

Metody optymalizacji

Wykład nr 1

Paweł Zieliński

Katedra Podstaw Informatyki,
Wydział Informatyki i Telekomunikacji,
Politechnika Wrocławska



Problemy optymalizacyjne

Problemy optymalizacyjne możemy podzielić na:

- problemy optymalizacyjne, których zmienne decyzyjne są ciągłe (zmienne decyzyjne o ciągłych dziedzinach) – **problemy optymalizacyjne ciągłe**,
- problemy optymalizacyjne, których zmienne decyzyjne są dyskretne (zmienne decyzyjne o dyskretnych dziedzinach) – **problemy optymalizacyjne dyskretne**.



Problemy optymalizacyjne...

Definicja

Egzemplarz problemu optymalizacyjnego, oznaczonego przez $\mathcal{I}$, jest parą $(\mathbb{X}, f)$, gdzie $\mathbb{X}$ jest zbiorem rozwiązań dopuszczalnych, natomiast f jest funkcją celu (kosztu),

$$f : \mathbb{X} \rightarrow \mathbb{R}.$$

Zatem, szukamy rozwiązania dopuszczalnego $\mathbf{x}$, $\mathbf{x} \in \mathbb{X}$, takiego, że $f(\mathbf{x}) \leq f(\mathbf{y})$ dla wszystkich $\mathbf{y} \in \mathbb{X}$, czyli:

$$f(\mathbf{x}) = \min_{\mathbf{y} \in \mathbb{X}} f(\mathbf{y}).$$

*Rozwiązanie $\mathbf{x}$ jest rozwiązaniem **globalnie optymalnym** dla egzemplarza $\mathcal{I}$ lub po prostu **rozwiązaniem optymalnym**.*



Problemy optymalizacyjne...

Definicja

Problem optymalizacyjny jest zbiorem egzemplarzy problemu optymalizacyjnego – jest kolekcją egzemplarzy.



Minimalne drzewo rozpinające (MST)

MINIMUM SPANNING TREE:

Dane: Graf nieskierowany spójny $G = (V, E)$, gdzie V jest zbiorem wierzchołków, $|V| = n$, E jest zbiorem krawędzi, $|E| = m$, koszty krawędzi $c_e \in \mathbb{Z}_+$, $e \in E$. Zbiór rozwiązań dopuszczalnych jest następującej postaci:

$$\mathbb{X} = \{T = (V, E') : E' \subseteq E, T \text{ jest spójny i acykliczny}\}.$$

Funkcja celu (kosztu) $f : \mathbb{X} \rightarrow \mathbb{Z}_+$, $f(T) = \sum_{e \in T} c_e$.

Wyjście: Drzewo rozpinające T^* takie, że

$$f(T^*) = \min_{T \in \mathbb{X}} f(T).$$



Liniowe programowanie (LP)

LINEAR PROGRAMMING:

Dane: liczba wierszy (ograniczeń) m , liczba kolumn (zmiennych decyzyjnych) n , wektor prawych stron $\mathbf{b} \in \mathbb{R}^m$, wektor współczynników funkcji celu $\mathbf{c} \in \mathbb{R}^n$, macierz ograniczeń $\mathbf{A} = (a_{ij}) \in \mathbb{R}^{m \times n}$. Zbiór rozwiązań dopuszczalnych jest następującej postaci:

$$\mathbf{X} = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}.$$

Funkcja celu (kosztu) $f : \mathbb{X} \rightarrow \mathbb{R}$, $f(\mathbf{x}) = \mathbf{c}^T \mathbf{x} = \sum_{j \in [n]} c_j x_j$,

Wyjście: Rozwiązanie dopuszczalne x^* takie, że

$$f(\mathbf{x}^*) = \min_{\mathbf{x} \in \mathbb{X}} f(\mathbf{x}).$$



Egzemplarze problemów

Egzemplarz problemu liniowego programowania:

$$m = 1, n = 3, \mathbf{A} = [1, 1, 1], \mathbf{b} = [2], \mathbf{c}^T = [c_1, c_2, c_3]^T,$$

$$\begin{aligned} \mathbf{c}^T \mathbf{x} &= c_1 x_1 + c_2 x_2 + c_3 x_3 \rightarrow \min \\ \mathbb{X} &= \begin{cases} x_1 + x_2 + x_3 & = 2 \\ x_1, x_2, x_3 \geq 0 \end{cases} \end{aligned}$$

Egzemplarz minimalnego drzewa rozpinającego:

dane są: graf pełny o trzech wierzchołkach $n = 3$, $m = 3$, koszty krawędzi c_{12} , c_{13} , c_{23} .

