Podstawy optymalnego projektowania konstrukcji

Metody analityczne

Marcin Rodak
marcin.rodak@put.poznan.pl

Zakład Wytrzymałości i Konstrukcji
Instytut Mechaniki Stosowanej
Politechnika Poznańska

wrzesień 2015

Spis treści

1 Wprowadzenie

2 Optymalizacja funkcji jednej zmiennej bez ograniczeń
- Sformułowanie problemu
- Warunek konieczny istnienia ekstremum funkcji
- Warunek dostateczny istnienia ekstremum funkcji

3 Optymalizacja funkcji wielu zmiennych bez ograniczeń
- Sformułowanie problemu
- Warunek konieczny
- Warunek dostateczny

4 Optymalizacja funkcji wielu zmiennych z ograniczeniami równościowymi
- Sformułowanie problemu
- Bezpośrednie podstawienie
- Tytył
- Metoda wyznaczników Lagrange'a

5 Optymalizacja funkcji wielu zmiennych z ograniczeniami nierównościowymi
- Sformułowanie problemu
- Warunki konieczne

Optymalizacja to akt, czynność, której celem jest otrzymanie najlepszych rezulatatów ze względu na wybrane kryterium w danych warunkach.

Problem optymalizacyjny można przedstawić matematycznie następująco

$$\min_{\boldsymbol{x}\in\mathbb{R}^n} f(\boldsymbol{x}),$$

pod warunkiem, że

$$\begin{aligned} g_i(\boldsymbol{x}) &\leqslant 0 \qquad i = 1, 2, \ldots, m_n, \\ h_j(\boldsymbol{x}) &\leqslant 0 \qquad j = 1, 2, \ldots, m_e, \end{aligned}$$

gdzie $\boldsymbol{x} = [x_1, x_2, \ldots, x_n]^{\mathrm{T}}$.

Przy poszukiwaniu optimum funkcji ciągłej i różniczkowalnej przydatne są metody analityczne (częściowo poznane na kursie analizy matematycznej).

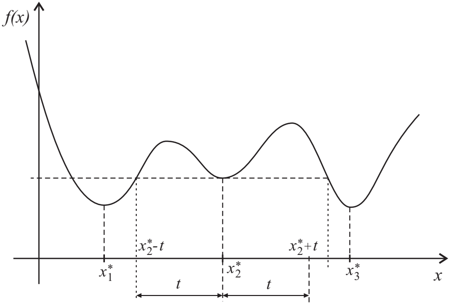
Metody analityczne nie mają praktycznego zastosowania w przypadku problemów bardziej złożonych, natomiast stanowią podstawę teoretyczną dla metod komputerowych przedstawionych w dalszej części wykładów.

Sformułowanie problemu optymalizacyjnego

Sformułowanie problemu optymalizacyjnego

$$\min_{x\in[a,b]} f(x) \tag{2.1}$$

gdzie $x \in \mathbb{R}$ oraz $f : \mathbb{R} \to \mathbb{R}$



Rysunek: Funkcja jednej zmiennej

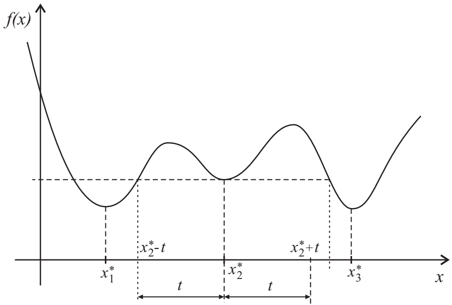
Sformułowanie problemu optymalizacyjnego

Definicja (minimum/maksimum lokalne funkcji jednej zmiennej)

Funkcja jednej zmiennej $f : \mathbb{R} \to \mathbb{R}$ ma w punkcie $x = x^*$

- minimum lokalne, gdy

$$\exists_{t>0} \forall_{|h|<t} f(x^*) \leqslant f(x^* + h)$$



Rysunek: Funkcja jednej zmiennej

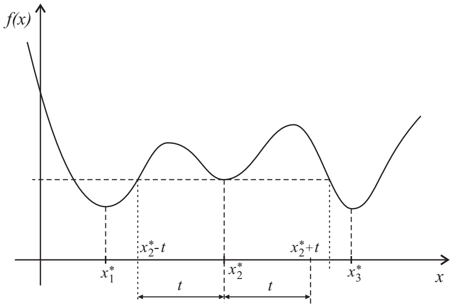
Sformułowanie problemu optymalizacyjnego

Definicja (minimum/maksimum lokalne funkcji jednej zmiennej)

Funkcja jednej zmiennej $f : \mathbb{R} \to \mathbb{R}$ ma w punkcie $x = x^*$

- maksimum lokalne, gdy

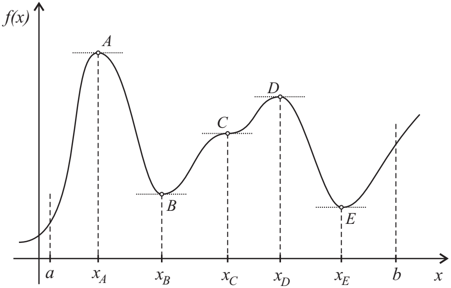
$$\exists_{t>0} \forall_{|h|<t} f(x^*) \geqslant f(x^* + h)$$



Rysunek: Funkcja jednej zmiennej

Twierdzenie (Warunek konieczny ekstremum funkcji)

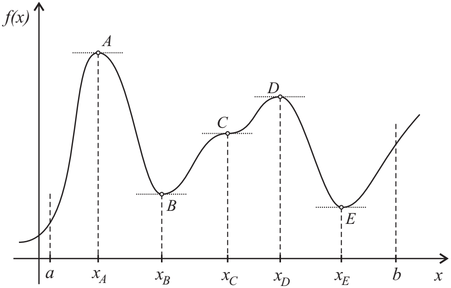
Funkcja jednej zmiennej $f : \mathbb{R} \to \mathbb{R}$ określona na przedziale $[a, b]$ osiąga ekstremum w punkcie $x = x^*$ takim, że $a < x^* < b$ oraz pochodna $f'(x) = \frac{\mathrm{d}f(x)}{\mathrm{d}x}$ istnieje w punkcie $x = x^*$ to $f'(x^*) = 0$.



