

Optymalizacja systemów

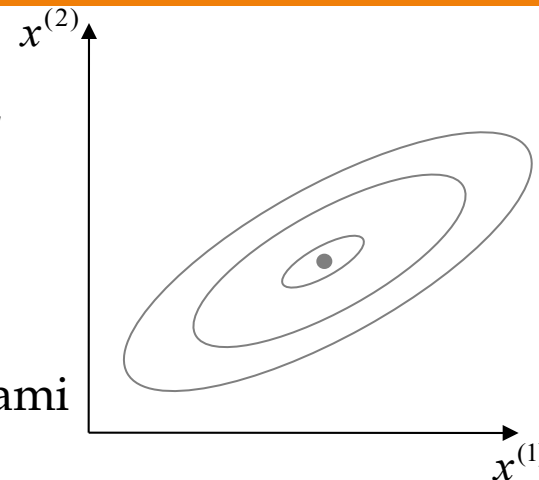
<https://dlaczegologia.pl/>



Wykład 2. Analityczne metody optymalizacji c.d.

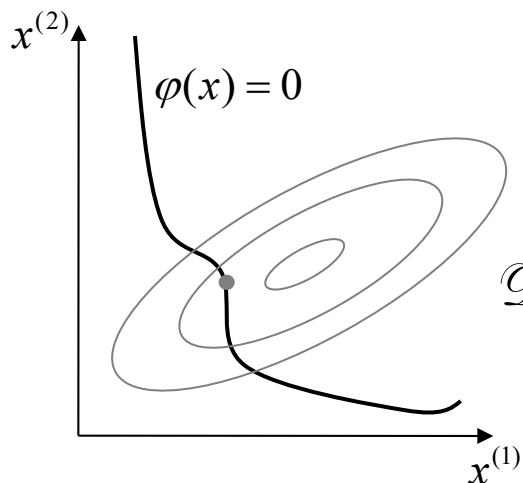
Typowe zadania decyzyjne

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



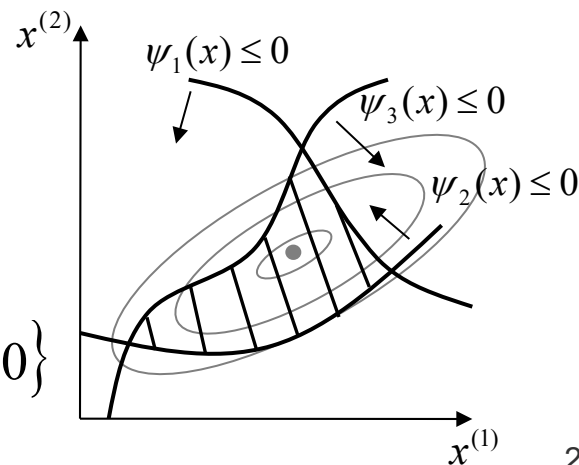
Zadanie decyzyjne z ograniczeniami równościowymi:

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



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 nierównościowymi



2.3 Metoda Kuhna-Tuckera

Zadanie optymalizacji z ograniczeniami nierównościami

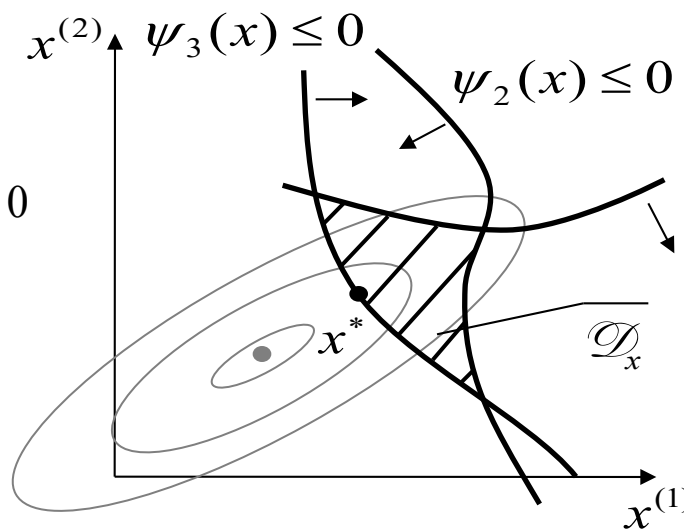
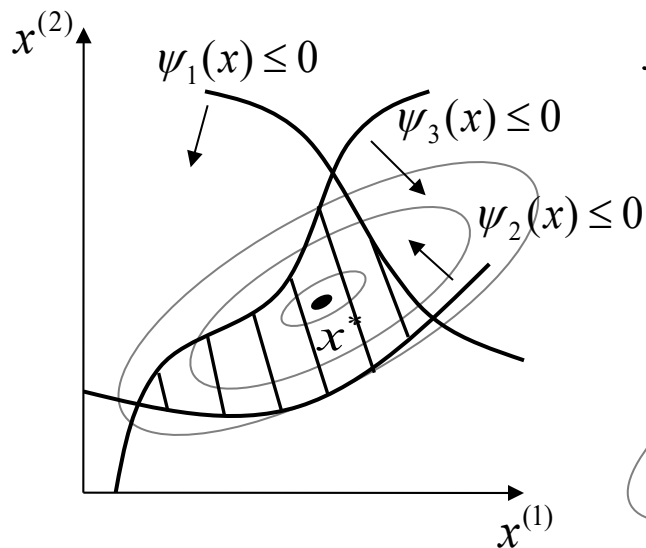
Metoda Kuhna – Tucker'a

Zadanie optymalizacji:

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

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

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



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

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

Nie ma takiej
nierówności !!!

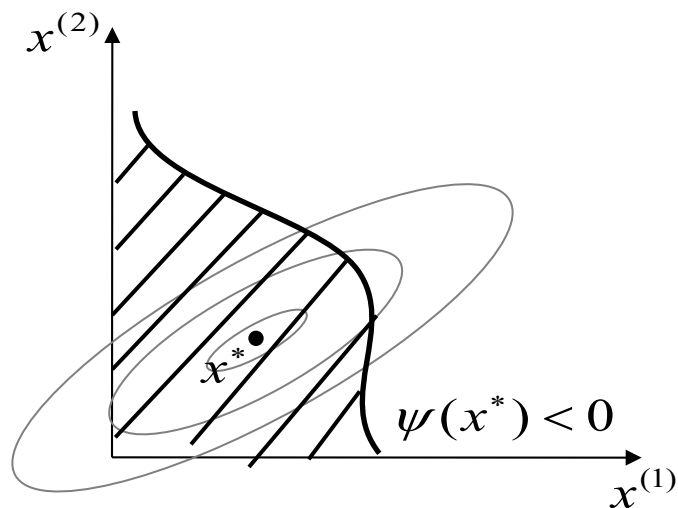
Tak się umawiamy

Zadanie optymalizacji z ograniczeniami nierównościowymi

Metoda Kuhna – Tucker'a

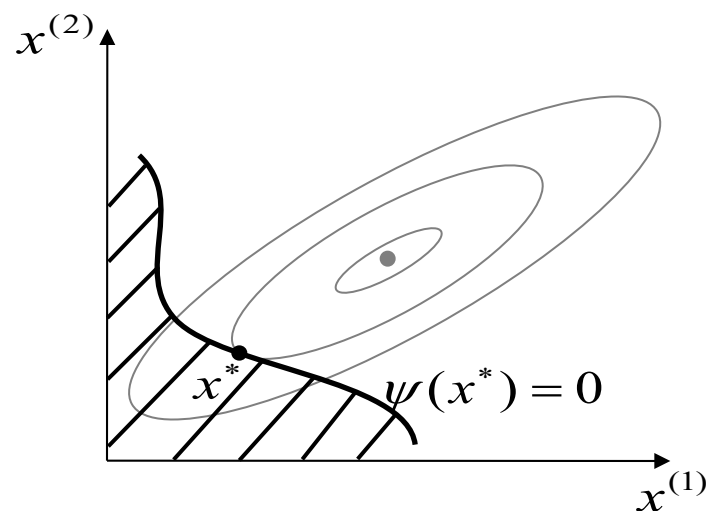
Ograniczenia nieaktywne

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



Ograniczenia aktywne

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



Zadanie optymalizacji z ograniczeniami nierównościami

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

Warunki konieczne:

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

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

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

$$\mu \geq 0_M$$

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

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

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

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

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

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

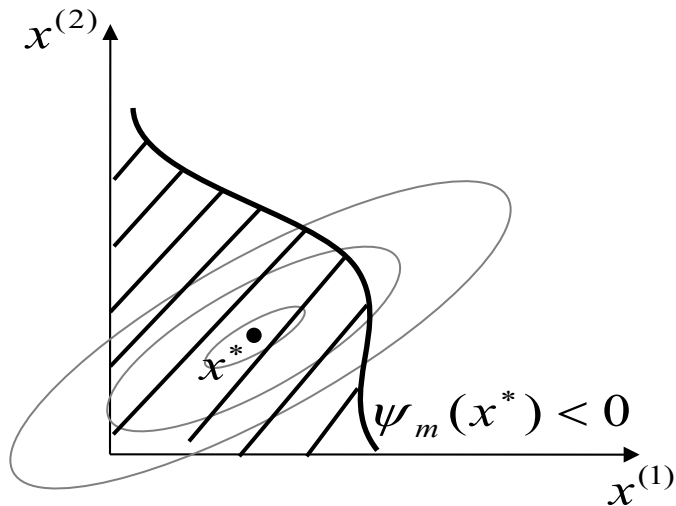
$$\mu_m^* \geq 0 \quad m = 1, 2, \dots, M$$

Zadanie optymalizacji z ograniczeniami nierównościowymi

Metoda Kuhna – Tucker'a

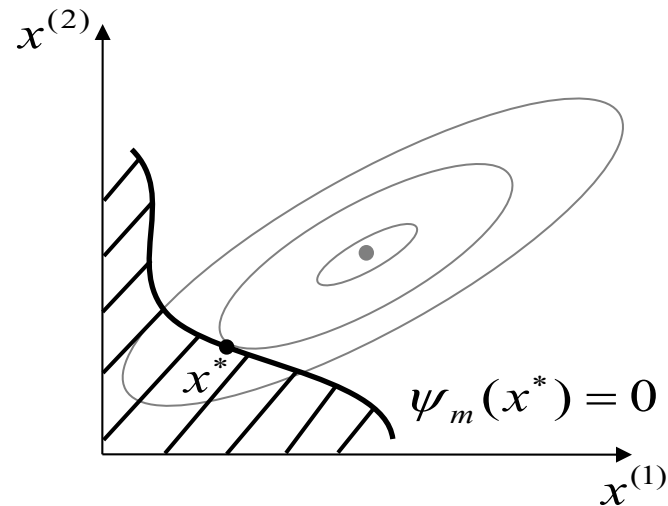
Ograniczenia nieaktywne

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



Ograniczenie aktywne

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



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

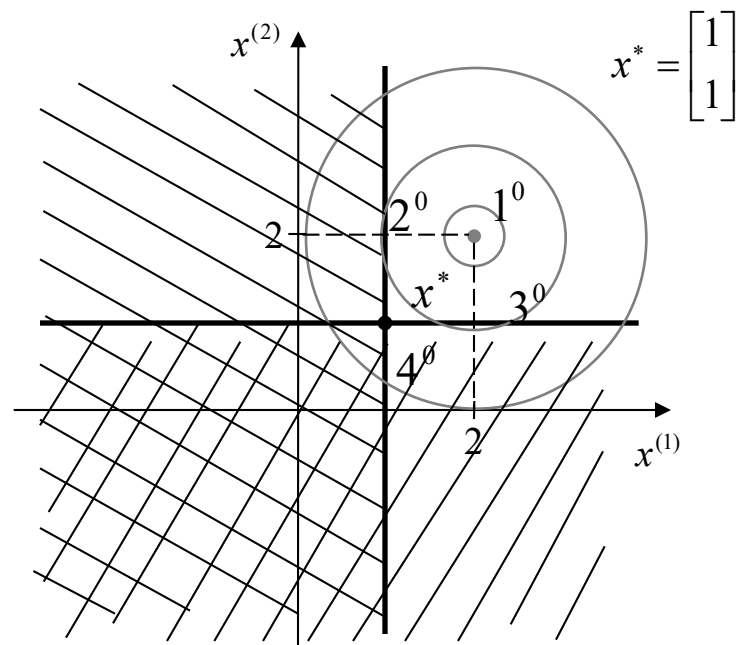
Metoda Kuhna – Tucker'a

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.

