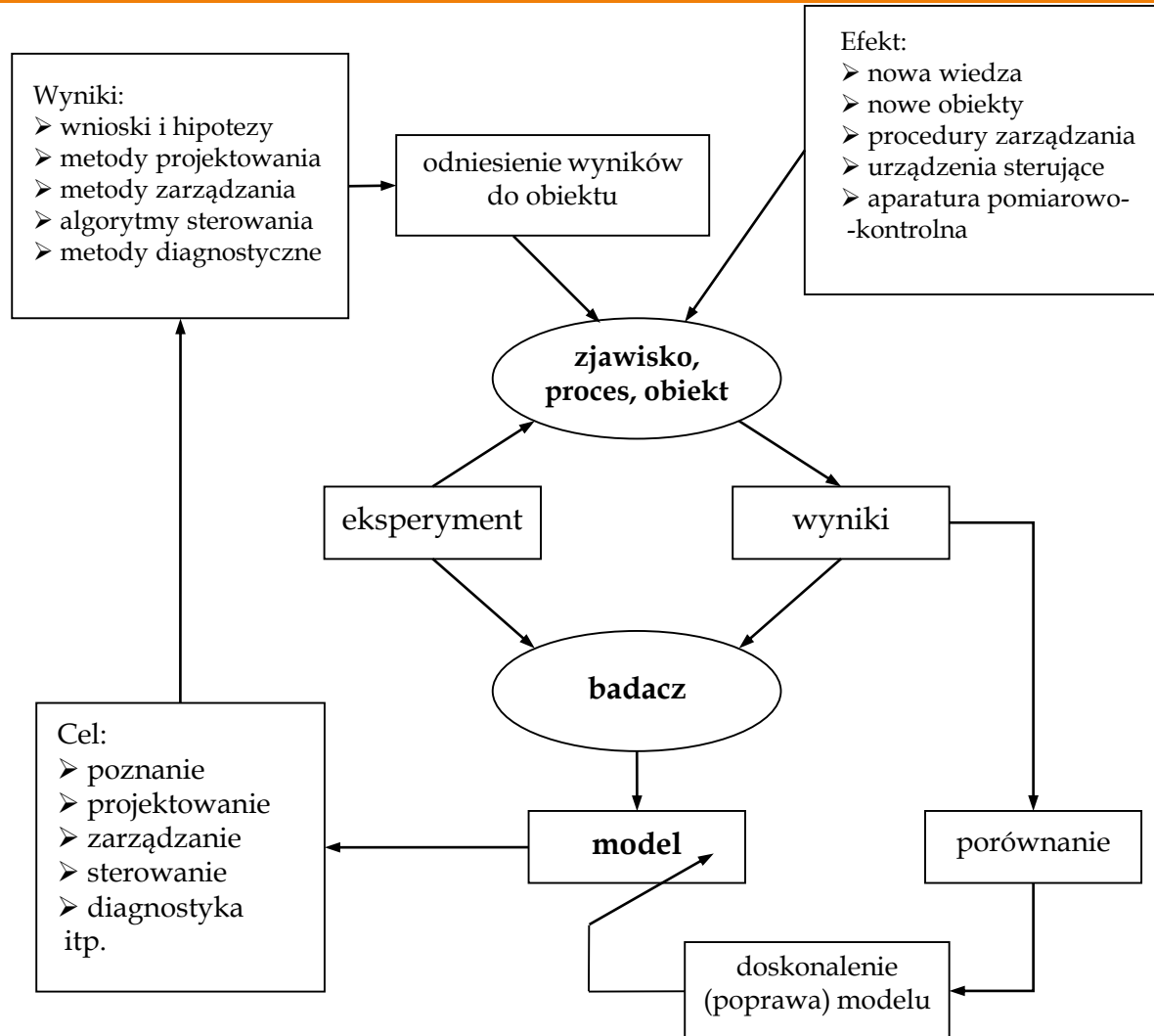


Metody Systemowe i Decyzyjne w Informatyce



Wykład 3. Elementy wspomagania decyzji

Model w badaniach systemowych



- ⌘ Sformułowanie zadania decyzyjnego
- ⌘ Typowe zadania decyzyjne
- ⌘ Analityczne metody optymalizacji
 - Zadanie optymalizacji bez ograniczeń
 - Zadanie optymalizacji z ograniczeniami równościowymi metoda mnożników Lagrange' a
 - Zadanie optymalizacji z ograniczeniami nierównościowymi metoda Kuhna – Tucker'a
- ⌘ Numeryczne metody optymalizacji - wprowadzenie

Sformułowanie zadania wspomagania decyzji

Decyzja: różne typy elektrowni



Obrazy:

<http://ziemianarozdrozu.pl/encyklopedia/67/hydroenergetyka/>

http://kresy24.pl/showNews/news_id/5871/

<http://windy-future.info/2009/10/13/large-wind-turbine/>

Elektrownia wodna

$x^{(1)}$

Elektrownia atomowa

$x^{(2)}$

Elektrownia wiatrowa

$x^{(3)}$

$x^{(1)}, x^{(2)}, x^{(3)}$ - obciążenie elektrowni –
zmienne decyzyjne

c_1, c_2, c_3 – koszt jednostkowy
wytworzenia energii

Cel: minimalizacja kosztów wytworzenia: $F(x^{(1)}, x^{(2)}, x^{(3)}) = c_1 x^{(1)} + c_2 x^{(2)} + c_3 x^{(3)}$

Ograniczenia: – spełnienie zapotrzebowania na energię: $x^{(1)} + x^{(2)} + x^{(3)} \geq \beta$
– możliwości poszczególnych elektrowni: $0 \leq x^{(n)} \leq \alpha_n, n = 1, 2, 3$

Sformułowanie zadania wspomagania decyzji

Zmienne decyzyjne: $x = \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(s)} \end{bmatrix}$

Funkcja celu: $y = F(x)$

Zbiór rozwiązań dopuszczalnych (zwykle określony przez wskazanie ograniczeń):

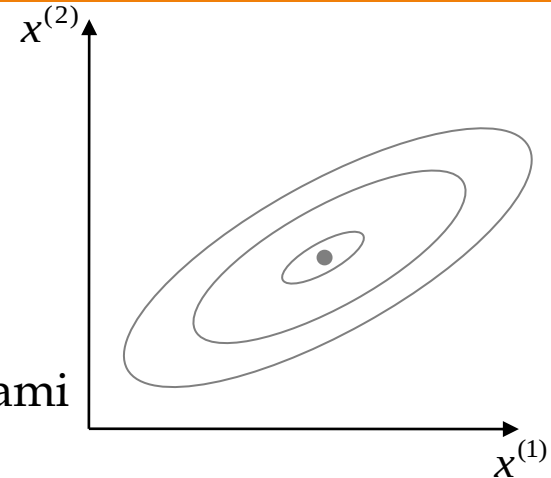
$$x \in \mathcal{D}_x$$

Decyzja optymalna: $x^* \rightarrow F(x^*) = \min_{x \in \mathcal{D}_x} F(x), \quad \min F(x) = -\max(-F(x))$

Decyzja zadowalająca: $\tilde{x} \rightarrow F(\tilde{x}) \leq F(x^*) + \alpha = \gamma \wedge x \in \mathcal{D}_x$

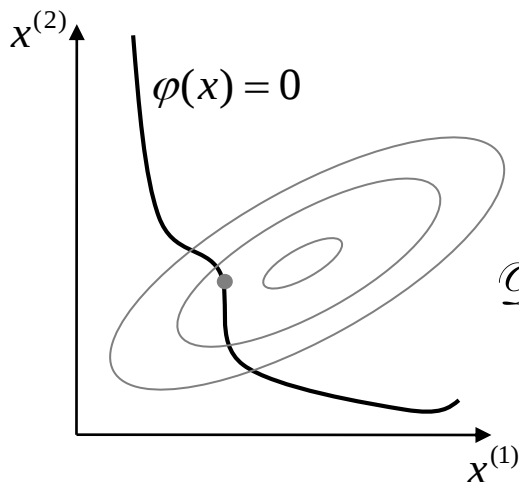
Typowe zadania decyzyjne

Zadanie decyzyjne bez ograniczeń: $\mathcal{D}_x = \mathcal{R}^S$



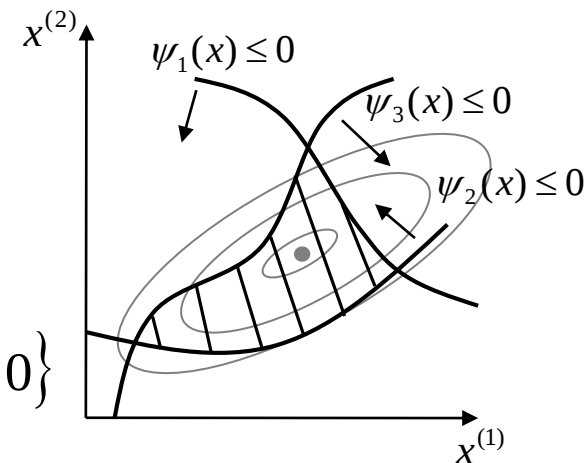
Zadanie decyzyjne z ograniczeniami równościowymi:

$$\mathcal{D}_x = \{x \in \mathcal{R}^S : \varphi_1(x) = 0, \varphi_2(x) = 0, \dots, \varphi_L(x) = 0, L \leq S\}$$



Zadanie decyzyjne z ograniczeniami nierównościami:

$$\mathcal{D}_x = \{x \in \mathcal{R}^S : \psi_1(x) \leq 0, \psi_2(x) \leq 0, \dots, \psi_M(x) \leq 0\}$$



Analityczne metody optymalizacji

- ⌘ Zadanie optymalizacji bez ograniczeń
- ⌘ Zadanie optymalizacji z ograniczeniami równościowymi – metoda współczynników Lagrange’a
- ⌘ Zadanie optymalizacji z ograniczeniami nierównościowymi – metoda Kuhna-Tuckera

Typowe zadania decyzyjne

Programowanie liniowe

Zmienne decyzyjne: $x \in \mathcal{D}_x \subseteq \mathcal{R}^S$

Funkcja celu: $F(x) = c^T x = \sum_{s=1}^S c_s x^{(s)}$ $c = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_S \end{bmatrix}$

Ograniczenia:

$$\varphi_l(x) = a_l^T x - \alpha_l = 0, \quad a_l = \begin{bmatrix} a_l^{(1)} \\ a_l^{(2)} \\ \vdots \\ a_l^{(S)} \end{bmatrix}$$
$$l = 1, 2, \dots, L$$

$$\psi_m(x) = b_m^T x - \beta_m \leq 0, \quad b_m = \begin{bmatrix} b_m^{(1)} \\ b_m^{(2)} \\ \vdots \\ b_m^{(S)} \end{bmatrix}$$
$$m = 1, 2, \dots, M$$

Typowe zadania decyzyjne

⌘ Programowanie kwadratowe

Zmienne decyzyjne: $x \in \mathcal{D}_x \subseteq \mathcal{R}^S$

Funkcja celu: $F(x) = x^T A x + b^T x + c$ $A \in \mathcal{R}^{S \times S}$, $b \in \mathcal{R}^S$, $c \in \mathcal{R}$

Ograniczenia:

$$\begin{aligned} \varphi_l(x) &= d_l^T x - \alpha_l = 0, \\ l &= 1, 2, \dots, L \end{aligned} \quad d_l = \begin{bmatrix} d_l^{(1)} \\ d_l^{(2)} \\ \vdots \\ d_l^{(S)} \end{bmatrix} \quad \begin{aligned} \psi_m(x) &= e_m^T x - \beta_m \leq 0, \\ m &= 1, 2, \dots, M \end{aligned} \quad e_l = \begin{bmatrix} e_m^{(1)} \\ e_m^{(2)} \\ \vdots \\ e_m^{(S)} \end{bmatrix}$$

Typowe zadania decyzyjne

∞ Programowanie ilorazowe

Zmienne decyzyjne: $x \in \mathcal{D}_x \subseteq \mathcal{R}^S$

Funkcja celu: $F(x) = \frac{a^T x + b}{c^T x + d}$

$$a \in \mathcal{R}^S, b \in \mathcal{R}, c \in \mathcal{R}^S, d \in \mathcal{R}$$

Ograniczenia:

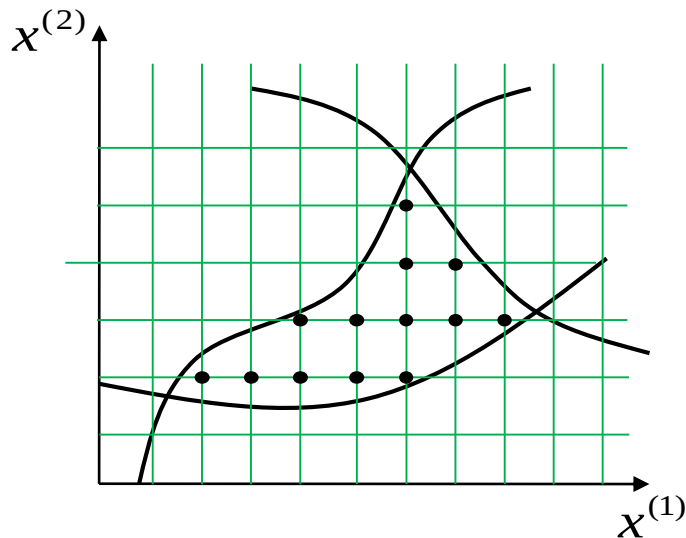
$$\varphi_l(x) = p_l^T x - \alpha_l = 0, \quad l = 1, 2, \dots, L$$
$$p_l = \begin{bmatrix} p_l^{(1)} \\ p_l^{(2)} \\ \vdots \\ p_l^{(S)} \end{bmatrix}$$

$$\psi_m(x) = q_m^T x - \beta_m \leq 0, \quad m = 1, 2, \dots, M$$
$$q_l = \begin{bmatrix} q_m^{(1)} \\ q_m^{(2)} \\ \vdots \\ q_m^{(S)} \end{bmatrix}$$

Typowe zadania decyzyjne

⌘ Programowanie całkowitoliczbowe

Zmienne decyzyjne: $x \in \overline{\mathcal{D}}_x = \mathcal{D}_x \cap \{x^{(s)} \in \mathcal{C}, s = 1, 2, \dots, S\}$

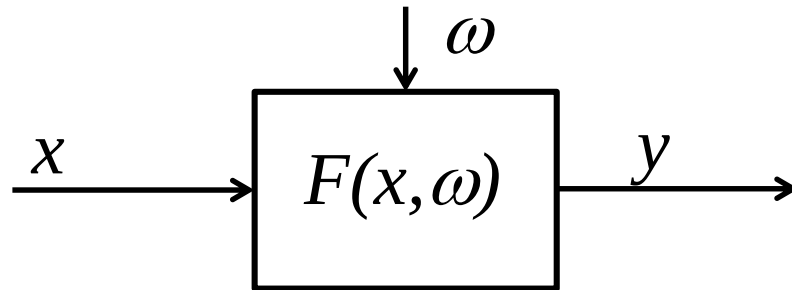


Inne równoważne

$x \in \overline{\mathcal{D}}_x = \{x_1, x_2, \dots, x_M\}$ - typoszereg

$x \in \overline{\mathcal{D}}_x = \{x^{(s)} \in \{0, 1\}, s = 1, 2, \dots, S\}$ - programowanie
zero - jedynkowe (Boolowskie)

Podjmowanie decyzji w warunkach niepewności



$$F(x) = E_{\underline{\omega}}[F(x, \underline{\omega})]$$

$$\mathcal{D}_x = E_{\underline{\omega}}[\mathcal{D}_x(\underline{\omega})] =$$

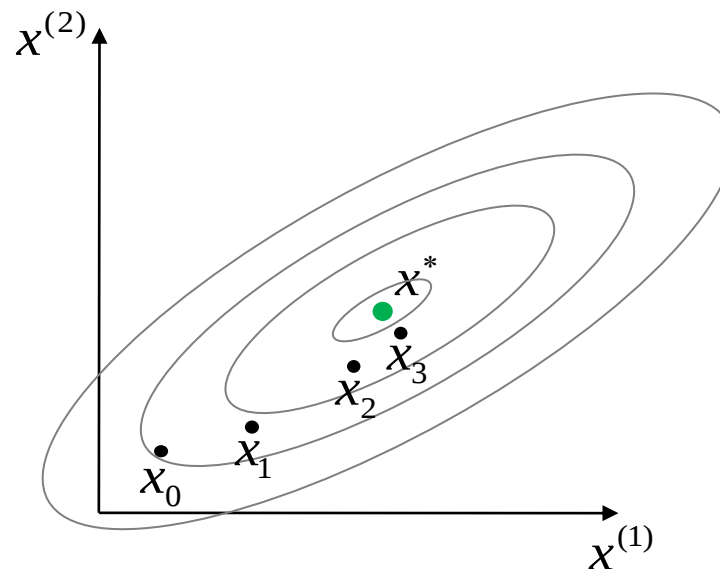
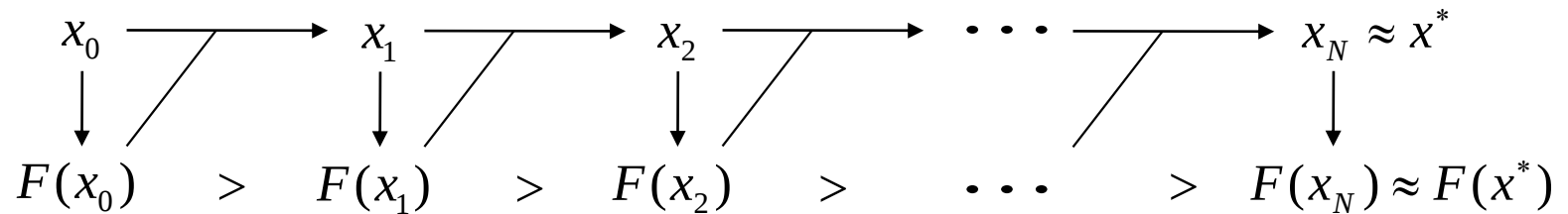
$$= \left\{ x \in \mathcal{R}^S; E_{\underline{\omega}}[\varphi_l(x, \underline{\omega})] = 0, l = 1, \dots, L, E_{\underline{\omega}}[\psi_m(x, \underline{\omega})] \leq 0, m = 1, \dots, M \right\}$$

$$x^* \rightarrow F(x^*) = \min_{x \in D_x} F(x)$$

Metody numeryczne

$$x^* \rightarrow F(x^*) = \min_{x^* \in \mathcal{D}_x} F(x),$$

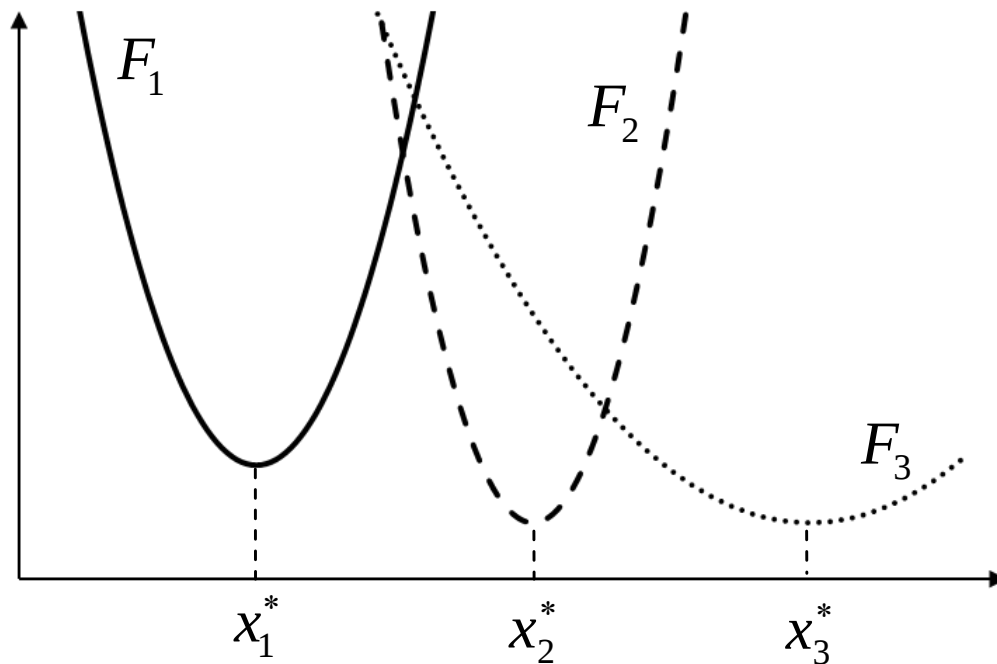
Konstruujemy ciąg przybliżeń na podstawie wartości funkcji $F(x)$ w danym punkcie x



