

László Lovász

Je nový algoritmus lineárního programování lepší nebo horší než simplexová metoda?

Pokroky matematiky, fyziky a astronomie, Vol. 26 (1981), No. 4, 193--202

Persistent URL: <http://dml.cz/dmlcz/139010>

Terms of use:

© Jednota českých matematiků a fyziků, 1981

Institute of Mathematics of the Academy of Sciences of the Czech Republic provides access to digitized documents strictly for personal use. Each copy of any part of this document must contain these *Terms of use*.



This paper has been digitized, optimized for electronic delivery and stamped with digital signature within the project *DML-CZ: The Czech Digital Mathematics Library* <http://project.dml.cz>

Je nový algoritmus lineárního programování lepší nebo horší než simplexová metoda?*)

L. Lovász

Sovětský matematik L. G. Čačijan uveřejnil v únoru 1979 v Dokladech Akademie věd SSSR nový algoritmus pro řešení problémů lineárního programování. Výsledek vyvolal senzaci a zprávu o něm přinesly mimo jiné *Science magazine*, *The New York Times* a *The Manchester Guardian*. Taková publicita nemá u matematického výsledku téměř obdoby. Pro řešení lineárních programů však známe poměrně efektivní metodu, tzv. simplexovou metodu. Její vylepšení nebo nalezení trochu lepší alternativní metody by přece nemohly vyvolat takovou senzaci. V čem je tedy příčina tak velkého nadšení?

Problém

Kdybyste si sestavili statistiku toho, jak je ve světě rozložen počítačový čas mezi jednotlivé matematické problémy, pak (nebudete-li uvažovat problémy manipulace s databází jako je třídění a vyhledávání) by se na první místo zřejmě dostalo lineární programování. Problém je skutečně velmi přirozený a jednoduchý: je-li dána množina nerovností (omezení)

$$(1) \quad a_i x \leq b_i \quad (i = 1, \dots, m),$$

{kde všechna a_i jsou reálné n -rozměrné vektory, všechna b_i jsou reálná čísla a $x = (x_1, \dots, x_n)^T$ je neznámá), nalezněte maximum výrazu

$$(2) \quad cx = c_1 x_1 + \dots + c_n x_n,$$

který splňuje uvedená omezení.

V dalším textu budeme předpokládat, že koeficienty a_{ij} , b_i a c_j jsou celá čísla. Pro a_{ij} , b_i , c_j racionální můžeme tento předpoklad učinit bez újmy na obecnosti (a uvažovat jiné než racionální koeficienty nemá ve výpočtech smysl). Vektor x , který splňuje (1), se nazývá *přípustný*. Jednodušší verze problému požaduje pouze nalezení přípustného vektoru. Obě dvě verze (a několik dalších) jsou ekvivalentní v tom smyslu, že je-li známa metoda pro řešení jedné z nich, lze jednoduše nalézt obdobně efektivní metodu pro řešení druhé, která používá první metodu jako podprogram.

*) L. LOVÁSZ: *A New Linear Programming Algorithm — Better or Worse Than the Simplex Method?* The Mathematical Intelligencer, Vol. 2, No. 3, 1980, pp. 141—146. Přeložila H. NEŠETŘILOVÁ. Copyright © Springer-Verlag Berlin—Heidelberg—New York 1980.

Širší souvislosti problematiky tohoto článku jsou uvedeny v článku J. NEŠETŘILA a S. POLJAKA: *Algebraické a geometrické souvislosti kombinatorické optimalizace*. Sborník SOFSEM '80, str. 35—77.

Některé odpovědi a další problémy

Je překvapivé, že tak přirozený a elementární problém, jako je třeba nalezení řešení (1), nebyl formulován už před dvěma sty nebo třemi sty lety. Nutnou a postačující podmínku pro řešitelnost (1) našel J. Farkas v roce 1902, viz [7]. S tím související výsledek pro optimalizační problém, známá věta o dualitě, byl dokázán Kantorovičem (1939) a von Neumannem (1947), viz [9] a [14]. Jeden způsob, jak dokázat větu o dualitě, vede přes analýzu simplexového algoritmu. Tento algoritmus lze použít k řešení obou problémů, optimalizace i přípustnosti; našel jej Dantzig (1951), viz [4]. Od té doby bylo navrženo mnoho různých vylepšení algoritmu, v tomto směru se však zde nebudeme pouštět do podrobností.

