

Optymalizacja systemów

<https://dlaczegologia.pl/>



Wykład 2. Analityczne metody optymalizacji

Podstawy matematyczne



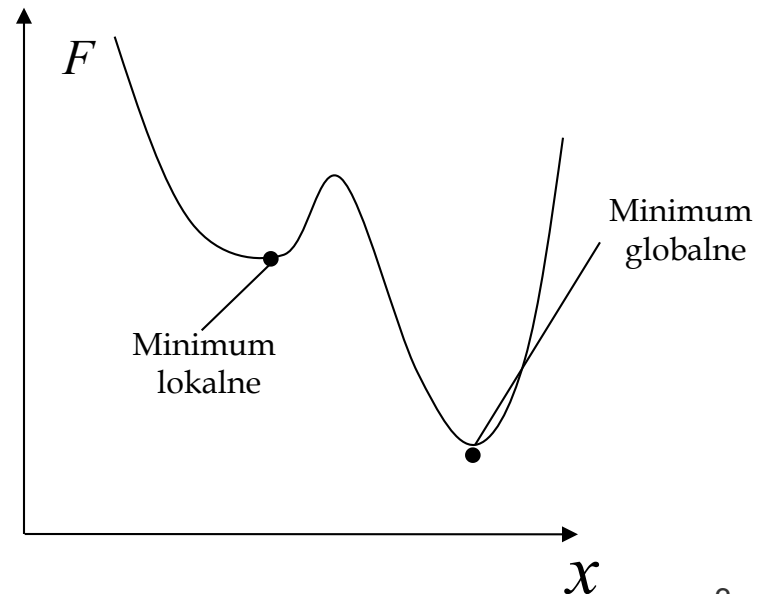
2.0 Przydatne definicje, własności

Podstawy matematyczne

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

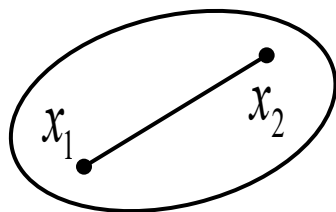
Minimum lokalne: $\forall_{\varepsilon > 0} \exists_{x \in O(x^*, \varepsilon)} F(x^*) < F(x)$

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

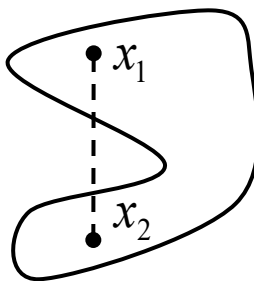


Podstawy matematyczne

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



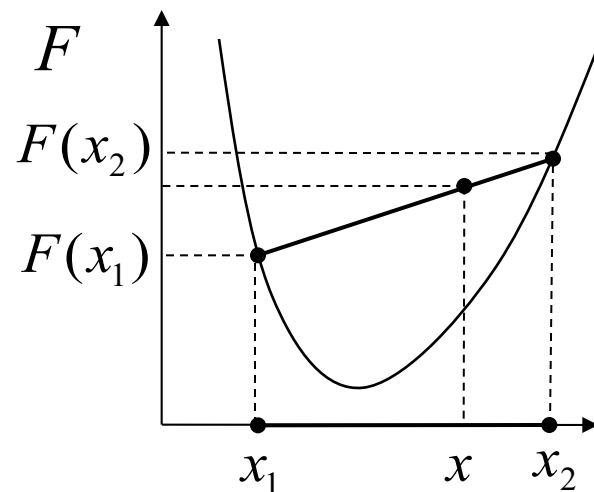
- zbiór wypukły



- 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:

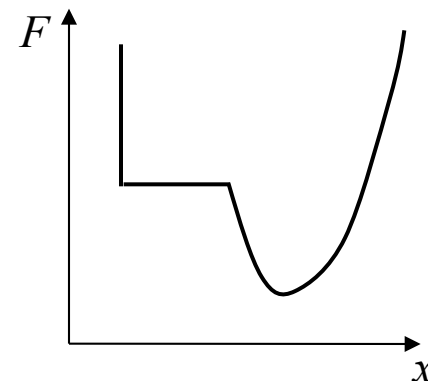
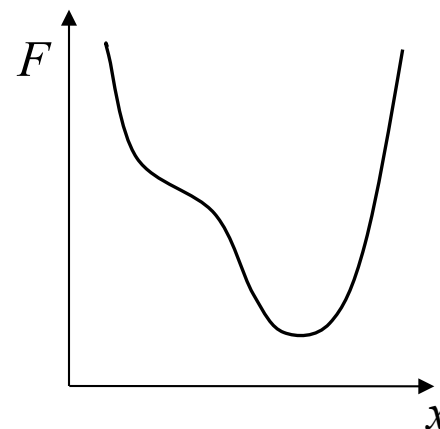
$$(x - x_0)^T [\nabla_x F(x_0)] \geq 0 \Rightarrow F(x) > F(x_0)$$

Zgodnie z rozwinięciem Taylor' a funkcji mamy:

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

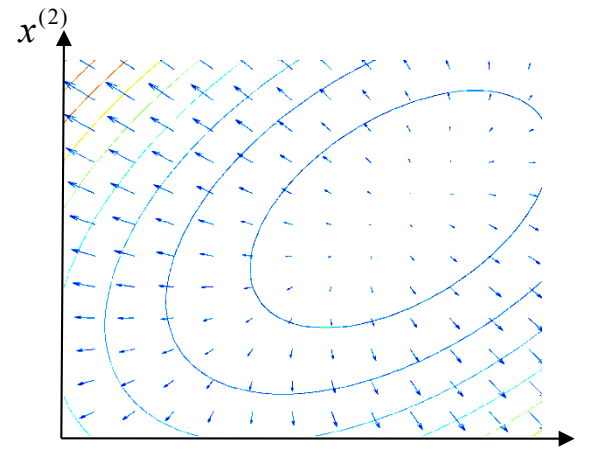
Funkcja quasi - wypukła:

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



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:

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

Określoność macierzy (Hessjanu)

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:

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

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

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

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

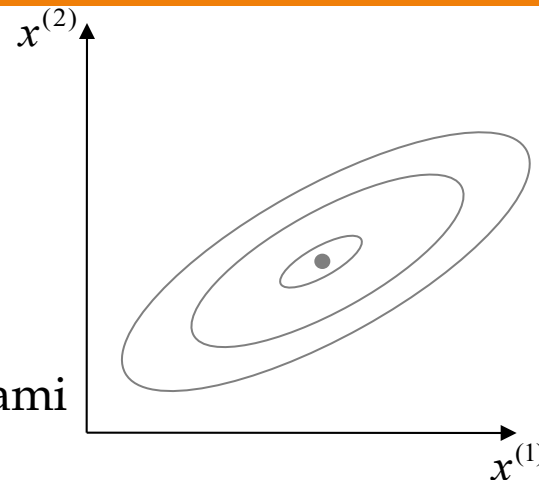
$$H(x) = \nabla_{xx}^2 F(x) = \begin{bmatrix} \frac{\partial^2 F}{\partial (x^{(1)})^2} & \frac{\partial^2 F}{\partial x^{(1)} \partial x^{(2)}} & \cdots & \frac{\partial^2 F}{\partial x^{(1)} \partial x^{(s)}} \\ \frac{\partial^2 F}{\partial x^{(2)} \partial x^{(1)}} & \frac{\partial^2 F}{\partial (x^{(2)})^2}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

