Počítačová fyzika XI

Minimalizácia

Peter Papp, Ján Ďurian

(Peter Markoš)

Katedra experimentálnej fyziky F2-81

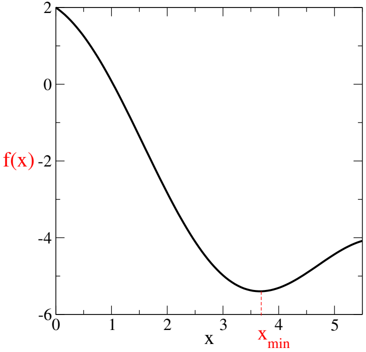
Obsah

1. Funkcia jednej premennej

Daná je funkcia $f(x)$
Chceme nájsť hodnotu $x_{\min}$, v ktorej $f(x)$ nadobúda minimálnu hodnotu.
Ak je funkcia analytická, môžem skúsiť

$$\frac{\partial f(x)}{\partial x} = 0$$

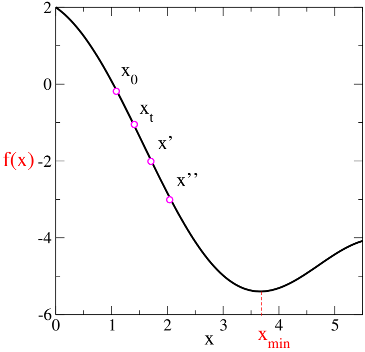
Riešením rovnice je $x_{\min}$
Funkcia ale môže byť zložitá, možno aj nederivovateľná.
Riešenie rovnice pre deriváciu môže byť ťažké.



Východisko: algoritmus minimalizácie funkcie.

1. Algoritmus

1. Zvolíme:
   štartovací bod x_0
   a
   druhý bod x_t.
2. Vypočítame
   $f_0 = f(x_0)$ a $f_t = f(x_t)$
3. ak $f_0 < f_t$ tak vymením
   $x_0 \longleftrightarrow x_t$
   a samozrejme $f_0 \longleftrightarrow f_t$
4. Po kroku 3 je určite $f_0 > f_t$
5. Zvolíme ďalší bod v smere kam funkcia klesá:
   $x' = x_t + (x_t - x_0)$
   a nájdem
   $f' = f(x')$



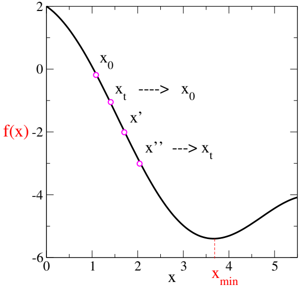
1. Algoritmus – prípad, keď stále klesám

(a) Ak $f' < f_0$ tak sa pohybujeme správnym smerom k minimu.
Skúsime ešte bod $x'' = x_t + 2(x_t - x_0)$ a $f'' = f(x'')$

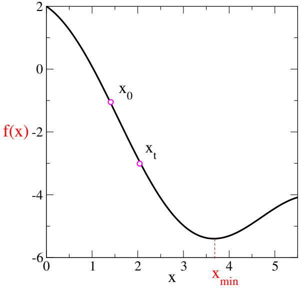
- Ak $f'' < f_0$ tak $x_0 = x''$ a $f_0 = f''$
- Ak $f'' > f_0$ tak $x_0 = x'$ a $f_0 = f'$
- ideme na bod 3.

Čo sa zmenilo po opätovnom vykonaní bodu 3:

pôvodné body x_0 a x_t



nové body:



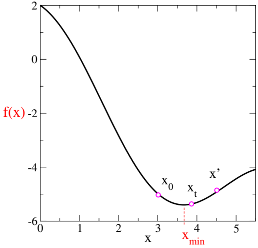
1. Algoritmus – ak sa priblížim k minimu

(b) Ak $f' > f_0$ tak zvolíme nový bod x_0:

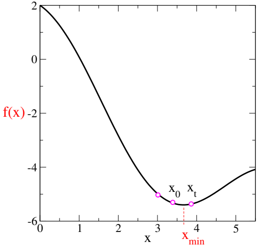
$$x_0 = \frac{x_0 + x_t}{2}, \qquad f_0 = f(x_0)$$

a opať ideme na bod 3.
Všimnime si, že vzdialenosť bodov x_0 a x_t sa zmenšila na polovicu - v okolí minima musíme krok, ktorym sa pohybujeme, zmenšovať, aby sme sa trafili.

pôvodné body x_0 a x_t



nové body:



1. Algoritmus – zastavenie

Algoritmus teda pozostáva z opakovaného hľadania dvojice bodov x_0 a x_t.