Zadanie polioptymalizacji

X – wektor zmiennych decyzyjnych

$F_1(x), F_2(x), \dots, F_M(x)$ – kilka różnych funkcji celu – ocena wielokryterialna

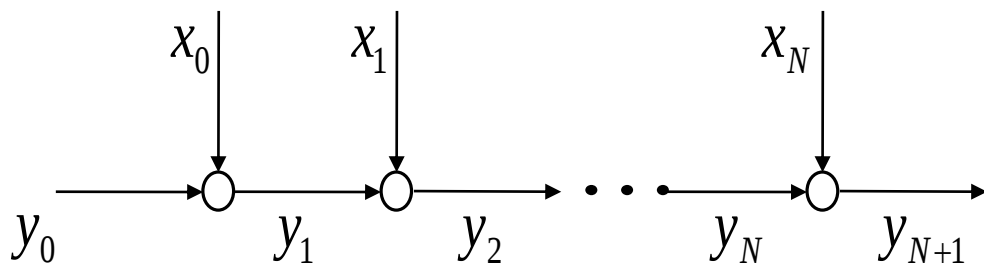


Programowanie dynamiczne

Proces dynamiczny: $y_{n+1} = P(y_n, x_n)$ n – takt

x_n – decyzja w n – tym takcie

y_n – stan procesu w n – tym takcie



15 Saving Tips by
RateCatcher.com



<http://www.all-freeware.com/>

Zadanie decyzyjne: znaleźć ciąg: $x_0^*, x_1^*, \dots, x_N^*$,

dla których wskaźnik $Q(x_0, x_1, \dots, x_N)$ przyjmuje wartość minimalną

Podstawy matematyczne



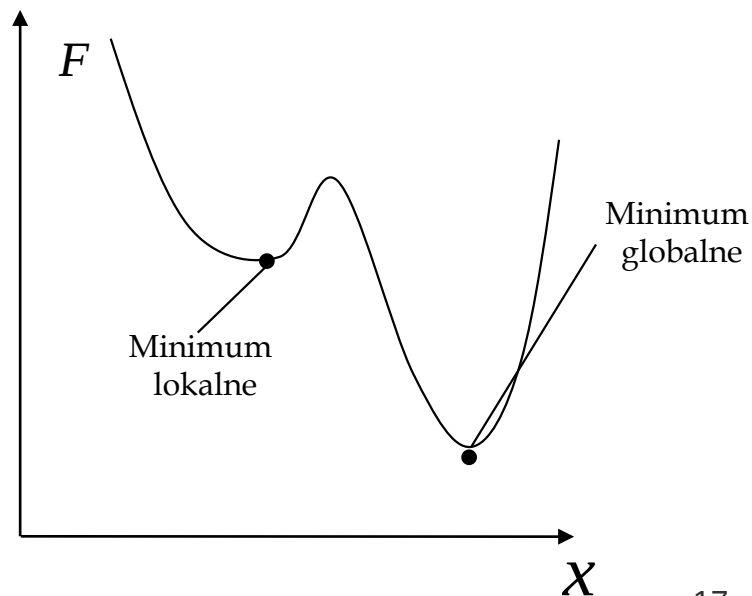
3.0 Przydatne definicje, własności

Podstawy matematyczne

Zadanie optymalizacji: $x^* \rightarrow F(x^*) = \min_{x \in \mathcal{D}_x} F(x)$

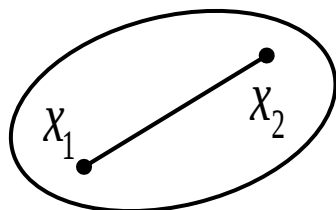
Minimum lokalne: $\forall_{\varepsilon > 0} \exists_{x \in O(x^*, \varepsilon)} F(x^*) < F(x)$

Minimum globalne: $\forall_{x \in \mathcal{D}_x} F(x^*) < F(x)$

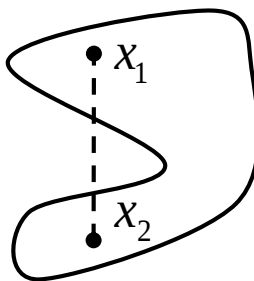


Podstawy matematyczne

Zbiór wypukły: $\forall_{x_1, x_2 \in \mathcal{D}_x} \lambda x_1 + (1-\lambda)x_2 \in \mathcal{D}_x, \quad \lambda \in \langle 0, 1 \rangle$



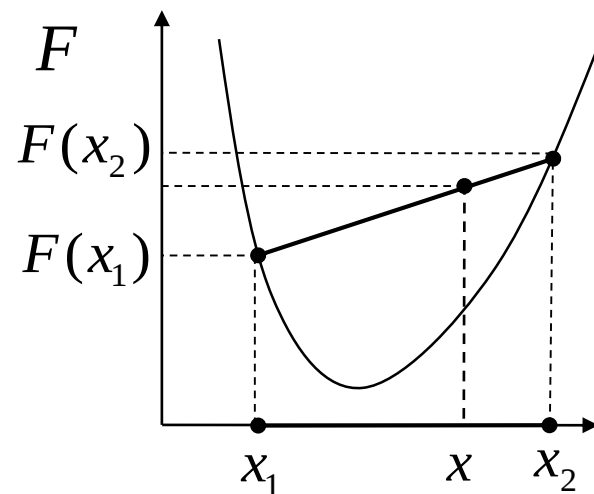
- zbiór wypukły



- zbiór niewypukły

Funkcja wypukła:

$$F(\lambda x_1 + (1-\lambda)x_2) \leq \lambda F(x_1) + (1-\lambda)F(x_2), \quad \lambda \in \langle 0, 1 \rangle$$



Podstawy matematyczne

Funkcja pseudo - wypukła:

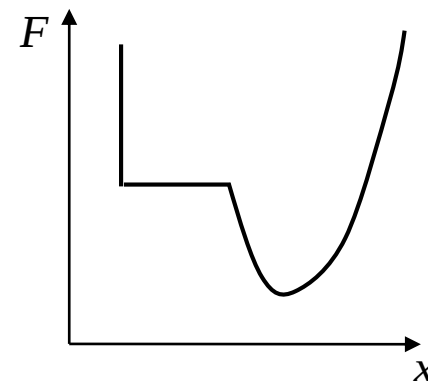
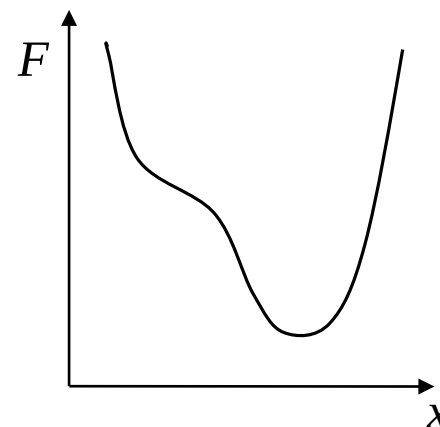
Zgodnie z rozwinięciem Taylor' a funkcji mamy:

$$F(x) = F(x_0) + (x - x_0)^T [\nabla_x F(x_0)] + O_2(\|x - x_0\|)$$

$$(x - x_0)^T [\nabla_x F(x_0)] \geq 0 \Rightarrow F(x) > F(x_0)$$

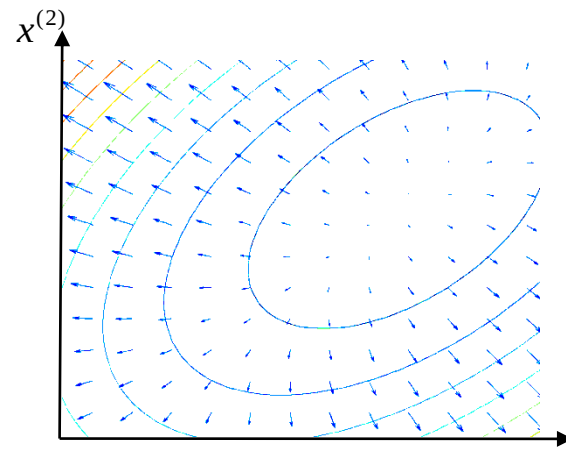
Funkcja quasi - wypukła:

$$\mathcal{D}_\alpha = \{x \in \mathcal{D}_x : F(x) \leq \alpha\} \text{ - zbiór wypukły}$$



Podstawy matematyczne

Gradient: $\nabla_x F(x) = \begin{bmatrix} \frac{\partial F}{\partial x^{(1)}} \\ \frac{\partial F}{\partial x^{(2)}} \\ \vdots \\ \frac{\partial F}{\partial x^{(s)}} \end{bmatrix}$



Hessian: $H(x) = \nabla_{xx}^2 F(x) = \begin{bmatrix} \frac{\partial^2 F}{\partial (x^{(1)})^2} & \frac{\partial^2 F}{\partial x^{(1)} \partial x^{(2)}} & \cdots & \frac{\partial^2 F}{\partial x^{(1)} \partial x^{(s)}} \\ \frac{\partial^2 F}{\partial x^{(2)} \partial x^{(1)}} & \frac{\partial^2 F}{\partial (x^{(2)})^2} & \cdots & \frac{\partial^2 F}{\partial x^{(2)} \partial x^{(s)}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 F}{\partial x^{(s)} \partial x^{(1)}} & \frac{\partial^2 F}{\partial x^{(s)} \partial x^{(2)}} & \cdots & \frac{\partial^2 F}{\partial (x^{(s)})^2} \end{bmatrix}$

Podstawy matematyczne

Własności Hessjanu:

$$\frac{\partial^2 F}{\partial x^{(i)} \partial x^{(j)}} = \frac{\partial^2 F}{\partial x^{(j)} \partial x^{(i)}} \Rightarrow H \text{ jest macierzą symetryczną}$$

Jeżeli $\forall_{x \neq 0_s} x^T H x > 0$ to H jest dodatnio określony

Jeżeli $\forall_{x \neq 0_s} x^T H x < 0$ to H jest ujemnie określony

Jeżeli $\forall_{x \neq 0_s} x^T H x \geq 0$ to H jest dodatnio pół określony

Jeżeli $\forall_{x \neq 0_s} x^T H x \leq 0$ to H jest ujemnie pół określony

Podstawy matematyczne

Kryterium Sylwestra:

$H = [h_{ij}]_{\substack{i=1,2,\dots,S \\ j=1,2,\dots,S}}$ - macierz Hessa

Jeżeli $\forall s = 1, 2, \dots, S \quad \det(H_{ss}) = \det \left([h_{ij}]_{\substack{i=1,2,\dots,s \\ j=1,2,\dots,s}} \right) > 0$ to macierz H jest dodatnio określona

Jeżeli $\forall \{i_1, i_2, \dots, i_s\} \in \{1, 2, \dots, S\} \quad \det \left([h_{ij}]_{\substack{i \in \{i_1, i_2, \dots, i_s\} \\ j \in \{i_1, i_2, \dots, i_s\}}} \right) \geq 0$ to macierz H jest dodatnio pół określona

Wartości własne macierzy H

$\det(H - hI) = 0 \quad h_1, h_2, \dots, h_S$ - wartości własne macierzy H

Jeżeli $\forall s = 1, 2, \dots, S \quad h_s > 0$ to macierz H jest dodatnio określona

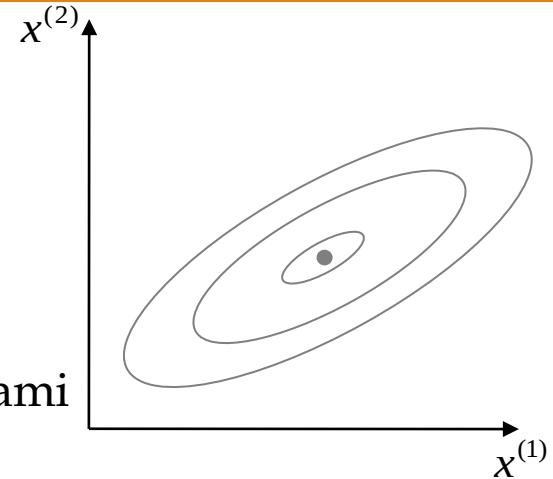
Jeżeli $\forall s = 1, 2, \dots, S \quad h_s \geq 0$ to macierz H jest dodatnio pół określona

Typowe zadania decyzyjne



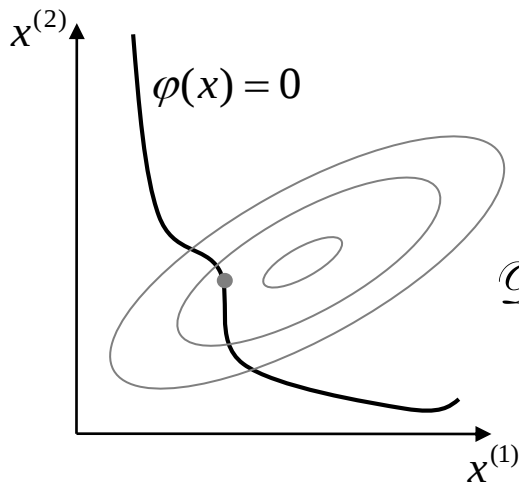
Typowe zadania decyzyjne

Zadanie decyzyjne bez ograniczeń: $\mathcal{D}_x = \mathcal{R}^S$



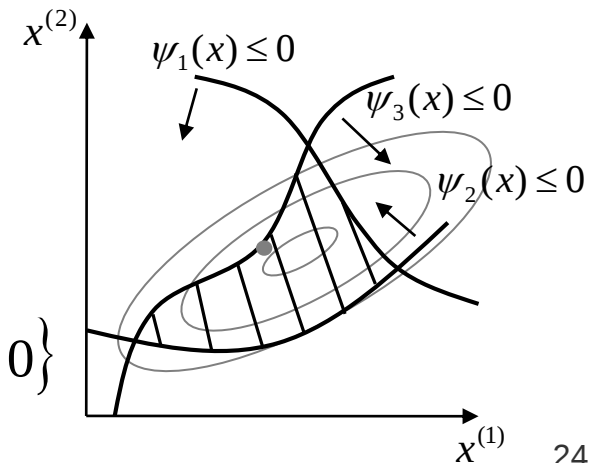
Zadanie decyzyjne z ograniczeniami równościowymi:

$$\mathcal{D}_x = \{x \in \mathcal{R}^S : \varphi_1(x) = 0, \varphi_2(x) = 0, \dots, \varphi_L(x) = 0, L \leq S\}$$



Zadanie decyzyjne z ograniczeniami nierównościami:

$$\mathcal{D}_x = \{x \in \mathcal{R}^S : \psi_1(x) \leq 0, \psi_2(x) \leq 0, \dots, \psi_M(x) \leq 0\}$$



Analityczne metody optymalizacji

- ⌘ Zadanie optymalizacji bez ograniczeń
- ⌘ Zadanie optymalizacji z ograniczeniami równościowymi – metoda współczynników Lagrange’a
- ⌘ Zadanie optymalizacji z ograniczeniami nierównościowymi – metoda Kuhna-Tuckera

Zadanie optymalizacji bez ograniczeń



3.1 Analityczne warunki optymalności

Zadanie optymalizacji bez ograniczeń

Zadanie optymalizacji: $x^* \rightarrow F(x^*) = \min_{x \in \mathcal{D}_x = \mathcal{R}^S} F(x)$

Założenie: $F(x)$ jest funkcją ciągłą i różniczkowalną.

