

Układy równań nieliniowych (wielowymiarowa metoda Newtona-Raphsona)

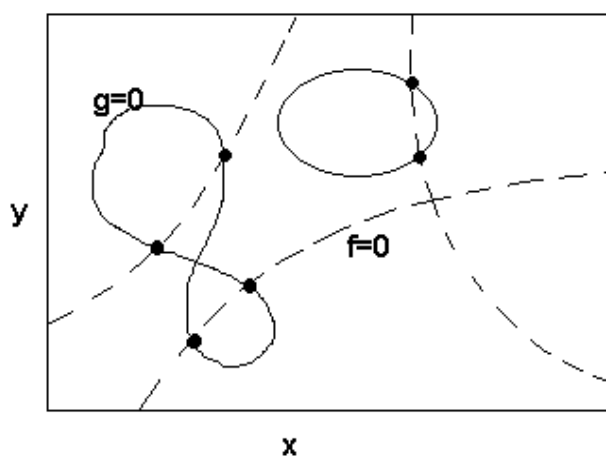
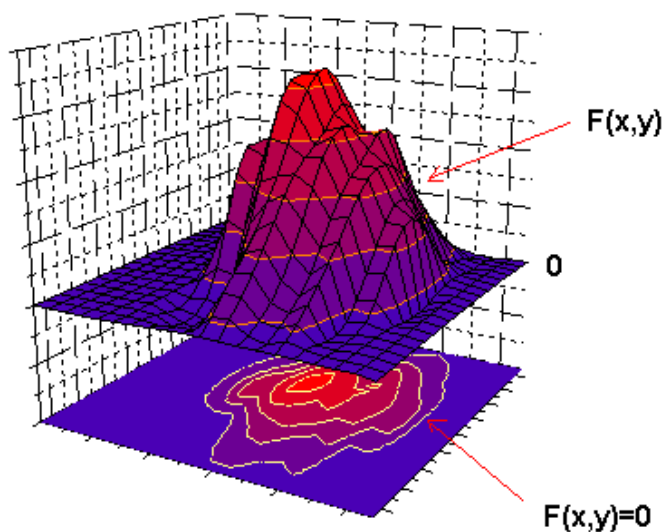
$$\mathbf{f}(\mathbf{x}) = \mathbf{0}, \quad \text{gdzie}$$

$$\mathbf{f} = (f_1, f_2, \dots, f_n) \quad \mathbf{x} = (x_1, x_2, \dots, x_n)$$

dla $n=2$ np.

$$f(x, y) = 0$$

$$g(x, y) = 0$$



dla każdej wielowymiarowej f_i stosujemy rozwinięcie w szereg Taylora

$$f_i(\mathbf{X} + \delta\mathbf{X}) = f_i(\mathbf{X}) + \sum_{j=1}^N \frac{\partial f_i}{\partial x_j} \delta x_j + O(\delta\mathbf{X}^2) \quad (**)$$

zostawiając tylko wyrazy liniowe i żądając

$$f_i(\mathbf{X} + \delta\mathbf{X}) = 0$$

(tzn.

chcemy o tyle przesunąć się od początkowo wybranego \mathbf{X} , żeby w $\mathbf{X} + \delta\mathbf{X}$, $f_i = 0$, z dokładnością do przybliżenia **)

dostajemy układ algebraicznych równań liniowych niejednorodnych na $\delta\mathbf{X}$, tzn. szukamy takiej zmiany \mathbf{X} , która prowadzi do $f_i(\mathbf{X} + \delta\mathbf{X}) = 0$

$$\sum_{j=1}^N \alpha_{ij} \delta x_j = \beta_i \quad (*)$$

gdzie

$$\alpha_{ij} \equiv \frac{\partial f_i}{\partial x_j} \quad \beta_i \equiv -f_i$$

rozwiązując układ (*) **równań algebraicznych (!)** budujemy

$$\mathbf{X}^{\text{new}} = \mathbf{X}^{\text{old}} + \delta\mathbf{X} \quad (***)$$

co daje nam nowe $\beta_i \equiv -f_i(\mathbf{X}^{\text{new}})$

[pamiętamy, że poprzednie $\beta_i = f_i(\mathbf{X}^{\text{old}})$]

wyznaczamy pochodne α_{ij} w \mathbf{X}^{new}
i proces (**) powtarzamy aż do otrzymania zbieżności z żadaną dokładnością.