Rysunek: Funkcja jednej zmiennej

Uwagi

1. Punkt, w którym pochodna funkcji ciągłej i różniczkowalnej równa się zero nazywamy *punktem stacjonarnym*.
2. Punkt stacjonarny, w którym funkcja jednej zmiennej nie osiąga ekstremum nazywamy *punktem przegięcia*.



Rysunek: Funkcja jednej zmiennej

Ekstremum funkcji – warunek konieczny istnienia – dowód

Dowód

Pochodną funkcji oblicza się następująco

$$f'(x^*) = \lim_{h\to 0} \frac{f(x^*+h) - f(x^*)}{h} \tag{i}$$

Jeśli w punkcie x^* funkcja f osiąga minimum to

$$f(x^*+h) - f(x^*) \geqslant 0$$

dla dostatecznie małej wartości $|h|$. Zatem

$$\frac{f(x^*+h) - f(x^*)}{h} \geqslant 0 \quad \text{dla} \quad h \geqslant 0 \tag{ii}$$

$$\frac{f(x^*+h) - f(x^*)}{h} \leqslant 0 \quad \text{dla} \quad h \leqslant 0 \tag{iii}$$

Z równań (ii) i (iii) wynika, że w wyrażeniu (i) granica prawostronna $f'(x^*) \geqslant 0$, natomiast lewostronna $f'(x^*) \leqslant 0$. Z istnienia pochodnej wnosimy, że

$$f'(x^{*-}) = f'(x^{*+}) = f'(x^*) = 0.$$

Podobne rozumowanie możemy przeprowadzić przy założeniu, że w $x = x^*$ funkcja osiąga maksimum lokalne. □

Warunek konieczny istnienia ekstremum funkcji – uwagi

Uwagi

- Twierdzenie (**??**) nie określa co dzieje się w przypadku, gdy funkcja $f(x)$ w punkcie $x = x^*$ nie posiada pochodnej, np.

$$\lim_{h \to 0^-} \frac{f(x^* + h) - f(x^*)}{h} = m^- < 0$$
$$\lim_{h \to 0^+} \frac{f(x^* + h) - f(x^*)}{h} = m^+ > 0$$

- Twierdzenie (**??**) nie mówi co dzieje się na krańcach przedziału $[a, b]$.
- Warunek $f'(x^*) = 0$ nie jest warunkiem dostatecznym ekstremum lokalnego funkcji $f(x)$, np. dla funkcji $f(x) = x^3$ w punkcie $x = 0$ pochodna zeruje się natomiast funkcja ta nie osiaga w tym punkcie ekstremum lokalnego.

Warunek dostateczny istnienia ekstremum funkcji – twierdzenie

Twierdzenie (Warunek dostateczny istnienia ekstremum funkcji)

Niech $f'(x^*) = f''(x^*) = f'''(x^*) = \ldots = f^{(n-1)}(x^*) = 0$ oraz $f^{(n)}(x^*) \neq 0$ wtedy funkcja $f(x)$ ma w punkcie $x = x^*$

1. minimum lokalne, gdy $f^{(n)}(x^*) > 0$ i n jest liczbą parzystą,
2. maksimum lokalne, gdy $f^{(n)}(x^*) < 0$ i n jest liczbą parzystą,
3. punkt przegięcia, gdy n jest liczbą nieparzystą.

Warunek dostateczny istnienia ekstremum funkcji – dowód

Dowód

Korzystając z szeregu Taylora z resztą otrzymujemy równość

$$f(x^*+h)=f(x^*)+\frac{h}{1!}f'(x^*)+\frac{h^2}{2!}f''(x^*)+\frac{h^3}{3!}f'''(x^*)+\ldots+$$
$$+\frac{h^{n-1}}{(n-1)!}f^{(n-1)}(x^*)+\frac{h^n}{n!}f^{(n)}(x^*+\theta h)\quad \text{dla}\quad 0\leqslant\theta\leqslant 1$$

Stąd, że $f'(x^*)=f''(x^*)=f'''(x^*)=\ldots=f^{(n-1)}(x^*)=0$ mamy

$$f(x^*+h)-f(x^*)=\frac{h^n}{n!}f^{(n)}(x^*+\theta h)\quad \text{dla}\quad 0<\theta<1$$

W przypadku, gdy $f^{(n)}(x^*)>0$ to dla dostatecznie małej wartości $|h|$ wartość n-tej pochodnej funkcji nie zmienia znaku (ciągłość funkcji), czyli $f^{(n)}(x^*+\theta h)>0$ dla $0<\theta<1$ dodatkowo, gdy n jest liczbą parzystą to wyrażenie po prawej stronie przyjmuje wartość dodatnią niezależnie od znaku h. Zatem $f(x^*+h)\geqslant f(x^*)$ dla dostatecznie małych wartości $|h|$, czyli w punkcie $x=x^*$ funkcja $f(x)$ osiąga minimum lokalne. Podobnie dowód przeprowadza się w przypadku, gdy $f^{(n)}(x^*)<0$ i n jest liczbą parzystą. Natomiast, gdy n jest liczbą nieparzystą wyrażenie $\frac{h^n}{n!}$ ma taki sam znak jak wartość h zatem nie można określić znaku wyrażenia $f(x^*+h)-f(x^*)$ w otoczeniu punktu $x=x^*$. □