Program je potrebné ukončiť. Podmienky zastavenia si môžeme zvoliť:

1. ak f_0 a f_t sú (takmer) rovnaké:

   Ak $|f_0 - f_t| < \varepsilon$ skonči
2. Ak x_0 a x_t sú takmer identické: Ak $|x_0 - x_t| < \varepsilon$ skonči
3. Ak program prekročil povolený počet krokov $N_{\max}$.

Hodnoty ϵ a $N_{\max}$ si zvolíme sami, napr. $\epsilon = 10^{-10}$, $N_{\max} \sim 100$.

1. Ukážka minimalizácie

Program testujeme na najjednoduchšej funkcii s jedným minimom (parabola):

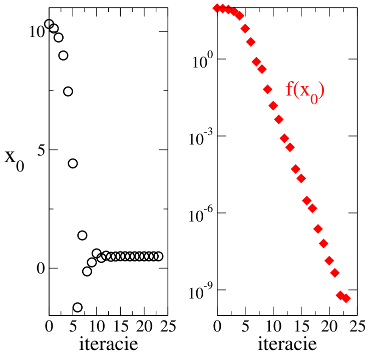
$$f(x) = x(x-1) + \frac{1}{4}$$

má minimum pre

$$x_{\min} = 1/2 \qquad f(x_{\min}) = 0$$

Program končí, ak $|f_0 - f_t| < 10^{-10}$. Nájde minimum po 24 krokoch.

Obrázok ukazuje, ako sa zlepšuje presnosť výsledku s každou ďalšou iteráciou.



1. Komentáre

- Program určite nájde minimum.
- Ak má funkcia viac miním, nie je jasné, ktoré nájde:
  - niektoré môže "preskočiť"
  - môže padnúť do najbližšieho, ktoré ale nie je absolútnym minimom (program sa "nevyhrabe" z jamy)

Riešenie: pustiť program viackrát s rôznymi vstupnými parametrami x_0 a x_t a porovnať, či sme dosiahli to isté minimum.

Iné riešenie: metódy simulovaného žíhania.

2. Funkcia N premenných

Budeme minimalizovať funkciu N premenných,

$$f(\mathbf{s}), \quad \mathbf{s} = (x_1, x_2, \ldots, x_N)$$

Algoritmus je založený na podobnej filozofii, ako pre jednu premennú:
Na začiatku potrebujeme zvoliť:

- vstupný vektor $\mathbf{s}_0$
- N vektorov Δ, ktoré definujú ďalších N bodov $\mathbf{s}_i = \mathbf{s}_0 + \mathbf{\Delta}_i$

$$\mathbf{\Delta}_i(j) = \delta_{ij}\Delta_i$$

- presnosť ϵ a maximálny počet iterácií $N_{\max}$.

Napríklad pre $N = 2$ zvolím $\mathbf{s}_0 = (x_1, x_2)$ a minimalizáciu odštartujem z bodov

$$\mathbf{s}_0 = (x_1, x_2), \quad \mathbf{s}_1 = (x_1 + \Delta_1, x_2), \quad \mathbf{s}_2 = (x_1, x_2 + \Delta_2)$$

2. Funkcia N premenných - Tvorba simplexu

Zvolím $\mathbf{s}_0 = (x_1, x_2, \dots x_N)$, a minimalizáciu odštartujem z bodov

$$\begin{aligned}
\mathbf{s}_0 &= (x_1, x_2, \dots) \\
\mathbf{s}_1 &= (x_1 + \Delta_1, x_2, \dots) \\
\mathbf{s}_2 &= (x_1, x_2 + \Delta_2 \dots) \\
\dots & \quad \dots \\
\mathbf{s}_N &= (x_1, x_2, \dots x_N + \Delta_N)
\end{aligned}$$

2. Funkcia N premenných

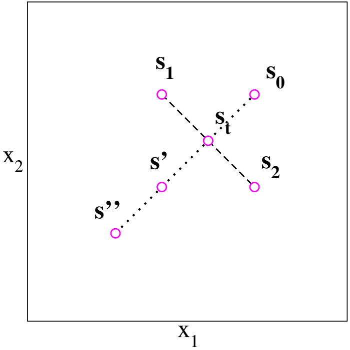
1. Body preorganizujeme tak, aby

$$f(\mathbf{s}_0) > f(\mathbf{s}_1),\ f(\mathbf{s}_2), \ldots f(\mathbf{s}_N)$$

ukážka pre $N = 2$
$\mathbf{s}_0$ je najhorší bod,
ideme od neho preč

2. Nájdeme ťažisko bodov $\mathbf{s}_1$, $\mathbf{s}_2$, $\ldots \mathbf{s}_N$

$$\mathbf{s}_t = \frac{1}{N}\left[\mathbf{s}_1 + \mathbf{s}_2 + \cdots + \mathbf{s}_N\right]$$



