

Sformułowanie zadania optymalizacji nieliniowej bez ograniczeń:

Funkcja celu $f(x)$:

$$f: \mathbb{R}^n \rightarrow \mathbb{R}^1$$

Zadanie optymalizacji polega na znalezieniu wektora zmiennych decyzyjnych x , należącego do zbioru rozwiązań dopuszczalnych \mathbb{R}^n

takiego, że dla $\forall x \in \mathbb{R}^n$

$$f(\hat{x}) \leq f(x)$$

Co jest równoznaczne zapisowi:

$$\min_{x \in \mathbb{R}^n} f(x) = f(\hat{x})$$

Zadanie programowania nieliniowego bez ograniczeń

DEFINICJA.

Kierunkiem d w przestrzeni \mathbb{R}^n nazywamy dowolny n -wymiarowy wektor kolumnowy. Niech będzie dany punkt $x \in \mathbb{R}^n$ oraz skalar $\tau \in [0; +\infty)$. Dowolny punkt $y \in \mathbb{R}^n$ leżący na półprostej wychodzącej z

punktu x w kierunku $d \neq 0$ będzie wówczas określony zależnością

$$y = x + \tau d$$

LEMAT.

Niech $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^1$ będzie funkcją różniczkowalną w punkcie $x^0 \in X$. Załóżmy, że istnieje d , dla którego:

$$\langle \nabla f(x^0), d \rangle < 0,$$

Wówczas istnieje takie $\sigma > 0$, że dla wszystkich $\tau \in (0, \sigma]$ zachodzi

$$f(x^0 + \tau d) < f(x^0).$$

Dowód: wynika z własności różniczki Gateaux.

Zadanie programowania nieliniowego bez ograniczeń

Twierdzenie. Niech $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^1$ będzie funkcją różniczkowalną. Jeśli $\hat{x} \in X$ minimalizuje funkcję $f(x)$ tzn.

$$f(\hat{x}) \leq f(x), \quad \forall x \in X,$$

to

$$\nabla f(\hat{x}) = 0$$

Dowód: nie wprost.

Punkt \hat{x} jest nazywany punktem stacjonarym.

Twierdzenie. Niech $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^1$ będzie funkcją wypukłą i różniczkowalną. Punkt $\hat{x} \in X$ stanowi minimum funkcji $f(x)$, tzn.

$$f(\hat{x}) \leq f(x),$$

dla każdego $x \in X$ wtedy i tylko wtedy, gdy \hat{x} spełnia warunek

$$\nabla f(\hat{x}) = 0$$

Jest to warunek konieczny istnienia ekstremum lokalnego $f(x)$ w punkcie \hat{x} .

Minimum lokalne i globalne funkcji $f(x)$

Punkt \hat{x} stanowi minimum lokalne funkcji $f(x)$ w przestrzeni \mathbb{R}^n , jeżeli istnieje takie otwarte otoczenie $E \subset \mathbb{R}^n$ punktu \hat{x} , że

$$\forall x \in E \quad f(\hat{x}) \leq f(x)$$

Przy czym jeśli zachodzi $f(\hat{x}) < f(x)$ dla $x \neq \hat{x}$ to istnieje wtedy ściśle minimum lokalne.

Punkt \hat{x} stanowi minimum globalne funkcji $f(x)$ w przestrzeni \mathbb{R}^n , jeżeli istnieje takie otwarte otoczenie \mathbb{R}^n punktu \hat{x} , że

$$\forall x \in \mathbb{R}^n \quad f(\hat{x}) \leq f(x)$$

Przy czym jeśli zachodzi $f(\hat{x}) < f(x)$ dla $x \neq \hat{x}$ to ten punkt stanowi ściśle minimum globalne.

Minimum globalne funkcji $f(x)$

Twierdzenie:

Jeśli $f: X \subset \mathbb{R}^n \rightarrow \mathbb{R}^1$ będzie funkcją ściśle wypukłą i różniczkowalną, to wektor $\hat{x} \in X$ spełniający warunek konieczny $\nabla f(\hat{x}) = 0$ jest jedynym minimum globalnym funkcji $f(x)$.