Example

Znajdź lokalne ekstrema funkcji

$$f(x) = 12x^5 - 45x^4 + 40x^3 + 5.$$

Rozwiązanie. Pochodna funkcji wynosi $f'(x) = 60(x^4 - 3x^3 + 2x^2)$ i przyjmuje wartość zero w punktach $x_1 = 0$, $x_2 = 1$, $x_3 = 2$. Są to punkty, w których możemy podejrzewać istnienie ekstremów lokalnych wymaga to jednak dalszych badań. Druga pochodna jest równa $f''(x) = 60(4x^3 - 9x^2 + 4x)$. Dla $x = 0$ $f''(0) = 0$ i wymaga to obliczenia trzeciej pochodnej. Dla $x = 1$ $f''(1) = -60 < 0$ i funkcja osiąga w tym punkcie maksimum lokalne

$$f_{max} = f(1) = 12. \tag{2.2}$$

Dla $x = 2$ $f''(2) = 240 > 0$ i funkcja osiąga w tym punkcie minimum lokalne

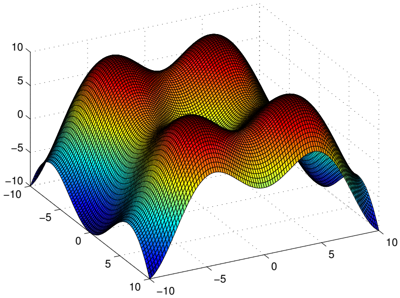
$$f_{min} = f(2) = -11. \tag{2.3}$$

Trzecia pochodna wynosi $f'''(x) = 120(6x^2 - 9x + 2)$ i $f'''(0) = 240 \neq 0$. Zatem w $x = 0$ funkcja nie osiąga ani minimum ani maksimum lokalnego, jest to punkt przegięcia.

Sformułowanie problemu optymalizacyjnego

$$\min_{\boldsymbol{x} \in \mathbb{R}^n} f(\boldsymbol{x}) \tag{3.1}$$

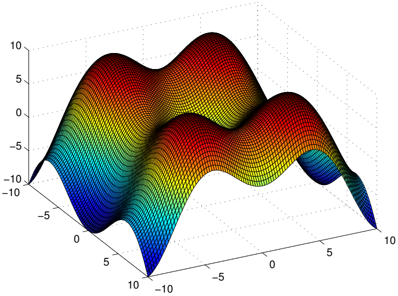
gdzie $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, n – liczba zmiennych decyzyjnych.



Rysunek: Funkcja $f(\boldsymbol{x}) = x_1 \sin(\frac{\pi x_1}{10}) + x_2 \sin(\frac{3\pi x_2}{20})$

Definicja

Funkcja wielu zmiennych $f : \mathbb{R}^n \to \mathbb{R}$ ma w punkcie $\boldsymbol{x} = \boldsymbol{x}^*$ minimum lokalne, gdy istnieje liczba $r > 0$ taka, że $f(\boldsymbol{x}^*) \leqslant f(\boldsymbol{x})$ dla każdego $|\boldsymbol{x} - \boldsymbol{x}^*| < r$.



Rysunek: Funkcja $f(\boldsymbol{x}) = x_1 \sin(\frac{\pi x_1}{10}) + x_2 \sin(\frac{3\pi x_2}{20})$

Różniczka r-tego rzędu

Definicja

Różniczką r-tego rzędu funkcji $f : \mathbb{R}^n \to \mathbb{R}$ w kierunku wektora $h = [h_1, h_2, \ldots, h_n]^T$ nazywamy wielomian postaci

$$\mathrm{d}^r f(\boldsymbol{x}) = \underbrace{\sum_{i=1}^{n} \sum_{j=1}^{n} \cdots \sum_{k=1}^{n}}_{r \text{ razy}} h_i h_j \ldots h_k \frac{\partial^r f(\boldsymbol{x})}{\partial x_i \partial x_j \ldots \partial x_k}$$

Różniczka r-tego rzędu – przykład

Przykład

Różniczka drugiego rzędu

Szereg Taylora

Szereg Taylora

$$f(\boldsymbol{x}+\boldsymbol{h}) = f(\boldsymbol{x}) + \mathrm{d}f(\boldsymbol{x}) + \frac{1}{2}\mathrm{d}^2 f(\boldsymbol{x})\frac{1}{3!}\mathrm{d}^3 f(\boldsymbol{x}) + \\ + \ldots + \frac{1}{r!}\mathrm{d}^r f(\boldsymbol{x}) + R_r(\boldsymbol{x},\boldsymbol{h}),$$

gdzie

$$R_r(\boldsymbol{x},\boldsymbol{h}) = \frac{1}{(r+1)!}\mathrm{d}^{(r+1)} f(\boldsymbol{x}+\theta\boldsymbol{h})$$

i $0 < \theta < 1$.

Przykład

Szereg taylora danej funkcji

Twierdzenie (warunek konieczny)

Funkcja $f : \mathbb{R}^n \to \mathbb{R}$ klasy C^1 ma ekstremum lokalne w punkcie $\boldsymbol{x} = \boldsymbol{x}^*$ jeśli

$$\frac{\partial f(\boldsymbol{x}^*)}{\partial x_1} = \frac{\partial f(\boldsymbol{x}^*)}{\partial x_2} = \ldots = \frac{\partial f(\boldsymbol{x}^*)}{\partial x_n} = 0$$

Uwaga

Punkt, w którym wszystkie pochodne cząstkowe funkcji ciągłej i różniczkowalnej równają się zero nazywamy *punktem stacjonarnym*.