(*) i (***) można inaczej zapisać:

$$\mathbf{X}^{k+1} = \mathbf{X}^k - \alpha^{-1}(\mathbf{X}^k)\mathbf{f}(\mathbf{X}^k)$$

przypomnijmy, że w jednym wymiarze mieliśmy:

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

czyli znów mamy $\mathbf{X}_{k+1} = \mathcal{F}(\mathbf{X}_k)$

zauważmy, że znów kolejną metodę sprowadziliśmy do iteracyjnego procesu rozwiązywania układów liniowych równań algebraicznych !

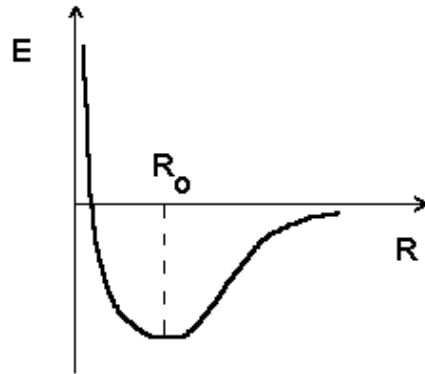
SZUKANIE EKSTREMÓW

(minimalizacja i maksymalizacja)

przykład:

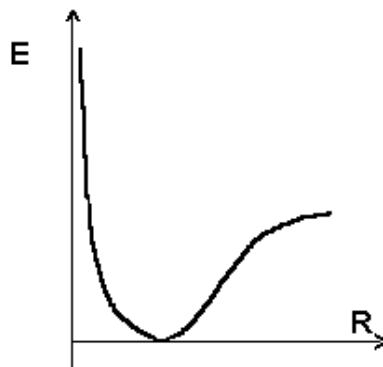
- szukanie minimum powierzchni energetycznej molekuly jako funkcji wielu zmiennych (położeń jąder) - optymalizacja geometrii

w jednym wymiarze (dla cząsteczki dwuatomowej)



energia potencjalna dla ruchu jąder w molekule dwuatomowej
w funkcji odległości między jądrami

zagadnienie podobne do znajdowania zer



Jak szukać minimum?

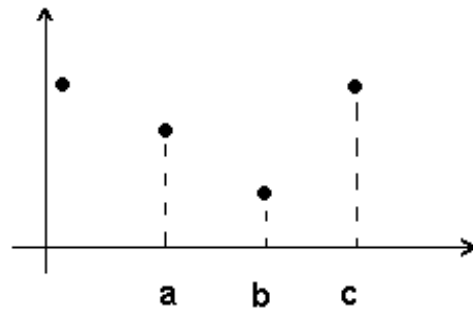
funkcja jednej zmiennej:

Metoda złotego podziału

szukamy kolejnych punktów $X_0, X_1, X_2, \dots, X_n, \dots$

aż znajdziemy dwa takie $X_{i-1}, X_i, X_{i+1} = a, b, c$, takie, że dla $a < b < c$ ($a = X_{i-1}, b = X_i, c = X_{i+1}$)
 (*)

$$f(a) > f(b) \quad \text{oraz} \quad f(c) > f(b)$$



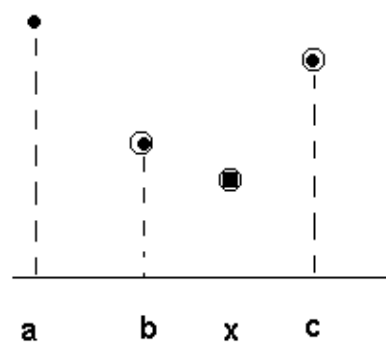
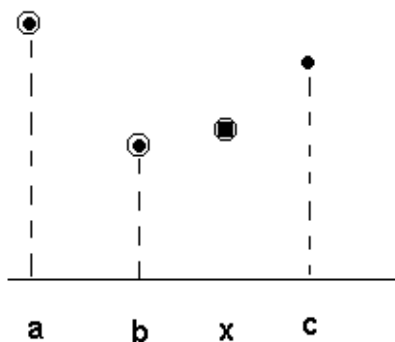
punkty nie muszą być równoodległe;