Abychom vysvětlili přednosti a nedostatky simplexové metody a věty o dualitě, formulujeme a rozebereme nejprve větu o dualitě. Sestavme nový program, *duální* k (1) až (2): uvažme m neurčitých y_1, \dots, y_m , omezení

$$(3) \quad y_1, \dots, y_m \geq 0$$
$$y_1 a_1 + \dots + y_m a_m = c$$

a výraz

$$(4) \quad by = b_1 y_1 + \dots + b_m y_m,$$

který splňuje tato omezení. Věta o dualitě potom tvrdí:

Věta: Minimální hodnota výrazu (4), který splňuje (3), je rovna maximální hodnotě výrazu (2), který splňuje (1).

Program (3)–(4) vypadá trochu jinak než náš původní program (1)–(2), to však není podstatné. Oba programy se dají snadno přepsat do tvaru, který je obdobou druhého z nich. Co nového nám však v tom případě dává věta o dualitě? Proč bychom ji prostě neměli považovat za přeformulování původního problému?

Odpověď se dá ozřejmit pomocí tzv. principu šéfa. Představte si, že za vámi přijde váš šéf a požádá vás, abyste našli maximum výrazu (2), který splňuje (1). Nepospíchá na vás, dokonce i způsob řešení si můžete zvolit sami; můžete hádat, používat počítač, provádět fyzikální experimenty atd. Když jste hotovi, oznámíte svému šéfovi, že hledané maximum je nějaké číslo z . Jeho však nezajímá, jak jste číslo z našli, chce mít pouze jistotu, že odpověď, kterou jste mu dali, je správná. Uvedete-li současně také přípustný vektor x takový, že $cx = z$, pak má alespoň tu jistotu, že maximum výrazu (2) není menší, než tvrdíte vy. Nemůže se však bezprostředně přesvědčit o tom, že neexistuje žádné lepší řešení. Věta o dualitě mu však umožní rychlou kontrolu: uvedete-li zároveň také řešení y , které splňuje (3) a minimalizuje výraz (4), tj. takové, že $b_1 y_1 + \dots + b_m y_m = z$, pak se může o správnosti vašeho řešení přesvědčit tak, že si ověří, je-li x přípustné pro (1), y přípustné pro (3) a $cx = by = z$.

Formulujme nyní princip šéfa přesně.

Racionální funkci $f(X)$ definovanou na konečných posloupnostech X nul a jedniček nazveme *dobře charakterizovanou*, můžeme-li rychle *dokázat* $f(X) = y$ (v případě, že tvrzení platí). „Rychle“ přitom znamená, že délka důkazu musí být omezena (pevnou)

mocninou délky posloupnosti X . Zakódujeme-li parametry problému lineárního programování tak, že každému konkrétnímu zadání odpovídá určitá posloupnost X nul a jedniček (přesný způsob, jak to udělat, je více méně nepodstatný, ale předpokládáme, že taková korespondence je vzájemně jednoznačná), pak maximální hodnotu výrazu (2) dosažitelnou při podmínkách (1) můžeme označit $f(X)$, kde X je příslušná posloupnost nul a jedniček. Na základě výše uvedené testovací metody (užité šéfem) se snadno ukáže, že *funkce $f(X)$ je dobře charakterizována*.

Princip šéfa říká, že je účelné hledat dobrou charakterizaci (má-li šéf pochybnost o správnosti řešení a navíc málo času).