Dowód

Załóżmy, że funkcja $f(x)$ w punkcie $\boldsymbol{x} = \boldsymbol{x}^*$ ma ekstremum lokalne oraz, że pochodne cząstkowe pierwszego rzędu nie zerują się w tym punkcie. Szereg Taylora możemy zapisać w postaci

$$f(\boldsymbol{x}^* + \boldsymbol{h}) = f(\boldsymbol{x}^*) + \mathrm{d}f(\boldsymbol{x}^*) + R_1(\boldsymbol{x}^*, \boldsymbol{h}),$$

gdzie reszta $R_1(\boldsymbol{x}, \boldsymbol{h})$ jest sumą iloczynów postaci $h_i h_j$. Dla dostatecznie małej wartości $|\boldsymbol{h}|$ drugi składnik po prawej stronie powyższego równania dominuje nad składnikami wyższego rzędu. Zatem wyrażenie

$$f(\boldsymbol{x}^* + \boldsymbol{h}) - f(\boldsymbol{x}^*)$$

przyjmuje taki sam znak jak wyrażenie $\mathrm{d}f(\boldsymbol{x}^*)$. Jeśli $\mathrm{d}f(\boldsymbol{x}^*) > 0$ dla wektora $\boldsymbol{h}$ to $\mathrm{d}f(\boldsymbol{x}^*) < 0$ dla wektora $-\boldsymbol{h}$. Co przeczy założeniu, że funkcja ma w punkcie $\boldsymbol{x} = \boldsymbol{x}^*$ ekstremum lokalne. Stąd $\mathrm{d}f(\boldsymbol{x}^*) = 0$ dla dowolnego wektora $\boldsymbol{h}$, więc pochodne cząskowe są równe zeru.

Warunek dostateczny

Twierdzenie (warunek dostateczny)

Warunkiem dostatecznym na to, aby punkt stacjonarny x^* był ekstremum lokalnym jest to, aby macierz drugich pochodnych cząskowych (hesjan) funkcji $f(x)$ w punkcie $x = x^*$ była

(i) dodatnio określona dla minimum lokalnego lub

(ii) ujemnie określona dla maksimum lokalnego.

Warunek dostateczny

Dowód

Szereg Taylora możemy zapisać w postaci

$$f(\boldsymbol{x}^* + \boldsymbol{h}) = f(\boldsymbol{x}^*) + \sum_{i=1}^{n} h_i \frac{\partial f(\boldsymbol{x}^*)}{\partial x_i} + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j \frac{\partial^2 f(\boldsymbol{x}^* + \theta \boldsymbol{h})}{\partial x_i \partial x_j} \qquad \text{(i)}$$

dla $0 < \theta < 1$. Stąd, że $\frac{\partial f(\boldsymbol{x}^*)}{\partial x_i} = 0$ dla $i = 1, 2, \ldots, n$ to o znaku wyrażenia

$$f(\boldsymbol{x}^* + \boldsymbol{h}) - f(\boldsymbol{x}^*) \qquad \text{(ii)}$$

decyduje ostatni składnik wyrażenia po prawej stronie równania (i).
Z ciągłości drugich pochodnych w punkcie $\boldsymbol{x}^*$ wynika, że $(\partial f(\boldsymbol{x})/\partial x_i \partial x_j)|_{\boldsymbol{x}=\boldsymbol{x}^*}$ ma taki sam znak jak $(\partial f(\boldsymbol{x})/\partial x_i \partial x_j)|_{\boldsymbol{x}=\boldsymbol{x}^*+\boldsymbol{h}}$ dla dostatecznie małej wartości $|h|$. Zatem wyrażenie (ii) jest dodatnie, więc funkcja $f(\boldsymbol{x})$ w punkcie $\boldsymbol{x} = \boldsymbol{x}^*$ ma minimum, gdy

$$Q = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} h_i h_j \frac{\partial^2 f(\boldsymbol{x}^*)}{\partial x_i \partial x_j} > 0. \qquad \text{(iii)}$$

Warunek dostateczny

Dowód (cd.)

Wyrażenie (iii) możemy zapisać w postaci

$$Q = \frac{1}{2}\boldsymbol{h}^T \mathbf{H}(\boldsymbol{x}^*)\boldsymbol{h},$$

gdzie

$$\mathbf{H}(\boldsymbol{x}^*) = \left.\begin{pmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \frac{\partial^2 f}{\partial x_3 \partial x_1} & \frac{\partial^2 f}{\partial x_3 \partial x_2} & \frac{\partial^2 f}{\partial x_3^2} & \cdots & \frac{\partial^2 f}{\partial x_3 \partial x_n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \frac{\partial^2 f}{\partial x_n \partial x_3} & \cdots & \frac{\partial^2 f}{\partial x_n^2} \end{pmatrix}\right|_{\boldsymbol{x}=\boldsymbol{x}^*}$$

Różniczka Q jest większa od zera dla dowolnego wektora $\boldsymbol{h}$, gdy macierz $\mathbf{H}(\boldsymbol{x}^*)$ jest dodatnio określona. Podobne rozumowanie przeprowadzamy dla maksimum lokalnego.

Warunek dostateczny

Uwagi

Macierz $\mathbf{A} = [a_{ij}]$ jest

- dodatnio określona, gdy jej wszystkie wartości własne, czyli wszystkie pierwiastki równania charakterystycznego

$$|\mathbf{A} - \lambda \mathbf{I}| = 0$$

  przyjmują wartości dodatnie,
- ujemnie określona, gdy jej wszystkie wartości własne przyjmują wartości ujemne.

Warunek dostateczny

Uwagi

Macierz $\mathbf{A} = [a_{ij}]$ jest

- dodatnio określona, gdy jej wszystkie wyznaczniki

$$A_1 = |a_{11}|, \qquad A_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix},$$

$$A_3 = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix}, \qquad \ldots, \qquad A_n = \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \multicolumn{4}{c}{\dotfill} \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}$$

  przyjmują wartości dodatnie,
- ujemnie określona, gdy każdy z wyznaczników A_j dla $j = 1, 2, \ldots, n$ przyjmuje wartość o tym samym znaku co liczba $(-1)^j$.