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

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

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

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

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

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

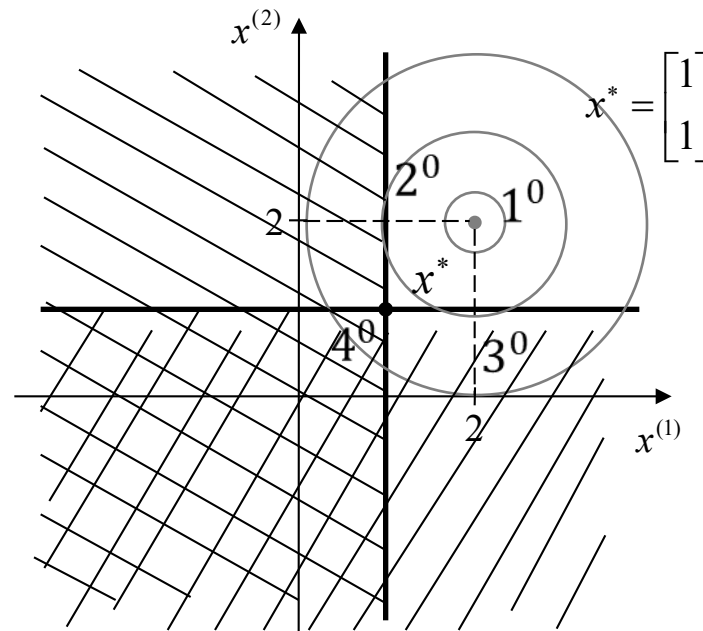
z (1) $\rightarrow 2(1 - 2) + \mu_1 = 0 \rightarrow \mu_1 = 2$

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

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

Przykład 1. c.d.



Zadanie optymalizacji z ograniczeniami nierównościami

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

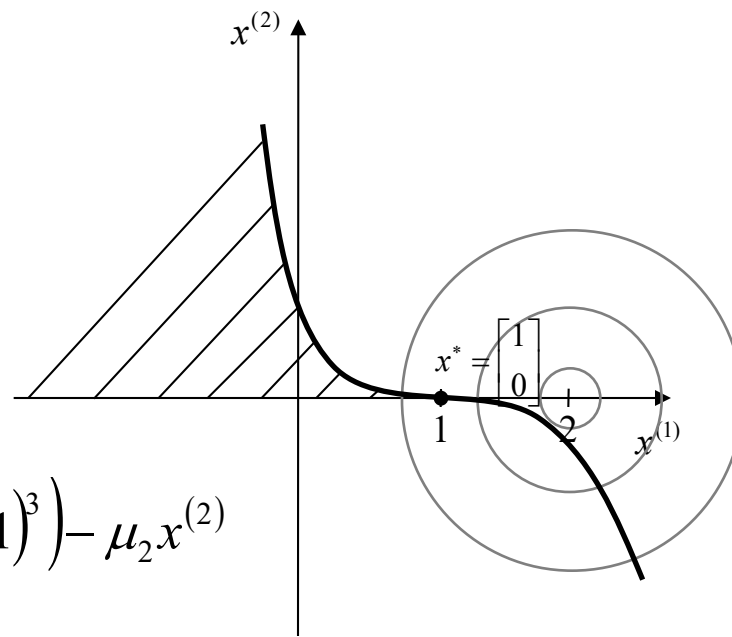
Przykład 2 – rozwiązanie nieregularne

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

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

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

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



Przykład 2.

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

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

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

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

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

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

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

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

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

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

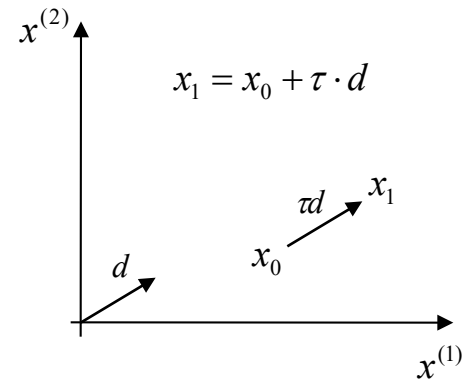
$$\text{☞ } \mu = \begin{cases} \mu_1 \geq 0 & (7) \\ \mu_2 \geq 0 & (8) \end{cases}$$