Tento aspekt, který byl ilustrován na principu šéfa, je třeba vzít v úvahu. Zůstává však jasné, že dobrá charakterizace nenahradí algoritmus. Známa simplexová metoda dokáže nalézt řešení (1), pro které výraz (2) nabývá svého maxima a současně i řešení (3), pro které výraz (4) nabývá minima. *Empirická data* ukazují, že čas potřebný k výpočtu pomocí simplexové metody je úměrný mn^3 , a je tedy možné efektivně používat tuto metodu i na poměrně velké lineární programy. Navíc je to metoda velmi pružná, na několika místech máme možnost volit různá pokračování a zavedeme-li další speciální pravidla, dostaneme různé verze simplexové metody. Vzhledem k těmto dobrým vlastnostem se mělo zato, že čas potřebný k výpočtu pomocí simplexové metody (nebo alespoň některé její verze) by mohl být *ve všech případech* kratší než nějaká mocnina n a m . V. Klee a G. Minty však v roce 1972 zkonstruovali příklad (viz [11]), na který simplexová metoda potřebuje více než 2^n kroků při $m = 2n$ nerovnostech a n proměnných. Zdá se, že všechny verze simplexové metody, které byly od té doby navrženy, spotřebují v některém nepříznivém případě exponenciálně rostoucí čas. (Je však také možné, že existuje nějaká verze, o které se zatím neuvažovalo a která vyžaduje pouze polynomiální čas.)

Příklad, který zkonstruovali Klee a Minty tedy ukazuje, že i když je simplexová metoda efektivní v převážné většině případů, v některých umělých případech efektivní být nemusí. Máme se však vůbec pokoušet zrychlit algoritmus i pro tyto výjimečné případy? Jsou pro to jiné důvody než čistě teoretické?

Na uvedené otázku lze odpovědět hned dvěma způsoby:

1. Skutečnost, že simplexová metoda je efektivní „téměř ve všech“ případech, se vztahuje na tu distribuci lineárních programů, která vyplývá ze *současné* praxe. Je však možné, že brzy budeme chtít řešit lineární programy jiných matematických modelů a tím distribuci lineárních programů změňme. „Špatné“ případy, které jsou dnes řídké, se možná začnou vyskytovat častěji. Je například známo, že simplexová metoda není příliš vhodná pro lineární programy týkající se úloh o „rozkladech množin“ (koeficienty v (1) jsou v tom případě nuly a jedničky). Problémy tohoto typu jsou strukturovány a zdá se, že jejich struktura není pro simplexovou metodu vhodná. Člověk, který často řeší lineární programy pro úlohy týkající se rozkladů množin, určitě ocení rozdíl mezi algoritmem, který je „téměř vždycky rychlý“ a „vždycky rychlý“.

2. Druhá odpověď na výše uvedené otázky souvisí s vědou o počítačích. Existuje řada „dobře charakterizovaných funkcí“, které vznikly na základě různých zajímavých

kombinatorických problémů. Jako příklad můžeme uvést párovací číslo grafu. V daném grafu je definováno jako maximální počet hran, které nemají společný vichol. Výsledek, který získali Tutte (1947) a Berge (1958), tvrdí, že (vyjádřeno pomocí naší terminologie) toto číslo je dobře charakterizováno (viz [16] a [1]). V době o více než deset let později našel J. Edmonds [5] algoritmus, který určí párovací číslo n -bodového grafu v nejvýše $O(n^3)$ krocích. Nalezení Edmondsova algoritmu bylo podstatně obtížnější než důkaz Bergeova vzorce, ale hlavní myšlenka, o kterou se opírá, tzv. střídavé cesty, byla použita už při důkazu Bergeova vzorce a dokonce ještě dřív.

Velmi podobnou historii měly i další dobře charakterizované funkce. Nejdříve byla dokázána charakterizační věta. Její důkaz byl zpravidla obtížný, ale velmi užitečný pro studium dané funkce. Potom se hledaly algoritmy pro výpočet takové funkce a ve všech známých případech tyto algoritmy vyžadovaly exponenciální čas. Klademe si tedy přirozenou otázku: *je možné spočítat každou dobře charakterizovanou funkci v polynomiálním čase?*

Kladná odpověď na tuto otázku by měla obrovský význam. Znamenala by totiž, že při hledání efektivního způsobu výpočtu dané funkce stačí nalézt dobrou charakterizaci. Ze zkušenosti víme, že tento úkol je mnohem jednodušší. Tak obecný a silný pozitivní výsledek by v teorii algoritmů neměl téměř obdoby.