Punkt siodłowy

Problem

Problem optymalizacji funkcji wielu zmiennych z ograniczeniami równościowymi można przedstawić w postaci

$$\min_{\boldsymbol{x} \in \mathbb{X}} f(\boldsymbol{x}),$$

gdzie $\mathbb{X}$ to zbiór rozwiązań dopuszczalnych równy

$$\mathbb{X} = \{\boldsymbol{x} \in \mathbb{R}^n : h_i(\boldsymbol{x}) = 0, i = 1, 2, \ldots, m_e\}$$

$h_i : \mathbb{R}^n \to \mathbb{R}$ oraz m_e liczba ograniczeń równościowych

Bezpośrednie podstawienie

Sformułowany problem można doprowadzić do problemu optymalizacji funkcji $n - m_e$ zmiennych bez ograniczeń.
Rozwiązuje się układ równań

$$h_i(\boldsymbol{x}) = 0 \qquad \text{dla} \qquad i = 1, 2, \ldots, m_e.$$

Przedstawiamy dowolny zbiór m_e zmiennych decyzyjnych jako funkcje pozostałych $n - m_e$ zmiennych decyzyjnych

Bezpośrednie podstawienie

Przykład [lit]

Oblicz wymiary prostopadłościanu o maksymalnej objętości wpisanego w sferę o promieniu $r = 1$

Rozwiązanie.

Środki prostopadłościanu i sfery umieśćmy w początku kartezjańskiego układu wpółrzędnych $Ox_1x_2x_3$. Krawędzie prostopadłościanu niech będą równoległe do odpowiednich osi układu wspórzędnych, a jego wymiary równe odpowiednio $2x_1$, $2x_2$ i $2x_3$. Zatem mamy znaleźć maksimum funkcji

$$f(\boldsymbol{x}) = 8x_1x_2x_3$$

pod warunkiem, że

$$x_1^2 + x_2^2 + x_3^2 = 1.$$

Przykład cd. [lit]

Z równania możemy obliczyć, że

$$x_3 = \pm\sqrt{1 - x_1^2 - x_2^2}$$

Zatem funkcja celu przyjmuje postać

$$f(\boldsymbol{x}) = \pm 8 x_1 x_2 \sqrt{1 - x_1^2 - x_2^2}$$

Jej pochodne cząstkowe wynoszą odpowiednio

$$\frac{\partial f}{\partial x_1} = \pm 8 \frac{x_2 \left(1 - 2x_1^2 - x_2^2\right)}{\sqrt{1 - x_1^2 - x_2^2}},$$
$$\frac{\partial f}{\partial x_2} = \pm 8 \frac{x_1 \left(1 - x_1^2 - 2x_2^2\right)}{\sqrt{1 - x_1^2 - x_2^2}}.$$

Przyrównując je do zera otrzymuje się układ równań

$$x_2 \left(1 - 2x_1^2 - x_2^2\right) = 0,$$
$$x_1 \left(1 - x_1^2 - 2x_2^2\right) = 0,$$

Przykład cd. [lit]

który ma następujące rozwiązanie (po odrzuceniu rozwiązań ujemnych oraz nie należących do dziedziny)

$$x_1^* = \frac{1}{\sqrt{3}}, \qquad x_2^* = \frac{1}{\sqrt{3}}, \qquad x_3^* = \frac{1}{\sqrt{3}}.$$

Drugie pochodne cząstkowe są odpowiednio równe

$$\left.\frac{\partial^2 f}{\partial x_1^2}\right|_{\boldsymbol{x}=\boldsymbol{x}^*} = -\frac{32}{\sqrt{3}},$$

$$\left.\frac{\partial^2 f}{\partial x_2^2}\right|_{\boldsymbol{x}=\boldsymbol{x}^*} = -\frac{32}{\sqrt{3}},$$

$$\left.\frac{\partial^2 f}{\partial x_1 \partial x_2}\right|_{\boldsymbol{x}=\boldsymbol{x}^*} = -\frac{16}{\sqrt{3}}.$$

Przykład cd. [lit]

Hesjan funkcji $f(\boldsymbol{x})$ jest ujemnie określony w punkcie $\boldsymbol{x} = \boldsymbol{x}^*$, ponieważ

$$\left.\frac{\partial^2 f}{\partial x_1^2}\right|_{\boldsymbol{x}=\boldsymbol{x}^*} = -\frac{32}{\sqrt{3}} < 0 \qquad \text{oraz} \qquad \left.\frac{\partial^2 f}{\partial x_1^2}\frac{\partial^2 f}{\partial x_2^2} - \left(\frac{\partial^2 f}{\partial x_1^2}\right)^2\right|_{\boldsymbol{x}=\boldsymbol{x}^*} = 256 > 0.$$

Maksymalna wartość funkcji wynosi

$$f_{\max} = \frac{8}{3\sqrt{3}}$$

Rozpatrzmy problem optymalizacji funkcji dwóch zmiennych z jednym ograniczeniem równościiwym (tzn. $n = 2$, $m_e = 1$, czyli

$$\min_{\boldsymbol{x} \in \mathbb{R}^2} f(x_1, x_2)$$

pod warunkiem, że

$$h(x_1, x_2) = 0$$

Rozwiązanie

Jeżeli funkcja $f(x_1, x_2)$ ma w punkcie (x_1^*, x_2^*) miniumum to spełniony jest warunek konieczny optimum postaci

$$\left. \frac{\partial f}{\partial x_1}\mathrm{d}x_1 + \frac{\partial f}{\partial x_2}\mathrm{d}x_2 \right|_{\boldsymbol{x}=\boldsymbol{x}^*} = 0.$$