Rozwiązanie nieregularne

Zadanie optymalizacji z ograniczeniami nierównościowymi

Metoda Kuhna – Tucker'a

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

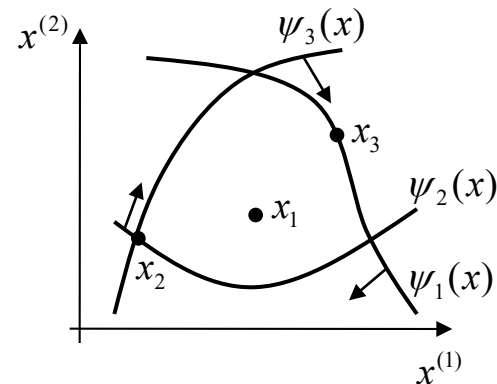
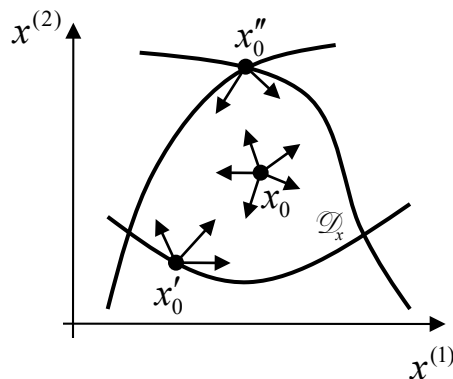


Zbiór kierunków dopuszczalnych

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

Zbiór ograniczeń aktywnych

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



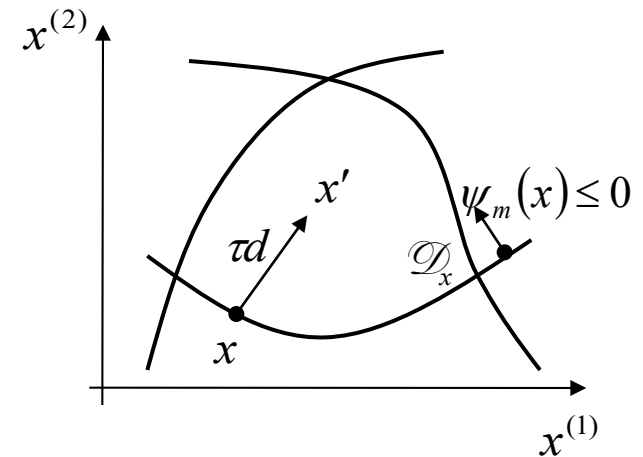
$$\begin{aligned} I(x_1) &= \emptyset \\ I(x_2) &= \{2, 3\} \\ I(x_3) &= \{1\} \end{aligned}$$

Zadanie optymalizacji z ograniczeniami nierównościowymi

Metoda Kuhna – Tucker'a

Jaki warunek musi spełniać kierunek dopuszczalny na ograniczeniu aktywnym?

$$\begin{aligned}\forall m \in I(x) \\ \text{tj.}: \psi_m(x) = 0 \\ x' = x + \tau d \in \mathcal{D}_x, \quad \tau > 0 \\ \psi_m(x') \leq 0\end{aligned}$$