Nic však obvykle nejde tak pěkně a jednoduše, jak bychom si přáli. Všeobecně se má proto spíše zato, že odpověď je negativní. Avšak už jenom výběr vhodné funkce, která by měla naději posloužit jako protipříklad, je obtížný. Zdá se, že dnešní metody *k důkazu* polynomiální neřešitelnosti nějakého problému nestačí, mohou však alespoň naznačit, že je to *pravděpodobné*. Ještě nedávno byly známy dvě dobře charakterizované funkce, pro které nebyl znám polynomiální algoritmus: optimum lineárního programu a charakteristická funkce prvočísel.

V druhém případě jde o nalezení algoritmu, který by pro dané číslo dokázal rozhodnout, je-li prvočíslo, v čase, který je polynomiální funkcí počtu míst v zápise čísla. Tento problém se však jako protipříklad výše uvedeného obecného problému nehodí, byly pro něj totiž vypracovány algoritmy (G. Miller [13]), které jsou pravděpodobně polynomiální, jenomže my to o nich nevíme (přesněji řečeno, algoritmus je polynomiální, platí-li zobecněná Riemannova hypotéza).

Nedávný Čačijanův výsledek tvrdí, že každý lineární program je řešitelný pomocí algoritmu, jehož čas je polynomiální vzhledem k velikosti vstupu. I tento zamýšlený protipříklad na výše uvedenou hypotézu je tedy polynomiálně řešitelný.

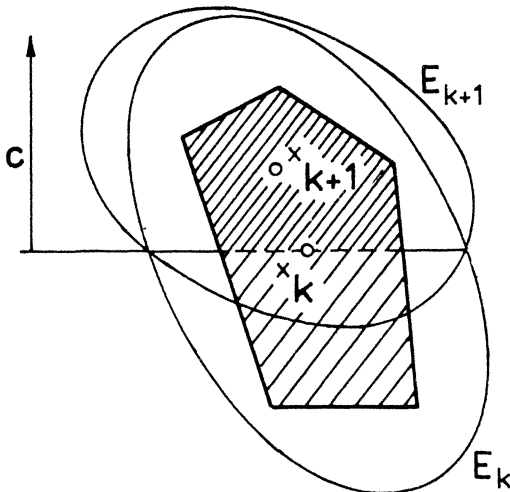
To vysvětluje, proč je Čačijanův výsledek z teoretického hlediska tak důležitý: zdá se, že dobře charakterizované problémy jsou (případně za nějakých rozumných předpokladů) polynomiálně řešitelné. V současné době alespoň neznáme žádný problém, který by mohl být protipříkladem.

Metoda

Nová metoda, které budeme říkat elipsoidová metoda, je adaptací dřívějších prací Šora [15] a Judina a Nemirovského [8] na lineární programování.

Označme P množinu řešení nerovností (1). Chceme nalézt maximum lineární účelové funkce cx na P . Předpokládejme, že P je kompaktní množina dimenze n a předpokládejme dále, že známe horní odhad K norem prvků množiny P (tyto předpoklady rozebereme později). Nechť $x_0 = (0, \dots, 0) \in R^n$, označme E_0 kouli o poloměru K se středem x_0 , tedy $E_0 \supset P$. Algoritmus je založen na konstrukci posloupnosti bodů x_0, x_1, \dots v R^n a posloupnosti elipsoidů E_0, E_1, \dots , které získáme následujícím rekursivním postupem: předpokládejme, že x_k a E_k jsou již zkonstruovány. Ověříme, je-li x_k přípustné. V případě že není, vezmeme tu podmínku $a_i x \leq b_i$, která byla porušena, a uvážíme polovinu elipsoidu, který je průnikem E_k a poloprostoru $a_i x \leq a_i x_k$. Tuto polovinu elipsoidu vnoříme do nového elipsoidu s nejmenším možným objemem. Nový elipsoid označíme E_{k+1} a jeho střed x_{k+1} .

Pro x_k přípustné postupujeme stejně jen s tím rozdílem, že uvážíme průnik E_k s poloprostorem $cx \geq cx_k$ (viz obr. 1).



Obr. 1. Konstrukce nového elipsoidu.