Punkt $\boldsymbol{x}^*$ musi spełniać ponadto ograniczenie

$$h(x_1^*, x_2^*) = 0,$$

a kierunek $(\mathrm{d}x_1, \mathrm{d}x_2)$

$$h(x_1^* + \mathrm{d}x_1, x_2^* + \mathrm{d}x_2) = 0.$$

Rozwijając to wyrażenie w szereg Taylora otrzymujemy

$$h(x_1^* + \mathrm{d}x_1, x_2^* + \mathrm{d}x_2) = h(x_1^*, x_2^*) + \frac{\partial h}{\partial x_1}(x_1^*, x_2^*)\mathrm{d}x_1 + \frac{\partial h}{\partial x_2}(x_1^*, x_2^*)\mathrm{d}x_2.$$

Zatem

$$\frac{\partial h}{\partial x_1}(x_1^*, x_2^*)\mathrm{d}x_1 + \frac{\partial h}{\partial x_2}(x_1^*, x_2^*)\mathrm{d}x_2 = 0.$$

Rozwiązanie cd.

Jeżeli funkcja $\frac{\partial h}{\partial x_2}(x_1^*, x_2^*) \neq 0$ to

$$\mathrm{d}x_2 = -\left(\frac{\partial h}{\partial x_1} \Big/ \frac{\partial h}{\partial x_2}\right)_{\boldsymbol{x}=\boldsymbol{x}^*} \mathrm{d}x_1.$$

Podstawiając to wyrażenie do warunku koniecznego otrzymujemy

$$\left(\frac{\partial f}{\partial x_1}\frac{\partial h}{\partial x_2} - \frac{\partial f}{\partial x_2}\frac{\partial h}{\partial x_1}\right)_{\boldsymbol{x}=\boldsymbol{x}^*} \mathrm{d}x_1 = 0.$$

czyli

$$\begin{vmatrix} \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} \\ \frac{\partial h}{\partial x_1} & \frac{\partial h}{\partial x_2} \end{vmatrix} = 0.$$

Jest to warunek konieczny istnienia ekstremum funkcji.

W przypadku dowolnej wartości n i m_e warunki konieczne są następujące

$$J\left(\frac{f, h_1, h_2, \ldots, h_{m_e}}{x_k, x_1, x_2, \ldots, x_{m_e}}\right) = \begin{vmatrix} \frac{\partial f}{\partial x_k} & \frac{\partial f}{\partial x_1} & \frac{\partial f}{\partial x_2} & \cdots & \frac{\partial f}{\partial x_{m_e}} \\ \frac{\partial h_1}{\partial x_k} & \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \cdots & \frac{\partial h_1}{\partial x_{m_e}} \\ \frac{\partial h_2}{\partial x_k} & \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} & \cdots & \frac{\partial h_2}{\partial x_{m_e}} \\ \multicolumn{4}{l}{\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots} & \\ \frac{\partial h_{m_e}}{\partial x_k} & \frac{\partial h_{m_e}}{\partial x_1} & \frac{\partial h_{m_e}}{\partial x_2} & \cdots & \frac{\partial h_{m_e}}{\partial x_{m_e}} \end{vmatrix} = 0$$

dla $k = m_e + 1, m_e + 2, \ldots, n$. Ponadto musi być spełniony warunek

$$J\left(\frac{h_1, h_2, \ldots, h_{m_e}}{x_1, x_2, \ldots, x_{m_e}}\right) \neq 0.$$

Warunek konieczny ($n = 2$, $m_e = 1$)

Rozpatrzmy problem optymalizacji funkcji dwóch zmiennych z jednym ograniczeniem równościwym (tzn. $n = 2$, $m_e = 1$, czyli

$$\min_{\boldsymbol{x} \in \mathbb{R}^2} f(x_1, x_2)$$

pod warunkiem, że

$$h(x_1, x_2) = 0$$

Warunek konieczny ($n = 2$, $m_e = 1$)

Rozwiązanie

Z wcześniejszych rozważań wynika, że funkcja f ma w punkcie $\boldsymbol{x}^* = (x_1^*, x_2^*)$, gdy

$$\left.\frac{\partial f}{\partial x_1} - \frac{\partial f/\partial x_2}{\partial h/\partial x_2}\frac{\partial h}{\partial x_1}\right|_{\boldsymbol{x}=\boldsymbol{x}^*} = 0. \tag{i}$$

Definiując

$$\lambda = -\left.\frac{\frac{\partial f}{\partial x_2}}{\frac{\partial h}{\partial x_2}}\right|_{\boldsymbol{x}=\boldsymbol{x}^*}. \tag{ii}$$

Zatem równanie (i) możemy zapisać

$$\left.\frac{\partial f}{\partial x_1} + \lambda\frac{\partial h}{\partial x_1}\right|_{\boldsymbol{x}=\boldsymbol{x}^*} = 0, \tag{iii}$$

natomiast równanie (ii)

$$\left.\frac{\partial f}{\partial x_2} + \lambda\frac{\partial h}{\partial x_2}\right|_{\boldsymbol{x}=\boldsymbol{x}^*} = 0. \tag{iv}$$

Ponadto

$$h(\boldsymbol{x})|_{\boldsymbol{x}=\boldsymbol{x}^*} = 0. \tag{v}$$

Równania (iii)-(v) tworzą układ warunków koniecznych istnienia ekstremum funkcji.