Typowe zadania decyzyjne



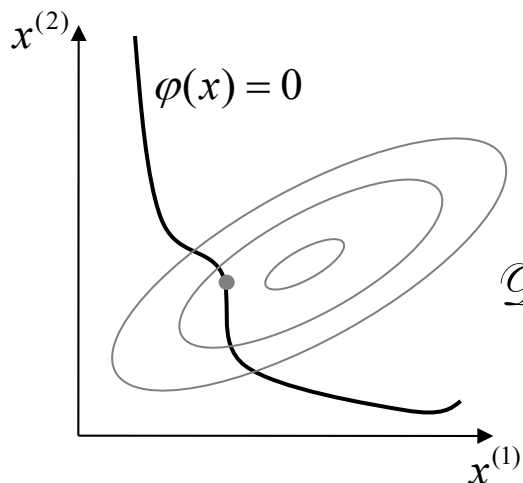
Typowe zadania decyzyjne

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



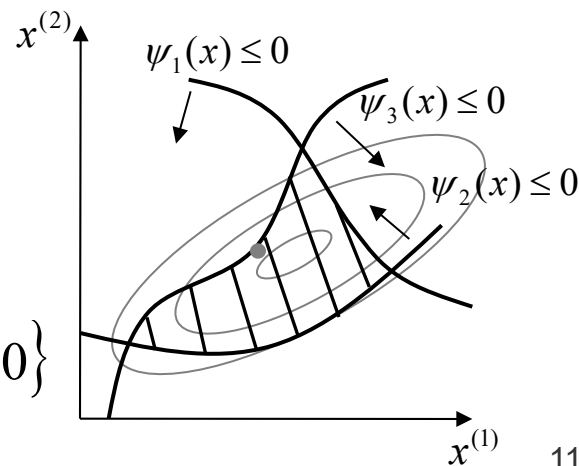
Zadanie decyzyjne z ograniczeniami równościowymi:

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



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



2.1 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:

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

$$0_S = \left\{ \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right\} S \text{ zer}$$

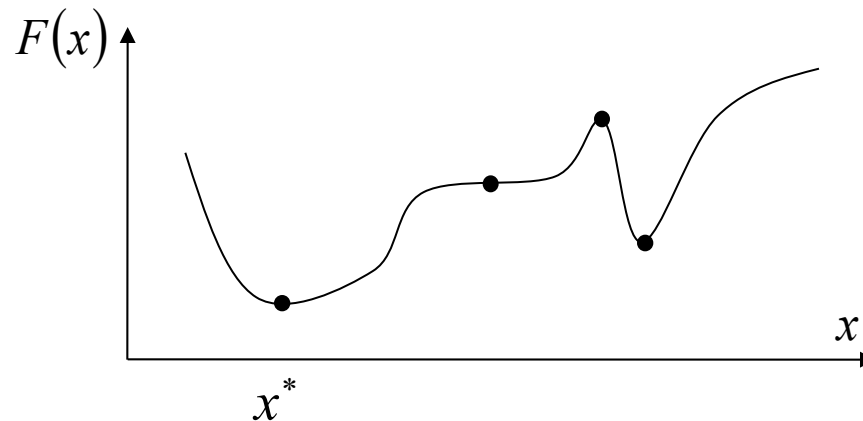
Jeżeli $F(x)$ jest funkcją pseudo - wypukłą, powyższe równanie jest warunkiem koniecznym i wystarczającym aby x^* było minimum globalnym

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

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

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

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

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

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

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

$$\text{☞ } \det H_{22} = \det \begin{bmatrix} 10 & 4 \\ 4 & 2 \end{bmatrix} = 20 - 16 = 4 > 0$$

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

Przykład 2.1.2

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

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

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

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

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

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

$$\ni \det H_{22} = \det \begin{bmatrix} 2\alpha & 4 \\ 4 & 2 \end{bmatrix} = 8\alpha - 16 > 0 \rightarrow \alpha > 2$$

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

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.

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

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

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

Przykład 2.1.3 c.d.

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

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

Przykład 2.1.3 c.d.

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

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

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

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

Przykład 2.1.3 c.d.

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

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

Przykład 2.1.3 c.d.