Přípustné body x_k aproximují optimální řešení. Přesněji: necht' t_p je nejlepší hodnota nalezená uvedeným způsobem v prvních p krocích, tj.

$$t_p = \max \{cx_k; 0 \leq k \leq p, x_k \text{ přípustné}\}.$$

Není těžké nahlédnout, že elipsoid E_p obsahuje mnohostěn $\{x \in P : cx \geq t_p\}$. Protože objem $\lambda(E_p)$ elipsoidu E_p se zmenšuje exponenciálně,

$$\lambda(E_p) < (2K)^n \exp\left(-\frac{p}{2n+2}\right),$$

musí se exponenciálně zmenšovat i objem tohoto „odřezku“. Z toho plyne, že $t_p \rightarrow z = \max \{cx : c \in P\}$ exponenciálně rychle.

Jak z posloupnosti aproximací dostaneme optimální řešení? Nebudeme zde zabíhat do podrobností, které sice nejsou obtížné, ale jsou poněkud zdoluhavé. Uvedeme jeden způsob. Pozměňme nepatrně účelovou funkci tak, aby výraz (2) nabýval optimální

hodnoty v jediném vrcholu v mnohostěnu P^*). Souřadnice tohoto vrcholu jsou racionální čísla. Není také těžké nalézt horní odhad Q pro jmenovatele těchto racionálních čísel. To plyne ze skutečnosti, že tito jmenovatelé jsou determinanty vytvořené z vektorů a_i . Nalezneme-li tedy přípustné x_k , pro které $|v - x_k| < 1/2Q^2$, pak vrchol v dostaneme zaokrouhlením jednotlivých složek x_k na nejbližší racionální číslo, jehož jmenovatel je nejvýše Q .

Vrátíme se nyní k předpokladům, které jsme museli udělat o P . Všimněte si, že souřadnice vrcholů P se dají zapsat jako poměr dvou determinantů vytvořených z koeficientů systému nerovností (1). Na základě toho lze snadno určit horní odhad T těchto souřadnic.

Vytvoříme-li nyní průnik mnohostěnu P s krychlí $C = \{(x_1, \dots, x_n) : 0 \leq x_i \leq T\}$, dostaneme omezený mnohostěn P' takový, že každý vrchol P je zároveň také vrcholem P' . Necht' cx je účelová funkce, kterou chceme maximalizovat na P . Zase nepatrně pozměníme c tak, abychom mohli předpokládat, že nabývá-li vůbec účelová funkce maxima, pak ho nabývá v jediném vrcholu P . Vyšetřujeme-li nyní stejnou účelovou funkci na P' , zjistíme buď, že nabývá maxima ve vrcholu mnohostěnu P' , který leží uvnitř krychle C (v tom případě je to také vrchol mnohostěnu P a příslušné maximum je zároveň maximum účelové funkce na P), nebo že vrchol mnohostěnu P' leží na povrchu C (v tom případě není účelová funkce na P omezená). Je-li tedy možné vyřešit problém lineárního programování pro omezené mnohostěny, je možné jej vyřešit i pro libovolný mnohostěn.

Požadavek, aby P mělo dimenzi n , lze splnit podobným trikem; podrobnosti přenecháme čtenáři.

Může mít metoda praktický význam?

Na první pohled se zdá, že elipsoidová metoda má mnoho nepraktických rysů. Zkusme tedy prozkoumat některé otázky, které se bezprostředně nabízejí:

Není výpočetní stránka použitých geometrických konstrukcí příliš komplikovaná? Jak provádět zaokrouhlovací postupy, které vedou k nalezení optimálního řešení? Algoritmus vyžaduje, aby výpočty byly prováděny s vysokou přesností (i pro nepříliš velké problémy je třeba počítat s přesností několika set míst). Není to příliš těžkopádné?

Při každé iteraci musíme do paměti uložit údaje o elipsoidu, to však pokaždé představuje n^2 čísel, z nichž každé má mnoho míst. Není to mnoho?

Doba, kterou algoritmus potřebuje k výpočtu, je $O(n^5 m)$, zatímco pro simplexovou metodu je v praxi tato doba $O(n^3 m)$. V čem je tedy výhoda?