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

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

$$d^T \nabla_x \psi_m(x) \leq 0 \quad - \text{warunek analityczny}$$

Zadanie optymalizacji z ograniczeniami nierównościovymi Metoda Kuhna – Tucker’a

Przykład 1

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

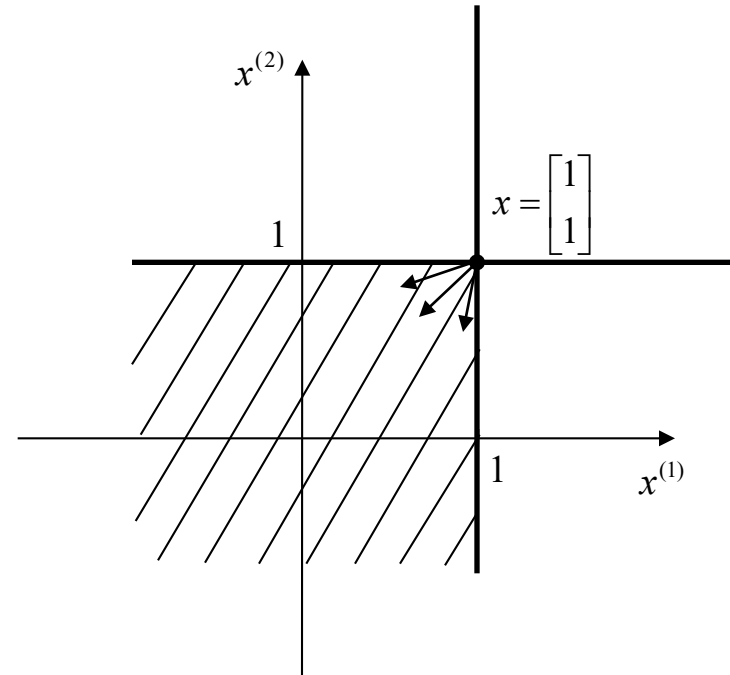
$$\psi_2(x) = x^{(2)} - 1 \leq 0$$

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. $\mathcal{D}(x) \neq D(x)$ Może to prowadzić do rozwiązania nieregularnego

$$\mathcal{D}(x) = \{d \in \mathbb{R}^s : \forall_m \in I(x), d^T \nabla_x \psi_m(x) \leq 0\} \quad D(x) = \{d \in \mathbb{R}^s : \exists \tau \quad x + \tau d \in \mathcal{D}_x\}$$

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

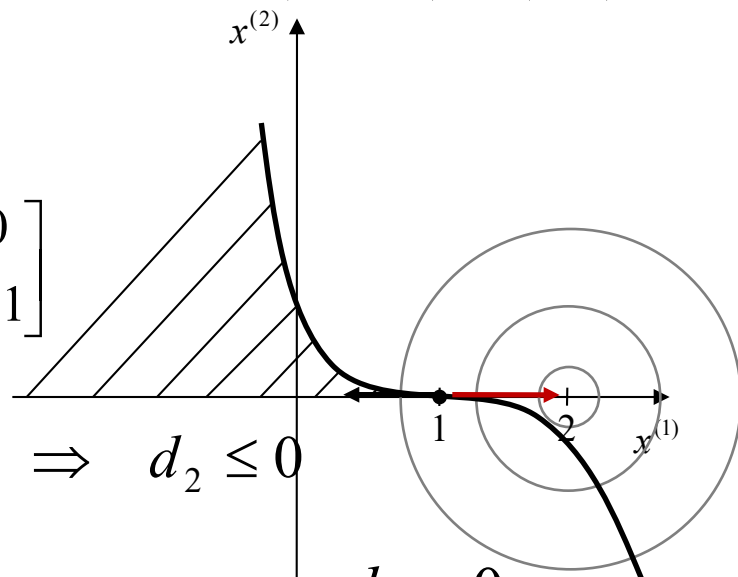
W punkcie $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $I\left(x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \{1, 2\}$

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

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

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

$d_2 = 0$
 $d_1 - \text{dowolne}$



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

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

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

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

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

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

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:

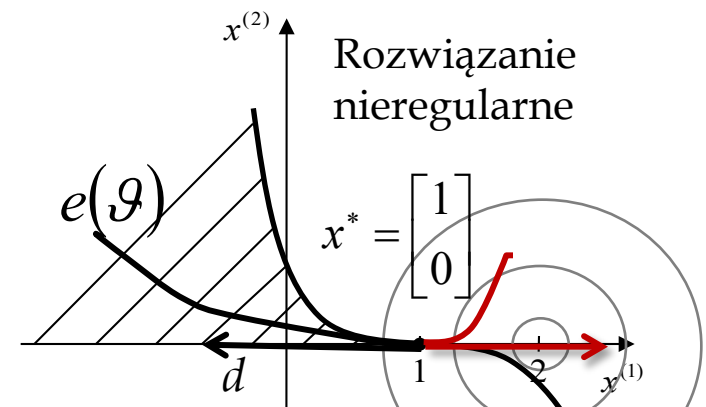
$$\forall m \in I(x^*) \quad \nabla_x \psi_m(x^*) \Big|_{x=x^*} \text{ są liniowo niezależne}$$