Warunki wystarczające optymalizacji dla zadania bez ograniczeń

Funkcja $f(x)$ jest funkcją ciągłą i dwukrotnie różniczkowalną. Posiada macierz drugich pochodnych (hesjan) - A

Macierz A posiada ciąg podwyznaczników głównych $|A_i|$

$$|A_1| = \left| \frac{\partial^2 f}{\partial x_1^2}(x) \right|$$

$$|A_2| = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \dots & \frac{\partial^2 f}{\partial x_2^2}(x) \end{vmatrix}$$

$$|A_n| = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \dots & \dots & \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \dots & \frac{\partial^2 f}{\partial x_2^2}(x) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x) \\ \dots & \dots & \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) & \dots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x) & \dots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{vmatrix}$$

Warunki stacjonarności dla zadania optymalizacji nieliniowej bez ograniczeń cd.

Twierdzenie:

Założono, że \hat{x} jest punktem stacjonarnym funkcji $f(x)$. Wówczas zachodzą poniższe zależności:

1. Jeśli hesjan A jest dodatnio określony tzn: $\left|A(\hat{x})\right| > 0$ dla $i=1, \dots, n$ to funkcja $f(x)$ ma minimum lokalne w tym punkcie

2. Jeśli hesjan A jest ujemnie określony tzn: $(-1)\left|A(\hat{x})\right| > 0$ dla $i=1, \dots, n$ to funkcja $f(x)$ ma maksimum lokalne w tym punkcie

3. Jeśli hesjan A jest pół-dodatnio określony tzn: $\left|A(\hat{x})\right| \geq 0$ dla $i=1, \dots, n-1$ oraz $\left|A(\hat{x})\right|_n = 0$ bądź hesjan pół-ujemnie określony

$$(-1)\left|A(\hat{x})\right| \geq 0 \text{ dla } i=1, \dots, n-1 \text{ oraz } \left|A(\hat{x})\right|_n = 0$$

to nie można rozstrzygnąć o typie ekstremum funkcji $f(x)$ w tym punkcie

4. Jeśli nie są spełnione warunki 1 i 2 z nieostryimi nierównościami (wówczas hesjan A nie jest określony) to funkcja $f(x)$ nie ma ekstremum w punkcie \hat{x}

Warunek stacjonarności: poprawić gradient $\nabla f(x)$ i hesjan A

TWIERDZENIE. Jeśli funkcja $f(x)$ jest dwukrotnie różniczkowalna, to w każdym jej minimum lokalnym bez ograniczeń spełnione są następujące warunki konieczne optymalności zadania ZPN bez ograniczeń.

$$\nabla f(\hat{x}) = 0 \quad \text{warunek I rzędu}$$

$$d^T A(\hat{x}) d > 0 \quad \text{dla } \forall d \neq 0 \quad \text{warunek II rzędu}$$

$A = \nabla^2 f(\hat{x})$ Macierz A jest macierzą ściśle dodatnio określoną

• **Warunek I rzędu** jest często nazywany warunkiem stacjonarności, ponieważ oznacza zerowanie się pierwszej pochodnej.

• **Warunek II rzędu** dla funkcji dwukrotnie różniczkowalnych implikuje lokalną wypukłość minimalizowanej funkcji celu.

Sformułowanie zadania optymalizacji nieliniowej z ograniczeniami PN:

Znaleźć \hat{x} takie, że:

$$f(\hat{x}) = \min_{x \in X} f(x),$$

gdzie:

$$X = \{x : g_i(x) \leq 0, i=1, \dots, m\},$$

$$f: X \subset R^n \rightarrow R^1$$

$$\text{oraz } g_i: X \subset R^n \rightarrow R^1.$$

Sformułowanie zadania optymalizacji nieliniowej z ograniczeniami

DEFINICJA. Kierunek d poprowadzony z punktu x jest dopuszczalny, jeśli istnieje takie $\sigma > 0$ że dla dowolnego

$$\tau \in [0; \sigma], x + \tau d \in X,$$

gdzie X_0 oznacza zbiór określony jako PN.

Zbiór wszystkich kierunków dopuszczalnych:

$$D(x) = \{d : \exists \sigma > 0 \text{ takie, że}$$

$$\tau \in [0; \sigma] \Rightarrow x + \tau d \in X\}.$$

Funkcja Lagrange'a

DEFINICJA.

