

Optimalizačné a výpočtové metódy – motivácia a obsah

Obsah predmetu je zložený z dvoch pomerne samostatných častí: efektívne výpočty v Rku a stručný prehľad rôznych optimalizačných metód.

Výpočtové metódy

Najprv si ukážeme, ako pomocou simulácií odhadovať rôzne hodnoty (napr. odhad chyby testu, ktorú nevieme vypočítať analyticky), keďže je to častá úloha v bakalárkach a diplomkách, ale je to aj užitočný prístup v praxi.

Potom prejdeme na efektívne výpočty: programovacie jazyky “vyššej úrovne” (high-level programming languages)¹ ako R alebo Python umožňujú omnoho jednoduchšie programovanie ako napr. C-čko, ale cenou za to je, že veľa výpočtov realizujú zbytočne pomaly. Typickým príkladom sú cykly: ak v Rku niečo počítame s cyklami, tak to zvyčajne spôsobuje výrazné spomalenie oproti ne-cyklovému zápisu. Vyskúšajte si napríklad na počítači pustiť príkaz:

```
sum(runif(1e+6))
```

a porovnať ho s rovnakým výpočtom:

```
s <- 0
for(i in 1:1e+6) {
  s <- s + runif(1)
}
```

Uvidíte citelné spomalenie (ak máte rýchly počítač a oba výpočty zbehnú okamžite, nastavte vyššie číslo, napr. $1e+7$).

Na predmete sa teda stručne povenujeme rôznym prístupom, ako písať kódy efektívnejšie (aby bežali rýchlejšie), resp. akým postupom sa vyvarovať. Všetko bude prezentované v Rku, ale podobné správanie je typické aj pre iné jazyky vyššej úrovne.

Optimalizačné metódy

Numerická optimalizácia (teda riešenie maximalizačných alebo minimalizačných úloh na počítači) je kľúčový nástroj, na ktorom stojí obrovské množstvo metód štatistiky (strojového učenia) a poistnej matematiky. Príkladmi sú:

- hľadanie modelu, ktorý čo najlepšie popisuje dáta (minimalizácia chyby; modelovanie úmrtnosti),
- hľadanie modelu, ktorý čo najlepšie predikuje/klasifikuje (minimalizácia predikčnej/klasifikačnej chyby; predikcia výšky škody),
- odhadovanie parametrov rozdelenia (metóda maximálnej vierohodnosti).

Aj keď optimalizačné metódy typicky nepotrebujeme manuálne programovať a stačí sa spoľahnúť na predprogramované funkcie, je dôležité vedieť, ako fungujú, lebo:

- Je zvyčajne na výber veľké množstvo optimalizačných metód a treba si vedieť dobre zvoliť.
- Metódy nie sú stopercentne spoľahlivé a niekedy dajú zvláštny výsledok. Je teda dôležité vedieť správne pochopiť a interpretovať ich výstupy.

¹Nie je úplná zhoda, či jazyky ako R a Python volajú “high level” alebo “very high level”, čo je ale pre potreby tohto predmetu irelevantné.

- Metódy majú rôzne nastavenia, ktoré je niekedy nutné meniť – a nedávajú zmysel bez toho, že by človek chápal, ako metóda funguje.
- Aj predprogramované funkcie, ktoré spúšťajú rôzne štatistické metódy, typicky umožňujú užívateľovi určiť rôzne nastavenia týkajúce sa použitej optimalizačnej funkcie. Ak štatistická metóda nerobí to, čo má, môže to byť riešiteľné aj (niekedy iba) zmenou týchto technických nastavení.

Zhrnieme si, aké sú triedy optimalizačných úloh a aké typické metódy sa používajú na ich riešenie. Stručne si tiež povieme, ako jednotlivé metódy fungujú a ukážeme si, ako sa spúšťajú v Rku.