szukając dalej,

nowy punkt próbny X możemy wybrać w $[a, b]$ lub $[b, c]$;
 weźmy X w $[b, c]$,

jeśli teraz $f(b) < f(x)$ to nowy zestaw trzech punktów gwarantujących (*) jest: $a < b < x$

jeśli natomiast $f(b) > f(x)$ to następane trzy punkty wybieramy: $b < x < c$



Pytanie:

gdzie w przedziale $[a,b]$ lub $[b,c]$ wybrać następny próbny punkt?

albo inaczej: jak dzielić punktem b odcinek $[a,c]$, żeby zoptymalizować i zunifikować procedurę poszukiwania ekstremum?

wprowadźmy oznaczenia:

$$\frac{b-a}{c-a} = W \quad \text{i} \quad \frac{c-b}{c-a} = 1-W$$

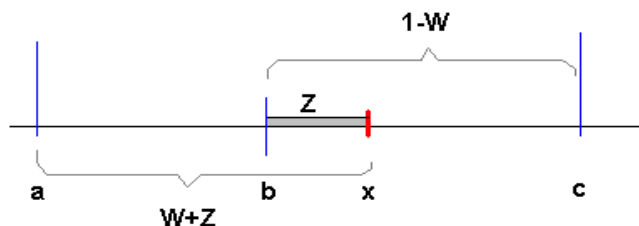
i załóżmy, że X wybrano w $[b,c]$,

oznaczmy

$$\frac{x-b}{c-a} = Z$$

widać (z rysunku), że następny segment zawierający trzy punkty będzie albo

$$W + Z \quad \text{albo} \quad 1 - W$$



W stanowi ułamek całego przedziału $[a,c]$

Z też jest ułamkiem nowego przedziału zawierającego albo (a, b, x) albo (b, x, c) ;

chcemy wyznaczyć ten "wewnętrzny punkt"

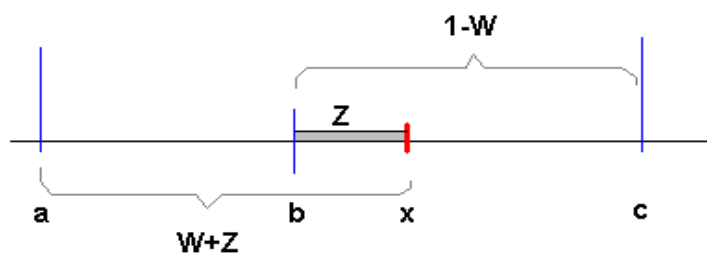
(b dla przedziału $[a,c]$, albo x dla przedziału $[b,c]$ zawsze tak samo, zatem:

$$(**) \quad \frac{Z}{(1-W) \text{ lub } (W+Z)} = \frac{W}{1}$$

z drugiej strony chcemy, żeby X było wybrane tak względem b , żeby obie możliwości

$$W + Z \quad \text{albo} \quad 1 - W$$

były „równo cenne”, tzn: $W + Z = 1 - W$

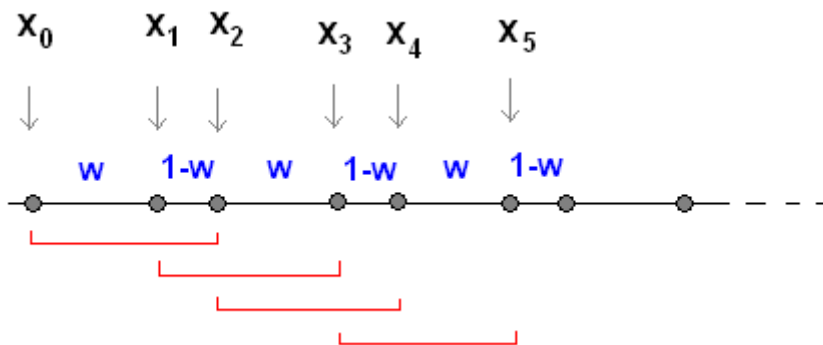


to prowadzi do $Z = 1 - 2W$ i w powiązaniu z (**)
daje:

$$W^2 - 3W + 1 = 0 \quad \Rightarrow \quad W = \frac{3 - \sqrt{5}}{2} \quad (***)$$

ostatecznie:

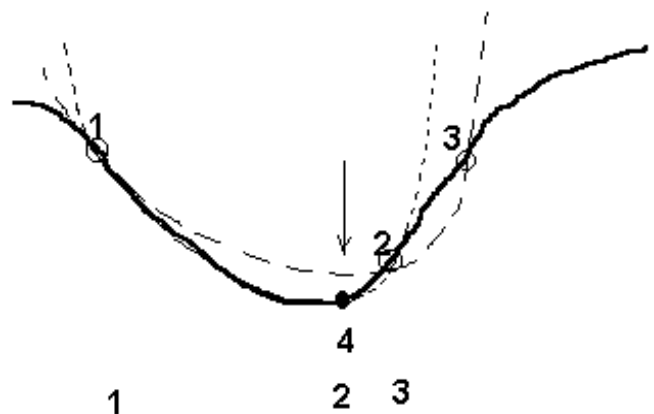
1. trzy kolejne punkty a, b, c wybieramy zawsze wg (***)



2. jeśli spełnione jest (*) to X wybieramy w przedziale $(1-w)$,
tak, żeby a, b, x oraz b, x, c spełniały regułę złotego cięcia (**)

Metoda Brent'a

(interpolacja parabolą)



...program BRENT_GA ...

Jeszcze inny sposób:

szukanie zer funkcji pochodnej

MINIMALIZACJA W WIELU WYMIARACH
Metody gradientowe (najszybszego spadku)

niech $F: \mathbb{R}^n \rightarrow \mathbb{R}^1$

$$F(\mathbf{x}) = F(x_1, x_2, \dots, x_n)$$

oraz $\mathbf{h} = (h_1, h_2, \dots, h_n)$

rozwijając $F(\mathbf{x} + \mathbf{h})$ w szereg Taylora

$$F(\mathbf{x} + \mathbf{h}) = F(\mathbf{x}) + \mathbf{G}(\mathbf{x})^T \mathbf{h} + \frac{1}{2} \mathbf{h}^T \mathbf{H}(\mathbf{x}) \mathbf{h} + \dots$$

gdzie wektor gradientu

$$\mathbf{G}(\mathbf{x})^T = \left(\frac{\partial F(\mathbf{x})}{\partial x_1}, \frac{\partial F(\mathbf{x})}{\partial x_2}, \dots, \frac{\partial F(\mathbf{x})}{\partial x_n} \right)$$

a $\mathbf{H}_{ij}(x) = \frac{\partial^2 F}{\partial x_i \partial x_j}$ elementy macierzy Hessianu

załóżmy, że \mathbf{h} jest wektorem jednostkowym, wówczas

$$\frac{d}{dt} F(\mathbf{x} + t\mathbf{h})|_{t=0}$$

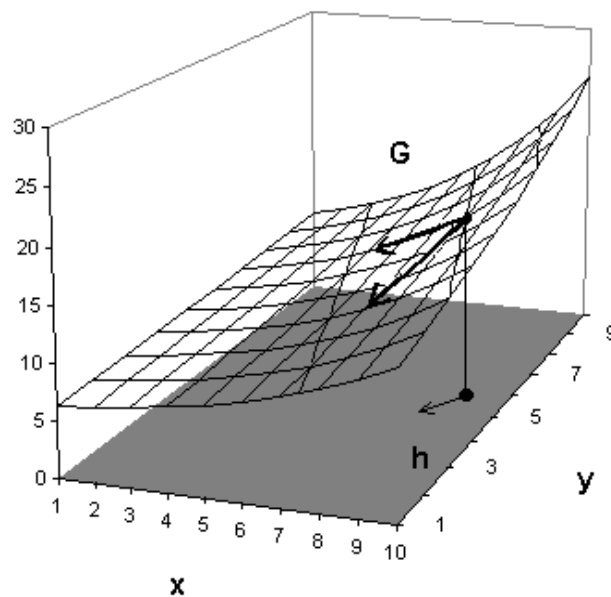
mierzy szybkość zmiany $F(\mathbf{x})$ idąc w kierunku \mathbf{h} ,
i jest równa

$$\mathbf{G}(\mathbf{x})^T \mathbf{h} \quad (*)$$

zapytajmy:

w kierunku jakiego wektora jedn. \mathbf{h} szybkość zmian

F jest największa ?