Warunkiem koniecznym aby x^* było minimum lokalnym jest:

$$\nabla_x F(x) \Big|_{x^*} = 0_S$$

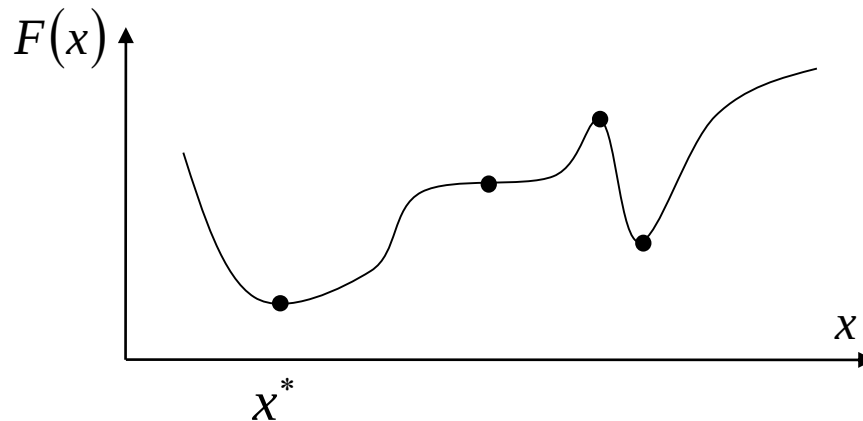
Jeżeli $F(x)$ jest funkcją pseudo - wypukłą, powyższe równanie jest warunkiem koniecznym i wystarczającym aby x^* było minimum globalnym

Jeżeli $H(x)$ jest dodatnio pół określona $\forall x \in \mathcal{R}^S$ to rozwiązanie powyższego równania jest minimum globalnym

Jeżeli $H(x)$ jest dodatnio określona $\forall x \in \mathcal{R}^S$ powyższe równanie ma jedno rozwiązanie x^* i jest ono minimum globalnym

Zadanie optymalizacji bez ograniczeń

Równanie $\nabla_x F(x)|_{x^*} = 0_S$ może mieć wiele rozwiązań



Warunki optymalności drugiego rzędu:

Jeżeli $H(x^*)$ jest dodatnio pół określona w punkcie x^*
to x^* jest minimum lokalnym

Jeżeli $H(x^*)$ jest ujemnie pół określona w punkcie x^*
to x^* jest maksimum lokalnym

Przykład 2.1.1

$$\text{☞ } F(x^{(1)}, x^{(2)}) = 5(x^{(1)})^2 + (x^{(2)})^2 - 4x^{(1)}x^{(2)} - 2x^{(1)} + 3$$

$$\text{☞ } \nabla_x F(x^{(1)}, x^{(2)})|_{x=x^*} = \left[\begin{array}{c} \frac{\partial F(x^{(1)}, x^{(2)})}{\partial x^{(1)}} \\ \frac{\partial F(x^{(1)}, x^{(2)})}{\partial x^{(2)}} \end{array} \right] \bigg|_{x=x^*} = \begin{bmatrix} 10x^{(1)*} - 4x^{(2)*} - 2 \\ 2x^{(2)*} - 4x^{(1)*} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \begin{matrix} (1) \\ (2) \end{matrix}$$

$$\text{☞ } z \ (2) \rightarrow x^{(2)*} = 2x^{(1)*}$$

$$\text{☞ } z \ (1) \rightarrow 10x^{(1)*} - 8x^{(1)*} = 2 \rightarrow x^{(1)*} = 1, x^{(2)*} = 2$$

$$\text{☞ } H(x) = \nabla_{xx} F(x^{(1)}, x^{(2)}) = \begin{bmatrix} 10 & -4 \\ -4 & 2 \end{bmatrix}$$

$$\text{☞ } \det H_{11} = \det[10] = 10 > 0$$

$$\text{☞ } \det H_{11} = \det \begin{bmatrix} 10 & 4 \\ 4 & 2 \end{bmatrix} = 20 - 16 = 4 > 0$$

$$\text{☞ } \text{Macierz } H(x) \text{ jest dodatnio określona zatem punkt } x^* = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \text{ - minimum}$$

$$F(x^{(1)}, x^{(2)}) = \alpha(x^{(1)})^2 + (x^{(2)})^2 - 4x^{(1)}x^{(2)} - 2x^{(1)} + 3$$

$$\partial x^{(2)}$$

Przykład 2.1.2

$$\infty F(x^{(1)}, x^{(2)}) = \alpha(x^{(1)})^2 + (x^{(2)})^2 - 4x^{(1)}x^{(2)} - 2x^{(1)} + 3$$

$$\infty \nabla_x F(x^{(1)}, x^{(2)})|_{x=x^*} = \begin{bmatrix} \frac{\partial F(x^{(1)}, x^{(2)})}{\partial x^{(1)}} \\ \frac{\partial F(x^{(1)}, x^{(2)})}{\partial x^{(2)}} \end{bmatrix} \bigg|_{x=x^*} = \begin{bmatrix} 2\alpha x^{(1)*} - 4x^{(2)*} - 2 \\ 2x^{(2)*} - 4x^{(1)*} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \begin{matrix} (1) \\ (2) \end{matrix}$$

$$\infty z \ (2) \rightarrow x^{(2)*} = 2x^{(1)*}$$

$$\infty z \ (1) \rightarrow 2\alpha x^{(1)*} - 8x^{(1)*} = 2 \rightarrow x^{(1)*} = \frac{1}{\alpha-4}, x^{(2)*} = \frac{2}{\alpha-4}, \alpha \neq 4$$

$$\infty H(x) = \nabla_{xx} F(x^{(1)}, x^{(2)}) = \begin{bmatrix} 2\alpha & -4 \\ -4 & 2 \end{bmatrix}$$

$$\infty \det H_{11} = \det[2\alpha] = 2\alpha > 0 \rightarrow \alpha > 0$$

$$\infty \det H_{11} = \det \begin{bmatrix} 2\alpha & 4 \\ 4 & 2 \end{bmatrix} = 8\alpha - 16 > 0 \rightarrow \alpha > 2$$

$$\infty \text{ Dla } \alpha > 2 \text{ macierz } H(x) \text{ jest dodatnio określona a punkt } x^* = \begin{bmatrix} \frac{1}{\alpha-4} \\ \frac{2}{\alpha-4} \end{bmatrix} \text{ - minimum}$$

Przykład 2.1.2 c.d.

☞ Dla $\alpha > 2$ macierz $H(x)$ jest dodatnio określona

☞ Punkt $x^* = \begin{bmatrix} \frac{1}{\alpha-4} \\ 2 \\ \frac{1}{\alpha-4} \end{bmatrix}$ - minimum $\alpha \neq 4$ funkcji

☞ $F(x^{(1)}, x^{(2)}) = \alpha(x^{(1)})^2 + (x^{(2)})^2 - 4x^{(1)}x^{(2)} - 2x^{(1)} + 3$

Przykład 2.1.3

☞ $F(x) = x^T A x + b^T x + c$

☞ A – macierz symetryczna, dodatnio określona

☞ $A = [a_{ij}]_{\substack{i=1,2,\dots,S \\ j=1,2,\dots,S}} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1S} \\ a_{21} & a_{22} & \cdots & a_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ a_{S1} & a_{S2} & \cdots & a_{SS} \end{bmatrix}$

☞ $x = \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(S)} \end{bmatrix}, b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_S \end{bmatrix},$ - S -wymiarowe wektory

☞ $\nabla_x F(x)|_{x=x^*} = \nabla_x (x^T A x + b^T x + c)|_{x=x^*} = 0_S \quad |_{x=x^*}$

Przykład 2.1.3 c.d.

$$\infty \quad x^T A x = [x^{(1)} \quad x^{(2)} \quad \dots \quad x^{(S)}] \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1S} \\ a_{21} & a_{22} & \dots & a_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ a_{S1} & a_{S2} & \dots & a_{SS} \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(S)} \end{bmatrix}$$

$$\infty \quad x^T A x = \sum_{i=1}^S \sum_{j=1}^S a_{ij} x^{(i)} x^{(j)}$$

$$\infty \quad \nabla_x (x^T A x) = \begin{bmatrix} \frac{\partial (x^T A x)}{\partial x^{(1)}} \\ \frac{\partial (x^T A x)}{\partial x^{(2)}} \\ \vdots \\ \frac{\partial (x^T A x)}{\partial x^{(S)}} \end{bmatrix} = \begin{bmatrix} \frac{\partial}{\partial x^{(1)}} \left(\sum_{i=1}^S \sum_{j=1}^S a_{ij} x^{(i)} x^{(j)} \right) \\ \frac{\partial}{\partial x^{(2)}} \left(\sum_{i=1}^S \sum_{j=1}^S a_{ij} x^{(i)} x^{(j)} \right) \\ \vdots \\ \frac{\partial}{\partial x^{(S)}} \left(\sum_{i=1}^S \sum_{j=1}^S a_{ij} x^{(i)} x^{(j)} \right) \end{bmatrix}$$

Przykład 2.1.3 c.d.

$$\infty \nabla_x (x^T A x) = \begin{bmatrix} \frac{\partial}{\partial x^{(1)}} \left(\sum_{i=1}^S \sum_{j=1}^S a_{ij} x^{(i)} x^{(j)} \right) \\ \frac{\partial}{\partial x^{(2)}} \left(\sum_{i=1}^S \sum_{j=1}^S a_{ij} x^{(i)} x^{(j)} \right) \\ \vdots \\ \frac{\partial}{\partial x^{(S)}} \left(\sum_{i=1}^S \sum_{j=1}^S a_{ij} x^{(i)} x^{(j)} \right) \end{bmatrix}$$

$$\infty \nabla_x (x^T A x) = \begin{bmatrix} \sum_{j=1}^S a_{1j} x^{(j)} + \sum_{i=1}^S a_{i1} x^{(i)} \\ \sum_{j=1}^S a_{2j} x^{(j)} + \sum_{i=1}^S a_{i2} x^{(i)} \\ \vdots \\ \sum_{j=1}^S a_{Sj} x^{(j)} + \sum_{i=1}^S a_{iS} x^{(i)} \end{bmatrix} =$$

Przykład 2.1.3 c.d.

$$\infty \quad \infty \quad \nabla_x(x^T Ax) = \begin{bmatrix} \sum_{j=1}^S a_{1j}x^{(j)} \\ \sum_{j=1}^S a_{2j}x^{(j)} \\ \vdots \\ \sum_{j=1}^S a_{Sj}x^{(j)} \end{bmatrix} + \begin{bmatrix} \sum_{i=1}^S a_{i1}x^{(i)} \\ \sum_{i=1}^S a_{i2}x^{(i)} \\ \vdots \\ \sum_{i=1}^S a_{S1}x^{(i)} \end{bmatrix}$$

$$\nabla_x(x^T Ax) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1S} \\ a_{21} & a_{22} & \cdots & a_{2S} \\ \vdots & \vdots & \ddots & \vdots \\ a_{S1} & a_{S2} & \cdots & a_{SS} \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(S)} \end{bmatrix} + \begin{bmatrix} a_{11} & a_{21} & \cdots & a_{S1} \\ a_{12} & a_{22} & \cdots & a_{S2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{S1} & a_{S2} & \cdots & a_{SS} \end{bmatrix} \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(S)} \end{bmatrix}$$

$$\nabla_x(x^T Ax) = Ax + A^T x$$

$$\text{dla } A = A^T \quad \nabla_x(x^T Ax) = 2Ax$$

Przykład 2.1.3 c.d.

$$\infty \quad b^T x = [b_1 \quad b_2 \quad \dots \quad b_S] \begin{bmatrix} x^{(1)} \\ x^{(2)} \\ \vdots \\ x^{(S)} \end{bmatrix} = \sum_{i=1}^S b_i x^{(i)}$$

$$\infty \quad \nabla_x (b^T x) = \begin{bmatrix} \frac{\partial}{\partial x^{(1)}} \left(\sum_{i=1}^S b_i x^{(i)} \right) \\ \frac{\partial}{\partial x^{(2)}} \left(\sum_{i=1}^S b_i x^{(i)} \right) \\ \vdots \\ \frac{\partial}{\partial x^{(S)}} \left(\sum_{i=1}^S b_i x^{(i)} \right) \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_S \end{bmatrix} = b$$

Przykład 2.1.3 c.d.

$$\ni \nabla_x (x^T A x + b^T x + c) |_{x=x^*} = 2Ax^* + b = 0_S$$

$$\ni x^* = -\frac{1}{2}A^{-1}b$$

$$\ni H(x) = \nabla_{xx} (x^T A x + b^T x + c) = \nabla_x (2Ax + b) = 2A$$

$\ni$ Macierz Hessa dodatnio określona bo A jest dodatnio określona

Zadanie optymalizacji z ograniczeniami równościowymi



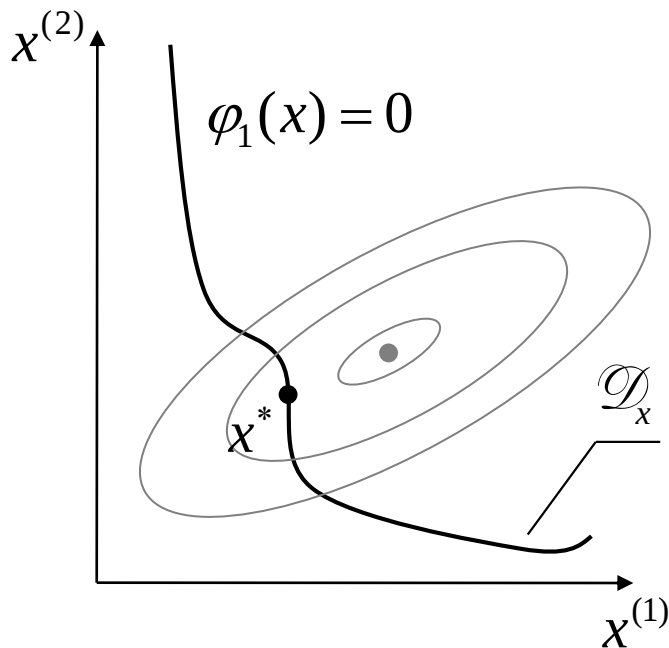
3.2 Metoda mnożników Lagrange'a

Zadanie optymalizacji z ograniczeniami równościowymi

Metoda mnożników Lagrange' a

Zadanie optymalizacji: $x^* \rightarrow F(x^*) = \min_{x^* \in \mathcal{D}_x} F(x)$

$$\mathcal{D}_x = \left\{ x \in \mathcal{R}^S : \varphi_1(x) = 0, \varphi_2(x) = 0, \dots, \varphi_L(x) = 0, L \leq S \right\}$$



$$\mathcal{D}_x = \left\{ x \in \mathcal{R}^S : \varphi(x) = 0_L, L \leq S \right\}$$

$$\varphi(x) = \begin{bmatrix} \varphi_1(x) \\ \varphi_2(x) \\ \vdots \\ \varphi_L(x) \end{bmatrix} \quad 0_L = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \left. \vphantom{\begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}} \right\} L - \text{zer}$$

Zadanie optymalizacji z ograniczeniami równościowymi

Metoda mnożników Lagrange' a

∞ Metoda współczynników Lagrange' a

Funkcja Lagrange' a:

$$L(x, \lambda) = F(x) + \sum_{l=1}^L \lambda_l \varphi_l(x) = F(x) + \lambda^T \varphi(x)$$

gdzie: $\varphi(x) = \begin{bmatrix} \varphi_1(x) \\ \varphi_2(x) \\ \vdots \\ \varphi_L(x) \end{bmatrix}$, $\lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_L \end{bmatrix}$ - wektor współczynników Lagrange' a

Warunki konieczne optymalności:

$$\nabla_x L(x, \lambda) \Big|_{x^*, \lambda^*} = 0_S$$

$$\nabla_\lambda L(x, \lambda) \Big|_{x^*, \lambda^*} = 0_L$$

Rozwiązanie regularne

$$\Leftrightarrow \text{rank } G(x) = \text{rank} \begin{bmatrix} G(x) & \vdots & -\nabla_x F(x) \end{bmatrix},$$

$$G(x) = \begin{bmatrix} \nabla_x \varphi_1(x) & \vdots & \nabla_x \varphi_2(x) & \vdots & \dots & \vdots & \nabla_x \varphi_L(x) \end{bmatrix}$$

Zadanie optymalizacji z ograniczeniami równościowymi

Metoda mnożników Lagrange' a

Powyższy układ równań może mieć wiele rozwiązań

Warunki optymalności drugiego rzędu:

Oznaczmy: $H_L(x) = \nabla_{xx} L(x, \lambda)$

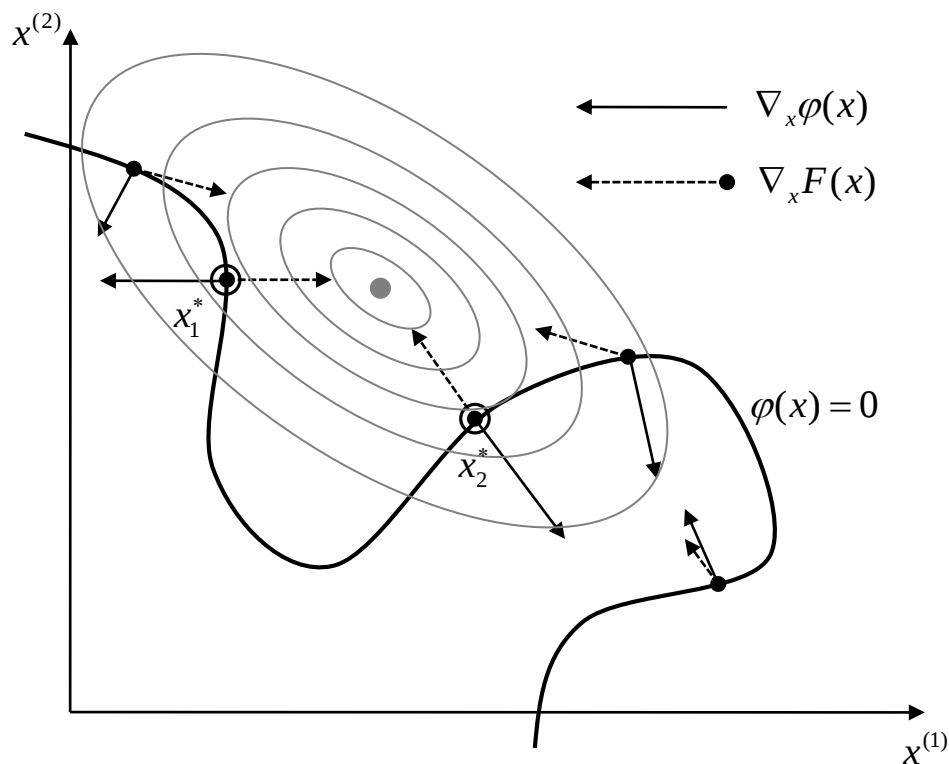
Jeżeli $H_L(x^*)$ jest dodatnio określona w punkcie x^*
to x^* jest minimum lokalnym

Jeżeli $H_L(x^*)$ jest ujemnie określona w punkcie
to x^* jest maksimum lokalnym x^*

Jeżeli funkcja $F(x)$ jest wypukła, a ograniczenia są liniowe czyli mają postać $\varphi_l(x) = p_l^T x - \alpha_l = 0$, $l = 1, 2, \dots, L$ to powyższy układ równań ma jedno rozwiązanie i jest ono rozwiązaniem optymalnym

Zadanie optymalizacji z ograniczeniami równościowymi

Metoda mnożników Lagrange' a



$$L(x, \lambda) = F(x) + \lambda \varphi(x)$$

Warunek konieczny optymalności:

$$\nabla L(x, \lambda) = \nabla_x F(x) + \lambda \nabla_x \varphi(x) = 0$$

gdzie:

$\lambda \in \mathcal{R}$ - współczynnik Lagrange' a

Wyjaśnienie warunków koniecznych

$$\ni x_1^*, x_2^* \rightarrow F(x_1^*, x_2^*) = \min_{x_1, x_2} F(x_1, x_2)$$

$$\ni \text{Przy ograniczeniu } \varphi(x_1, x_2) = 0$$

$$\ni \varphi(x_1, x_2) = 0 \rightarrow x_2 = \psi(x_1)$$

$$\ni x_1^* \rightarrow F(x_1^*, \psi(x_1^*)) = \min_{x_1} F(x_1, \psi(x_1))$$

$$\ni \frac{dF(x_1, \psi(x_1))}{dx_1} = \frac{\partial F(x_1, x_2)}{\partial x_1} + \frac{\partial F(x_1, x_2)}{\partial x_2} \frac{d\psi(x_1)}{dx_1} = 0$$

$$\ni \text{Pochodna funkcji rozwikłanej}$$

$$\ni \frac{d\psi(x_1)}{dx_1} = - \frac{\frac{\partial \varphi(x_1, x_2)}{\partial x_1}}{\frac{\partial \varphi(x_1, x_2)}{\partial x_2}}$$

