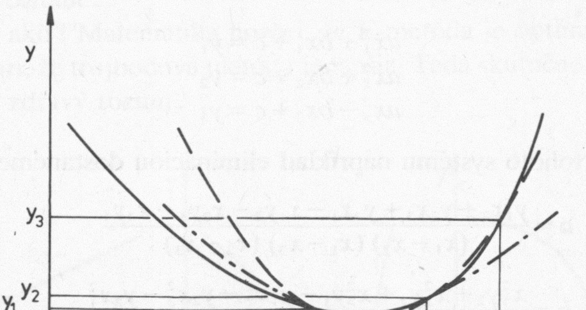
O MINIMÁCH A MAXIMÁCH A O ICH HĽADANÍ

PAVOL BRUNOVSKÝ

IV. Matematika kontra zdravý rozum?

Súhlasím s vami, že po troch pokračovaniach je už na čase skončiť. To som skutočne aj mal v úmysle a určite by ma nebolo napadlo pokračovať, nebyť poznámky Dr. V. Černého (povolaním teoretického fyzika) k druhej časti článku (teda aspoň jeden čitateľ okrem recenzenta!).

O čo išlo? V druhej časti článku sme dokazovali, že istá metóda hľadania minima funkcie využívajúca Fibonacciho čísla (budeme ju ďalej nazývať F-metódou) je v istom presne definovanom zmysle pre istú triedu funkcií optimálna. Obsah poznámky Dr. Černého spočíval v tom, že on používa inú metódu hľadania minima, ktorá sa mu vidí byť lepšia.



Obr. 1. ——— graf minimalizovanej funkcie, - - - - - - graf prvej aproximácie, –.–.–.–.– graf druhej aproximácie

A tak sme urobili zopár experimentov s náhodne vybratými funkciami a ukázalo sa, že skutočne metóda, ktorú spomínal Dr. Černý, dávala pri rovnakom počte krokov asi tri razy lepšiu presnosť.

Tak teda ako? O tamtej metóde sme dokázali, že je optimálna a táto nám dáva lepšie výsledky?

Ide o tzv. trojbodovú metódu, alebo tiež metódu prekladania funkcie. Začne sa tým, že sa hodnota funkcie f spočíta v troch rozličných bodoch x_1, x_2, x_3. Potom sa bodmi (x_i, y_i), $y_i = f(x_i)$, $i = 1, 2, 3$ preloží parabola, ktorá je grafom kvadratickej funkcie

$$y = ax^2 + bx + c \tag{1}$$

a nájde sa bod x_4, v ktorom táto funkcia nadobúda minimum. Zistí sa hodnota $f(x_4)$ funkcie f v bode x_4. Bod x_4 sa doplní k nemu najbližšími dvoma bodmi z trojice x_1, x_2, x_3 a s touto novou trojicou sa procedúra zopakuje; takto sa pokračuje, kým sa nedosiahne požadovaná presnosť *(obr. 1)*.

Na prvý pohľad táto metóda vyzerá dosť komplikovane (musíme prekladať parabolu, počítať minimum atď.). Tieto veci však možno urobiť raz prevždy všeobecne a pre x_4 odvodiť jednoduchý vzorček. Z podmienky, že parabola (1) má prechádzať bodmi (x_i, y_i), $i = 1, 2, 3$, dostávame 3 rovnice s 3 neznámymi a, b, c

$$\begin{aligned} ax_1^2 + bx_1 + c &= y_1 \\ ax_2^2 + bx_2 + c &= y_2 \\ ax_3^2 + bx_3 + c &= y_3 \end{aligned}$$

Riešením tohoto systému napríklad elimináciou dostaneme

$$a = \frac{y_1x_2 + y_2x_3 + y_3x_1 - x_1y_2 - x_2y_3 - x_3y_1}{(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)}$$

$$b = \frac{x_1^2y_2 + x_2^2y_3 + x_3^2y_1 - y_1x_2^2 - y_2x_3^2 - y_3x_1^2}{(x_1 - x_2)(x_1 - x_3)(x_2 - x_3)}$$

konštantu c nebudeme v ďalšom potrebovať.