Warunek konieczny

Funkcja Lagrange'a

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \boldsymbol{\lambda}^{\mathrm{T}} \boldsymbol{h}(\boldsymbol{x}),$$

gdzie

$$\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \qquad \boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_{m_e} \end{bmatrix} \qquad \boldsymbol{h} = \begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_{m_e} \end{bmatrix}$$

lub

$$L(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \lambda_1 h_1(\boldsymbol{x}) + \lambda_1 h_1(\boldsymbol{x}) + \cdots + \lambda_{m_e} h_{m_e}(\boldsymbol{x}),$$

Warunki konieczne

Warunki konieczne

Warunki konieczne możemy zapisać w postaci

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m_e} \lambda_j \frac{\partial h_j}{\partial x_i} = 0 \qquad i = 1, 2, \ldots, n,$$

$$\frac{\partial L}{\partial \lambda_j} = h_j(\boldsymbol{x}) = 0 \qquad j = 1, 2, \ldots, m_e,$$

lub

$$\nabla L(\boldsymbol{x}, \boldsymbol{\lambda}) = 0,$$
$$\nabla_\lambda L(\boldsymbol{x}, \boldsymbol{\lambda}) = 0,$$

gdzie

$$\nabla = \begin{bmatrix} \frac{\partial}{\partial x_1} \\ \frac{\partial}{\partial x_2} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{bmatrix} \qquad \nabla_\lambda = \begin{bmatrix} \frac{\partial}{\partial \lambda_1} \\ \frac{\partial}{\partial \lambda_2} \\ \vdots \\ \frac{\partial}{\partial \lambda_{m_e}} \end{bmatrix}$$

Warunek dostateczny

Warunek dostateczny

Warunkiem dostatecznym istnienia minimum funkcji w punkcie stacjonarnym $\boldsymbol{x} = \boldsymbol{x}^*$ jest przyjmowanie wartości dodatniej formy kwadratowej

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 L}{\partial x_i \partial x_j} \mathrm{d}x_i \mathrm{d}x_j$$

dla każdej wartości wektora $\mathrm{d}\boldsymbol{x}$ spełniającej ograniczenia, tzn. musi zachodzić warunek

$$h_j(\boldsymbol{x} + \mathrm{d}\boldsymbol{x}) = 0, \qquad j = 1, 2, \ldots, m_e.$$

Warunek dostateczny

Uwagi

Warunkiem dostatecznym istnienia maksimum funkcji w punkcie stacjonarnym $\boldsymbol{x} = \boldsymbol{x}^*$ jest przyjmowanie wartości ujemnej formy kwadratowej $Q(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$

$$Q = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 L(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)}{\partial x_i \partial x_j} \mathrm{d}x_i \mathrm{d}x_j < 0$$

dla każdego dopuszczalnego wektora $\mathrm{d}\boldsymbol{x}$.

Warunek dostateczny

Uwagi

Warunkiem koniecznym przyjmowania wartości dodatniej (ujemnej) formy kwadratowej $Q(\boldsymbol{x}^*, \boldsymbol{\lambda}^*)$ dla dopuszczalnych kierunków $\mathrm{d}\boldsymbol{x}$ jest to, aby wszystkie pierwiastki wielomianu

$$\begin{vmatrix} L_{11}-z & L_{12} & \dots & L_{1n} & h_{11} & h_{21} & \dots & h_{m_e1} \\ L_{21} & L_{22}-z & \dots & L_{2n} & h_{12} & h_{22} & \dots & h_{m_e2} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ L_{n1} & L_{n2} & \dots & L_{nn}-z & h_{1n} & h_{2n} & \dots & h_{m_en} \\ h_{11} & h_{12} & \dots & h_{1n} & 0 & 0 & \dots & 0 \\ h_{21} & h_{22} & \dots & h_{2n} & 0 & 0 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ h_{m_e1} & h_{m_e2} & \dots & h_{m_en} & 0 & 0 & \dots & 0 \end{vmatrix} = 0,$$

gdzie

$$L_{ij} = \frac{\partial^2 L}{\partial x_i \partial x_j}(\boldsymbol{x}^*, \boldsymbol{\lambda}^*),$$

$$h_{ij} = \frac{\partial h_i}{\partial x_j}(\boldsymbol{x}^*),$$

były dodatnie (ujemne).

Problem

Problem optymalizacji funkcji wielu zmiennych z ograniczeniami nierównościowymi można przedstawić w postaci

$$\min_{\boldsymbol{x} \in \mathbb{X}} f(\boldsymbol{x}),$$

gdzie $\mathbb{X}$ to zbiór rozwiązań dopuszczalnych równy

$$\mathbb{X} = \{\boldsymbol{x} \in \mathbb{R}^n : g_i(\boldsymbol{x}) \leqslant 0, i = 1, 2, \dots, m_n\}$$

$g_i : \mathbb{R}^n \to \mathbb{R}$ oraz m_n liczba ograniczeń nierównościowych.

Wyprowadzenie warunków koniecznych

Wprowadźmy nowe zmienne decyzyjne y_j dla $j = 1, 2, \ldots, m_n$ i przekształćmy ograniczenia do postaci

$$g_j + y_j^2 = 0,$$

Otrzymaliśmy nowy problem – optymalizację funkcji z ograniczeniami równościowymi. Wymiar problemu wzrósł o liczbę ograniczeń nierównościowych.

Funkcja Lagrange'a wygląda następująco

$$L(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) + \sum_{j=1}^{m_n} \lambda_j \left[g_j(\boldsymbol{x}) + y_j^2 \right].$$

Wyprowadzenie warunków koniecznych

Warunki konieczne istnienia ekstremum funkcji z ograniczeniami równościowymi mają postać

$$\frac{\partial L}{\partial x_i} = \frac{\partial f}{\partial x_i} + \sum_{j=1}^{m_n} \lambda_j \frac{\partial g_j}{\partial x_i} = 0 \qquad i = 1, 2, \ldots, n, \qquad \text{(i)}$$

$$\frac{\partial L}{\partial \lambda_j} = g_j(\boldsymbol{x}) + y_j^2 = 0 \qquad j = 1, 2, \ldots, m_n, \qquad \text{(ii)}$$

$$\frac{\partial L}{\partial y_j} = 2\lambda_j y_j = 0 \qquad j = 1, 2, \ldots, m_n. \qquad \text{(iii)}$$

- Równanie (ii) zapewnia, że $g(\boldsymbol{x}^*) \leqslant 0$.
- Równanie (iii) zapewnia, że $\lambda_j = 0$ lub $y_j = 0$.
  - Jeżeli $\lambda_j = 0$ i $y_j \neq 0$ to $g_j < 0$ (ograniczenie nieaktywne).
  - Jeżeli $\lambda_j \neq 0$ i $y_j = 0$ to $g_j = 0$ (ograniczenie aktywne).