3. Nájdeme body

$$\mathbf{s}' = \mathbf{s}_t + (\mathbf{s}_t - \mathbf{s}_0)$$

$$\mathbf{s}'' = \mathbf{s}_t + 2(\mathbf{s}_t - \mathbf{s}_0)$$

a pokračujeme tak, ako pre funkciu jednej premennej:

2. Funkcia N premenných – smerujeme k minimu

V ďalších krokoch porovnávame hodnoty funkcie v bode $\mathbf{s}_0$ a v ťažisku $\mathbf{s}_t$ rovnako ako v prípade $N = 1$:

Ak

$$f(\mathbf{s}') < f(\mathbf{s}_0)$$

tak nahradíme $\mathbf{s}' \to \mathbf{s}_0$

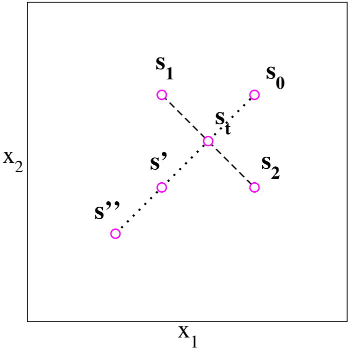
Ak je aj

$$f(\mathbf{s}'') < f(\mathbf{s}')$$

tak nahradíme $\mathbf{s}'' \to \mathbf{s}_0$

Vrátime sa na (1)

$\mathbf{s}_0$ je najhorší bod,
skúsme ho nahradiť
$\mathbf{s}'$ alebo $\mathbf{s}''$



2. Funkcia N premenných

Ak

$$f(\mathbf{s}') > f(\mathbf{s}_0)$$

tak skúsime nový bod

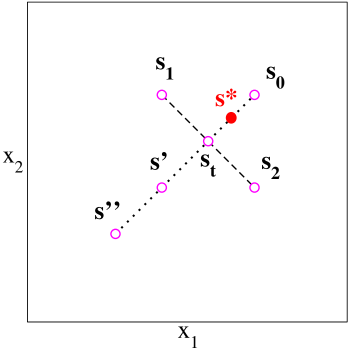
$$\mathbf{s}^* = \frac{\mathbf{s}_0 + \mathbf{s}_t}{2}$$

Ak

$$f(\mathbf{s}^*) < f(\mathbf{s}_0)$$

tak nahradíme $\mathbf{s}^* \rightarrow \mathbf{s}_0$

Vrátime sa na (1)



2. Funkcia N premenných

Ak neuspejeme ani s bodom $\mathbf{s}^*$, zmeníme stratégiu:
Preorganizujeme body $\mathbf{s}_0$, $\mathbf{s}_i$ tak, aby $f(\mathbf{s}_0) < f(\mathbf{s}_1),\ f(\mathbf{s}_2), \ldots f(\mathbf{s}_N)$

1. Nájdeme ťažisko bodov $\mathbf{s}_1$, $\mathbf{s}_2$, $\ldots \mathbf{s}_N$

$$\mathbf{s}_t = \frac{1}{N}\left[\mathbf{s}_1 + \mathbf{s}_2 + \cdots + \mathbf{s}_N\right]$$

2. Nájdeme bod

$$\mathbf{s}' = \mathbf{s}_0 + (\mathbf{s}_0 - \mathbf{s}_t)$$

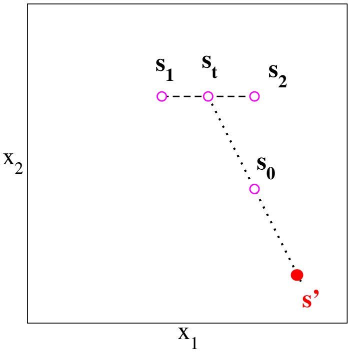
Ak

$$f(\mathbf{s}') < f(\mathbf{s}_0)$$

tak nahradíme $\mathbf{s}' \rightarrow \mathbf{s}_0$

Vrátime sa na (1)

$\mathbf{s}_0$ je najlepší bod
skúsime ísť v smere minima



2. Funkcia N premenných

Ak ani teraz neuspejeme, teda ak

$$f(\mathbf{s}') > f(\mathbf{s}_0)$$

tak položíme

$$\Delta_i \to \Delta_i/2$$

a vrátime sa späť na slide "Tvorba simplexu" - začneme nový cyklus okolo najlepšieho bodu $\mathbf{s}_0$.