Egzemplarz problemu wyrażony za pomocą programowania liniowego:

$$\begin{aligned} c_1 x_1 + c_2 x_2 + c_3 x_3 &\rightarrow \min \\ \mathbb{X} &= \begin{cases} x_1 + x_2 + x_3 & = 2 \\ x_1 \leq 1, x_2 \leq 1, x_3 \leq 1 \\ x_1, x_2, x_3 \geq 0 \end{cases} \end{aligned}$$

gdzie $c_1 = c_{12}$, $c_2 = c_{13}$, $c_3 = c_{23}$.



Otoczenie

Dla danego rozwiązania $\mathbf{x} \in \mathbb{X}$ otoczenie jest zbiorem rozwiązań dopuszczalnych, które leżą "blisko" $\mathbf{x}$.

Definicja

*Dla dowolnego problemu optymalizacyjnego, będącego zbiorem egzemplarzy $(\mathbb{X}, f)$, **otoczenie** jest odwzorowaniem*

$$N : \mathbb{X} \rightarrow 2^{\mathbb{X}}$$

zdefiniowanym dla każdego egzemplarza.



Lokalne i globalne optimum

Definicja

*Dla danego egzemplarza $(\mathbb{X}, f)$ problemu optymalizacyjnego i otoczenia N , rozwiązanie $\mathbf{x} \in \mathbb{X}$ jest nazywane **optimum lokalnym (minimum lokalnym)** względem otoczenia N , jeśli*

$$f(\mathbf{x}) \leq f(\mathbf{y}) \text{ dla wszystkich } \mathbf{y} \in N(\mathbf{x}).$$

Definicja

*Dany jest problem optymalizacyjny, zbiór rozwiązań dopuszczalnych $\mathbb{X}$ i otoczenie N . Mówimy, że otoczenie N jest **dokładne** (ang. exact), jeśli optimum lokalne $\mathbf{x} \in \mathbb{X}$ względem N jest również optimum globalnym.*

UWAGA: Otoczenie minimalnego drzewa rozpinającego jest dokładne.



Zbiory wypukłe i funkcje wypukłe

Definicja

Dowolnych punktów $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, punkt postaci

$$\mathbf{z} = \lambda \mathbf{x} + (1 - \lambda) \mathbf{y}, \quad \lambda \in [0, 1]$$

nazywamy *wypukłą kombinacją* $\mathbf{x}$ i $\mathbf{y}$.

Jeśli $\lambda \in (0, 1)$, $\mathbf{z}$ nazywamy *ściłą wypukłą kombinacją* $\mathbf{x}$ i $\mathbf{y}$.

Definicja

Zbiór $\mathbb{U} \subseteq \mathbb{R}^n$ jest *wypukły*, jeśli zawiera wszystkie wypukłe kombinacje par punktów $\mathbf{x}$ i $\mathbf{y}$ należących do $\mathbb{U}$.

UWAGA: Zbiór $\mathbb{R}^n$ jest wypukły. Każdy przedział jest wypukły.



Zbiory wypukłe i funkcje wypukłe...

Lemat (o przekroju zbiorów wypukłych)

Przekrój zbiorów wypukłych jest zbiorem wypukłym.

Dowód.

Oczywisty.



Zbiory wypukłe i funkcje wypukłe...

Definicja

Niech $\mathbb{U} \subseteq \mathbb{R}^n$ będzie zbiorem wypukłym. Funkcja $f : \mathbb{U} \rightarrow \mathbb{R}$ jest **wypukła** w zbiorze $\mathbb{U}$, jeśli dla dowolnych dwóch punktów $\mathbf{x}, \mathbf{y} \in \mathbb{U}$ zachodzi nierówność:

$$f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}) \leq \lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y}), \quad \lambda \in [0, 1].$$

Jeśli $\mathbb{U} = \mathbb{R}^n$, to mówimy, że f jest wypukła.

UWAGA: Dowolna liniowa funkcja jest wypukła w dowolnym zbiorze wypukłym $\mathbb{U}$.



Zbiory wypukłe i funkcje wypukłe...