Wyjaśnienie warunków koniecznych

$$\begin{aligned} \text{☞} & \frac{\partial F(x_1, x_2)}{\partial x_1} + \frac{\partial F(x_1, x_2)}{\partial x_2} \left(-\frac{\frac{\partial \varphi(x_1, x_2)}{\partial x_1}}{\frac{\partial \varphi(x_1, x_2)}{\partial x_2}} \right) = 0 \\ \text{☞} & \end{aligned}$$

$$\text{☞ oznaczmy } \lambda = -\frac{\frac{\partial F(x_1, x_2)}{\partial x_2}}{\frac{\partial \varphi(x_1, x_2)}{\partial x_2}}$$

$$\text{☞ } \frac{\partial F(x_1, x_2)}{\partial x_1} + \lambda \frac{\partial \varphi(x_1, x_2)}{\partial x_1} = 0$$

$$\text{☞ } \frac{\partial F(x_1, x_2)}{\partial x_2} + \lambda \frac{\partial \varphi(x_1, x_2)}{\partial x_2} = 0$$

$$\text{☞ } x_2 = \psi(x_1) \rightarrow \varphi(x_1, x_2) = 0$$

Wyjaśnienie warunków koniecznych

$$\propto L(x_1, x_2, \lambda) = F(x_1, x_2) + \lambda \varphi(x_1, x_2)$$

$$\propto \frac{\partial L(x_1, x_2, \lambda)}{\partial x_1} = 0 \rightarrow \frac{\partial F(x_1, x_2)}{\partial x_1} + \lambda \frac{\partial \varphi(x_1, x_2)}{\partial x_1} = 0$$

$$\propto \frac{\partial L(x_1, x_2, \lambda)}{\partial x_2} = 0 \rightarrow \frac{\partial F(x_1, x_2)}{\partial x_2} + \lambda \frac{\partial \varphi(x_1, x_2)}{\partial x_2} = 0$$

$$\propto \frac{\partial L(x_1, x_2, \lambda)}{\partial \lambda} = 0 \rightarrow \varphi(x_1, x_2) = 0$$

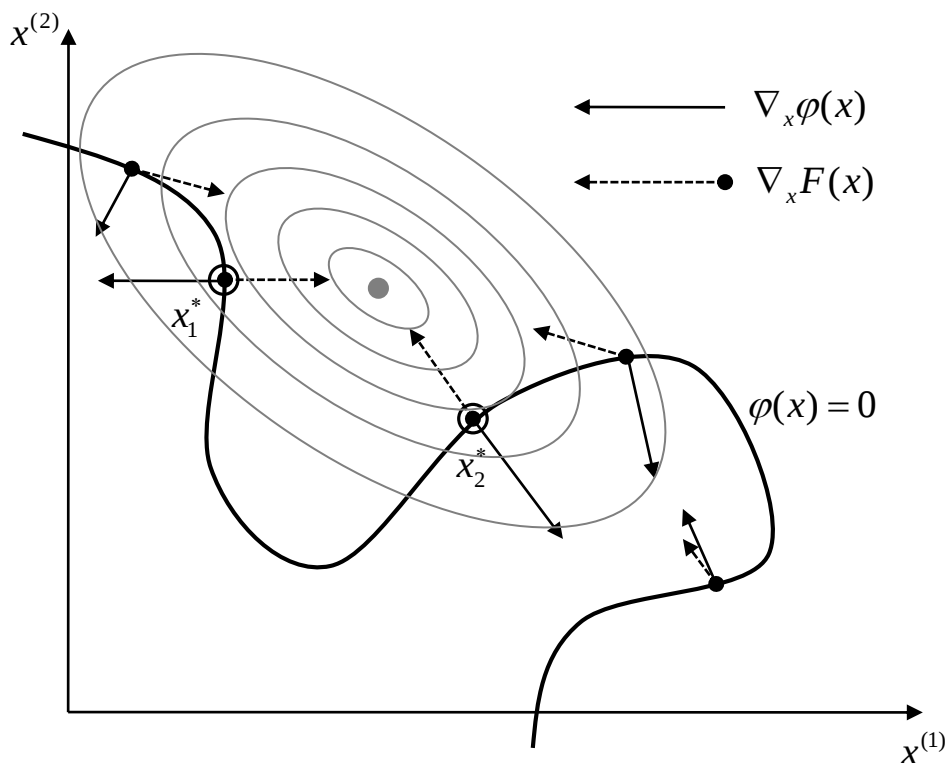
$\propto$ Ogólniej

$$\nabla_x L(x, \lambda) \big|_{x^*, \lambda^*} = 0_S$$

$$\nabla_\lambda L(x, \lambda) \big|_{x^*, \lambda^*} = 0_L$$

Zadanie optymalizacji z ograniczeniami równościowymi

Metoda mnożników Lagrange' a



$$L(x, \lambda) = F(x) + \lambda \varphi(x)$$

gdzie:

$\lambda \in \mathcal{R}$ - współczynnik Lagrange' a

Warunek konieczny optymalności:

$$\nabla_x L(x, \lambda) = \nabla_x F(x) + \lambda \nabla_x \varphi(x) = 0$$

$$\nabla_x F(x) = -\lambda \nabla_x \varphi(x)$$

$$\nabla_\lambda L(x, \lambda) = \frac{\partial L(x_1, x_2, \lambda)}{\partial \lambda} = 0$$

$$\varphi(x_1, x_2) = 0$$

Zadanie optymalizacji z ograniczeniami równościowymi

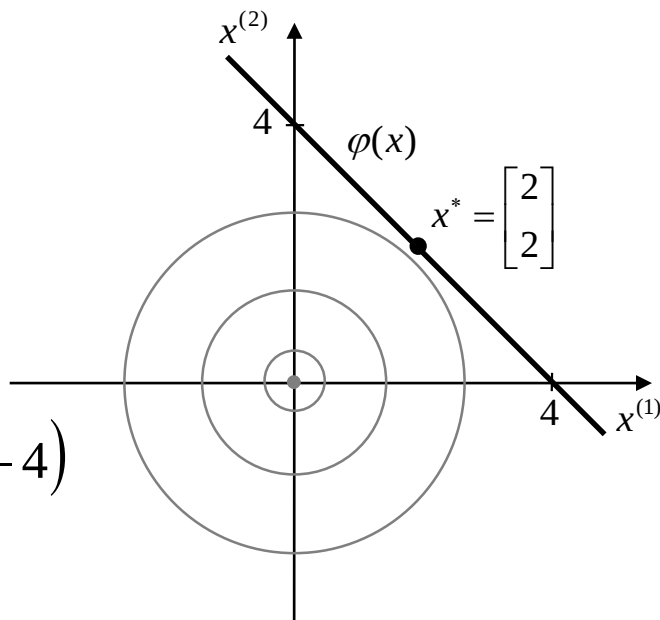
Metoda mnożników Lagrange' a

☞ Metoda Lagrange' a - Przykład 1

$$F(x) = (x^{(1)})^2 + (x^{(2)})^2$$

$$\varphi(x) = x^{(1)} + x^{(2)} - 4 = 0$$

$$L(x, \lambda) = (x^{(1)})^2 + (x^{(2)})^2 + \lambda(x^{(1)} + x^{(2)} - 4)$$



Przykład 2.2.1

$$\text{☞ } L(x, \lambda) = (x^{(1)})^2 + (x^{(2)})^2 + \lambda(x^{(1)} + x^{(2)} - 4)$$

$$\text{☞ } \nabla_x L(x, \lambda) = \begin{bmatrix} \frac{\partial L}{\partial x^{(1)}} \\ \frac{\partial L}{\partial x^{(2)}} \end{bmatrix} = \begin{bmatrix} 2x^{(1)} + \lambda \\ 2x^{(2)} + \lambda \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \begin{matrix} (1) \\ (2) \end{matrix}$$

$$\text{☞ } \nabla_\lambda L(x, \lambda) = \frac{\partial L}{\partial \lambda} = x^{(1)} + x^{(2)} - 4 = 0 \quad (3)$$

$$\text{☞ } z \ (1) \rightarrow x^{(1)} = -\frac{\lambda}{2}, \quad z \ (2) \rightarrow x^{(2)} = -\frac{\lambda}{2}$$

$$\text{☞ } z \ (3) \rightarrow \left(-\frac{\lambda}{2}\right) + \left(-\frac{\lambda}{2}\right) - 4 = 0 \quad \text{czyli } \lambda = -4$$

$$\text{☞ } x^{(1)} = -\frac{\lambda}{2} = -\frac{-4}{2} = 2, \quad x^{(2)} = -\frac{\lambda}{2} = -\frac{-4}{2} = 2$$

Przykład 2.2.1 c.d.

$$\begin{aligned} \& \nabla_x L(x, \lambda) &= \begin{bmatrix} \frac{\partial L}{\partial x^{(1)}} \\ \frac{\partial L}{\partial x^{(2)}} \end{bmatrix} = \begin{bmatrix} 2x^{(1)} + \lambda \\ 2x^{(2)} + \lambda \end{bmatrix} \end{aligned}$$

$$\& H_L = \nabla_{xx} L(x, \lambda) = \begin{bmatrix} \frac{\partial^2 L}{\partial^2 x^{(1)}} & \frac{\partial^2 L}{\partial x^{(1)} \partial x^{(2)}} \\ \frac{\partial^2 L}{\partial x^{(2)} \partial x^{(1)}} & \frac{\partial^2 L}{\partial^2 x^{(2)}} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

$$\& \det H_{L11} = \det[2] = 2 > 0, \quad \det H_{L22} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = 2 \times 2 = 4 > 0$$

$$\& \text{Macierz } H_L = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \text{ jest dodatnio określona}$$

$$\& \text{Punkt } x = \begin{bmatrix} 2 \\ 2 \end{bmatrix} - \text{minimum}$$

Zadanie optymalizacji z ograniczeniami równościowymi

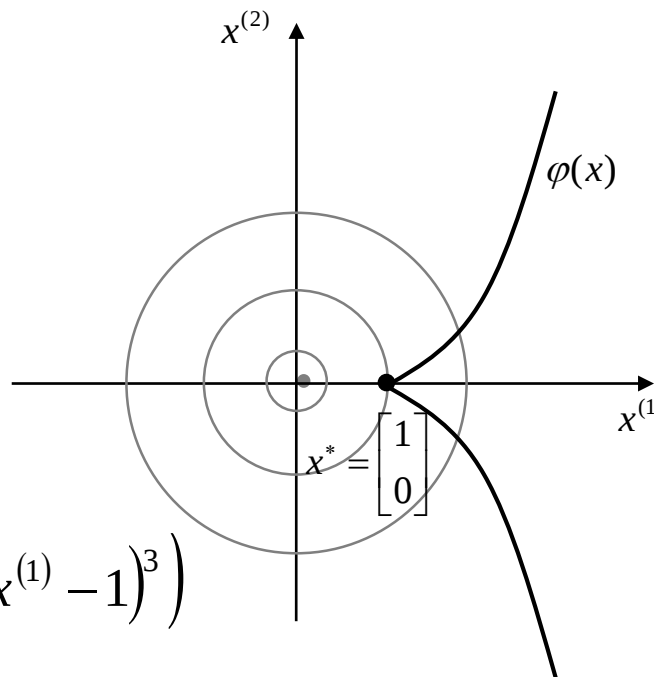
Metoda mnożników Lagrange' a

∞ Metoda współczynników Lagrange' a – Przykład 2
rozwiązanie nieregularne

$$F(x) = (x^{(1)})^2 + (x^{(2)})^2$$

$$\varphi(x) = (x^{(2)})^2 - (x^{(1)} - 1)^3 = 0$$

$$L(x, \lambda) = (x^{(1)})^2 + (x^{(2)})^2 + \lambda \left((x^{(2)})^2 - (x^{(1)} - 1)^3 \right)$$



Przykład 2.2.2

$$\text{☞ } L(x, \lambda) = (x^{(1)})^2 + (x^{(2)})^2 + \lambda((x^{(2)})^2 + (x^{(1)} - 1)^3)$$

$$\text{☞ } \nabla_x L(x, \lambda) = \begin{bmatrix} \frac{\partial L}{\partial x^{(1)}} \\ \frac{\partial L}{\partial x^{(2)}} \end{bmatrix} = \begin{bmatrix} 2x^{(1)} - 3\lambda(x^{(1)} - 1)^2 \\ 2x^{(2)} + 2\lambda x^{(2)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \begin{matrix} (1) \\ (2) \end{matrix}$$

$$\text{☞ } \nabla_\lambda L(x, \lambda) = \frac{\partial L}{\partial \lambda} = (x^{(2)})^2 + (x^{(1)} - 1)^3 = 0 \quad (3)$$

$$\text{☞ z (2)} \rightarrow 2(1 + \lambda)x^{(2)} = 0 \text{ czyli } x^{(2)} = 0,$$

$$\text{☞ z (3)} \rightarrow (0)^2 + (x^{(1)} - 1)^3 = 0, \text{ czyli } x^{(1)} = 1,$$

$$\text{☞ z (1)} \rightarrow 2x^{(1)} - 3\lambda(x^{(1)} - 1)^2 = 2 \times 1 - 3\lambda(1 - 1)^2 = 2 \neq 0$$

☞ Sprzeczność ??

Zadanie optymalizacji z ograniczeniami równościowymi

Metoda mnożników Lagrange' a

☞ Metoda współczynników Lagrange' a - wyjaśnianie

$$\nabla_x L(x, \lambda) = \nabla_x F(x) + \sum_{l=1}^L \lambda_l \nabla_x \varphi_l(x) = 0_S$$

$$G(x) = [\nabla_x \varphi_1(x) \quad \vdots \quad \nabla_x \varphi_2(x) \quad \vdots \quad \dots \quad \vdots \quad \nabla_x \varphi_L(x)]$$

$$\nabla_x F(x) + G(x)\lambda = 0 \quad G(x)\lambda = -\nabla_x F(x)$$

Rozwiązanie powyższego równania liniowego istnieje jeżeli i jest jednoznaczne gdy:

$$\text{rank } G(x) = \text{rank} [G(x) \quad \vdots \quad -\nabla_x F(x)] = L$$

oraz niejednoznaczne gdy:

$$\text{rank } G(x) = \text{rank} [G(x) \quad \vdots \quad -\nabla_x F(x)] < L$$

Zadanie optymalizacji z ograniczeniami równościowymi

Metoda mnożników Lagrange' a

Jeżeli $F(x)$ jest funkcją ciągłą, różniczkowalną i wypukłą oraz ograniczenia $\varphi_1(x), \varphi_2(x), \dots, \varphi_L(x)$ są liniowe to układ równań:

$$\nabla_x L(x, \lambda) \Big|_{x^*, \lambda^*} = 0_S$$

$$\nabla_\lambda L(x, \lambda) \Big|_{x^*, \lambda^*} = 0_L$$

ma jedno rozwiązanie i jest ono rozwiązaniem zadania optymalizacji z ograniczeniami równościowymi.

Powyższy układ równań w tym przypadku jest warunkiem koniecznym i wystarczającym

Zadanie optymalizacji z ograniczeniami równościowymi

Metoda mnożników Lagrange' a

Uogólniona metoda współczynników Lagrange' a

Uogólniona funkcja Lagrange' a:

$$L(x, \lambda, \lambda_0) = \lambda_0 F(x) + \sum_{l=1}^L \lambda_l \varphi_l(x)$$

Warunki konieczne optymalności:

$$\nabla_x L(x, \lambda, \lambda_0) \Big|_{x^*, \lambda^*, \lambda_0} = 0_S$$

$$\nabla_\lambda L(x, \lambda, \lambda_0) \Big|_{x^*, \lambda^*, \lambda_0} = 0_L$$

Zadanie optymalizacji z ograniczeniami równościowymi

Metoda mnożników Lagrange' a

∞ Uogólniona metoda współczynników Lagrange' a

$$\nabla_x L(x, \lambda, \lambda_0) = \lambda_0 \nabla_x F(x) + \sum_{l=1}^L \lambda_l \nabla_x \varphi_l(x) = 0_S$$

$$1^\circ \quad \lambda_0 \neq 0 \quad \nabla_x F(x) + \sum_{l=1}^L \frac{\lambda_l}{\lambda_0} \nabla_x \varphi_l(x) = 0_S \Rightarrow \nabla_x F(x) + \sum_{l=1}^L \lambda'_l \nabla_x \varphi_l(x) = 0_S$$

$$\lambda_0 = 1 \quad \nabla_x F(x) + \sum_{l=1}^L \lambda_l \nabla_x \varphi_l(x) = 0_S \quad \text{- z tego warunku otrzymamy}$$

rozwiązania regularne

$$2^\circ \quad \lambda_0 = 0 \quad \sum_{l=1}^L \lambda_l \nabla_x \varphi_l(x) = 0_S \quad \text{- z tego warunku otrzymamy rozwiązania}$$

nieregularne

Podobnie jak poprzednio otrzymane rozwiązania wymagają zbadania warunków drugiego rzędu czyli zbadania określoności macierzy:

$$H_L(x, \lambda, \lambda_0) = \nabla_{xx}^2 L(x, \lambda, \lambda_0).$$

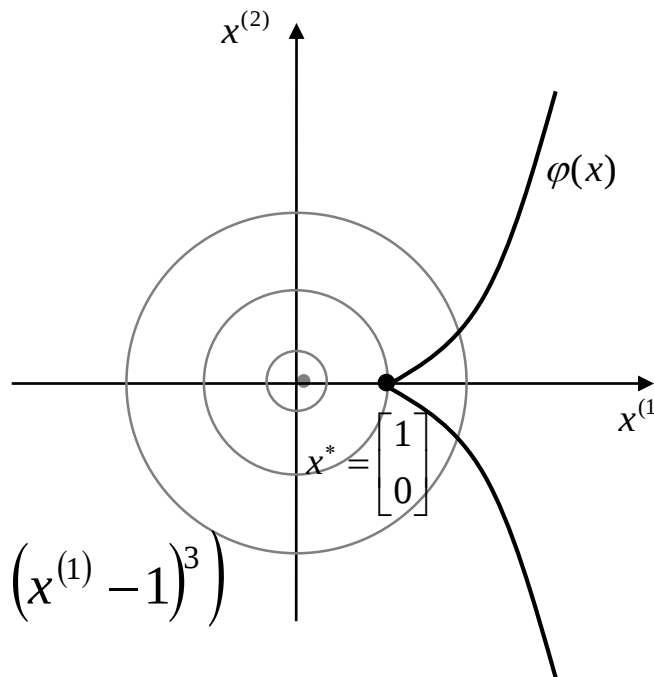
Zadanie optymalizacji z ograniczeniami równościowymi