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

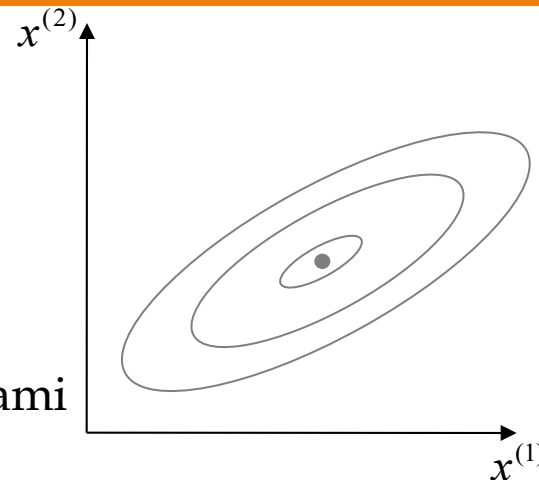
$$\ni x^* = -\frac{1}{2} A^{-1} b$$

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

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

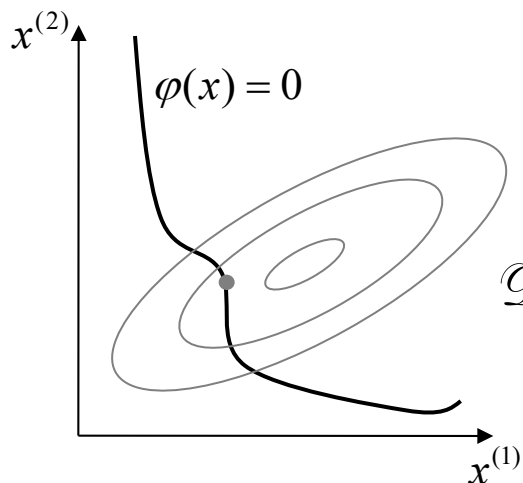
Typowe zadania decyzyjne

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



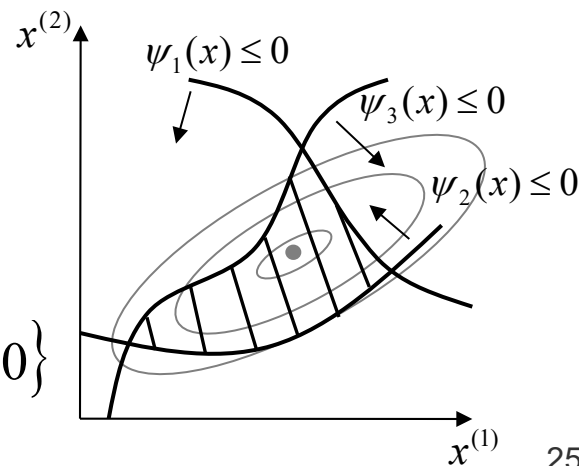
Zadanie decyzyjne z ograniczeniami równościowymi:

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



Zadanie decyzyjne z ograniczeniami nierównościami:

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



Zadanie optymalizacji z ograniczeniami równościowymi



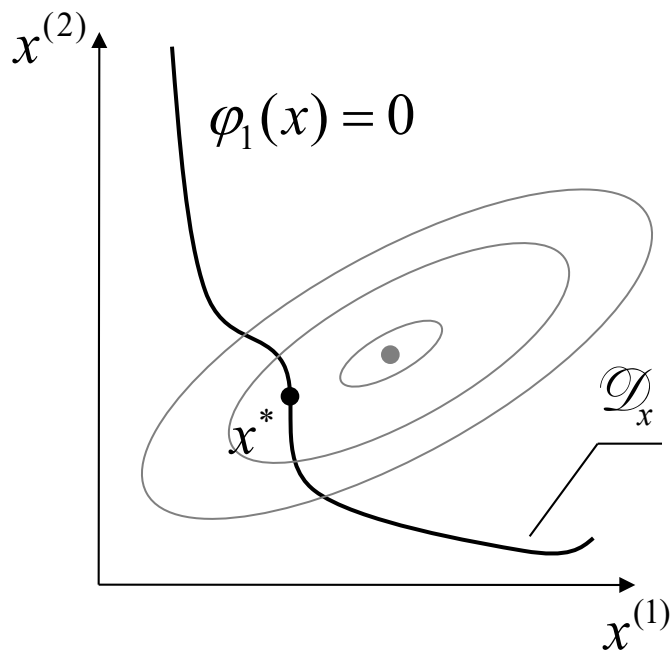
2.2 Metoda mnożników Lagrange'a

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:

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

$$\nabla_\lambda L(x, \lambda) \Big|_{x^*, \lambda^*} = 0_L \Leftrightarrow \text{rank } G(x) = \text{rank} \begin{bmatrix} G(x) & \vdots & -\nabla_x F(x) \end{bmatrix},$$

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

Zadanie optymalizacji z ograniczeniami równościowymi

Metoda mnożników Lagrange' a

Powyższy układ równań może mieć wiele rozwiązań

Warunki optymalności drugiego rzędu:

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

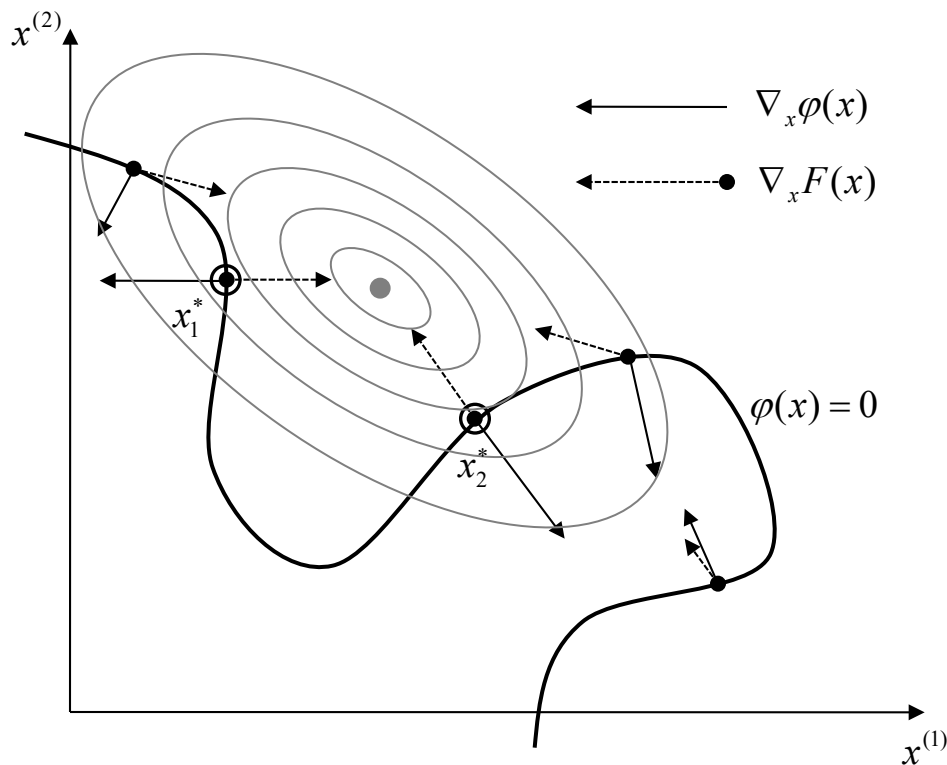
Jeżeli $H_L(x^*)$ jest dodatnio określona w punkcie x^*
to x^* jest minimum lokalnym

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



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

gdzie:

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

Warunek konieczny optymalności:

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

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

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

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

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

$$\infty \frac{\partial F(x_1, x_2)}{\partial x_1} + \frac{\partial F(x_1, x_2)}{\partial x_2} \left(-\frac{\frac{\partial \varphi(x_1, x_2)}{\partial x_1}}{\frac{\partial \varphi(x_1, x_2)}{\partial x_2}} \right) = 0$$