Ak má kvadratická funkcia (1) mať minimum, musí byť, samozrejme, $a > 0$. To budeme predpokladať a za tohto predpokladu môžeme písať

$$y = ax^2 + bx + c = a(x + b/2a)^2 + c - b^2/4a$$

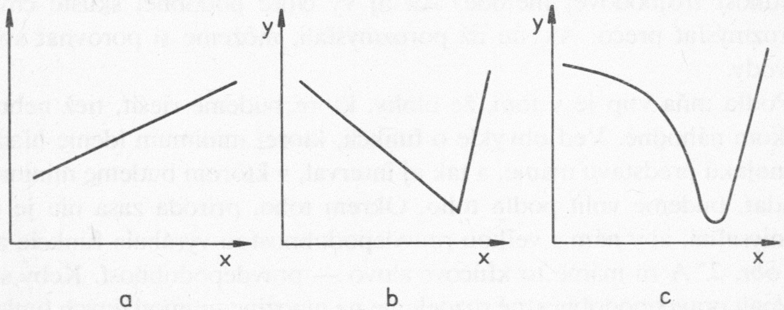
z čoho ihneď vidno, že minimum sa dosahuje pri $x = -b/2a$, a teda

$$x_4 = \frac{x_1^2 y_2 + x_2^2 y_3 + x_3^2 y_1 - y_1 x_2^2 - y_2 x_3^2 - y_3 x_1^2}{2(x_1 y_2 + x_2 y_3 + x_3 y_1 - y_1 x_2 - y_2 x_3 - y_3 x_1)}$$

Nebudeme teraz dokazovať, že trojbodová metóda konverguje k minimu, a že pri dostatočne veľkom počte výpočtov funkcie f sa môžeme k nemu ľubovoľne priblížiť. Je zrejmé, že na to treba urobiť nejaké predpoklady o funkcii f a prvých troch bodoch a dá to kúsok práce. Skôr sa pozrime na to, prečo je táto metóda intuitívne taká príťažlivá, prečo očakávame, že by mala rýchlo viesť k minimu.

Vezmite si kus papiera a načrtnite si naň len tak odruky bez rozmýšľania funkciu, ktorá má minimum. Dopadne to asi nejako tak, ako na *obr. 1.* Ak si na tejto funkcii vymedzíte nie veľmi veľké okolie bodu, v ktorom f dosahuje minimum (označíme ho ako obyčajne $\hat{x}$), bude sa na ňom f veľmi ponášať na parabolu. To značí, že ak si body x_1, x_2. x_3 zvolíte nie veľmi ďaleko od bodu $\hat{x}$, bude funkcia (1) veľmi dobre vystihovať funkciu f, a teda bod x_4 bude veľmi blízko bodu $\hat{x}$. Ďalšia trojica bodov bude teda ešte bližšie k bodu $\hat{x}$, a tak v ďalšom kroku bude kvadratická funkcia (1) vystihovať funkciu f ešte lepšie — a tak ďalej. Na *obr. 1* si môžete pozrieť, ako to bude vyzerať — paraboly na nich nie sú iba tak načrtnuté, ale skutočne spočítané.

Tak teda ako? Matematika hovorí, že F-metóda je optimálna, zdravý rozum hovorí, že trojbodová metóda je lepšia. Teda skutočne — matematika kontra zdravý rozum?



Obr. 2

Ako matematici dúfame, že nie, a preto ideme hľadať chybu. Prvá vec, ktorá nás napadne je, či trojbodová metóda patrí medzi tie, spomedzi ktorých je F-metóda optimálna. Tu však veľmi nepochodíme, lebo si ľahko overíme, že ona medzi ne skutočne patrí. Hodnota n-tej aproximácie x_n sa totiž určuje čiste z hodnôt prvých $n-1$ aproximácií a hodnôt funkcie f v týchto aproximáciách.