Lemat (o ograniczeniu wypukłym)

Niech $f, f : \mathbb{U} \rightarrow \mathbb{R}$, będzie funkcją wypukłą w zbiorze $\mathbb{U}$. Wtedy dla dowolnego $t \in \mathbb{R}$ zbiór

$$\mathbb{U}_t = \{\mathbf{x} \in \mathbb{U} : f(\mathbf{x}) \leq t\}$$

jest wypukły.



Zbiory wypukłe i funkcje wypukłe...

Dowód.

Niech $\mathbf{x}$ i $\mathbf{y}$ będą dowolnymi punktami $\mathbf{x}, \mathbf{y} \in \mathbb{U}_t$. Zatem $\mathbf{x}, \mathbf{y} \in \mathbb{U}$. Z wypukłości $\mathbb{U}$ mamy, że $\lambda \mathbf{x} + (1 - \lambda) \mathbf{y} \in \mathbb{U}$, $\lambda \in [0, 1]$. Ponieważ f jest wypukła w $\mathbb{U}$ następujące nierówności są prawdziwe:

$$f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}) \leq \lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y}) \leq \lambda t + (1 - \lambda) t = t.$$

Stąd $\lambda \mathbf{x} + (1 - \lambda) \mathbf{y} \in \mathbb{U}_t$.



Zbiory wypukłe i funkcje wypukłe...

Definicja

Niech $U \subseteq \mathbb{R}^n$ będzie zbiorem wypukłym. Funkcja $f : U \rightarrow \mathbb{R}$ jest *wklęsła* w zbiorze U , jeśli $-f$ jest *wypukła* w zbiorze U .

UWAGA: Dowolna liniowa funkcja jest wklęsła w dowolnym zbiorze wypukłym U .



Zbiory wypukłe i funkcje wypukłe...

Twierdzenie (Charakteryzacja funkcji wypukłej - gradient)

Niech $f \in C^1$ (f ma ciągłe pochodne cząstkowe) w otwartym zbiorze zawierającym $\mathbb{U} \subseteq \mathbb{R}^n$. Wtedy f jest wypukła w zbiorze $\mathbb{U}$ wtedy i tylko wtedy, gdy

$$f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x})$$

dla wszystkich $\mathbf{x}, \mathbf{y} \in \mathbb{U}$, gdzie $\nabla f(\mathbf{x})$ jest gradientem w punkcie $\mathbf{x}$.



Zbiory wypukłe i funkcje wypukłe...

Dowód.

($\Rightarrow$) Niech $\mathbf{x}, \mathbf{y} \in \mathbb{U}$, $\mathbf{x} \neq \mathbf{y}$. Wtedy dla każdego $\lambda \in (0, 1]$

$$f(\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x})) = f(\lambda\mathbf{y} + (1 - \lambda)\mathbf{x}) \leq \lambda f(\mathbf{y}) + (1 - \lambda)f(\mathbf{x}).$$

Rozwijamy lewą stronę w szereg Taylor'a w $\mathbf{x}$

$$f(\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x})) = f(\mathbf{x}) + \nabla f(\boldsymbol{\eta})^T (\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x}) - \mathbf{x}) = f(\mathbf{x}) + \nabla f(\boldsymbol{\eta})^T (\mathbf{y} - \mathbf{x})\lambda,$$

gdzie $\boldsymbol{\eta}$ jest kombinacją wypukłą $\mathbf{x}$ i $\mathbf{x} + \lambda(\mathbf{y} - \mathbf{x})$. Zatem

$$\begin{aligned} f(\mathbf{x}) + \nabla f(\boldsymbol{\eta})^T (\mathbf{y} - \mathbf{x})\lambda &\leq \lambda f(\mathbf{y}) + (1 - \lambda)f(\mathbf{x}) = f(\mathbf{x}) + \lambda(f(\mathbf{y}) - f(\mathbf{x})) \\ \nabla f(\boldsymbol{\eta})^T (\mathbf{y} - \mathbf{x}) &\leq f(\mathbf{y}) - f(\mathbf{x}). \end{aligned}$$

Przechodząc do granicy $\lambda \rightarrow 0$, $\nabla f(\boldsymbol{\eta}) \rightarrow \nabla f(\mathbf{x})$.

Zbiory wypukłe i funkcje wypukłe...

Dowód.

$(\Leftarrow) f(\mathbf{y}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x})$ zachodzi dla wszystkich $\mathbf{x}, \mathbf{y} \in \mathbb{U}$.

