

METODY OPTIMALIZACJI

Wydział Elektroniki
Kier. Elektronika
Studia II st. - magisterskie

Wykład III
Własności funkcji

Właściwe minimum lokalne:

Funkcja $f(x)$ ma w punkcie \hat{x} właściwe minimum lokalne, jeżeli istnieje $\delta > 0$ takie, że: $\forall x \in E$

$$f(\hat{x}) < f(x)$$

przy czym: $E = X \cap \Delta$

$$\Delta = \{x: 0 < \|x - \hat{x}\| < \delta\}$$

Słabe minimum lokalne

Funkcja $f(x)$ ma w punkcie \hat{x} słabe minimum lokalne, jeżeli istnieje $\delta_i > 0$ takie, że $\forall x \in E_i$

$$f(\hat{x}) \leq f(x_i)$$

przy czym:

$$E_i = X \cap \Delta_i$$

$$\Delta_i = \{x: 0 < \|x - \hat{x}\| < \delta_i\}$$

Minimum lokalne i globalne funkcji $f(x)$

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

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

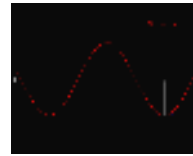
Przy czym jeśli zachodzi $f(\hat{x}) < f(x)$ dla $x \neq \hat{x}$ to istnieje wtedy właściwe minimum lokalne.

Punkt \hat{x} stanowi minimum globalne funkcji $f(x)$ w przestrzeni R^n , jeżeli

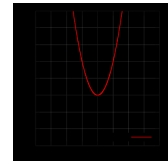
$$\forall x \in R^n \quad f(\hat{x}) \leq f(x)$$

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

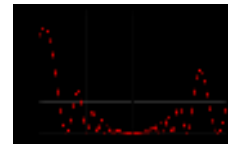
Przykłady funkcji



Rys.1



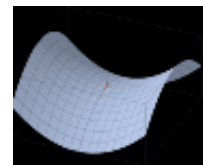
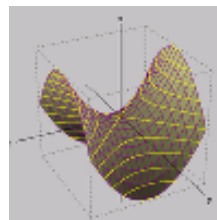
Rys.2



Rys.3

- Funkcja cosinus osiąga maksimum dla każdej parzystej wielokrotności π , oraz minimum dla każdej nieparzystej wielokrotności π , czyli są to lokalne ekstrema właściwe i jednocześnie ekstrema globalne.
- Funkcja kwadratowa $f(x)=x^2$ osiąga właściwe minimum (lokalne i globalne) dla $x=0$. Nie ma maksimum, nawet lokalnego. Dla każdego argumentu można w jego bezpośrednim sąsiedztwie wskazać punkt, w którym funkcja przyjmuje większą wartość.
- Funkcja na rys.3 ma w punkcie $x=0$ minimum lokalne, jednak nie jest to minimum właściwe – w dowolnej bliskości tego punktu można znaleźć inne punkty, w których przyjmuje ona tę samą wartość (oprócz tego posiada nieskończoną liczbę minimów i maksimów właściwych).

Przykłady funkcji



Własności funkcji wypukłych

Definicja zbioru wypukłego:

Zbiór $X \subset R^n$ nazywamy wypukłym, jeżeli dla każdego dwóch punktów $x^1, x^2 \in X$ odcinek łączący te dwa punkty także należy do X tzn.:

$$X = \{x: x = \lambda x^1 + (1-\lambda)x^2, 0 \leq \lambda \leq 1\}$$

Zbiór wypukły – intuicyjnie, podzbiór pewnej przestrzeni euklidesowej, o tej własności, że dowolny odcinek, którego końce należą do tego zbioru, w całości się w nim zawiera.

Definicja funkcji wypukłej:

Niech $X \subset R^n$ będzie zbiorem wypukłym. Funkcję $f: X \rightarrow R^1$ będziemy nazywali wypukłą, jeżeli dla dowolnych dwóch punktów $x^1, x^2 \in X$ i dla każdego: $\lambda \in [0,1]$

$$f(\lambda x^1 + (1-\lambda)x^2) \leq \lambda f(x^1) + (1-\lambda)f(x^2)$$

Własności funkcji wypukłych

Definicja funkcji ściśle-wypukłej:

Niech $X \subset R^n$ będzie zbiorem wypukłym. Funkcję $f: X \rightarrow R^1$ będziemy nazywali ściśle-wypukłą, jeżeli dla dowolnych dwóch punktów $x^1, x^2 \in X$ i dla każdego: $\lambda \in [0,1]$

zachodzi: $f(\lambda x^1 + (1-\lambda)x^2) < \lambda f(x^1) + (1-\lambda)f(x^2)$

Definicja funkcji wklęsłej:

Niech $X \subset R^n$ będzie zbiorem wypukłym. Funkcję $f: X \rightarrow R^1$ będziemy nazywali wklęsłą, jeżeli dla dowolnych dwóch punktów $x^1, x^2 \in X$ i dla każdego: $\lambda \in [0,1]$

zachodzi: $f(\lambda x^1 + (1-\lambda)x^2) \geq \lambda f(x^1) + (1-\lambda)f(x^2)$

Wniosek: Funkcja $f(x)$ jest wklęsła wtedy i tylko wtedy gdy funkcja $-f(x)$ jest wypukła.