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

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

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

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

Wyjaśnienie warunków koniecznych

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

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

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

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

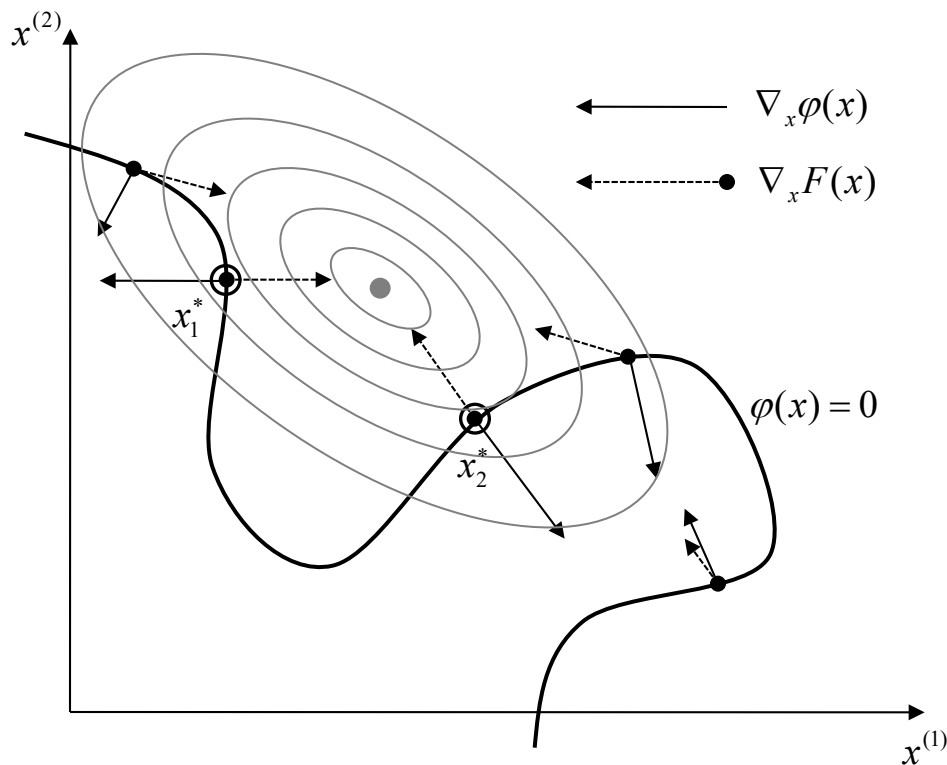
$\propto$ Ogólniej

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

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

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

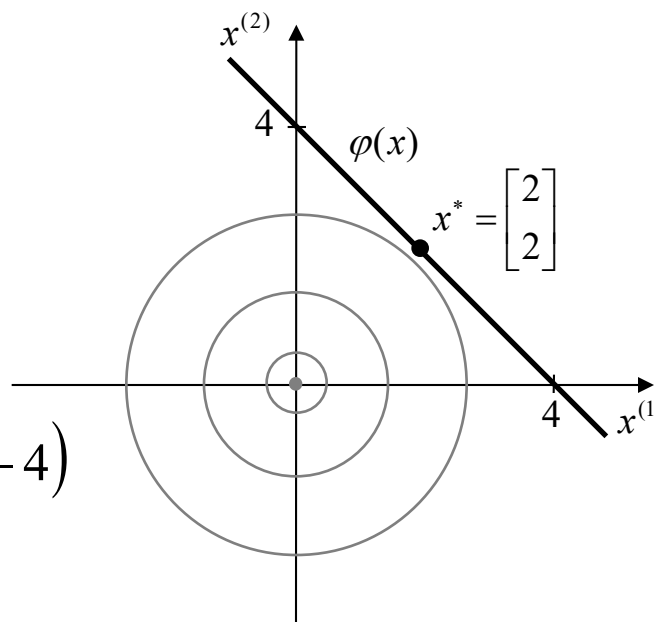
Metoda mnożników Lagrange' a

☞ Metoda Lagrange' a - Przykład 2.2.1

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

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

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



Przykład 2.2.1

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

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

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

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

$$\text{✂ } z \ (1) \rightarrow \left(-\frac{\lambda}{2}\right) + \left(-\frac{\lambda}{2}\right) - 4 = 0 \quad \text{czyli} \quad \lambda = -4$$

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

Przykład 2.2.1 c.d.