Metoda mnożników Lagrange' a

☞ Metoda współczynników Lagrange' a – Przykład 2

$$F(x) = (x^{(1)})^2 + (x^{(2)})^2$$

$$\varphi(x) = (x^{(2)})^2 - (x^{(1)} - 1)^3 = 0$$



$$L(x, \lambda) = \lambda_0 \left((x^{(1)})^2 + (x^{(2)})^2 \right) + \lambda \left((x^{(2)})^2 - (x^{(1)} - 1)^3 \right)$$

Przykład 2.2.2

$$\Leftrightarrow L(x, \lambda) = \lambda_0 \left((x^{(1)})^2 + (x^{(2)})^2 \right) + \lambda \left((x^{(2)})^2 + (x^{(1)} - 1)^3 \right)$$

$$\Leftrightarrow \nabla_x L(x, \lambda) = \begin{bmatrix} \frac{\partial L}{\partial x^{(1)}} \\ \frac{\partial L}{\partial x^{(2)}} \end{bmatrix} = \begin{bmatrix} 2\lambda_0 x^{(1)} - 3\lambda (x^{(1)} - 1)^2 \\ 2\lambda_0 x^{(2)} + 2\lambda x^{(2)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \begin{matrix} (1) \\ (2) \end{matrix}$$

$$\Leftrightarrow \nabla_\lambda L(x, \lambda) = \frac{\partial L}{\partial \lambda} = (x^{(2)})^2 + (x^{(1)} - 1)^3 = 0 \quad (3)$$

$\Leftrightarrow$ Dla $\lambda_0=1$

$$\Leftrightarrow \nabla_x L(x, \lambda) = \begin{bmatrix} \frac{\partial L}{\partial x^{(1)}} \\ \frac{\partial L}{\partial x^{(2)}} \end{bmatrix} = \begin{bmatrix} 2x^{(1)} - 3\lambda (x^{(1)} - 1)^2 \\ 2x^{(2)} + 2\lambda x^{(2)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \begin{matrix} (1) \\ (2) \end{matrix}$$

$\Leftrightarrow$ Jak poprzednio - sprzeczność

Przykład 2.2.2

$$\Leftrightarrow L(x, \lambda) = \lambda_0 \left((x^{(1)})^2 + (x^{(2)})^2 \right) + \lambda \left((x^{(2)})^2 + (x^{(1)} - 1)^3 \right)$$

$$\Leftrightarrow \nabla_x L(x, \lambda) = \begin{bmatrix} \frac{\partial L}{\partial x^{(1)}} \\ \frac{\partial L}{\partial x^{(2)}} \end{bmatrix} = \begin{bmatrix} 2\lambda_0 x^{(1)} - 3\lambda (x^{(1)} - 1)^2 \\ 2\lambda_0 x^{(2)} + 2\lambda x^{(2)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \begin{matrix} (1) \\ (2) \end{matrix}$$

$$\Leftrightarrow \nabla_\lambda L(x, \lambda) = \frac{\partial L}{\partial \lambda} = (x^{(2)})^2 + (x^{(1)} - 1)^3 = 0 \quad (3)$$

$$\Leftrightarrow \text{Dla } \lambda_0 = 0$$

$$\Leftrightarrow \nabla_x L(x, \lambda) = \begin{bmatrix} \frac{\partial L}{\partial x^{(1)}} \\ \frac{\partial L}{\partial x^{(2)}} \end{bmatrix} = \begin{bmatrix} 3\lambda (x^{(1)} - 1)^2 \\ \lambda x^{(2)} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \begin{matrix} (1) \\ (2) \end{matrix}$$

$$\Leftrightarrow z(1) \rightarrow 3\lambda (x^{(1)} - 1)^2 = 0 \text{ czyli } x^{(1)} = 1, z(2) \rightarrow x^{(2)} = 0$$

Przykład 2.2.2

$$\begin{aligned} \infty \quad \nabla_x L(x, \lambda) &= \begin{bmatrix} \frac{\partial L}{\partial x^{(1)}} \\ \frac{\partial L}{\partial x^{(2)}} \end{bmatrix} = \begin{bmatrix} 2\lambda_0 x^{(1)} - 3\lambda(x^{(1)} - 1)^2 \\ 2\lambda_0 x^{(2)} + 2\lambda x^{(2)} \end{bmatrix} \end{aligned}$$

$$\infty \quad H_L = \nabla_{xx} L(x, \lambda) = \begin{bmatrix} \frac{\partial^2 L}{\partial^2 x^{(1)}} & \frac{\partial^2 L}{\partial x^{(1)} \partial x^{(2)}} \\ \frac{\partial^2 L}{\partial x^{(2)} \partial x^{(1)}} & \frac{\partial^2 L}{\partial^2 x^{(2)}} \end{bmatrix} =$$

$$\infty \quad \begin{bmatrix} 2\lambda_0 - 6\lambda(x^{(1)} - 1) & 0 \\ 0 & 2\lambda_0 + 2\lambda \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 2\lambda \end{bmatrix}$$

Macierz H_L jest dodatnio pół określona

Punkt $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ - minimum

Zadanie optymalizacji z ograniczeniami nierównościowymi



2.3 Metoda Kuhna-Tuckera

Zadanie optymalizacji z ograniczeniami nierównościami

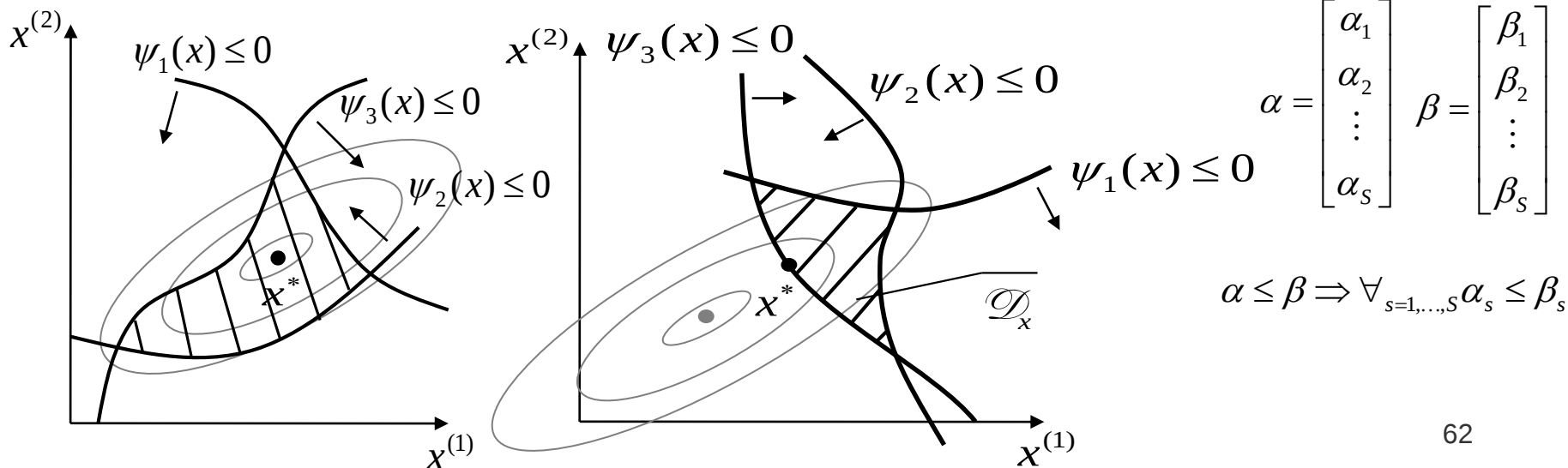
Metoda Kuhna – Tucker'a

Zadanie optymalizacji:

$$x^* \rightarrow F(x^*) = \min_{x \in \mathcal{D}_x} F(x)$$

$$\mathcal{D}_x = \left\{ x \in \mathbb{R}^S : \psi_1(x) \leq 0, \psi_2(x) \leq 0, \dots, \psi_M(x) \leq 0 \right\}$$

$$\mathcal{D}_x = \left\{ x \in \mathbb{R}^S : \psi(x) \leq 0_M \right\} \quad \text{gdzie:} \quad \psi(x) = \begin{bmatrix} \psi_1(x) \\ \psi_2(x) \\ \vdots \\ \psi_M(x) \end{bmatrix} \quad 0_M = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \left\} M$$

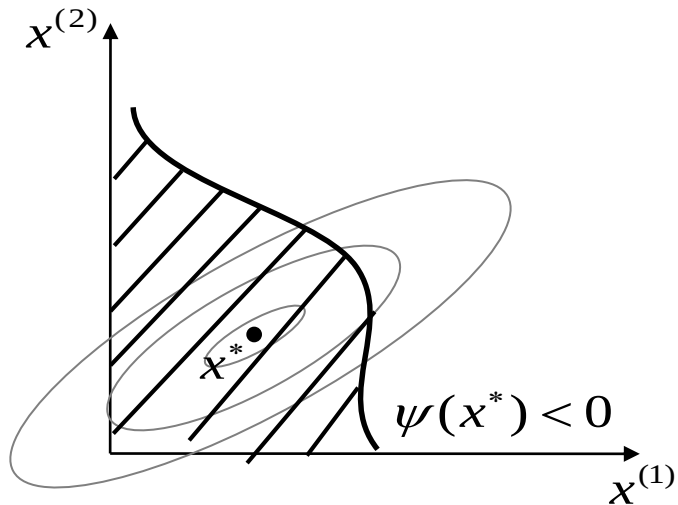


Zadanie optymalizacji z ograniczeniami nierównościowymi

Metoda Kuhna – Tucker'a

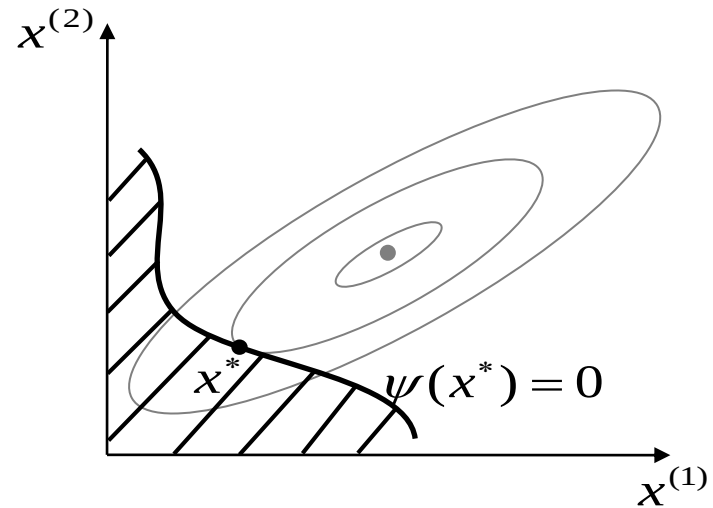
Ograniczenia nieaktywne

$$\psi(x^*) < 0$$



Ograniczenia aktywne

$$\psi(x^*) = 0$$



Zadanie optymalizacji z ograniczeniami nierównościowymi

Metoda Kuhna – Tucker'a

Funkcja Lagrange' a:

$$L(x, \mu) = F(x) + \mu^T \psi(x) \Leftrightarrow L(x, \mu) = F(x) + \sum_{m=1}^M \mu_m \psi_m(x) \quad \text{gdzie: } \mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_M \end{bmatrix} \quad \begin{array}{l} \text{- wektor} \\ \text{współczynników} \\ \text{Lgrange' a} \end{array}$$

Warunki konieczne optymalności:

$$\nabla_x L(x, \mu) \Big|_{x^*, \mu^*} = 0_S$$

$$\mu^T \nabla_\mu L(x, \mu) \Big|_{x^*, \mu^*} = 0$$

$$\nabla_\mu L(x, \mu) \Big|_{x^*, \mu^*} \leq 0_M$$

$$\mu^* \geq 0_M$$

$\Leftrightarrow$ gdy rozwiązanie jest regularne

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_s \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_s \end{bmatrix} \quad \alpha \leq \beta \Rightarrow \forall_{s=1, \dots, S} \alpha_s \leq \beta_s$$

Zadanie optymalizacji z ograniczeniami nierównościami

Metoda Kuhna – Tucker'a

Warunki konieczne:

$$\nabla_x L(x, \mu) = \nabla_x F(x) + \sum_{m=1}^M \mu_m \nabla_x \psi_m(x) = 0_S$$

$$\mu^T \nabla_\mu L(x, \mu) = \mu^T \psi(x) = \sum_{m=1}^M \mu_m \psi_m(x) = 0 \quad (*)$$

$$\nabla_\mu L(x, \mu) = \psi(x) \leq 0_M$$

$$\mu \geq 0_M$$

$$(*) \quad \mu_1 \psi_1(x) + \mu_2 \psi_2(x) + \dots + \mu_M \psi_M(x) = 0$$

$$\forall m \quad \psi_m(x) \leq 0 \quad \forall m \quad \mu_m \geq 0$$

$$\forall m \quad \mu_m \psi_m(x) = 0$$

$$\nabla_x L(x, \mu) \Big|_{x^*, \mu^*} = 0_S$$

$$\mu_m \psi_m(x) \Big|_{x^*, \mu^*} = 0 \quad m = 1, 2, \dots, M$$

$$\psi_m(x) \Big|_{x^*, \mu^*} \leq 0 \quad m = 1, 2, \dots, M$$

$$\mu_m^* \geq 0 \quad m = 1, 2, \dots, M$$

Zadanie optymalizacji z ograniczeniami nierównościami

Metoda Kuhna – Tucker'a

$$L(x, \mu) = F(x) + \mu_m \psi_m(x)$$

$$\nabla_x L(x, \mu) = \nabla_x F(x) + \mu_m \nabla_x \psi_m(x) = 0_s$$

$$\mu^T \nabla_\mu L(x, \mu) = \mu_m \psi_m(x) = 0$$

$$\nabla_\mu L(x, \mu) = \psi_m(x) \leq 0$$

$$\mu_m \geq 0$$

m -te ograniczenie nieaktywne

$$\mu_m = 0 \quad \psi_m(x) < 0$$

$$\nabla_x L(x, \mu) = \nabla_x F(x) = 0_s$$

$$\nabla_\mu L(x, \mu) = \psi_m(x) < 0$$

jak bez ograniczeń

m -te ograniczenie aktywne

$$\mu_m > 0 \quad \psi_m(x) = 0$$

$$\nabla_x L(x, \mu) = \nabla_x F(x) + \mu_m \nabla_x \psi_m(x) = 0_s$$

$$\nabla_\mu L(x, \mu) = \psi_m(x) = 0$$

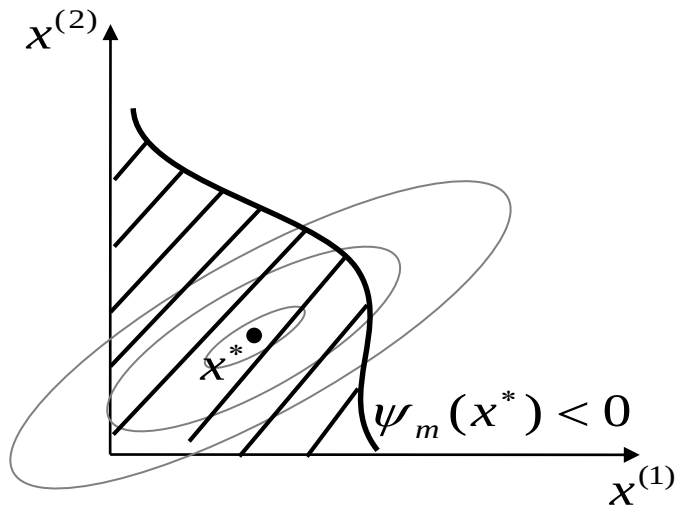
jak z ograniczeniami równościowymi

Zadanie optymalizacji z ograniczeniami nierównościowymi

Metoda Kuhna – Tucker'a

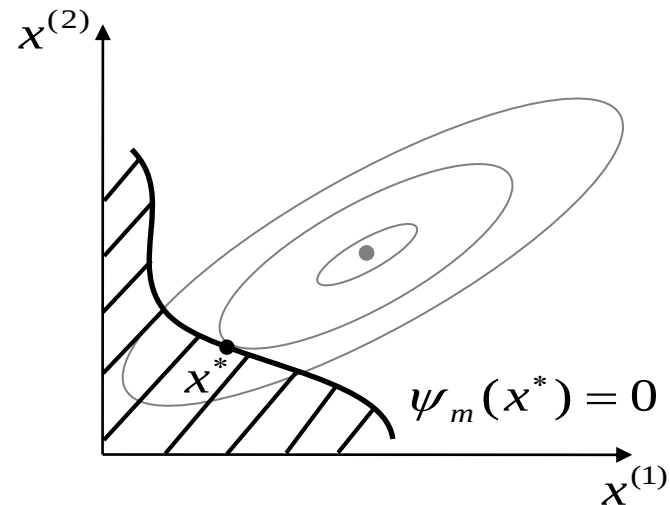
Ograniczenia nieaktywne

$$\psi(x^*)_m < 0$$



Ograniczenie aktywne

$$\psi_m(x^*) = 0$$



Zadanie optymalizacji z ograniczeniami nierównościowymi

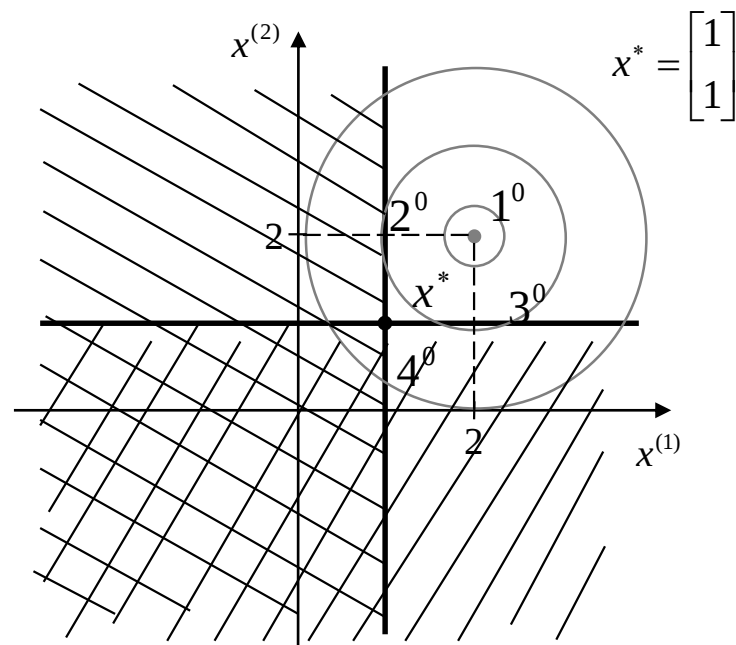
Metoda Kuhna – Tucker'a

Przykład 1

$$F(x) = (x^{(1)} - 2)^2 + (x^{(2)} - 2)^2$$

$$\psi_1(x) = x^{(1)} - 1 \leq 0$$

$$\psi_2(x) = x^{(2)} - 1 \leq 0$$



$$L(x, \lambda) = (x^{(1)} - 2)^2 + (x^{(2)} - 2)^2 + \mu_1(x^{(1)} - 1) + \mu_2(x^{(2)} - 1)$$

Przykład 1.

$$\text{☞ } L(x, \mu) = (x^{(1)} - 2)^2 + (x^{(2)} - 2)^2 + \mu_1(x^{(1)} - 1) + \mu_2(x^{(2)} - 1)$$

$$\text{☞ } \nabla_x L(x, \mu) = \begin{bmatrix} 2(x^{(1)} - 2) + \mu_1 \\ 2(x^{(2)} - 2) + \mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \begin{matrix} (1) \\ (2) \end{matrix}$$

$$\text{☞ } \mu^T \nabla_\mu L(x, \mu) \sim \begin{cases} \mu_1(x^{(1)} - 1) = 0 & (3) \\ \mu_2(x^{(2)} - 1) = 0 & (4) \end{cases}$$

$$\text{☞ } \nabla_\mu L(x, \mu) = \begin{cases} (x^{(1)} - 1) \leq 0 & (5) \\ (x^{(2)} - 1) \leq 0 & (6) \end{cases}$$

$$\text{☞ } \mu = \begin{cases} \mu_1 \geq 0 & (7) \\ \mu_2 \geq 0 & (8) \end{cases}$$

Przykład 1. c.d.