Niech $\mathbf{z} = \lambda \mathbf{x} + (1 - \lambda) \mathbf{y}$. Stąd

$$f(\mathbf{x}) \geq f(\mathbf{z}) + \nabla f(\mathbf{z})^T (\mathbf{x} - \mathbf{z}) = f(\mathbf{z}) + \nabla f(\mathbf{z})^T (\mathbf{x} - \lambda \mathbf{x} - (1 - \lambda) \mathbf{y}),$$

$$f(\mathbf{y}) \geq f(\mathbf{z}) + \nabla f(\mathbf{z})^T (\mathbf{y} - \mathbf{z}) = f(\mathbf{z}) + \nabla f(\mathbf{z})^T (\mathbf{y} - \lambda \mathbf{x} - (1 - \lambda) \mathbf{y}).$$

Mnożąc pierwszą nierówność przez λ a drugą przez $1 - \lambda$ i dodając stronami otrzymujemy

$$\lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y}) \geq f(\mathbf{z}) = f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}).$$



Zbiory wypukłe i funkcje wypukłe...

Twierdzenie (Charakteryzacja funkcji wypukłej - Hesjan)

Niech $f \in C^2$ (f ma ciągłe drugie pochodne cząstkowe) w otwartym zbiorze zawierającym $\mathbb{U} \subseteq \mathbb{R}^n$. Wtedy f jest wypukła w zbiorze $\mathbb{U}$ wtedy i tylko wtedy, gdy

$$\mathbf{s}^T \nabla^2 f(\mathbf{x}) \mathbf{s} \geq 0$$

dla wszystkich $\mathbf{s} \in \mathbb{R}^n$ i dla wszystkich $\mathbf{x} \in \mathbb{U}$ gdzie $\nabla^2 f(\mathbf{x})$ jest Hesjanem w punkcie $\mathbf{x}$ (macierzą drugich pochodnych cząstkowych).



Zbiory wypukłe i funkcje wypukłe...

Dowód.

($\Leftarrow$) Niech $\mathbf{x}, \mathbf{y} \in \mathbb{U}$. Z twierdzenia Taylor'a

$$f(\mathbf{y}) = f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x}) + \frac{1}{2} (\mathbf{y} - \mathbf{x})^T \nabla^2 f(\boldsymbol{\eta}) (\mathbf{y} - \mathbf{x}),$$

gdzie $\boldsymbol{\eta}$ jest kombinacją wypukłą $\mathbf{x}$ i $\mathbf{y}$, stąd $\boldsymbol{\eta} \in \mathbb{U}$.

Z założenia $\mathbf{s}^T \nabla^2 f(\mathbf{x}) \mathbf{s} \geq 0$, co daje

$$f(\mathbf{y}) = f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x}) + \frac{1}{2} (\mathbf{y} - \mathbf{x})^T \nabla^2 f(\boldsymbol{\eta}) (\mathbf{y} - \mathbf{x}) \geq f(\mathbf{x}) + \nabla f(\mathbf{x})^T (\mathbf{y} - \mathbf{x}).$$

Zatem f jest wypukła.



Zbiory wypukłe i funkcje wypukłe...

Dowód.

($\Rightarrow$) Niech $\mathbf{x} \in \mathbb{U}$, $\mathbf{s} \in \mathbb{R}^n$ i niech $t > 0$ (małe), takie, że $\mathbf{x} + t\mathbf{s} \in \mathbb{U}$.
Rozwijając w szereg Taylora mamy

$$f(\mathbf{x} + t\mathbf{s}) = f(\mathbf{x}) + \nabla f(\mathbf{x})^T t\mathbf{s} + \frac{1}{2}t^2 \mathbf{s}^T \nabla^2 f(\boldsymbol{\eta}) \mathbf{s},$$

gdzie $\boldsymbol{\eta}$ jest kombinacją wypukłą $\mathbf{x} + t\mathbf{s}$ i $\mathbf{x}$.
Z założenia f jest wypukła, co daje

$$f(\mathbf{x}) + \nabla f(\mathbf{x})^T t\mathbf{s} \leq f(\mathbf{x} + t\mathbf{s}) = f(\mathbf{x}) + \nabla f(\mathbf{x})^T t\mathbf{s} + \frac{1}{2}t^2 \mathbf{s}^T \nabla^2 f(\boldsymbol{\eta}) \mathbf{s}.$$

Stąd $\frac{1}{2}t^2 \mathbf{s}^T \nabla^2 f(\boldsymbol{\eta}) \mathbf{s} \geq 0$. Dzieląc przez $t^2/2$ i przechodząc do granicy $t \rightarrow 0$ dostajemy $\nabla^2 f(\boldsymbol{\eta}) \rightarrow \nabla^2 f(\mathbf{x})$.