4. 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 nierównościami Metoda Kuhna – Tucker'a

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

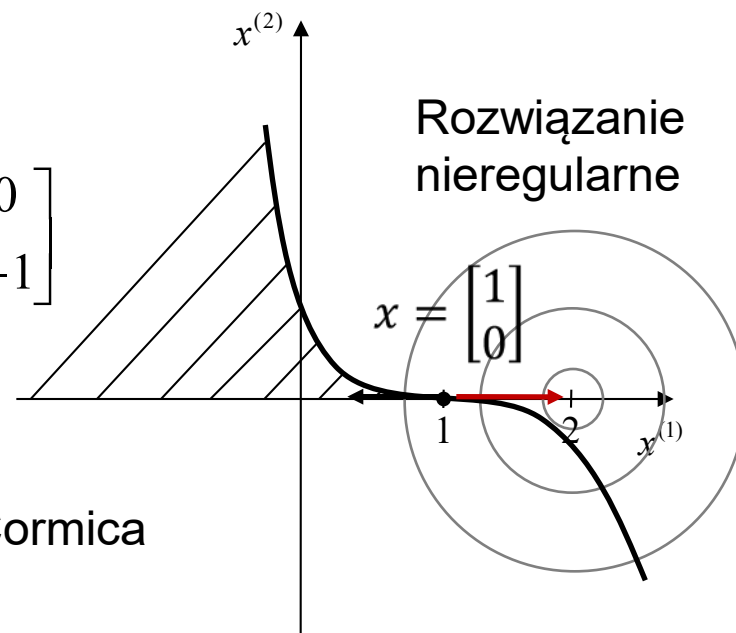
W punkcie $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $I\left(x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \{1, 2\}$

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 nierównościovymi Metoda Kuhna – Tucker’a

Warunki konieczna i wystarczające:

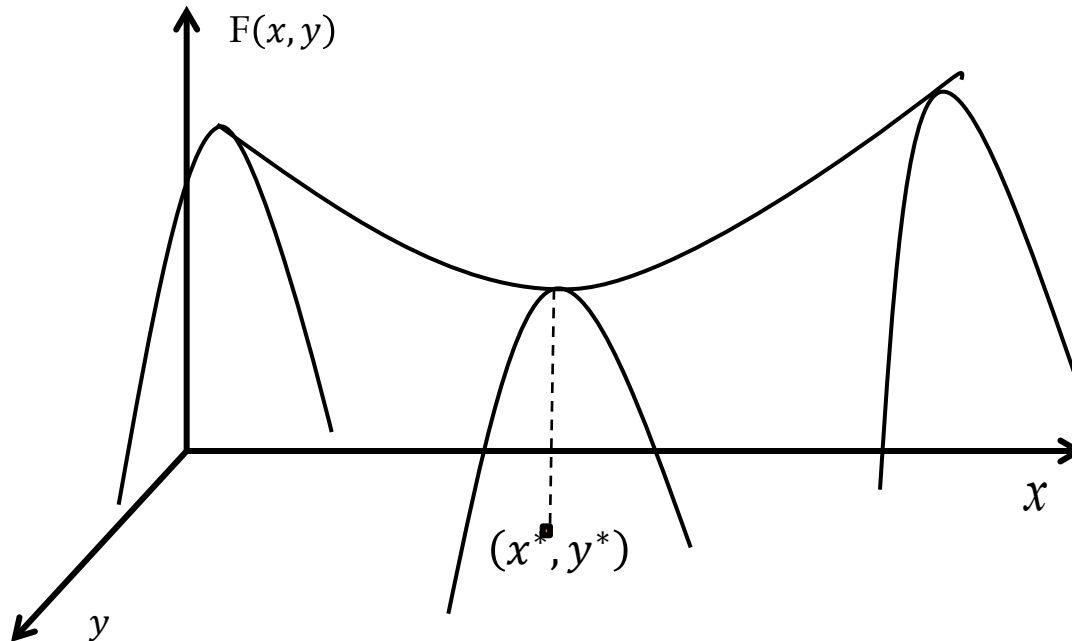
Jeżeli funkcje $F(x), \psi_1(x), \psi_2(x), \dots, \psi_M(x)$ są ciągłe i różniczkowalne oraz funkcja $F(x)$ jest funkcją pseudo – wypukłą, a ograniczenia $\psi_1(x), \psi_2(x), \dots, \psi_M(x)$ są funkcjami quasi – wypukłymi to układ równań i nierówności:

$$\begin{aligned}\nabla_x L(x, \mu) \Big|_{x^*, \mu^*} &= 0_S \\ \mu^T \nabla_\mu L(x, \mu) \Big|_{x^*, \mu^*} &= 0 \\ \nabla_\mu L(x, \mu) \Big|_{x^*, \mu^*} &\leq 0_M \\ \mu^* &\geq 0_M\end{aligned}$$

