \mathcal{X}$ takie, że $f(x_1) < f(x_0)$.

Dowód:

- Niech $x_0 \in \mathcal{X}$ – minimum lokalne. Załóżmy przeciwnie, że nie jest minimum globalne, tj. istnieje $x_1 \in \mathcal{X}$ takie, że $f(x_1) < f(x_0)$.
- Ponieważ x_0 jest minimum lokalnym, więc istnieje ϵ taki, że jeśli $\|x - x_0\| \leq \epsilon$, to $f(x_0) \leq f(x)$.

Dowód:

- Niech $x_0 \in \mathcal{X}$ – minimum lokalne. Załóżmy przeciwnie, że nie jest minimum globalne, tj. istnieje $x_1 \in \mathcal{X}$ takie, że $f(x_1) < f(x_0)$.
- Ponieważ x_0 jest minimum lokalnym, więc istnieje ϵ taki, że jeśli $\|x - x_0\| \leq \epsilon$, to $f(x_0) \leq f(x)$.
- Weźmy kombinację wypukłą $x_2 = \lambda x_1 + (1 - \lambda)x_0$.
Zauważmy, że:

$$\|x_2 - x_0\| = \|\lambda(x_1 - x_0)\| = \lambda\|x_1 - x_0\|$$

czyli dla odpowiednio małego λ , będziemy mieli

$$\|x_2 - x_0\| \leq \epsilon, \text{ a stąd } f(x_0) \leq f(x_2).$$

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

$$\|\mathbf{x}_2 - \mathbf{x}_0\| = \|\lambda(\mathbf{x}_1 - \mathbf{x}_0)\| = \lambda \|\mathbf{x}_1 - \mathbf{x}_0\|$$

czyli dla odpowiednio małego λ , będziemy mieli

$$\|\mathbf{x}_2 - \mathbf{x}_0\| \leq \epsilon, \text{ a stąd } f(\mathbf{x}_0) \leq f(\mathbf{x}_2).$$

- Z drugiej strony, z wypukłości funkcji $f(\mathbf{x})$:
$$f(\mathbf{x}_2) \leq \lambda f(\mathbf{x}_1) + (1 - \lambda)f(\mathbf{x}_0) < f(\mathbf{x}_0)$$

Dowód:

- Niech $x_0 \in \mathcal{X}$ – minimum lokalne. Załóżmy przeciwnie, że nie jest minimum globalne, tj. istnieje $x_1 \in \mathcal{X}$ takie, że $f(x_1) < f(x_0)$.
- Ponieważ x_0 jest minimum lokalnym, więc istnieje ϵ taki, że jeśli $\|x - x_0\| \leq \epsilon$, to $f(x_0) \leq f(x)$.
- Weźmy kombinację wypukłą $x_2 = \lambda x_1 + (1 - \lambda)x_0$.
Zauważmy, że:

$$\|x_2 - x_0\| = \|\lambda(x_1 - x_0)\| = \lambda\|x_1 - x_0\|$$

czyli dla odpowiednio małego λ , będziemy mieli

$$\|x_2 - x_0\| \leq \epsilon, \text{ a stąd } f(x_0) \leq f(x_2).$$

- Z drugiej strony, z wypukłości funkcji $f(x)$:
$$f(x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_0) < f(x_0)$$

sprzeczność!

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

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

- Problemy wypukłe mają tylko globalne minima
⇒ każda metoda szukająca lokalnych minimów, znajdzie również globalne minimum!
- Problemy niewypukłe są często dużo trudniejsze:
⇒ 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!

- Problemy wypukłe mają tylko globalne minima
⇒ każda metoda szukająca lokalnych minimów, znajdzie również globalne minimum!
- Problemy niewypukłe są często dużo trudniejsze:
⇒ 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.

- Problemy wypukłe mają tylko globalne minima
⇒ każda metoda szukająca lokalnych minimów, znajdzie również globalne minimum!
- Problemy niewypukłe są często dużo trudniejsze:
⇒ 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 ostatniej części przedmiotu (prof. Jaskiewicz) będziecie rozwiązywać problemy wyłącznie niewypukłe.