

Jídelníček pana Stiglera

Seminární práce EKO903

Jindra Lacko

Úvod & motivace

Problém jídelníčku pana Stiglera je jedním z klasických optimalizačních témat. Poprvé byl formulován během 2. světové války (1) – George J. Stigler, pozdější nositel Nobelovy ceny, se jako mladý a začínající výzkumník pokusil zodpovědět otázku, kterak co nejlevněji nakrmit amerického vojáka.

Vycházel přitom z ceny a nutričních informací o 77 potravinových komoditách – počínaje pšeničnou moukou, přes jablka, pomeranče, banány, steak z nízkého roštěnce či lososa v konzervě až k jahodové marmeládě.

Pro jídelníček jako celek přihlížel k doporučeným denním dávkám 9 nutričních složek (kalorie, bílkoviny, vápník, železo, vitamín A, vitamín B₁, vitamín B₂, vitamín B₃, vitamín C) z oběžníku No. 115 National Research Council.

Stojí za pozornost, že v čase, kdy Stigler problém definoval, byla optimalizační věda ještě v začátcích, a nebyl znám algoritmus pro obecnou minimalizaci. Místo obecného řešení proto Stigler aplikoval vlastní heuristiku, pomocí níž dospěl k hodnotě ročních nákladů ve výši 39,93 USD v cenách roku 1939. Jednalo se bezpochyby o nízkou částku, Stigler ale nebyl schopen prokázat, že hodnota představuje globální minimum (dovedl ale prokázat, že globální minimum nemůže být dramaticky nižší).

Když byla o několik let později formulována simplexová metoda, posloužil problém jídelníčku pana Stiglera jako modelový příklad pro ověření přínosu této metody (2). Tým devíti školených počtářů pod vedením J. Ladermana strávil přibližně 120 mandays (tedy necelé tři pracovní týdny) nad mechanickými sčítacími stroji, aby dospěl k hodnotě ročních nákladů 39,69 USD v cenách roku 1939 – tedy zlepšení optima o 24 US centů oproti Stiglerově výpočtu.

Dnes představuje problém jídelníčku pana Stiglera vedle historické zajímavosti středně obtížný problém pro úvod do optimalizačních metod; výpočet v prostředí Google OR Tools (3) s využitím Google Linear Optimization Package dobiehá v řádu středních milisekund. Optimalizace vede k ročním nákladům 39,66 USD – tedy o 3 US centy méně, nežli hodnota z

proof of concept simplexové metody. Tento rozdíl je s největší pravděpodobností způsoben zaokrouhlením.

Metodika

Nutriční hodnoty 77 potravin (v souboru `./OR/stigler-diet.csv`) jsou normalizovány cenou (i.e. jedná se o hodnotu veličiny vztažené k 1 USD komodity, nikoliv k naturální jednotce objemu).

Je tedy třeba minimalizovat funkci:

$$z = x_1 + x_2 + \dots + x_{77}$$

Ve které neznámé x_i představují náklad v USD na i -tou potravní komoditu, za obecných omezujících podmínek:

$$c_{1,j} * x_1 + c_{2,j} * x_2 + \dots + c_{77,j} * x_{77} \geq b_j$$

Kdy koeficienty $c_{i,j}$ představují hodnotu j -té nutriční metriky vztaženou na 1 USD i -té potravní komodity, a hodnota b_j představuje minimální denní dávku pro danou nutriční metriku.

V konkrétním případě minimálního příjmu energie (použité jednotky jsou kilokalorie, tj. jde o denní příjem 3000 cal) se jedná o omezující podmínku:

$$44,7 * x_1 + 11,6 * x_2 + \dots + 6,4 * x_{77} \geq 3$$

S tím, že nerovností pro omezující podmínky bude celkem devět (protože máme devět minimálních denních dávek).

Současně musí pro všechny hodnoty x_i platit:

$$x_i \geq 0$$

Jinými slovy ne všechny potraviny nutně musí být obsaženy ve finálním řešení, ale spotřeba žádné z nich nemůže jít do záporu.

Protože referenční hodnoty příjmu jsou formulovány v *denních* dávkách, a cílová metrika je minimální *roční* náklad, jsou finální výsledky po optimalizaci přepočteny na roční ekvivalent (tj. vynásobeny 365).

Vlastní řešení

Vlastní řešení je provedeno v jazyce python v souboru `./OR/stigler diet jla solution.py`, s tím že v `./OR/stigler diet master solution.py` je referenční zpracování ze stránek Google OR Tools.

Na rozdíl od referenčního řešení jsem:

- tabulku potravinových komodit s nutričními hodnotami na 1 USD jsem vyčlenil do samostatného csv souboru; mít nutriční hodnoty 77 potravin natvrdo v kódu jako array je nepraktické
- tabulku devíti obecných nutričních omezení jsem místo array natvrdo v kódu pro-jížděného for cyklem realizoval jako devítinásobné přidání obecné podmínky; devět podmínek není až *tak* moc (ve srovnání s 77 proměnnými) a duplikace části kódu je v mých očích přijatelná cena za větší přehlednost řešení (stačí mi jeden for cyklus místo dvou vnořených)
- pro kontrolu správnosti formulace problému jsem si uložil model v lp formátu do samostatného souboru; tento lidsky čitelný formát mi byl velkou pomocí pro kontrolu správnosti zadání (na první dobrou jsem to, nikoliv překvapivě, nedal...)

Závěr & poučení

Nalezení řešení problému jídelníčku pana Stiglera nepředstavuje revoluci – problém byl vyřešen před více než 60 lety.

Navíc celý problém představuje svého druhu hříčku s omezeným praktickým využitím. Vědomě abstrahuje od toho, že strava by mohla mít nějakou chuťovou pestrost. Rovněž nepřihlíží k tomu, že příprava jídla vyžaduje určité suroviny z technologických příčin (například tuk pro smažení). Kromě chuti v jídelníčku zcela chybí cukr a sůl.

Na stranu druhou jsem dokázal replikovat historicky významný optimalizační problém, dopočetl jsem se vlastní cestou na známé výsledky referenčního řešení – a naučil se přitom pracovat s Google OR-Tools coby optimalizačním frameworkem, GLOP coby solverem pro floating point optimalizaci, SCIP coby solverem pro celočíselnou optimalizaci a prohloubil jsem si své povědomí o programovacím jazyce python a práci s vývojovým prostředím Visual Studio Code.

Literatura

1. STIGLER, George J. The Cost of Subsistence. *Journal of Farm Economics*. Online. 1945. Vol. 27, no. 2, p. 303–314. [Accessed 2 January 2023]. DOI [10.2307/1231810](https://doi.org/10.2307/1231810).
2. DANTZIG, George B. The Diet Problem. *Interfaces*. Online. 1990. Vol. 20, no. 4, p. 43–47. [Accessed 2 January 2023]. Available from: <https://www.jstor.org/stable/25061369>
3. *Linear Optimization / OR-Tools / Google Developers*. Online. [Accessed 2 January 2023]. Available from: <https://developers.google.com/optimization/lp>