Własności różniczkowe funkcji

Jeżeli w punkcie x^0 istnieje wektor, który oznaczmy $\nabla f(x^0)$, taki, że: dla dowolnego punktu $x \in X$

$$f(x) = f(x^0) + \nabla f(x^0)^T (x - x^0) + o_1(x^0, x)$$

gdzie: $\lim_{x \rightarrow x^0} \frac{o_1(x^0, x)}{\|x - x^0\|} = 0$

Funkcja $f(x)$ jest różniczkowalna w punkcie x^0 , a wektor $\nabla f(x^0)$ jest gradientem funkcji $f(x)$.

Jeżeli w punkcie x^0 funkcja $f(x)$ jest różniczkowalna oraz istnieje symetryczna macierz kwadratowa $n \times n$ $\nabla^2 f(x^0)$ taka, że: dla dowolnego punktu $x \in X$ dla dowolnego punktu

$$f(x) = f(x^0) + \nabla f(x^0)^T (x - x^0) + 1/2(x - x^0)^T \nabla^2 f(x^0)(x - x^0) + o_2(x^0, x)$$

gdzie: $\lim_{x \rightarrow x^0} \frac{o_2(x^0, x)}{\|x - x^0\|^2} = 0$ $\nabla^2 f(x^0)$ - hesjan funkcji $f(x^0)$

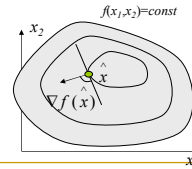
Wówczas funkcja jest dwukrotnie różniczkowalna w punkcie x^0 .

Hesjan funkcji $f(x)$

$$\nabla^2 f(x) = \frac{\partial^2 f(x)}{\partial x \partial x} = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_1 \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f(x)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x)}{\partial x_2 \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \frac{\partial^2 f(x)}{\partial x_n \partial x_2} & \dots & \frac{\partial^2 f(x)}{\partial x_n \partial x_n} \end{bmatrix} \quad \nabla^2 f(x) = A = H$$

Forma kwadratowa $f(x) = 1/2x^T Ax + b^T x + c$

Dla macierzy symetrycznej A – zachodzi: $\nabla^2 f(x) = A$



Kiedy funkcja $f(x)$ jest wypukła

Macierz A o wymiarze $n \times n$ nazywamy hesjanem, gdy jej elementami są drugie pochodne cząstkowe funkcji $f(x)$:

$$A(x) = \begin{bmatrix} \frac{\partial^2 f(x)}{\partial x_1 \partial x_1} & \dots & \frac{\partial^2 f(x)}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x)}{\partial x_n \partial x_1} & \dots & \frac{\partial^2 f(x)}{\partial x_n \partial x_n} \end{bmatrix}$$

Lemat 1

Niech $f: X \rightarrow R^1$ ma ciągle drugie pochodne cząstkowe oraz niech $X \subset R^n$ będzie zbiorem wypukłym. Funkcja $f(x)$ jest wypukła wtedy i tylko wtedy, gdy jej hesjan $A(x)$ jest dodatnio pół-określony dla każdego $x \in X$.

Definicja macierzy A dodatnio półokreślonej

Macierz A o wymiarze $n \times n$ jest dodatnio pół-określona, jeśli

dla każdego $x \in R^n$ $\langle x, Ax \rangle \geq 0$

Definicja macierzy A dodatnio określonej

Macierz A o wymiarze $n \times n$ jest dodatnio określona, jeśli

dla każdego niezerowego $x \in R^n$ $\langle x, Ax \rangle > 0$

Praktyczne sprawdzenie wypukłości funkcji $f(x)$ – czyli sprawdzenie ściślejszej dodatniej określoności rzeczywistej symetrycznej macierzy A

► **Kryterium Sylwestra** - kwadratowa macierz symetryczna A jest dodatnio pół-określona wtedy i tylko wtedy, gdy:

$$a_{11} \geq 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} \geq 0, \dots, \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} \geq 0$$

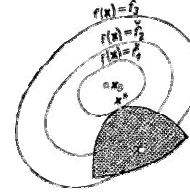
► **Kryterium Sylwestra** - kwadratowa macierz symetryczna A jest dodatnio określona wtedy i tylko wtedy, gdy:

$$a_{11} > 0, \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} > 0, \dots, \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} > 0$$

> Wszystkie wartości własne macierzy A są dodatnie

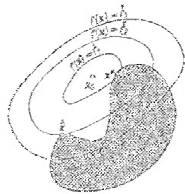
> W rozkładzie Choleskiego $A=L^T L$ wszystkie elementy leżące na głównej przekątnej są dodatnie.