Funkcją Lagrange'a zadania PN nazywamy skalarną funkcję

$$L(x, \lambda) = f(x) + \langle \lambda, g(x) \rangle,$$

gdzie $\lambda \in R^m$ jest wektorem mnożników Lagrange'a.

Ograniczenia aktywne:

$$1. \quad g_i(\hat{x} + \tau d) \leq 0, \quad \forall i \in A(\hat{x})$$

przy czym $\tau \in [0, \tau]$

$$2. \quad g_i(\hat{x} + \tau d) = g_i(\hat{x}) + \tau \langle \nabla g_i(\hat{x}), d \rangle + O(\tau) \leq 0$$

3. warunkiem koniecznym na to, aby kierunek d był dopuszczalny jest:

$$\langle \nabla g_i(\hat{x}), d \rangle \leq 0, \quad \forall i \in A(\hat{x})$$

Lemat Farkasa

Niech będzie dany w R^n zbiór n-wymiarowych wektorów

$\{b, a^i, i = 1, \dots, m\}$. Nierówność

$$\langle b, x \rangle \geq 0$$

zachodzi dla każdego $x \in R^n$, spełniającego

$$\langle -a^i, x \rangle \geq 0,$$

Wtedy i tylko wtedy, gdy istnieje $\lambda = [\lambda_1, \dots, \lambda_m]^T \geq 0$

takie, że

$$b + \sum_{i=1}^m \lambda_i a^i = 0.$$

Wzajemnie rozłączne zbiory kierunków

$$D_1(\hat{x}) = \{d : \langle \nabla g_i(\hat{x}), d \rangle \leq 0, i \in A(\hat{x}), \langle \nabla f(\hat{x}), d \rangle \geq 0\}$$

$$D_2(\hat{x}) = \{d : \langle \nabla g_i(\hat{x}), d \rangle \leq 0, i \in A(\hat{x}), \langle \nabla f(\hat{x}), d \rangle < 0\}$$

$$D_3(\hat{x}) = \{d : \langle \nabla g_i(\hat{x}), d \rangle > 0 \text{ dla pewnych } i \in A(\hat{x})\}$$

$$D(\hat{x}) = D_1(\hat{x}) \cup D_2(\hat{x}) \cup D_3(\hat{x})$$

Warunki konieczne Kuhn'a-Tucker'a-Karusch'a

TWIERDZENIE. Jeśli

a) funkcje f i g_i są różniczkowalne;

b) \hat{x} jest lokalnym minimum ZPN, to istnieje $\hat{\lambda} \geq 0, \dim \hat{\lambda} = m$.

Takie, że

$$\nabla f(\hat{x}) + \sum_{i=1}^m \hat{\lambda}_i \nabla g_i(\hat{x}) = 0$$

$$\hat{\lambda}_i g_i(\hat{x}) = 0, \quad i = 1, \dots, m,$$

Wtedy i tylko wtedy, gdy

$$D_2(\hat{x}) = \emptyset$$

Dowód warunków koniecznych Kuhn'a-Tucker'a-Karusch'a

DOWÓD. Niech $b = \nabla f(\hat{x})$ oraz

$$a^i = \nabla g_i(\hat{x}), \quad \forall i \in A(\hat{x})$$

wówczas zgodnie z lematem Farkasa, istnieje $\hat{\lambda}_i \geq 0, i \in A(\hat{x})$

takie, że $\nabla f(\hat{x}) = \sum_{i \in A(\hat{x})} \hat{\lambda}_i (-\nabla g_i(\hat{x}))$

wtedy i tylko wtedy, gdy

$$\langle \nabla f(\hat{x}), d \rangle \geq 0, \quad \forall d \in R^n,$$

Spełniającego warunek dla ograniczeń;

$$\langle \nabla g_i(\hat{x}), d \rangle \leq 0, \quad \forall i \in A(\hat{x})$$

tzn. wtedy i tylko wtedy, gdy

$$D_2(\hat{x}) = \emptyset$$

Dowód warunków koniecznych Kuhn'a-Tucker'a-Karusch'a

Dla $i \notin A(\hat{x})$ należy przyjąć $\hat{\lambda}_i = 0$

Więc są spełnione dwa równania:

$$\nabla f(\hat{x}) + \sum_{i=1}^m \hat{\lambda}_i \nabla g_i(\hat{x}) = 0$$

$$\hat{\lambda}_i g_i(\hat{x}) = 0, \quad i = 1, \dots, m,$$

Ckd

Warunki konieczne Kuhn'a-Tucker'a-Karusch'a z wykorzystaniem funkcji Lagrange'a

$$L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x).$$

warunki konieczne:

$$\nabla_x L(\hat{x}, \hat{\lambda}) = 0,$$

$$\langle \hat{\lambda}, \nabla_\lambda L(\hat{x}, \hat{\lambda}) \rangle = 0$$

$$\nabla_\lambda L(\hat{x}, \hat{\lambda}) \leq 0$$

$$\hat{\lambda} \geq 0$$

Warunki wystarczające optymalizacji określane jako warunki regularności ograniczeń:

1. Wszystkie funkcje ograniczeń $g_i(x)$ są liniowe – warunek regularności Karlina.
2. Wszystkie funkcje ograniczeń $g_i(x)$ są funkcjami wypukłymi oraz zbiór rozwiązań dopuszczalnych ma niepuste wnętrze – warunek regularności Slatera
3. Gradienty wszystkich ograniczeń aktywnych, a więc $\nabla g_i(\hat{x})$, dla $i \in A(\hat{x})$, są liniowo niezależne – warunek regularności Fiacco i McCormicka.

Teoria i metody optymalizacji
Dr inż. Ewa Szlachcic Wydział Elektroniki studia II st.
kier. Automatyka i Robotyka

Twierdzenie Kuhn'a-Tucker'a-Karusch'a – warunki konieczne i wystarczające dla zadania optymalizacji nieliniowej z ograniczeniami

TWIERDZENIE. Jeśli

- a) funkcje f i g_i są różniczkowalne;
- b) \hat{x} jest lokalnym minimum ZPN,
- c) Warunek regularności ograniczeń Kuhn'a-Tucker'a-Karusch'a jest spełniony w \hat{x}

to istnieje $\hat{\lambda} \geq 0, \dim \hat{\lambda} = m$.

Takie, że w punkcie \hat{x} zachodzą warunki konieczne Kuhn'a-Tucker'a.

$$\nabla f(\hat{x}) + \sum_{i=1}^m \hat{\lambda}_i \nabla g_i(\hat{x}) = 0$$

$$\hat{\lambda}_i g_i(\hat{x}) = 0, \quad i = 1, \dots, m,$$

Teoria i metody optymalizacji
Dr inż. Ewa Szlachcic Wydział Elektroniki studia II st.
kier. Automatyka i Robotyka

Ilustracja warunków koniecznych i wystarczających Kuhn'a-Tucker'a-Karusch'a

Przykład I

Minimalizacja funkcji $f(x)$ na zbiorze ograniczeń nierównościowych X

$$\min_{x \in X} f(x) = x_1^2 + x_1 * x_2 + 0.5 * x_2^2 - x_1 - x_2$$

$$X = \left\{ x : \begin{array}{l} x_1 + x_2 \leq 0 \\ x_2 \geq 2 \end{array} \right\}$$

$$\hat{x} = [-2, 2]^T \quad f(\hat{x}) = 2$$

Teoria i metody optymalizacji
Dr inż. Ewa Szlachcic Wydział Elektroniki studia II st.
kier. Automatyka i Robotyka

Sformułowanie zadania optymalizacji nieliniowej PN z ograniczeniami dodatkowo na zmienne decyzyjne x :

$$f(\hat{x}) = \min_{x \in X} f(x),$$

$$X = \{x : g_i(x) \leq 0, x \geq 0, i = 1, \dots, m\},$$

gdzie:

$$f(x) : R^n \rightarrow R^1$$

$$\text{oraz } g_i(x) : R^n \rightarrow R^1, i = 1, \dots, m$$

Funkcja Lagrange'a

$$L(x, \lambda) = f(x) + \sum_{i=1}^m \lambda_i g_i(x)$$

Teoria i metody optymalizacji
Dr inż. Ewa Szlachcic Wydział Elektroniki studia II st.
kier. Automatyka i Robotyka