☞ $1^0 \quad \mu_1 = 0 \ (x^{(1)} - 1 < 0 ??), \mu_2 = 0 \ (x^{(2)} - 1 < 0 ??),$

z (1) $\rightarrow 2(x^{(1)} - 2) = 0 \rightarrow x^{(1)} = 2$

z (2) $\rightarrow 2(x^{(2)} - 2) = 0 \rightarrow x^{(2)} = 2$

z (5) $\rightarrow (2 - 1) = 1 \geq 0$ sprzeczność z (5)

z (6) $\rightarrow (2 - 1) = 1 \geq 0$ sprzeczność z (6)

☞ $2^0 \quad \mu_1 > 0 \ (x^{(1)} - 1 = 0 ??), \mu_2 = 0 \ (x^{(2)} - 1 < 0 ??),$

z (3) $\rightarrow \mu_1(x^{(1)} - 1) = 0 / \mu_1 \rightarrow (x^{(1)} - 1) = 0 \rightarrow x^{(1)} = 1$

z (1) $\rightarrow 2(1 - 2) + \mu_1 = 0 \rightarrow \mu_1 = 2$

z (2) $\rightarrow 2(x^{(2)} - 2) = 0 \rightarrow x^{(2)} = 2$

z (6) $\rightarrow (2 - 1) = 1 \geq 0$ sprzeczność z (6)

Przykład 1. c.d.

3⁰ $\mu_1 = 0$ ($x^{(1)} - 1 < 0 ??$), $\mu_2 > 0$ ($x^{(2)} - 1 = 0 ??$),

z (1) $\rightarrow 2(x^{(1)} - 2) = 0 \rightarrow x^{(1)} = 2$

z (5) $\rightarrow (2 - 1) = 1 \geq 0$ sprzeczność z (5)

z (4) $\rightarrow \mu_2(x^{(2)} - 1) = 0 / \mu_1 \rightarrow (x^{(2)} - 1) = 0 \rightarrow x^{(2)} = 1$

z (2) $\rightarrow 2(1 - 2) + \mu_2 = 0 \rightarrow \mu_2 = 2$

4⁰ $\mu_1 > 0$ ($x^{(1)} - 1 = 0 ??$), $\mu_2 > 0$ ($x^{(2)} - 1 = 0 ??$),

z (3) $\rightarrow \mu_1(x^{(1)} - 1) = 0 / \mu_1 \rightarrow (x^{(1)} - 1) = 0 \rightarrow x^{(1)} = 1$

z (1) $\rightarrow 2(1 - 2) + \mu_1 = 0 \rightarrow \mu_1 = 2$

z (4) $\rightarrow \mu_2(x^{(2)} - 1) = 0 / \mu_1 \rightarrow (x^{(2)} - 1) = 0 \rightarrow x^{(1)} = 1$

z (2) $\rightarrow 2(1 - 2) + \mu_2 = 0 \rightarrow \mu_2 = 2$

Punkt $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ spełnia równania i jest rozwiązaniem zadania

Zadanie optymalizacji z ograniczeniami nierównościowymi

Metoda Kuhna – Tucker'a

Funkcja Lagrange' a:

$$L(x, \mu) = F(x) + \mu^T \psi(x) \Leftrightarrow L(x, \mu) = F(x) + \sum_{m=1}^M \mu_m \psi_m(x)$$

gdzie: $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_M \end{bmatrix}$ - wektor współczynników Lgrange' a

Warunki konieczne optymalności:

$$\nabla_x L(x, \mu) \Big|_{x^*, \mu^*} = 0_S$$

$$\mu^T \nabla_\mu L(x, \mu) \Big|_{x^*, \mu^*} = 0$$

$$\nabla_\mu L(x, \mu) \Big|_{x^*, \mu^*} \leq 0_M$$

$$\mu^* \geq 0_M$$

$\Leftrightarrow$ gdy rozwiązanie jest regularne

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_s \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_s \end{bmatrix}$$

$$\alpha \leq \beta \Rightarrow \forall_{s=1, \dots, S} \alpha_s \leq \beta_s$$

Zadanie optymalizacji z ograniczeniami nierównościami

Metoda Kuhna – Tucker'a

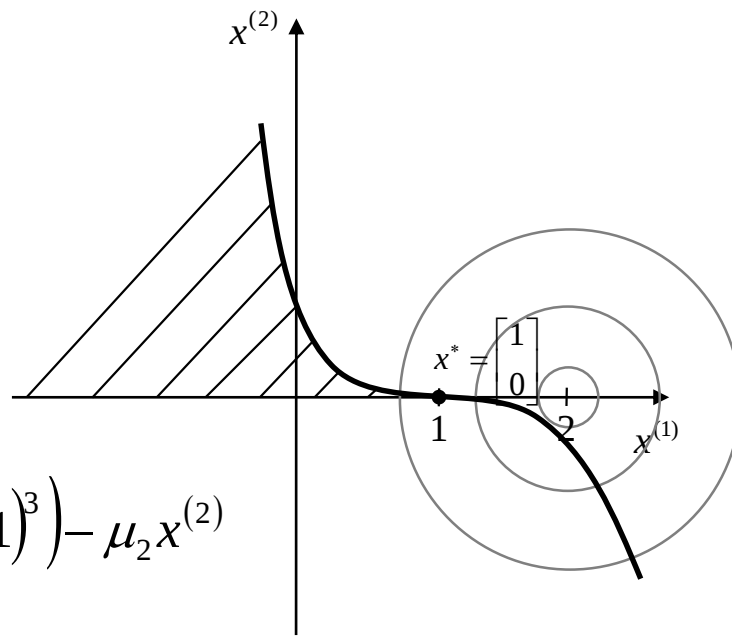
Przykład 2 – rozwiązanie nieregularne

$$F(x) = (x^{(1)} - 2)^2 + (x^{(2)})^2$$

$$\psi_1(x) = x^{(2)} + (x^{(1)} - 1)^3 \leq 0$$

$$\psi_2(x) = -x^{(2)} \leq 0$$

$$L(x, \lambda) = (x^{(1)} - 2)^2 + (x^{(2)})^2 + \mu_1 (x^{(2)} + (x^{(1)} - 1)^3) - \mu_2 x^{(2)}$$



Przykład 2.

$$\wp L(x, \mu) = (x^{(1)} - 2)^2 + (x^{(2)} - 2)^2 + \mu_1 (x^{(2)} + (x^{(1)} - 1)^3) - \mu_2 x^{(2)}$$

$$\wp \nabla_x L(x, \mu) = \begin{bmatrix} 2(x^{(1)} - 2) + 3\mu_1(x^{(1)} - 1)^2 \\ 2(x^{(2)} - 2) + \mu_1 - \mu_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad \begin{matrix} (1) \\ (2) \end{matrix}$$

$$\wp \mu^T \nabla_\mu L(x, \mu) \sim \begin{cases} \mu_1 (x^{(2)} + (x^{(1)} - 1)^3) = 0 & (3) \\ -\mu_2 x^{(2)} = 0 & (4) \end{cases}$$

$$\wp \nabla_\mu L(x, \mu) = \begin{cases} x^{(2)} + (x^{(1)} - 1)^3 \leq 0 & (5) \\ -x^{(2)} \leq 0 & (6) \end{cases}$$

$$\wp \mu = \begin{cases} \mu_1 \geq 0 & (7) \\ \mu_2 \geq 0 & (8) \end{cases}$$

Przykład 2.

- ✎ Powyższy układ równań i nierówności należy rozwiązywać jak poprzednio przyjmując odpowiednie μ_1 oraz μ_2 . Dla wszystkich przypadków otrzymamy sprzeczność.
- ✎ Pokażmy że rozwiązanie $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ widoczne na ilustracji graficznej nie spełnia układu równań i nierówności

Przykład 2.

☞ Dla $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ otrzymujemy

$$\text{☞ } \nabla_x L(x, \mu) = \begin{cases} 2(1 - 2) + 3\mu_1(1 - 1)^2 = -2 \neq 0 & \text{sprzeczność} \end{cases} \quad (1)$$

$$2(0 - 2) + \mu_1 - \mu_2 = 0 \quad (2)$$

$$\text{☞ } \mu^T \nabla_\mu L(x, \mu) \sim \begin{cases} \mu_1(0 + (1 - 1)^3) = 0 & (3) \\ -\mu_2 0 = 0 & (4) \end{cases}$$

$$\text{☞ } \nabla_\mu L(x, \mu) = \begin{cases} 0 + (1 - 1)^3 \leq 0 & (5) \\ -0 \leq 0 & (6) \end{cases}$$

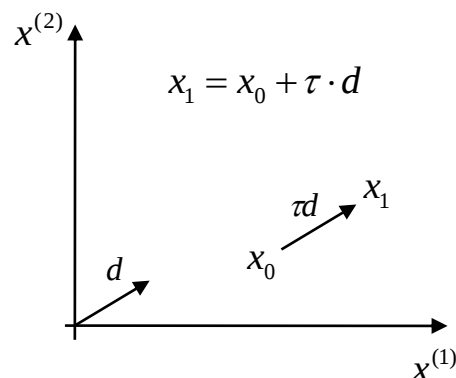
$$\text{☞ } \mu = \begin{cases} \mu_1 \geq 0 & (7) \\ \mu_2 \geq 0 & (8) \end{cases}$$

Rozwiązanie nieregularne

Zadanie optymalizacji z ograniczeniami nierównościami

Metoda Kuhna – Tucker'a

$$d = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_s \end{bmatrix} \quad - \text{ kierunek w } \mathcal{R}^s$$

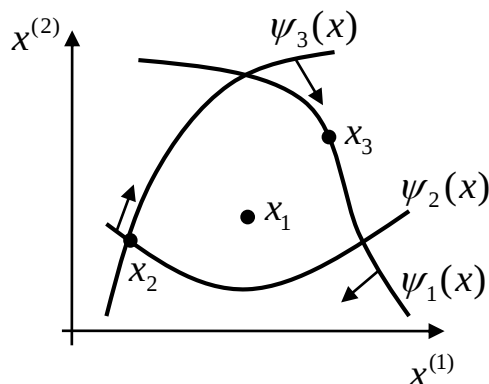
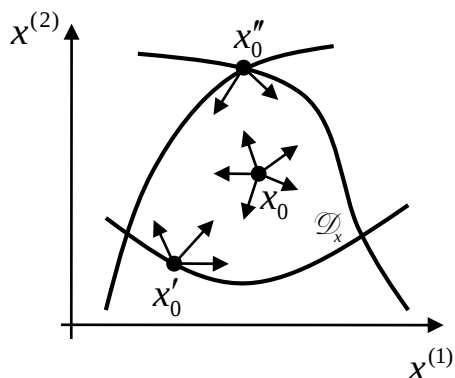


Zbiór kierunków dopuszczalnych

Zbiór ograniczeń aktywnych

$$D(x) = \{d \in \mathcal{R}^s : \exists \tau > 0 \quad x + \tau d \in \mathcal{D}_x\}$$

$$I(x) = \{m \in \{1, 2, \dots, M\} : \psi_m(x) = 0\}$$



$$\begin{aligned} I(x_1) &= \emptyset \\ I(x_2) &= \{2, 3\} \\ I(x_3) &= \{1\} \end{aligned}$$

Zadanie optymalizacji z ograniczeniami nierównościami

Metoda Kuhna – Tucker'a

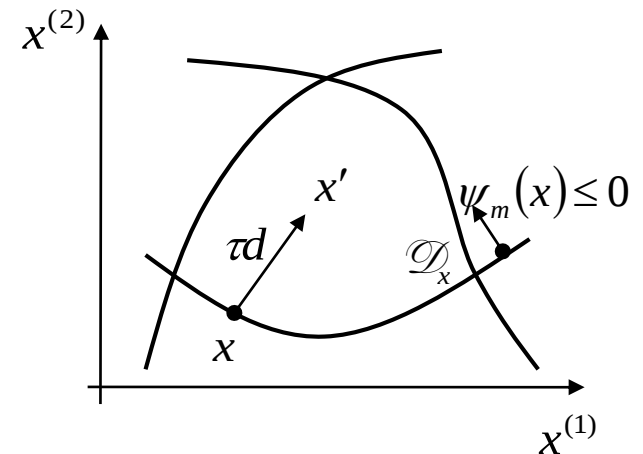
Jaki warunek musi spełniać na ograniczeniu aktywnym?

$$\forall d \in D(x) \wedge \forall m \in I(x)$$

$$\text{tj.: } \psi_m(x) = 0$$

$$x' = x + \tau d \in \mathcal{D}_x, \quad \tau > 0$$

$$\psi_m(x') \leq 0$$



$$\psi_m(x') = \psi_m(x + \tau d) = \psi_m(x) + \tau d^T \nabla_x \psi_m(x) + O_2(\|\tau d\|) \leq 0$$

$$\tau d^T \nabla_x \psi_m(x) \leq 0 \quad \tau > 0$$

$$\forall d \in D(x) \wedge \forall m \in I(x) \quad d^T \nabla_x \psi_m(x) \leq 0 \quad - \text{warunek analityczny}$$

Zadanie optymalizacji z ograniczeniami nierównościowymi

Metoda Kuhna – Tucker'a

Przykład 1

$$\psi_1(x) = x^{(1)} - 1 \leq 0$$

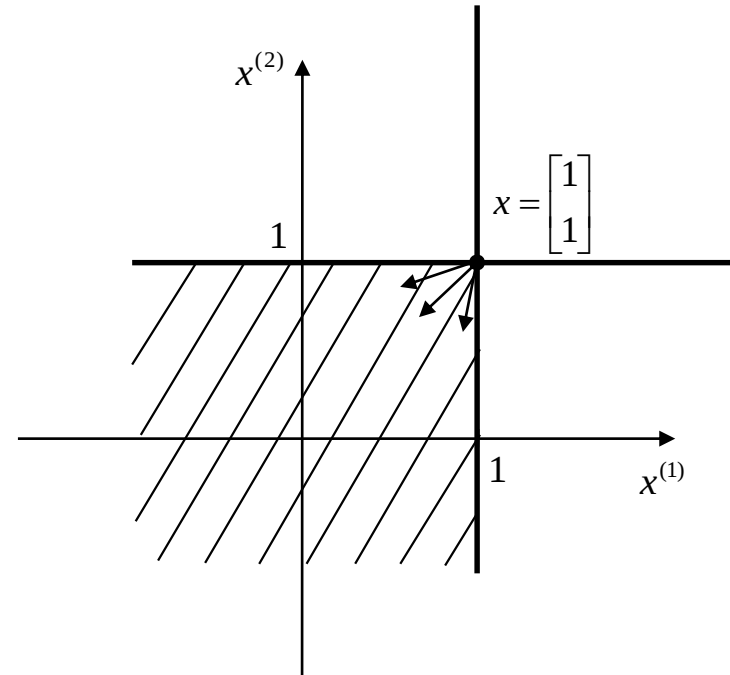
$$\psi_2(x) = x^{(2)} - 1 \leq 0$$

W punkcie $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $I\left(x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}\right) = \{1, 2\}$

$$\nabla_x \psi_1(x) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \nabla_x \psi_2(x) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad d = \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}$$

$$d^T \nabla_x \psi_1(x) = \begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1 \cdot d_1 + 0 \cdot d_2 \leq 0 \Rightarrow d_1 \leq 0$$

$$d^T \nabla_x \psi_2(x) = \begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \cdot d_1 + 1 \cdot d_2 \leq 0 \Rightarrow d_2 \leq 0$$



Zadanie optymalizacji z ograniczeniami nierównościami

Metoda Kuhna – Tucker'a

Uwaga: Nie każdy kierunek, który spełnia warunek $d^T \nabla_x \psi_m(x) \leq 0$ jest kierunkiem dopuszczalnym. Może to prowadzić do rozwiązania nieregularnego

$$\mathcal{D}(x) = \left\{ d \in \mathbb{R}^s : \forall_m \in I(x), d^T \nabla_x \psi_m(x) \leq 0 \right\} \quad \mathcal{D}(x) \neq D(x)$$

$$\psi_1(x) = x^{(2)} + (x^{(1)} - 1)^3 \leq 0 \quad \psi_2(x) = -x^{(2)} \leq 0 \quad F(x) = (x^{(1)} - 2)^2 + (x^{(2)})^2$$

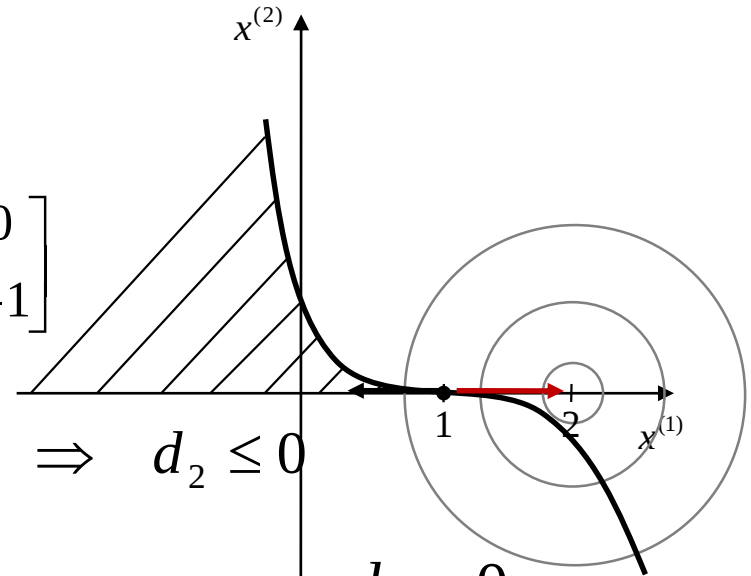
W punkcie $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $I\left(x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \{1, 2\}$

$$\nabla_x \psi_1(x) = \begin{bmatrix} 3(x^{(1)} - 1)^2 \\ 1 \end{bmatrix} \Big|_{x=\begin{bmatrix} 1 \\ 0 \end{bmatrix}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \nabla_x \psi_2(x) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

$$d^T \nabla_x \psi_1(x) = \begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \end{bmatrix} = 0 \cdot d_1 + 1 \cdot d_2 \leq 0 \Rightarrow d_2 \leq 0$$

$$d^T \nabla_x \psi_2(x) = \begin{bmatrix} d_1 & d_2 \end{bmatrix} \begin{bmatrix} 0 \\ -1 \end{bmatrix} = 0 \cdot d_1 - 1 \cdot d_2 \leq 0 \Rightarrow d_2 \geq 0$$