Rys.1 Zbiór X - wypukły



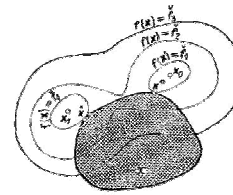
Funkcja $f(x)$ też jest wypukła

Rys.2 Zbiór X – nie jest wypukły



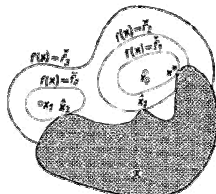
Funkcja $f(x)$ – funkcja wypukła.

Rys.3 Zbiór X – jest wypukły



Funkcja $f(x)$ nie jest wypukła,

Rys.4 Zbiór X – nie jest wypukły



Funkcja $f(x)$ też nie jest wypukła.

Funkcje wypukłe

Lemat 2

Niech funkcja $f(x)$ będzie funkcją różniczkowalną oraz zbiór $X \subset R^n$ będzie zbiorem wypukłym.

Funkcja $f(x)$ jest wypukła wtedy i tylko wtedy, gdy $\forall x, x^0 \in X$

$$f(x) \geq f(x^0) + \langle \nabla f(x^0), x - x^0 \rangle$$

Funkcja $f(x)$ jest wypukła w przedziale (a,b) wtedy i tylko wtedy, gdy wykres funkcji leży ponad wykresem stycznej dla każdego punktu x^0 z przedziału (a,b) dla $\forall a, b \in X$

Lemat 3

Niech funkcja $f: X \rightarrow R^1$, dla $X \subset R^n$ będzie funkcją wypukłą, wówczas dla dowolnego rzeczywistego α zbiór

$$X_\alpha = \{x: f(x) \leq \alpha\}$$

jest wypukły.

Lemat 4

Niech $X \subset \mathbb{R}^n$ będzie zbiorem wypukłym. Jeśli funkcje $f_i: X \rightarrow \mathbb{R}^1$ dla $i=1, \dots, k$ są funkcjami wypukłymi oraz jeśli wielkości skalarne są większe od zera dla $i=1, \dots, k$ $\alpha_i \geq 0$ to funkcja $f(x)$

$$f(x) = \sum_{i=1}^k \alpha_i f_i(x)$$

jest funkcją wypukłą.

Twierdzenie 2

Dowolne minimum lokalne funkcji wypukłej $f(x)$ na wypukłym zbiorze rozwiązań dopuszczalnych $X \subset \mathbb{R}^n$ jest na tym zbiorze jej minimum globalnym.

Dowód:

Niech w punkcie $x^1 \in X$ funkcja $f(x)$ ma swoje minimum lokalne. Oznacza to, że istnieje takie $\varepsilon > 0$ że:

$$f(x^1) = \min_{x \in V} f(x)$$

gdzie: $V = \{x \in X, \|x - x^1\| \leq \varepsilon\}$

Przyjmijmy $x^2 \in X$, $x^1 \neq x^2$. Wybierzmy $0 < \lambda < 1$ oraz

$$\lambda x^1 + (1 - \lambda)x^2 \in V$$

Ze względu na wypukłość $f(x)$:

$$\lambda f(x^1) + (1 - \lambda)f(x^2) \geq f(\lambda x^1 + (1 - \lambda)x^2)$$

$$f(x^2) \geq \frac{f(\lambda x^1 + (1 - \lambda)x^2) - \lambda f(x^1)}{1 - \lambda} \geq \frac{f(x^1) - \lambda f(x^1)}{1 - \lambda} = f(x^1)$$

ckd

Wniosek

Ścisłe wypukła funkcja $f(x)$ określona na wypukłym zbiorze rozwiązań dopuszczalnych X ma na tym zbiorze co najwyżej jedno minimum.

Przykład

Niech $c \in \mathbb{R}^n$, $d \in \mathbb{R}^1$. Wykazać, że funkcja $f: \mathbb{R}^n \rightarrow \mathbb{R}^1$ określona formułą liniową:

$$f(x) = c^T x + d$$

jest jednocześnie wypukła i wklęsła na \mathbb{R}^n .

Twierdzenie 3

Jeśli

$$f(x): \mathbb{R}^n \rightarrow \mathbb{R}$$

Jest funkcją wklęsłą i X jest zbiorem zwartym (domkniętym, ograniczonym) i wypukłym,

to istnieje punkt ekstremalny zbioru X ,
w którym $f(x)$ osiąga minimum na tym zbiorze.

Dowód: przez zaprzeczenie tezie

Wniosek

Funkcja liniowa $f(x)$ określona na wielościanie (wypukłym zbiorze rozwiązań dopuszczalnych X) osiąga swoje kresy w wierzchołkach tego wielościanu.