Warunki Lagrange'a dla ZPN z ograniczeniami na zmienne decyzyjne $x \geq 0$

$$\nabla_x L(\hat{x}, \hat{\lambda}) \geq 0 \qquad \nabla_x L(\hat{x}, \hat{\lambda}) \leq 0$$

$$\left\langle \hat{x}, \nabla_x L(\hat{x}, \hat{\lambda}) \right\rangle = 0 \qquad \left\langle \hat{\lambda}, \nabla_x L(\hat{x}, \hat{\lambda}) \right\rangle = 0$$

$$\hat{x} \geq 0 \qquad \hat{\lambda} \geq 0$$

Teoria i metody optymalizacji
Dr inż. Ewa Szlachcic Wydział Elektroniki studia II st.
kier. Automatyka i Robotyka

Ilustracja warunków koniecznych i wystarczających Kuhn'a-Tucker'a-Karusch'a

Przykład II

Minimalizacja funkcji $f(x)$ na zbiorze ograniczeń nierównościowych oraz ograniczeniach na znak zmiennej decyzyjnej

$$\min_{x \in X} f(x) = x_1^2 + x_1 * x_2 + 0.5 * x_2^2 - x_1 - x_2$$

$$X = \left\{ x : \begin{array}{l} x_1 + x_2 \leq 0 \\ x_2 \geq 0 \end{array} \right\}$$

$$\hat{x} = [2, 0]^T \quad f(\hat{x}) = 2$$

Teoria i metody optymalizacji
Dr inż. Ewa Szlachcic Wydział Elektroniki studia II st.
kier. Automatyka i Robotyka

Sformułowanie zadania optymalizacji nieliniowej PN z ograniczeniami mniejszościowymi i równościowymi:

$$f(\hat{x}) = \min_{x \in X} f(x)$$

gdzie:

$$X = \{x: g_i(x) \leq 0, i=1, \dots, p, h_i(x) = 0, i=p+1, \dots, m\},$$

gdzie

$$f(x): X = R^n \rightarrow R^1$$

$$g_i(x): X = R^n \rightarrow R^1, i=1, \dots, p$$

$$h_i(x): X = R^n \rightarrow R^1, i=p+1, \dots, m$$

Funkcja Lagrange'a

$$L(x, \lambda) = f(x) + \sum_{i=1}^p \lambda_i g_i(x) + \sum_{i=p+1}^m \lambda_i h_i(x).$$

Teoria i metody optymalizacji
Dr inż. Ewa Szlachcic

Wydział Elektroniki studia II st.
kier. Automatyka i Robotyka

Warunki Kuhn'a-Tuckera dla ZPN z ograniczeniami mniejszościowymi i równościowymi

Jeśli

- funkcje $f(x)$ i $g_i(x)$ dla $i=1, \dots, p$ oraz $h_i(x)$ dla $i=p+1, \dots, m$ są różniczkowalne;
- \hat{x} jest lokalnym minimum ZPN,

To istnieją $\hat{\lambda}_i, i=1, \dots, p$

oraz istnieją $\hat{\lambda}_i, i=p+1, \dots, m$ o nieograniczonym znaku, takie że:

$$\nabla f(\hat{x}) + \sum_{i=1}^p \hat{\lambda}_i \nabla g_i(\hat{x}) + \sum_{i=p+1}^m \hat{\lambda}_i \nabla h_i(\hat{x}) = 0$$

$$\hat{\lambda}_i g_i(\hat{x}) = 0, \quad i=1, \dots, p$$

$$\hat{\lambda}_i h_i(\hat{x}) = 0, \quad i=p+1, \dots, m$$

Teoria i metody optymalizacji
Dr inż. Ewa Szlachcic

Wydział Elektroniki studia II st.
kier. Automatyka i Robotyka

Ilustracja warunków koniecznych i wystarczających Kuhn'a-Tucker'a-Karusch'a

Przykład III

Minimalizacja funkcji $f(x)$ na zbiorze ograniczeń nierównościowych oraz zbiorze ograniczeń równościowych

$$\min_{x \in X} f(x) = x_1^2 + x_1 * x_2 + x_2^2$$

$$X = \left\{ x: \begin{array}{l} 2 * x_1 - x_2 \geq 4 \\ x_2 = 0 \end{array} \right\}$$

$$\hat{x} = [3, 2]^T \quad f(\hat{x}) = 19$$

Teoria i metody optymalizacji
Dr inż. Ewa Szlachcic

Wydział Elektroniki studia II st.
kier. Automatyka i Robotyka