Je tu však iná, vážnejšia námietka. Efektívnosť metódy sme testovali na „náhodne vybratých" funkciách. Bol to skutočne náhodný výber? Určite sa v našom výbere viac vyskytovali funkcie, ako z *obr. 1* než rozličné „zvrhlé" funkcie z *obr. 2.* (Je to niečo podobného, ako keby sme si náhodne na papier písali čísla: určite napíšeme viac čísel do stovky, než čísel nad milión.) Všimnime si, že trojbodová metóda nám pre lineárnu funkciu z *obr. 2a* zlyhá hneď v prvom kroku a pre funkcie z *obr. 2b* a *2c* pri istej voľbe počiatočných troch bodov nám zlyhá taktiež (prečo?).

Optimalitu F-metódy sme rozumeli v tom zmysle, že nám dáva pri danom počte výpočtov funkcie f najlepšiu zaručenú presnosť pri hocijako nepriaznivo vybratej unimodálnej funkcii. Ak sa lepšie pozriete do dôkazu optimality, zistíte, že najnepriaznivejšie sú práve tie funkcie, ktoré majú extrém na okraji intervalu. A to sú práve také funkcie, ktoré si „náhodne" človek volí zriedka.

Takže z hľadiska formálne matematického sme paradox dali do poriadku a sme náchylní v tomto momente zdravý rozum v titulnej otázke dať do úvodzoviek a zhovievavo sa pritom pousmiať...

Ale teraz ruku na srdce. Uspokojuje vás celkom tento argument? Mňa celkom nie a neviem, či by som predsa pri riešení konkrétnej úlohy nedal prednosť trojbodovej metóde. Ak aj vy cítite podobne, skúste chvíľu porozmýšľať prečo. Ak ste už porozmýšľali, môžeme si porovnať svoje dôvody.

Podľa mňa vtip je v tom, že úlohy, ktoré budeme riešiť, tiež nebudú celkom náhodné. Veď obvykle o funkcii, ktorej minimum ideme hľadať, už nejakú predstavu máme, a tak aj interval, v ktorom budeme minimum hľadať, budeme voliť podľa toho. Okrem toho, príroda zasa nie je tak zlomyseľná, aby nám s veľkou pravdepodobnosťou vyrábala funkcie ako na *obr. 2.* A tu máme to kľúčové slovo — pravdepodobnosť. Keby sme poznali pravdepodobnostné rozdelenie na množine unimodálnych funkcií, ktoré by nám udávalo, aká je početnosť výskytu jednotlivých typov funkcií v rozumných úlohách (myslím tým úlohy nie umelo vymýšľané), mohli by sme efektívnosť metód porovnávať podľa strednej hodnoty počtu výpočtov funkcie f, potrebných na dosiahnutie predpísanej presnosti a v tomto zmysle definovať optimalitu. Lenže túto pravdepodobnosť nevieme určiť; to nám však nemusí brániť, aby sme pre praktické účely do toho nevložili kus intuície. To je vlastne ten „zdravý rozum", ktorý však je tiež založený na matematickej úvahe. A čo keď nám trojbodová metóda nezaberie vôbec — napríklad keď f je lineárna, alebo konkávna? Nuž, v takom výnimočnom prípade vždy môžeme prejsť na F-metódu a strata, ktorá vznikne neúspešným použitím trojbodovej metódy, sa mnohonásobne vyváži jej úsporami pri iných použitiach.

Takto sa v numerike uvažuje veľmi často. Veď takmer každá numerická metóda v niektorých prípadoch nezaberie, alebo je neefektívna. Napokon, aj F-metóda nemusí viesť k cieľu, ak funkcia nie je unimodálna, čo vopred tiež nemôžeme obvykle s istotou tvrdiť.

Tak čo myslíte, patrí to „kontra" do nadpisu? Nebolo by ho skôr treba nahradiť plusom?