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

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

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

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

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

Zadanie optymalizacji z ograniczeniami równościowymi

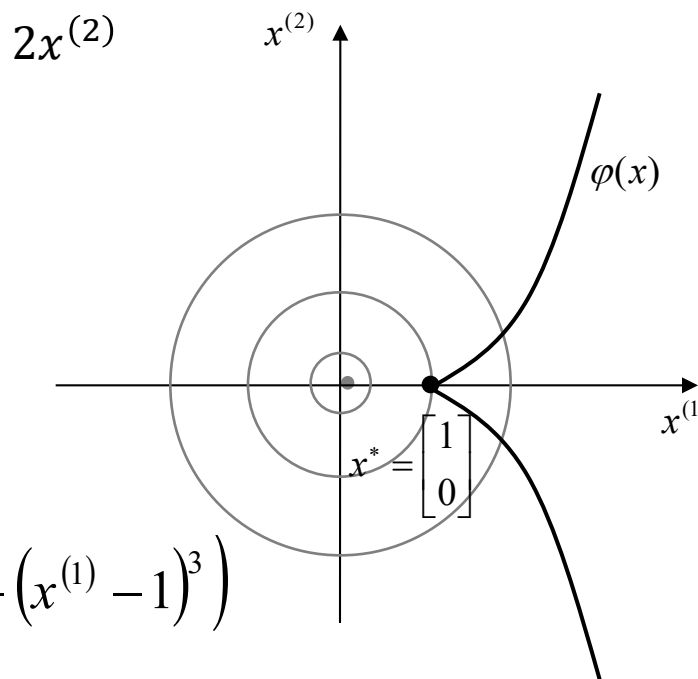
Metoda mnożników Lagrange' a

∞ Metoda współczynników Lagrange' a – Przykład 2.2.2
rozwiązanie nieregularne

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

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

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



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

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

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

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

Rozwiązanie powyższego równania liniowego istnieje jeżeli i jest jednoznaczne gdy:

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

oraz niejednoznaczne gdy:

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

⌘ $Ax = b$ gdzie A – macierz $(S \times S)$, b – S wymiarowy wektor

⌘ $\text{rank} A = \text{rank}[A : b] = S$ rozwiązanie jednoznaczne

⌘ $\text{rank} A = \text{rank}[A : b] < S$ rozwiązanie niejednoznaczne

⌘ $\text{rank} A \neq \text{rank}[A : b]$ układ równań sprzeczny

Zadanie optymalizacji z ograniczeniami równościowymi

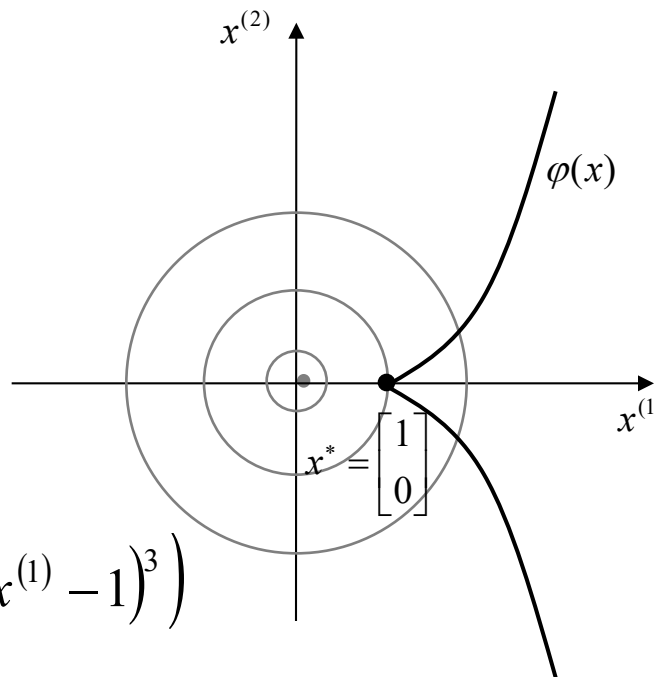
Metoda mnożników Lagrange' a

↻ Wracamy do przykładu 2.2.2

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

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

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



Przykład 2.2.2

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

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

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

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

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

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

☞ Sprzeczność ??

↻ Wracamy do przykładu 2.2.2

$$F(x) = \left(x^{(1)}\right)^2 + \left(x^{(2)}\right)^2 \quad \varphi(x) = \left(x^{(2)}\right)^2 - \left(x^{(1)} - 1\right)^3 = 0$$

$$G(x) = [\nabla_x \varphi(x)] = \begin{bmatrix} 3(x^{(1)} - 1)^2 \\ 2x^{(2)} \end{bmatrix} \Big|_{x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$[G(x) : -\nabla_x F(x)] = [\nabla_x \varphi(x) : -\nabla_x F(x)] = \begin{bmatrix} 3(x^{(1)} - 1)^2 & 2x^{(1)} \\ 2x^{(2)} & 2x^{(2)} \end{bmatrix} \Big|_{x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}} = \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}$$

$$\text{rzęd}[G(x) : -\nabla_x F(x)] = \text{rzęd} \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} = 1, \quad \text{rzęd} G(x) = \text{rzęd} \begin{bmatrix} 0 \\ 0 \end{bmatrix} = 0$$

$$\text{rzęd} G(x) \neq \text{rzęd}[G(x) : -\nabla_x F(x)] \rightarrow \text{punkt } x^* = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \text{rozwiązanie nieregularne}$$

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

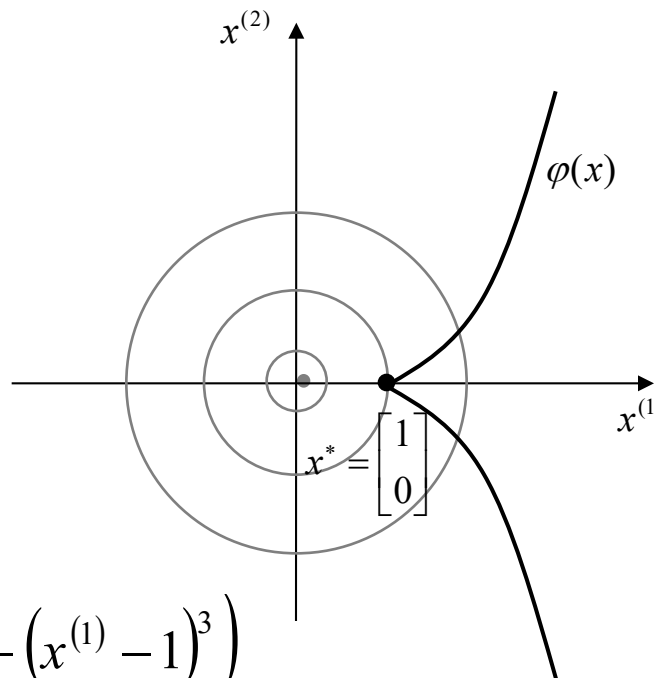
Metoda mnożników Lagrange' a

Uogólniona metoda współczynników Lagrange' a

Przykład 2.2.2

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

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



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

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

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

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

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

Macierz H_L jest dodatnio pół określona

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

Dziękuję za uwagę