Wyprowadzenie warunków koniecznych

Niech J_1 i J_2 będą zbiorami indeksów dla których ograniczenia są odpowiednio aktywne i nieaktywne, tzn.

$$\begin{aligned} g_j(\boldsymbol{x}^*) = 0 \qquad &\Longrightarrow \qquad j \in J_1, \\ g_j(\boldsymbol{x}^*) < 0 \qquad &\Longrightarrow \qquad j \in J_2. \end{aligned}$$

Równania (i-iii) można zapisać w postaci

$$\frac{\partial f}{\partial x_i} + \sum_{j \in J_1} \lambda_j \frac{\partial g_j}{\partial x_i} = 0 \qquad i = 1, 2, \ldots, n, \tag{i}$$

$$g_j(\boldsymbol{x}) = 0 \qquad j \in J_1, \tag{ii}$$

$$g_j(\boldsymbol{x}) + y_j^2 = 0 \qquad j \in J_2. \tag{iii}$$

Równanie (i) można przedstawić

$$-\nabla f(\boldsymbol{x}^*) = \lambda_{j_1} \nabla g_{j_1}(\boldsymbol{x}^*) + \lambda_{j_2} \nabla g_{j_2}(\boldsymbol{x}^*) + \cdots + \lambda_{j_p} \nabla g_{j_p}(\boldsymbol{x}^*) \tag{iv}$$

gdzie $j_1, j_2, \ldots, j_p \in J_1$ i $p = |J_1|$.

Wyprowadzenie warunków koniecznych

Weźmy dowolny kierunek dopuszczalny $\boldsymbol{S}$, tzn. taki, że

$$\boldsymbol{S}^{\mathrm{T}}\nabla g_j(\boldsymbol{x}^*) < 0, \qquad j \in J_1. \tag{v}$$

Jeżeli wymnożymy skalarnie równanie (iv) przez wektor $\boldsymbol{S}$ otrzymamy zależność

$$-\boldsymbol{S}^{\mathrm{T}}\nabla f(\boldsymbol{x}^*) = \lambda_{j_1}\boldsymbol{S}^{\mathrm{T}}\nabla g_{j_1}(\boldsymbol{x}^*) + \lambda_{j_2}\boldsymbol{S}^{\mathrm{T}}\nabla g_{j_2}(\boldsymbol{x}^*) + \cdots + \lambda_{j_p}\boldsymbol{S}^{\mathrm{T}}\nabla g_{j_p}(\boldsymbol{x}^*) \tag{vi}$$

W przypadku, gdy szukamy minimum funkcji wyrażenie $\boldsymbol{S}^{\mathrm{T}}\nabla f(\boldsymbol{x}^*)$ powinno przyjmować wartość dodatnią dla każdego dopuszczalnego wektora $\boldsymbol{S}$, w przeciwnym przypadku w kierunku tego wektora znajduje się punkt, w którym wartość funkcji jest mniejsza i w tym punkcie spełnione są wszystkie ograniczenia, co przeczy temu, że $f(\boldsymbol{x}^*)$ jest minimum lokalnym funkcji w obszarze dopuszczalnym.
Z powyższych rozważań wynika, że prawa strona równania (vi) musi być mniejsza od zera. Biorąc również pod uwagę warunki (v) wnosimy, że $\lambda_j > 0$ dla $j \in J_1$

Warunki konieczne (Kuhn-Tucker)

Warunki konieczne istnienia ekstremum funkcji z ograniczeniami nierównościowymi mają postać

$$\frac{\partial f(\boldsymbol{x}^*)}{\partial x_i} + \sum_{j=1}^{m_n} \lambda_j \frac{\partial g_j(\boldsymbol{x}^*)}{\partial x_i} = 0, \qquad i = 1, 2, \ldots, n, \tag{i}$$

$$\lambda_j g_j(\boldsymbol{x}^*) = 0, \tag{ii}$$

$$g_j(\boldsymbol{x}^*) \leqslant 0, \tag{iii}$$

$$\lambda_j \geqslant 0, \qquad j = 1, 2, \ldots, m_n. \tag{iv}$$

Warunki konieczne

Warunki konieczne (Kuhn-Tucker)

Podsumowując poprzednie rozważania warunki konieczne istnienia minimum lokalnego w punkcie $\boldsymbol{x}^*$ przedstawiają się następująco

$$\Delta f + \sum_{j=1}^{m_n} \lambda_j \Delta g_j + \sum_{k=1}^{m_e} \beta_k \Delta h_k = \mathbf{0}, \tag{5.1}$$

$$\lambda_j g_j(\boldsymbol{x}^*) = 0, \qquad j = 1, 2, \ldots, m_n, \tag{5.2}$$

$$g_j(\boldsymbol{x}^*) \leqslant 0, \qquad j = 1, 2, \ldots, m_n, \tag{5.3}$$

$$h_k(\boldsymbol{x}^*) = 0, \qquad k = 1, 2, \ldots, m_e, \tag{5.4}$$

$$\lambda_j \geqslant 0, \qquad j = 1, 2, \ldots, m_n. \tag{5.5}$$