$d_2 = 0$
 $d_1 - \text{dowo ln } e$



Zadanie optymalizacji z ograniczeniami nierównościami

Metoda Kuhna – Tucker'a

$$F(x) = F(x_0 + \tau d) = F(x_0) + \tau (\nabla_x F(x))^T d + O_3(\|\tau d\|)$$

Jeżeli kierunek d taki że: $(\nabla_x F(x))^T d < 0$ to $F(x) < F(x_0)$

czyli funkcja maleje w kierunku d

Podzielmy zbiór kierunków $\mathcal{D}(x) = \{d \in \mathcal{R}^s : \forall_m \in I(x), d^T \nabla_x \psi_m(x) \leq 0\}$ na:

$$\mathcal{D}_1(x) = \{d \in \mathcal{R}^s : \forall_m \in I(x), d^T \nabla_x \psi_m(x) \leq 0\} \wedge (\nabla_x F(x))^T d \geq 0$$

tj.: funkcja nie maleje we wskazanych kierunkach oraz

$$\mathcal{D}_2(x) = \{d \in \mathcal{R}^s : \forall_m \in I(x), d^T \nabla_x \psi_m(x) \leq 0\} \wedge (\nabla_x F(x))^T d < 0$$

tj.: funkcja maleje we wskazanych kierunkach oraz

Zadanie optymalizacji z ograniczeniami nierównościami

Metoda Kuhna – Tucker'a

Funkcja Lagrange' a:

$$L(x, \mu) = F(x) + \mu^T \psi(x) \Rightarrow L(x, \mu) = F(x) + \sum_{m=1}^M \mu_m \psi_m(x)$$

Twierdzenie Kuhna – Tuckera – warunki konieczne optymalności:

Jeżeli x^* jest minimum lokalnym zadania z ograniczeniami nierównościami, funkcje $F, \Psi_1, \Psi_2, \dots, \Psi_M$ są ciągłe oraz funkcja F jest różniczkowalna to istnieje zestaw współczynników Lagrange'a μ^* takich że wraz z x^* *spełnia*

$$\nabla_x L(x, \mu) \Big|_{x^*, \mu^*} = 0_S$$

$$\mu^T \nabla_\mu L(x, \mu) \Big|_{x^*, \mu^*} = 0$$

$$\nabla_\mu L(x, \mu) \Big|_{x^*, \mu^*} \leq 0_M$$

$$\mu^* \geq 0_M$$

$$\Leftrightarrow \mathcal{D}_x(x) = \left\{ d \in \mathcal{R}^s : \forall_m \in I(x), d^T \nabla_x \psi_m(x) \leq 0 \right\} \wedge (\nabla_x F(x))^T d < 0 = \Phi$$

Zadanie optymalizacji z ograniczeniami nierównościowymi

Metoda Kuhna – Tucker’a Warunki regularności rozwiązania

1. Karlina: ograniczenia $\psi_1(x), \psi_2(x), \dots, \psi_M(x)$ - liniowe
2. Slatera: ograniczenia $\psi_1(x), \psi_2(x), \dots, \psi_M(x)$ - wypukłe oraz zbiór rozwiązań dopuszczalnych ma niepuste wnętrze
3. Fiacco – Mac Cormica: w punkcie optymalnym gradienty wszystkich ograniczeń aktywnych są liniowo niezależne, czyli:

$$\forall m \in I(x^*) \quad \nabla_x \psi_m(x^*) \Big|_{x=x^*} \text{ są liniowo niezależne}$$

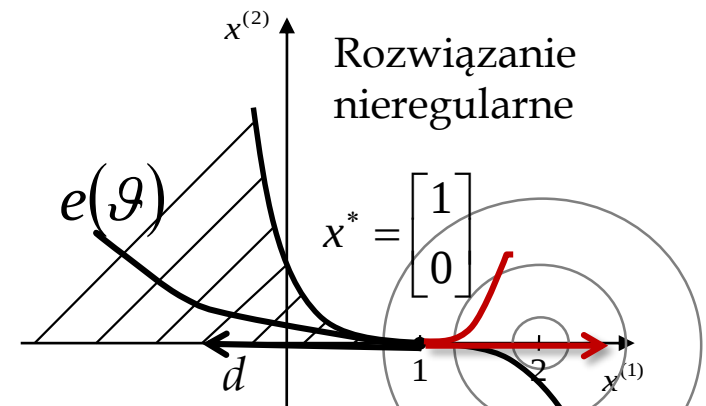
$$4. \text{Zangwila: } \mathcal{D}(x^*) = \overline{D}(x^*)$$

5. Kuhna – Tucker’a: dla każdego kierunku $d \in \mathcal{D}(x^*)$ istnieje krzywa regularna rozpoczynająca się w punkcie x^* i styczna do tego kierunku

$$\forall d \in \mathcal{D}(x^*) \quad \exists e(\vartheta), \quad \vartheta \in [0, 1]$$

- $e(0) = x^*$
- $e(\vartheta) \in D_x \quad \forall \vartheta \in [0, 1]$
- $\frac{de(\vartheta)}{d\vartheta} \Big|_{\vartheta=0} = \tau \cdot d$

$$e(\vartheta) = \begin{bmatrix} e_1(\vartheta) \\ e_2(\vartheta) \\ \vdots \\ e_s(\vartheta) \end{bmatrix}$$



Zadanie optymalizacji z ograniczeniami nierównościowymi

Metoda Kuhna – Tucker'a

$$\psi_1(x) = x^{(2)} + (x^{(1)} - 1)^3 \leq 0 \quad \psi_2(x) = -x^{(2)} \leq 0 \quad F(x) = (x^{(1)} - 2)^2 + (x^{(2)})^2$$

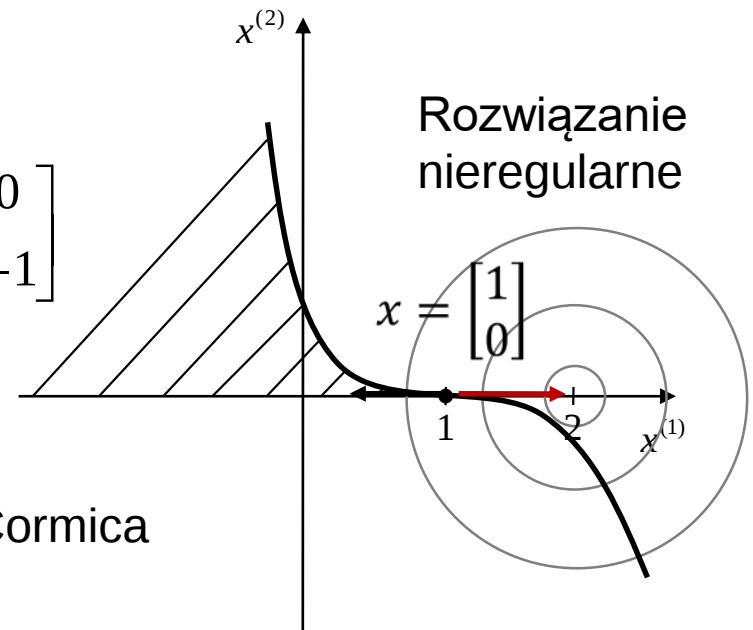
W punkcie $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $I\left(x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \{1, 2\}$

Ograniczenie 1 i 2 są aktywne

$$\nabla_x \psi_1(x) = \begin{bmatrix} 3(x^{(1)} - 1)^2 \\ 1 \end{bmatrix} \Big|_{x=\begin{bmatrix} 1 \\ 0 \end{bmatrix}} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \nabla_x \psi_2(x) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}$$

Gradienty ograniczeń są liniowo zależne

W punkcie $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ warunek Fiacco – Mac Cormica nie jest spełniony



Zadanie optymalizacji z ograniczeniami nierównościami

Metoda Kuhna – Tucker'a

Warunki konieczna i wystarczające:

Jeżeli funkcje $F(x)$, $\psi_1(x)$, $\psi_2(x)$, ..., $\psi_M(x)$ są ciągłe i różniczkowalne oraz funkcja $F(x)$ jest funkcją pseudo – wypukłą, a ograniczenia $\psi_1(x)$, $\psi_2(x)$, ..., $\psi_M(x)$ są funkcjami quasi – wypukłymi to wkład równań i nierówności:

$$\nabla_x L(x, \mu) \Big|_{x^*, \mu^*} = 0_S$$

$$\mu^T \nabla_\mu L(x, \mu) \Big|_{x^*, \mu^*} = 0$$

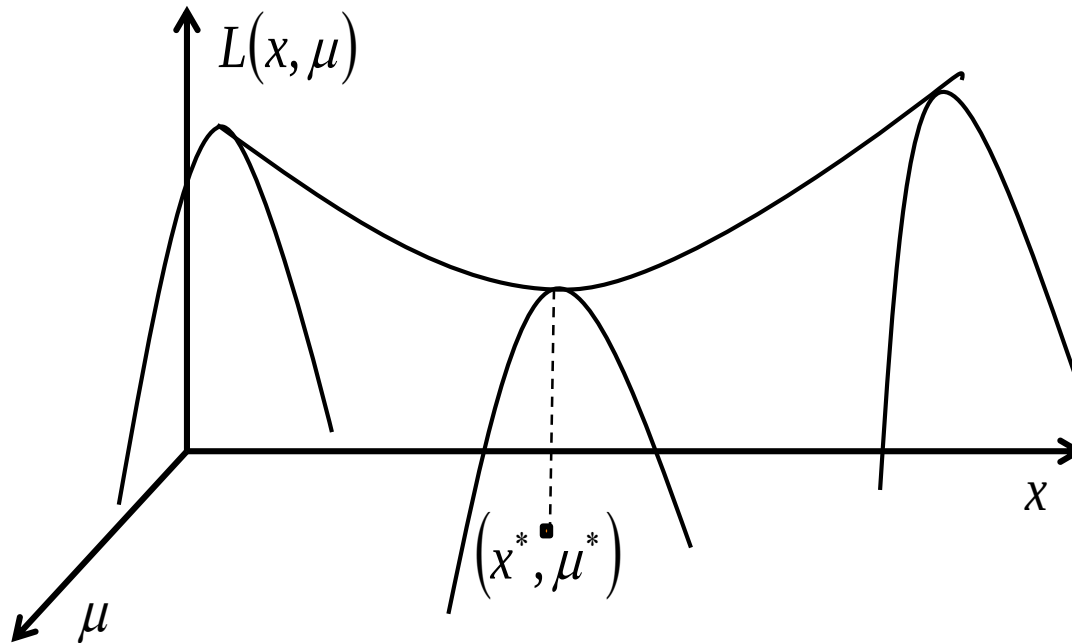
$$\nabla_\mu L(x, \mu) \Big|_{x^*, \mu^*} \leq 0_M$$

$$\mu^* \geq 0_M$$

ma jedno rozwiązanie i jest ono rozwiązaniem zadania optymalizacji z ograniczeniami nierównościami

Zadanie optymalizacji z ograniczeniami nierównościowymi - Punkt siodłowy

Punkt siodłowy (x^*, μ^*)



$$L(x^*, \mu^*) \leq L(x, \mu^*) \quad \forall x \in \mathcal{D}(x) \subseteq \mathcal{R}^s$$

$$L(x^*, \mu) \leq L(x^*, \mu^*) \quad \forall \mu \geq 0_M$$

$$L(x^*, \mu^*) = \min_{x \in \mathcal{D}(x)} \max_{\mu \geq 0_M} L(x, \mu)$$

Zadanie optymalizacji z ograniczeniami nierównościami - Punkt siodłowy

Punkt (x^*, μ^*) jest punktem siodłowym $(x \in \mathcal{D}(x), \mu \geq 0_M) \Leftrightarrow$

1. x^* – minimalizuje $L(x, \mu)$
2. $\psi_m(x^*) \leq 0 \quad m = 1, 2, \dots, M$
3. $\mu^* \psi_m(x^*) = 0 \quad m = 1, 2, \dots, M$

Jeżeli (x^*, μ^*) jest punktem siodłowym funkcji Lagrange'a $L(x, \mu)$ to jest rozwiązaniem zadania optymalizacji:

$$x^* \rightarrow F(x^*) = \min_{x \in \mathcal{D}_x} F(x)$$

$$\mathcal{D}_x = \{x \in \mathcal{R}^S : \psi_1(x) \leq 0, \psi_2(x) \leq 0, \dots, \psi_M(x) \leq 0\}$$

Szczególny przypadek

$$x^* \rightarrow F(x^*) = \min_{x \in \mathcal{D}_x} F(x)$$

$$\mathcal{D}_x = \{x \in \mathcal{R}^S : x \geq 0_S, \psi(x) \leq 0_M\}$$

$$L(x, \mu) = F(x) + \mu^T \psi(x)$$

$$\nabla_x L(x, \mu) \Big|_{x^*, \mu^*} \geq 0_S$$

$$\nabla_\mu L(x, \mu) \Big|_{x^*, \mu^*} \leq 0_M$$

$$x^T \nabla_x L(x, \mu) \Big|_{x^*, \mu^*} = 0$$

$$\mu^T \nabla_\mu L(x, \mu) \Big|_{x^*, \mu^*} = 0$$

$$x^* \geq 0_S$$

$$\mu^* \geq 0_M$$

$$\mathcal{D}_x = \left\{ x \in \mathcal{R}^S : x \geq 0_S, \psi(x) \leq 0_M \right\}$$

$$L(x, \mu) = F(x) + \mu^T \psi(x)$$

$$\mathcal{D}_x = \{x \in \mathcal{R}^S : \psi(x) \leq 0_M, -x \leq 0_S\}$$

$$L(x, \mu, \mu') = F(x) + \mu^T \psi(x) - \mu'^T x$$

Warunki Kuhna-Tuckera

$$L(x, \mu, \mu') = F(x) + \mu^T \psi(x) - \mu'^T x$$

$$\nabla_x L(x, \mu, \mu') = F(x) + \sum_{m=1}^M \mu_m \nabla_x \psi_m(x) - \mu' = 0_S$$

$$\mu^T \nabla_x L(x, \mu, \mu') = \mu^T \psi(x) = 0$$

$$\nabla_\mu L(x, \mu, \mu') = \psi(x) \leq 0_M$$

$$\nabla_{\mu'} L(x, \mu, \mu') = -x \leq 0_S$$

$$\mu \geq 0_M, \mu' \geq 0_S$$

Szczególny przypadek

$$x^* \rightarrow F(x^*) = \min_{x \in \mathcal{D}_x} F(x)$$

$$\mathcal{D}_x = \{x \in \mathcal{R}^S : x \geq 0_S, \psi(x) \leq 0_M\}$$

$$L(x, \mu) = F(x) + \mu^T \psi(x)$$

$$\nabla_x L(x, \mu) \Big|_{x^*, \mu^*} \geq 0_S$$

$$\nabla_\mu L(x, \mu) \Big|_{x^*, \mu^*} \leq 0_M$$

$$x^T \nabla_x L(x, \mu) \Big|_{x^*, \mu^*} = 0$$

$$\mu^T \nabla_\mu L(x, \mu) \Big|_{x^*, \mu^*} = 0$$

$$x^* \geq 0_S$$

$$\mu^* \geq 0_M$$

Szczególny przypadek

$$x^* \rightarrow F(x^*) = \min_{x \in \mathcal{D}_x} F(x)$$

$$\mathcal{D}_x = \{x \in \mathcal{R}^S : \varphi(x) = 0_L, \psi(x) \leq 0_M\}$$

$$L(x, \lambda, \mu) = F(x) + \lambda^T \varphi(x) + \mu^T \psi(x)$$

$$\nabla_x L(x, \lambda, \mu) \Big|_{x^*, \lambda^*, \mu^*} = 0_S$$

$$\nabla_\lambda L(x, \lambda, \mu) \Big|_{x^*, \lambda^*, \mu^*} = 0_L$$

$$\mu^T \nabla_\mu L(x, \lambda, \mu) \Big|_{x^*, \lambda^*, \mu^*} = 0$$

$$\nabla_\mu L(x, \lambda, \mu) \Big|_{x^*, \lambda^*, \mu^*} \leq 0_M$$

$$\mu^* \geq 0_S$$

$$\mathcal{D}_x = \{x \in \mathcal{R}^S : \varphi(x) = 0_L, \psi(x) \leq 0_M\}$$

$$L(x, \lambda, \mu) = F(x) + \lambda^T \varphi(x) + \mu^T \psi(x)$$

$$\varphi(x) = 0_L \equiv \varphi(x) \leq 0_L \cap -\varphi(x) \leq 0_L$$

$$\mathcal{D}_x = \{x \in \mathcal{R}^S : \varphi(x) \leq 0_L \cap -\varphi(x) \leq 0_L, \psi(x) \leq 0_M\}$$

$$L(x, \lambda', \lambda, \mu) = F(x) + \lambda^T \varphi(x) - \lambda'^T \varphi(x) + \mu^T \psi(x)$$

Warunki Kuhna-Tuckera

$$L(x, \lambda, \lambda', \mu) = F(x) + \lambda^T \varphi(x) - \lambda'^T \varphi(x) + \mu^T \psi(x)$$

$$\begin{aligned} \nabla_x L(x, \lambda, \lambda', \mu) &= \\ &= \nabla_x F(x) + \sum_{l=1}^L \lambda_l \nabla_x \varphi_l(x) - \sum_{l=1}^L \lambda'_l \nabla_x \varphi_l(x) + \sum_{m=1}^M \mu_m \nabla_x \psi_m(x) = 0_S \end{aligned}$$

$$\lambda^T \nabla_\lambda L(x, \lambda, \lambda', \mu) = \lambda^T \varphi(x) = 0$$

$$\lambda'^T \nabla_{\lambda'} L(x, \lambda, \lambda', \mu) = -\lambda'^T \varphi(x) = 0$$

$$\mu^T \nabla_\mu L(x, \lambda, \lambda', \mu) = \mu^T \psi(x) = 0$$

$$\nabla_\lambda L(x, \lambda, \lambda', \mu) = \varphi(x) \leq 0_L$$

$$\nabla_{\lambda'} L(x, \lambda, \lambda', \mu) = -\varphi(x) \leq 0_L$$

$$\nabla_\mu L(x, \lambda, \lambda', \mu) = \psi(x) \leq 0_M$$

$$\lambda \geq 0_L, \lambda' \geq 0_L, \mu \geq 0_M$$

Szczególny przypadek

$$x^* \rightarrow F(x^*) = \min_{x \in \mathcal{D}_x} F(x)$$

$$\mathcal{D}_x = \{x \in \mathcal{R}^S : \varphi(x) = 0_L, \psi(x) \leq 0_M\}$$

$$L(x, \lambda, \mu) = F(x) + \lambda^T \varphi(x) + \mu^T \psi(x)$$

$$\nabla_x L(x, \lambda, \mu) \Big|_{x^*, \lambda^*, \mu^*} = 0_S$$

$$\nabla_\lambda L(x, \lambda, \mu) \Big|_{x^*, \lambda^*, \mu^*} = 0_L$$

$$\mu^T \nabla_\mu L(x, \lambda, \mu) \Big|_{x^*, \lambda^*, \mu^*} = 0$$

$$\nabla_\mu L(x, \lambda, \mu) \Big|_{x^*, \lambda^*, \mu^*} \leq 0_M$$

$$\mu^* \geq 0_S$$

Metody analityczne

Problemy

Mogą pojawić się problemy analityczne:

F, φ, ψ - złożone funkcje nieliniowe

$\dim(x)$ - duży wymiar

F, φ, ψ - funkcje nieróżniczkowalne

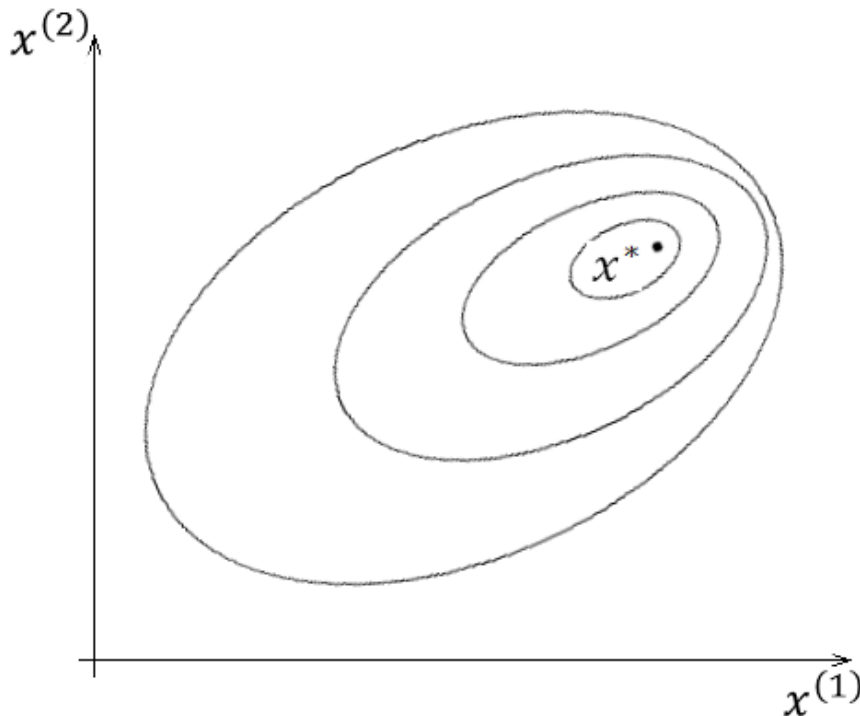
F - analityczna postać funkcji nie jest znana, a istnieje
możliwość pomiaru wartości funkcji w punkcie x

Powyższe przesłanki skłaniają do poszukiwania metod numerycznych

Numeryczne metody optymalizacji



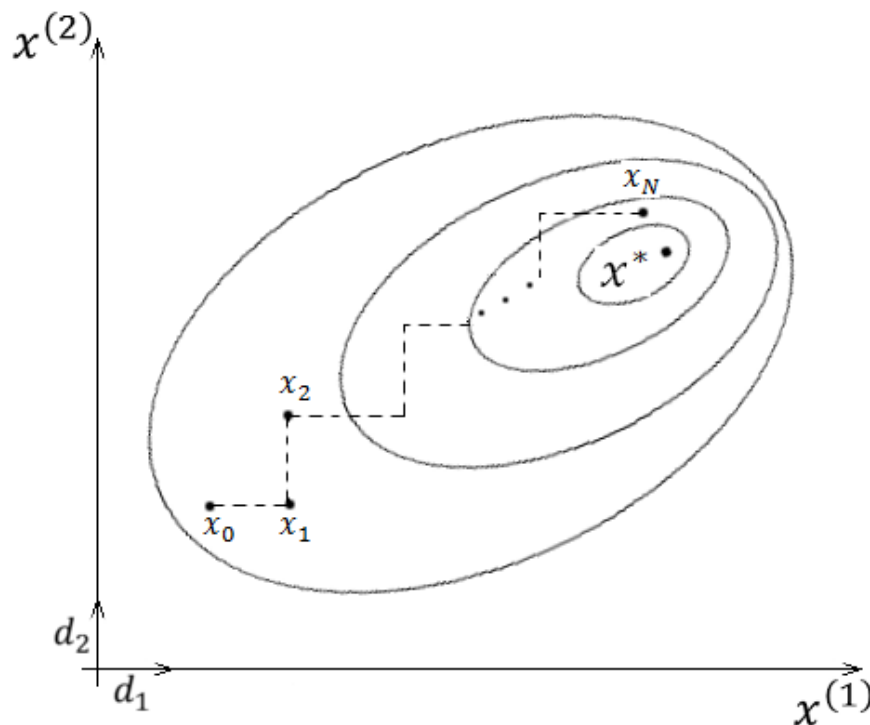
$$x^* \rightarrow F(x^*) = \min_{x \in D_x} F(x)$$



Problemy analityczne:

1. Nieliniowa złożona funkcja celu F i ograniczeń φ oraz ψ .
2. Nieróżniczkowalność funkcji F , φ oraz ψ .
3. Nie jest znana postać analityczna funkcji F , φ oraz ψ , można jedynie „zmierzyć” wartość funkcji
4. Duży wymiar wektora zmiennych decyzyjnych.

Numeryczne metody optymalizacji



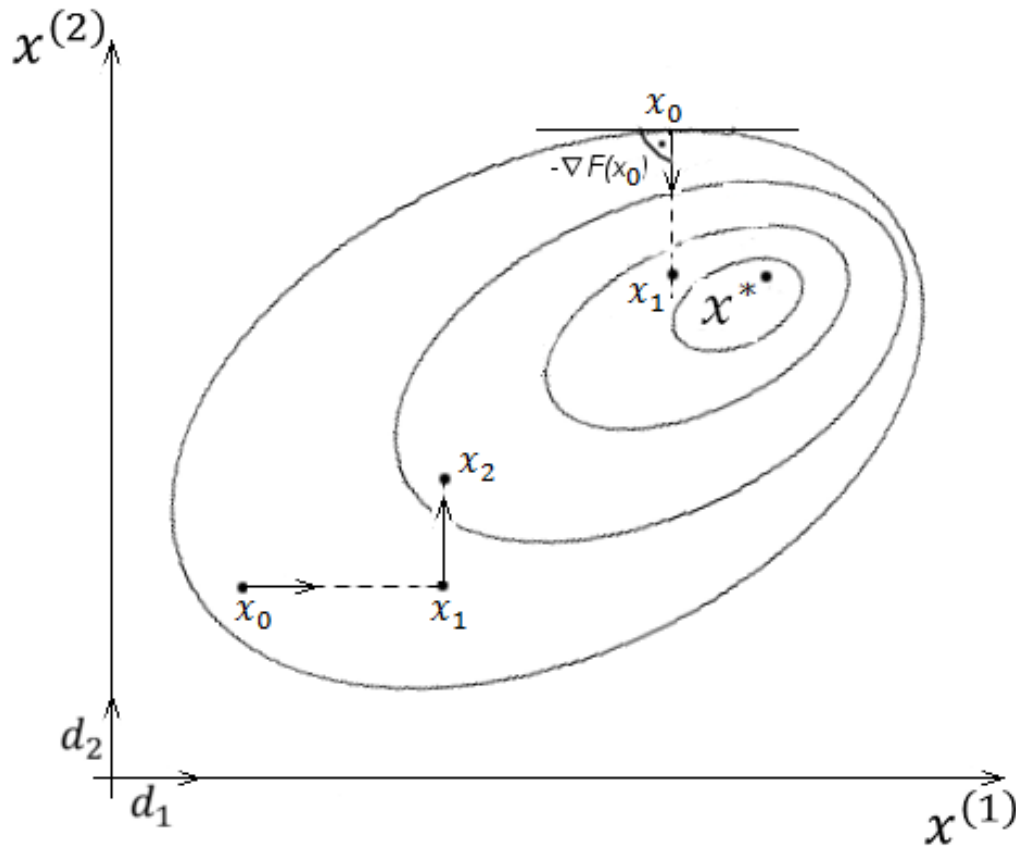
Algorytm

$$x_{n+1} = \Psi(x_n), x_0$$

- Wybór kierunku poszukiwań.
- Optymalizacja w kierunku.
- Warunki zatrzymania procedury.

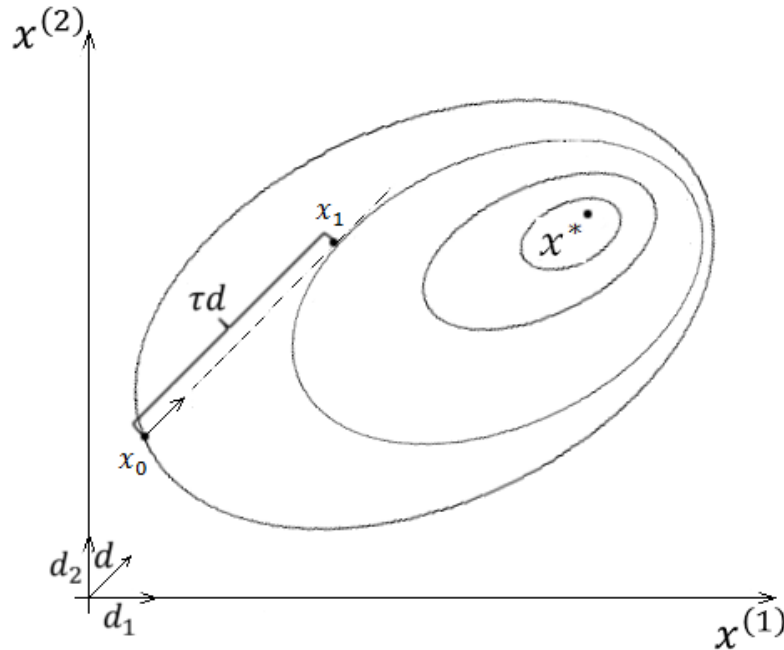
$$x_0, x_1, \dots, x_n, \dots, x_N \approx x^*$$
$$F(x_0) > F(x_1) > \dots > F(x_n) > \dots > F(x_N) \approx F(x^*)$$

Wybór kierunku poszukiwań



- Kierunki bazowe i ich modyfikacje – metody bez gradientowe.
- Kierunki oparte na gradiencie funkcji – metody gradientowe.
- Inne

Optymalizacja w kierunku



x_0 - punkt początkowy
 x_1 - punkt końcowy
 d - kierunek
 τ - długość kroku w kierunku

$$\tau^* \rightarrow F(x_0 + \tau^* d) = \min_{\tau} F(x_0 + \tau d)$$

x_0, d - ustalone $F(x_0 + \tau d) \triangleq f(\tau)$

$f(\tau)$ - funkcja jednej zmiennej (długości kroku τ)

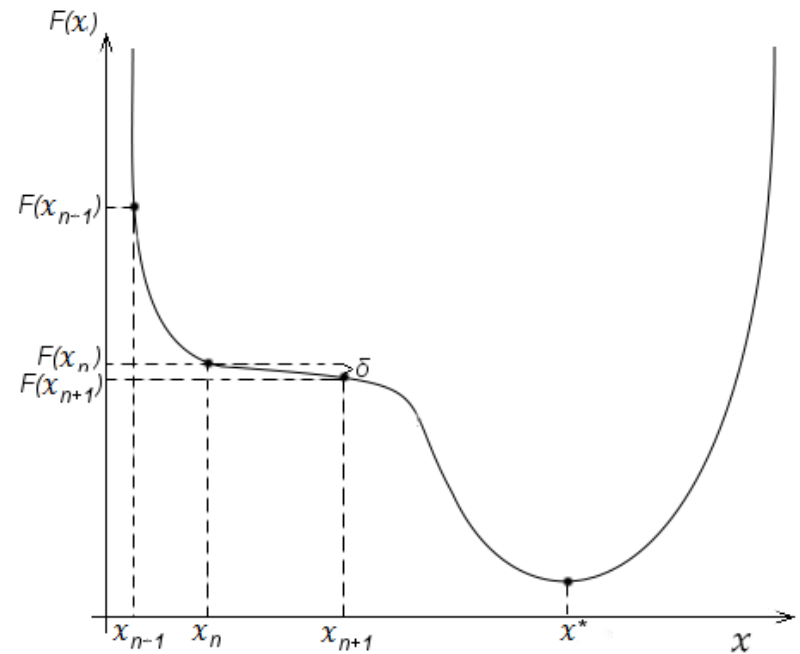
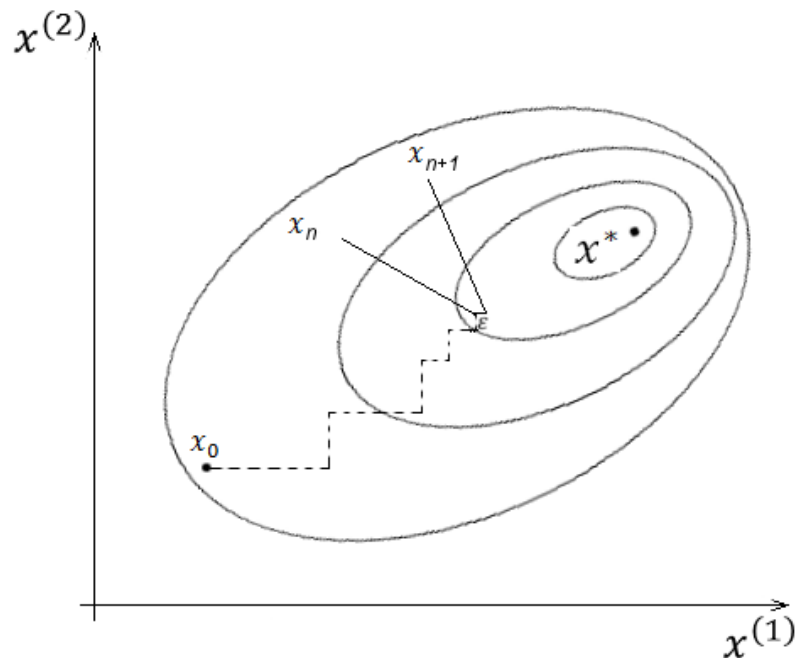
$$\tau^* \rightarrow f(\tau^*) = \min_{\tau} f(\tau)$$

optymalizacja w kierunku $\equiv$ optymalizacja funkcji jednej zmiennej

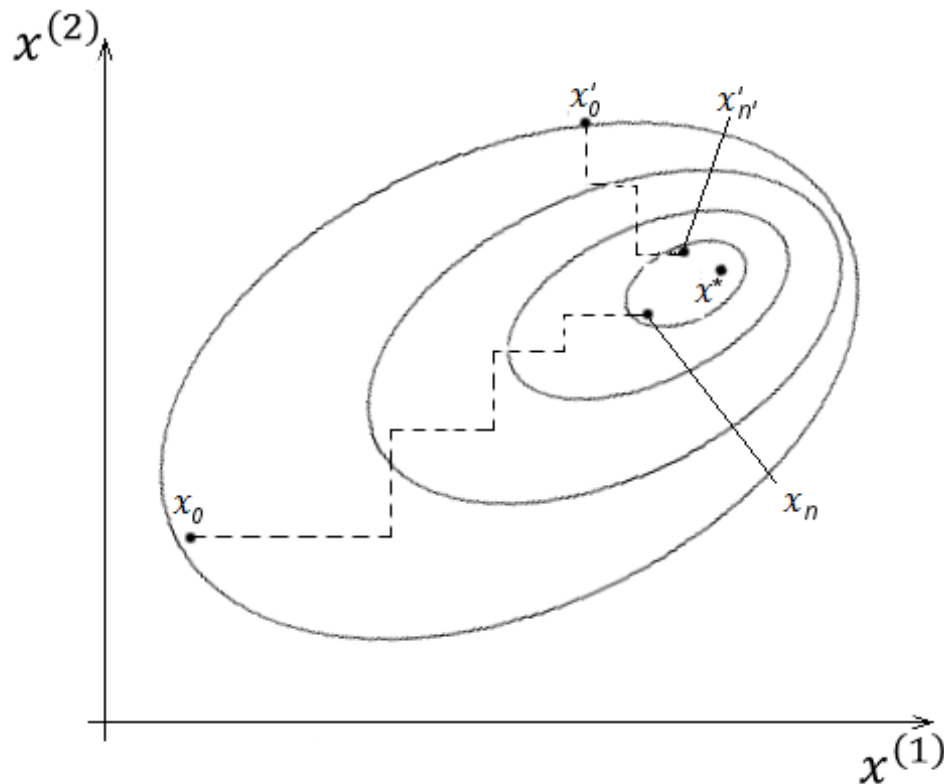
Warunki zatrzymania procedury

$$\|x_{n+1} - x_n\| < \varepsilon; \quad |F(x_{n+1}) - F(x_n)| < \delta; \quad \frac{|F(x_{n+1}) - F(x_n)|}{\|x_{n+1} - x_n\|} < \varrho$$

Uwaga!



Dobra rada



$$\|x_n - x'_{n'}\| < \varepsilon$$

x_0, x'_0 - różne punkty początkowe
 $x_n, x'_{n'}$ - odpowiednie punkty końcowe

∞ Metody optymalizacji w kierunku

- Podział równomierny
- Podział na połowę
- Złoty podział
- Aproksymacji kwadratowej
- Metoda pierwszej pochodnej
- Metoda znaku pochodnej
- Metoda Newtona
- Metoda Bolzano

∞ Bezgradientowe metody optymalizacji

- Hooka-Jeevesa (z krokiem dyskretnym i optymalnym)
- Rosenbrocka (z krokiem dyskretnym i optymalnym)
- Powella
- Neldera Meada

∞ Gradientowe metody optymalizacji

- Gradientu prostego
- Najszybszego spadku
- Newtona
- Gradientu sprzężonego
- Zmiennej metryki

∞ Metody optymalizacji z ograniczeniami

- Transformacji zmiennych
- Funkcji kary
 - Kary zewnętrznej (kary)
 - Kary wewnętrzne (bariery)
- Metoda kompleks
- Poszukiwania losowe

Dziękuję za uwagę