ponieważ

$$\mathbf{G}(\mathbf{x})^T \mathbf{h} = \sum_{i=1}^n G_i(\mathbf{x}) h_i \quad \text{a} \quad \sum_{i=1}^n h_i^2 = 1$$

to z nierówności Schwarz'a (lub z iloczynu skalarnego):

\mathbf{h} pokrywający się z \mathbf{G} wskazuje kierunek najszybsz. wzrostu

\mathbf{h} pokrywający się z $-\mathbf{G}$ wskazuje kierunek najszybsz. spadku

Szukanie ekstremum to

proces poszukiwania \mathbf{x} w którym $\mathbf{G}(\mathbf{x})=0$ (*)

można sprowadzić do układu równań:

$$\{G_i(x_1, x_2, \dots, x_n) = 0\}_{i=1}^n$$

i rozwiązać wspomnianą wcześniej metodą;

(konieczność obliczania drugich pochodnych cząstkowych z funkcji F)

W metodzie najszybszego spadku poszukiwanie (*) sprowadza się do minimalizacji jednowymiarowej

$$\Phi(t) = F(\mathbf{x} - t\mathbf{G}(\mathbf{x}))$$

dla funkcji Φ jednej zmiennej t

- dla każdego $\mathbf{x}^{(i)}$ znajdujemy $\mathbf{G}(\mathbf{x}^{(i)})$,
 i – numeruje tu krok (iterację) postępowania

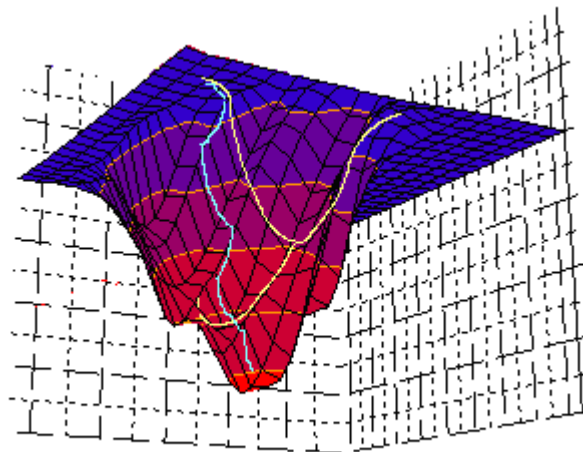
Φ jako funkcja t będzie miała *jakieś* minimum (w kierunku G),
(które nb. może leżeć na brzegu obszaru) – tak
znajdujemy kolejne $x^{(i+1)}$

możliwe są dwa scenariusze:

- A. startując z dowolnego X_i obliczamy $G(X_i)$ i w kierunku $G(x_i)$ poszukujemy ekstremum, np. w X_{i+1} (ozn. Z); dalej powtarzamy procedurę dla Z , tzn. z Z poszukujemy ekstremum w kierunku $G(Z)$, itd...

- B. startując z dowolnego X obliczamy $G(X)$ i w kierunku $G(x)$ posuwamy się $tG(X)$, t - małe, do X' , dalej powtarzamy procedurę dla X' tzn. w X' obliczamy $G(x')$ i w tym kierunku przesuwamy się o $tG(X')$; itd...

w obu przypadkach aż do osiągnięcia $G = 0$ z żadaną dokładnością



Jeśli każda składowa $G_i(x)$ jest liniową funkcją x ,

$$G_i(\mathbf{x}) = \sum_{j=1}^n a_{ij} x_j + b_i$$

to x dla którego $F(x)$ posiada min. znajdujemy od razu z żądania $G_i(x)=0$ co sprowadza się do rozwiązania układu równań liniowych

$$\mathbf{Ax} = \mathbf{b}$$