Sformułowanie zadania optymalizacji nieliniowej bez ograniczeń:

Funkcja celu f(x) :

$$f(\mathbf{x}): R^n \longrightarrow R^1$$

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

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

$$f(\hat{\mathbf{x}}) \le f(\mathbf{x})$$

Co jest równoznaczne zapisowi:

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

Zadanie programowania nieliniowego bez ograniczeń

DEFINICJA.

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

$$\mathbf{y} = \mathbf{x} + \tau\mathbf{d}$$

LEMAT.

Niech $f: \mathbf{X} = R^n \to R^1$ będzie funkcją różniczkowalną w punkcie $\mathbf{x^0} \in \mathbf{X}$

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 $f: \mathbf{X} = R^n \to R^1$ będzie funkcją różniczkowalną . Jeśli $\hat{x} \in X$ minimalizuje funkcję *f(x)* tzn.

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

to

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

Dowód: nie wprost.

Punkt $\hat{x}$ jest nazywany punktem stacjonarnym.

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

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

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 *R^n, jeżeli istnieje takie otwarte otoczenie* $E \subset R^n$ *punktu* $\hat{x}$*, że*

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

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

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

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

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

Minimum globalne funkcji f(x)

Twierdzenie:

Jeśli $f: X = R^n \to R^1$ będzie funkcją ściśle wypukłą i różniczkowalną, to wektor $\hat{\mathbf{x}} \in X$ 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 $|A|_i$

$$|A|_1 = \left| \frac{\partial^2 f}{\partial x_1^2}(x) \right|$$

$$|A|_i = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & ..... & \frac{\partial^2 f}{\partial x_1 x_i}(x) \\ ..... & & ...... \\ \frac{\partial^2 f}{\partial x_i x_1}(x) & ..... & \frac{\partial^2 f}{\partial x_i^2}(x) \end{vmatrix}$$

$$|A|_n = \begin{vmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 x_i}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 x_n}(x) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_i x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_i^2}(x) & \cdots & \frac{\partial^2 f}{\partial x_i x_n}(x) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_n x_1}(x) & \cdots & \frac{\partial^2 f}{\partial x_n x_i}(x) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{vmatrix}$$

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: $\left|A(\hat{x})\right|_i > 0 \quad dla \quad i = 1,\dots,n$ to funkcja f(x) ma minimum lokalne w tym punkcie

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

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

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

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

4. Jeśli nie są spełnione warunki 1 i 2 z nieostrymi 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.

$\nabla f\left(\hat{x}\right) = 0$ warunek I rzędu

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

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

- **Warunek I** rzędu jest często nazywamy 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 $\hat{\mathbf{x}}$ takie, że:

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

$$\text{gdzie :}$$

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

$$f : X = R^n \to R^1$$

$$\text{oraz } g_i : X = R^n \to R^1.$$

Sformułowanie zadania optymalizacji nieliniowej z ograniczeniami

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

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

gdzie X_0 oznacza zbiór określony jako PN.

Zbiór wszystkich kierunków dopuszczalnych :

$$D(x) = \{d : \exists\sigma > 0 \text{ takie, że}$$
$$\tau \in [0;\sigma] \Rightarrow \mathbf{x} + \tau\mathbf{d} \in X\ \}.$$

Funkcja Lagrange'a

DEFINICJA.

Funkcją Lagrange'a zadania PN nazywamy skalarną funkcję

$$L(\mathbf{x},\boldsymbol{\lambda}) = f(\mathbf{x}) + \langle \boldsymbol{\lambda}, g(\mathbf{x}) \rangle,$$

gdzie $\boldsymbol{\lambda} \in R^n$ jest wektorem mnożników Lagrange'a.

Ograniczenia aktywne:

1. $g_i(\hat{\mathbf{x}} + \tau\mathbf{d}) \leq 0, \qquad \forall i \in A(\hat{\mathbf{x}})$
przy czym $\tau \in [0,\tau]$

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

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

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

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

$$b + \sum_{i=1}^{m} \lambda_1 a^i = 0.$$

Wzajemnie rozłączne zbiory kierunków

$$D_1(\hat{x}) = \left\{ d : < \nabla g_i(\hat{x}), d > \leq 0, i \in A(\hat{x}), < \nabla f(\hat{x}), d > \geq 0 \right\}$$

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

$$D_3(\hat{x}) = \left\{ d : < \nabla g_i(\hat{x}), d > > 0 \; dla \; pewnych \; i \in A(\hat{x}) \right\}$$

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

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

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

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

Wtedy i tylko wtedy, gdy

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

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

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

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

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

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

wtedy i tylko wtedy, gdy

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

Spełniającego warunek dla ograniczeń:,

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

tzn. wtedy i tylko wtedy, gdy

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

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

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

Więc są spełnione dwa równania:

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

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

Ckd

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

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

warunki konieczne:

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

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

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

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

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.**

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

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

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

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

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

Przykład I

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

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

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

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

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

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

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

$$gdzie:$$

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

oraz $g_i(x) : R^n \to R^1, i = 1,\ldots,m$

Funkcja Lagrange'a

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

Warunki Lagrange'a dla ZPN z ograniczeniami na zmienne decyzyjne x≥0

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

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

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

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

Przykład II

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

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

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

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

Sformułowanie zadania optymalizacji nieliniowej PN z ograniczeniami mniejszościowymi i równościowymi:

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

gdzie :

$$X = \{\mathbf{x} : g_i(\mathbf{x}) \le 0,\ i = 1,..., p,\quad h_i(x) = 0, i = p+1,..., m\},$$

gdzie

$$f(x) : X = R^n \to R^1$$
$$g_i(x) : X = R^n \to R^1, i = 1,..., p$$
$$h_i(x) : X = R^n \to R^1, i = p+1,..., m$$

Funkcja Lagrange'a

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

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

Jeśli

a) funkcje $f(x)$ i $g_i(x)$ dla $i = 1,..., p$ oraz $h_i(x)$ *dla i=p+1,…,m* są różniczkowalne;

b) $\hat{x}$ jest lokalnym minimum ZPN,

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

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

$$\nabla f(\hat{\mathbf{x}}) + \sum_{i=1}^{p} \hat{\boldsymbol{\lambda}}_i \nabla g_i(\hat{\mathbf{x}}) + \sum_{i=p+1}^{m} \hat{\boldsymbol{\lambda}}_i \nabla h_i(\hat{\mathbf{x}}) = 0$$
$$\hat{\lambda}_i g_i(\hat{\mathbf{x}}) = 0, \quad \mathrm{i} = 1,..., \mathrm{p}$$
$$\hat{\lambda}_i h_i(\hat{\mathbf{x}}) = 0, \quad \mathrm{i} = p+1,..., \mathrm{m}$$

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

Przykład III

Minimalizacja funkcji *f(x)* na zbiorze ograniczeń nierówności owych oraz zbiorze ograniczeń równościowych

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

$$X = \left\{ x : \begin{matrix} 2 * x_1 - x_2 \ge 4 \\ x_2 = 0 \end{matrix} \right\}$$

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