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

Zadanie optymalizacji z ograniczeniami nierównościovymi - Punkt siodłowy

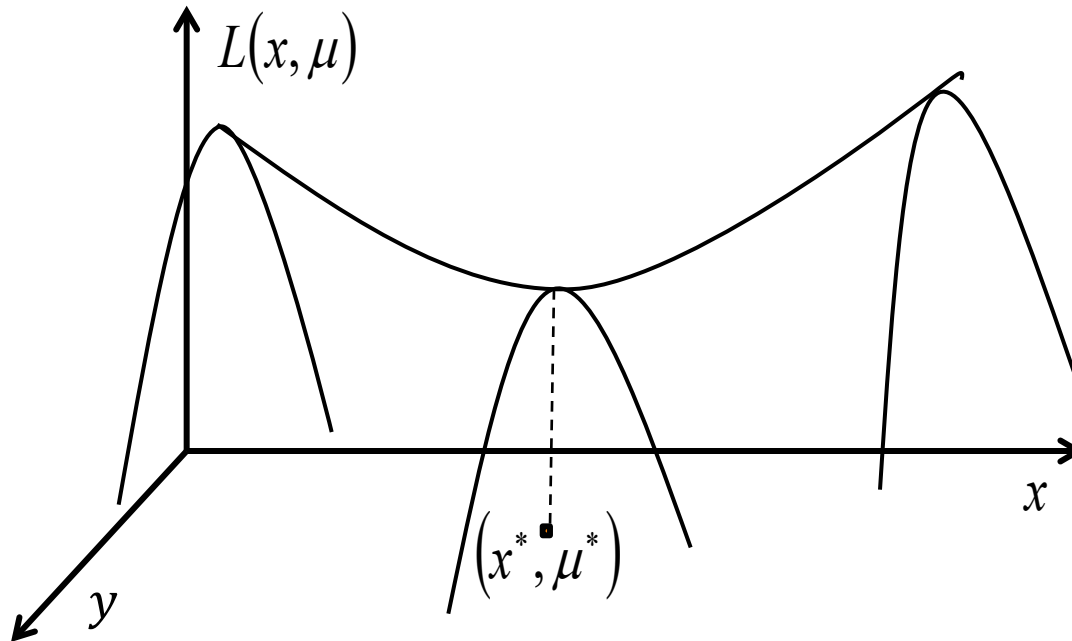
Punkt siodłowy (x^*, y^*)



$$\begin{aligned} F(x^*, y^*) &\leq F(x, y^*) \quad \forall x \in \mathbb{R}^S \\ F(x^*, \mu) &\leq F(x^*, y^*) \quad \forall y \in \mathbb{R}^L \\ F(x^*, y^*) &= \min_{x \in D(x)} \max_{y \geq 0_M} F(x, y) \end{aligned}$$

Zadanie optymalizacji z ograniczeniami nierównościami - Punkt siodłowy

Punkt siodłowy (x^*, μ^*)



$$\begin{aligned} L(x^*, \mu^*) &\leq L(x, \mu^*) \quad \forall x \in \mathcal{D}(x) \subseteq \mathcal{R}^s \\ L(x^*, \mu) &\leq L(x^*, \mu^*) \quad \forall \mu \geq 0_M \\ L(x^*, \mu^*) &= \min_{x \in \mathcal{D}(x)} \max_{\mu \geq 0_M} L(x, \mu) \end{aligned}$$

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

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

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

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

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

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

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

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

$$x^* \geq 0_S$$

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

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

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

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

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

Warunki Kuhna-Tuckera

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

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

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

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

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

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

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

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

Dziękuję za uwagę