2. Funkcia N premenných

Všimnime si, že nárast počtu premenných sa prejaví len na počte bodov, ktoré v každom kroku uvažujeme:

$$\mathbf{s}_0, \quad \mathbf{s}_1, \quad \mathbf{s}_2, \quad \ldots \mathbf{s}_N$$

a na výpočte ťažiska $\mathbf{s}_t$.

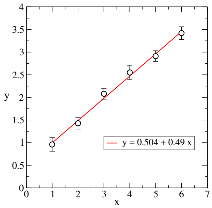
Kritérium, kedy ukončíme proces hľadania:

- ak sa nová funkčná hodnota líši od starej o menej ako ϵ
- ak počet iterácií prekročil nami stanovený limit

3. Fitovanie experimentálnych dát

- Experimentálne (alebo numericky) meriame veličinu $y(x)$.
- Namerali sme $\mathcal{N}$ hodnôt y_i s presnosťou Δ_i.
- Získané hodnoty chceme fitovať známou funkciou s N neznámymi parametrami.

Príklad:



dáta fitujeme lineárnou funkciou

$$y(x) = \alpha_0 + \alpha_1 x$$

s neznámymi parametrami

$$\mathbf{s} = (\alpha_0, \alpha_1)$$

Taký fit "urobí" každý grafický program. Ľahko ho urobíme aj sami metódou najmenších štvorcov.
Niekedy ale potrebujeme viac, napríklad zohľadniť neurčitosť dát, alebo fitovať oveľa komplikovanejšou funkciou.

3. Fitovanie experimentálnych dát metódou minimalizácie

Ak chceme nájsť fitovanú závislosť $y = f_{\mathbf{s}}(x)$ s N neznámymi parametrami $\mathbf{s}$, zostrojme funkciu

$$\mathcal{F}(\mathbf{s}) = \sum_{i=1}^{\mathcal{N}} \frac{[y_i - f_{\mathbf{s}}(x_i)\ ]^2}{\Delta_i^2}$$

a hľadajme, pre ktoré hodnoty $\mathbf{s}$ má funkcia $\mathcal{F}$ minimum.

Výhoda:

- Dátové body x_i, y_i, ktoré sú určené s najhoršou presnosťou (veľká hodnota Δ) ovplyvnia parametre fitu najmenej.
- Funkcia $f_{\mathbf{s}}(x)$ môže mať (takmer) ľubovoľný tvar, napr.

  $$f_{\mathbf{s}}(x) = s_1 + s_2(x - s_3)^{s_4} + s_5 x^{-s_6}$$

- Veličina y môže závisieť od ľuboľného počtu premenných: $y(T, p, x, B)$ (teplota, tlak, poloha, magentické pole ...), ale náročnosť fitovacej procedúry nenarastie.

3. Fitovanie experimentálnych dát metódou minimalizácie

Ak chceme odhadnúť neurčitosť parametrov **s**, môžeme vstupné dáta "zašumiť":
Mnohokrát (napr. $M - 10^4$-krát) zopakujeme proces:

- vygenerujeme náhodné čísla δ_i z intervalu $-\Delta_i < \delta_i < +\Delta_i$
- minimalizujme funkciu

$$\mathcal{F}(\mathbf{s}) = \sum_{i=1}^{\mathcal{N}} \frac{[y_i + \delta_i - f(x_i)\ ]^2}{\Delta_i^2}$$

- uložíme polohu minima **s** do štatistického súboru

Získame takto štatistický súbor parametrov **s**, z ktorých nájdeme strednú hodnotu $\langle \mathbf{s} \rangle$ a odchýlku.
Metóda vyžaduje viac CPU, čo ale pri dnešných počítačoch nie je neprekonateľný problém.

Úloha 12.1

- Napíšte program pre hľadanie minima funkcie jednej premennej
- Program testujte pre funkcie

$$f(x) = x(x-1) + 0.25$$

$$f(x) = \sin^2(\pi x)$$

Ukážte, ako voľba štartovacieho bodu

$$x_0 = 1,\ 2,\ 3,\ 4,\ 21.35\ldots$$

ovplyvní, ktoré minimum funkcie program nájde.

Úloha 12.2

- Napíšte program pre hľadanie minima funkcie N premenných simplexovou metódou
- Program overte pre niektoré jednoduché funkcie, napr.

$$f(x_1, x_2, \ldots, x_N) = \sum_{i=1}^{N} (x - a_i)^2$$

s ľubovoľne zvolenými konštantami a_i.