Zbiory wypukłe i funkcje wypukłe...

Twierdzenie (o otoczeniu dokładnym)

Niech $(f, \mathbb{X})$ będzie egzemplarzem problemu optymalizacyjnego, $\mathbb{X} \subseteq \mathbb{R}^n$ będzie zbiorem wypukłym a f będzie *funkcją wypukłą*. Wtedy otoczenie

$$N_\epsilon(\mathbf{x}) = \{\mathbf{y} : \mathbf{y} \in \mathbb{X}, \|\mathbf{x} - \mathbf{y}\|_2 \leq \epsilon\}$$

jest *dokładne* dla dowolnego $\epsilon > 0$.



Zbiory wypukłe i funkcje wypukłe...

Dowód.

Niech $\mathbf{x}$ będzie lokalnym optimum względem $N_\epsilon(\mathbf{x})$, $\epsilon > 0$, tj.

$$f(\mathbf{x}) \leq f(\mathbf{z}) \text{ dla wszystkich } \mathbf{z} \in N_\epsilon(\mathbf{x}).$$

Pokażemy, że $f(\mathbf{x}) \leq f(\mathbf{y})$ dla wszystkich $\mathbf{y} \in \mathbb{X}$. Oczywiście $N_\epsilon(\mathbf{x}) \subset \mathbb{X}$.

Wyberzmy dowolny $\mathbf{y} \in \mathbb{X} \setminus N_\epsilon(\mathbf{x})$.

Z wypukłości $\mathbb{X}$ i z faktu, że $\epsilon > 0$ istnieje $\mathbf{z} \in N_\epsilon(\mathbf{x})$, $\mathbf{z} \neq \mathbf{x}$, taki, że

$$\mathbf{z} = \lambda \mathbf{x} + (1 - \lambda) \mathbf{y} \text{ dla pewnego } \lambda \in (0, 1).$$

Z wypukłości f i lokalnej optymalności $\mathbf{x}$ mamy

$$f(\mathbf{x}) \leq f(\mathbf{z}) = f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}) \leq \lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y})$$

$$(1 - \lambda) f(\mathbf{x}) \leq (1 - \lambda) f(\mathbf{y}), \text{ gdzie } \lambda \in (0, 1)$$

$$f(\mathbf{x}) \leq f(\mathbf{y}).$$



Programowanie wypukłe

Definicja

Egzemplarz $(f, \mathbb{X})$ problemu optymalizacyjnego jest egzemplarzem problemu programowania wypukłego, jeśli f jest funkcją wypukłą, zbiór rozwiązań dopuszczalnych $\mathbb{X} \subseteq \mathbb{R}^n$ jest zdefiniowany następująco:

$$g_i(\mathbf{x}) \geq 0, \quad i \in [m],$$

gdzie $g_i : \mathbb{R}^n \rightarrow \mathbb{R}$ są funkcjami wklęsłymi.



Programowanie wypukłe...

Lemat

Zbiór rozwiązań dopuszczalnych $\mathbb{X}$ w problemie programowania wypukłego jest wypukły.

Dowód.

Wynika z lematu (o graniczeniu wypukłym) i lematu (o przekroju zbiorów wypukłych). □

Stąd i twierdzenia (o otoczeniu dokładnym) dostajemy:

Twierdzenie

W problemie programowania wypukłego każde rozwiązanie x lokalnie optymalne względem otoczenia $N_\epsilon(x)$ jest globalnie optymalne.



Programowanie wypukłe

Twierdzenie

Niech $f : \mathbb{X} \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$, $f \in C^1$ w zbiorze wypukłym $\mathbb{X}$ i f jest wypukła w $\mathbb{X}$. Jeśli dla $\mathbf{x}^* \in \mathbb{X}$,

$$\nabla f(\mathbf{x}^*) = \mathbf{0},$$

to $\mathbf{x}^*$ jest optymalny w $\mathbb{X}$.



Uwagi na temat treści wykładu

Treść wykładu w całości została przygotowana na podstawie książki



Christos H. Papadimitriou, Kenneth Steiglitz.

Combinatorial optimization: algorithms and complexity.

Dover Publications Inc., 1998.