I když na některé z těchto otázek bude možné odpovědět až po mnoha experimentech a dalším výzkumu, něco k nim přece jenom lze říci už nyní.

1. Potřebné geometrické konstrukce jsou po aritmetické stránce velmi jednoduché. Elipsoid E_k popíšeme tak, že udáme jeho střed x_k a pozitivně definitní symetrickou

*) Účelová funkce nabývá na P optimální hodnoty vždy alespoň v jednom vrcholu (pozn. překl.).

čtvercovou matici A_k řádu n takovou, že

$$E_k = \{x : (x - x_k)^T A_k^{-1} (x - x_k) \leq 1\} .*$$

Matice A_k lze vypočítat pomocí těchto rekurentních vztahů

$$(5) \quad x_{k+1} = x_k - \frac{1}{n+1} \cdot \frac{A_k a}{\sqrt{(a^T A_k a)}},$$

$$A_{k+1} = \frac{n^2}{n^2 - 1} \left(A_k - \frac{2}{n+1} \frac{(A_k a)(a^T A_k)}{a^T A_k a} \right),$$

(kde $a = -c$ pro x_k přípustné a $a = a_i$, když je pro x_k porušena podmínka $a_i x \leq b_i$). Všimněte si, že ve vzorcích (5) se neobjevují ani inverzní matice ani maticové násobení. K tomu, abychom z x_k a A_k dostali x_{k+1} a A_{k+1} , potřebujeme $O(n^2)$ aritmetických operací. Vzorce jsou tak jednoduché, že i na kapesní kalkulačce lze řešit problémy až s 6 proměnnými. (Zkuste si totéž pro simplexovou metodu!)

2. Přesné optimum budete v praxi potřebovat jen zřídka. Většinou postačí nalézt řešení, jehož chyba je menší než řekněme 10^{-10} . Stačí tedy vypočítat vektor x_k a celý trik se zaokrouhlováním bude zajímavý hlavně z teoretického hlediska. Pokud však skutečně potřebujete znát přesné optimální řešení, můžete zaokrouhlování provést pomocí metody řetězových zlomků.

3. V současnosti jsou pouze malé zkušenosti s chováním elipsoidové metody na počítačích, je proto obtížné odhadnout, jaká přesnost se při výpočtech *skutečně* potřebuje. Všimněte si, že vzhledem k druhé odmocnině ve vzorcích (5) je třeba zaokrouhlovat. Za současného stavu metoda funguje pouze tehdy, počítáme-li s přesností n^4 míst, podrobnější analýza však toto číslo určitě sníží. Osobně si myslím, že i když výsledek potřebujeme znát třeba jen na 10 míst, prvky matice A_k a další částečné výsledky je třeba spočítat a uložit do paměti s mnohem větší přesností, která se pro rostoucí n pravděpodobně blíží nekonečnu. Tohle vypadá jako seriózní nedostatek. Jestliže se však metoda jinak osvědčí, vyplatilo by se možná změnit počítačový hardware tak, aby výpočty s dlouhými čísly byly snazší. (Také další problémy, například různá nedávno navržená kryptografická schémata – viz článek G. H. Simmonse v tomto čísle *Intelligenceru***) – vyžadují, aby výpočty byly prováděny s čísly, která mají mnoho set míst.)

4. Možná, že nebude třeba ukládat do paměti údaje o příslušném elipsoidu v každém kroku. Mohli bychom to zkusit s jinými ekvivalentními údaji, tak například elipsoid E_k je jednoznačně určen posloupností $(K; x_0, \dots, x_k)$. V tomto směru je však ještě třeba dalšího studia.

5. Jak již bylo řečeno, máme zatím k dispozici jen velmi málo údajů o tom, jak se nová metoda chová v *praxi*. Odhad $O(n^5 m)$ je odhad pro nejhorší možný případ. Simplexová metoda pracuje v nejhorším případě mnohem pomaleji (s exponenciálním

*) Zde a ve vzorcích (5) jsou vektory pokládány za sloupce, tj. matice typu $(n, 1)$. (Pozn. překl.)

**) G. H. SIMMONS: *Cryptology: The Mathematics of Secure Communication*. M. I. vol. 1, No. 4, 1979, p. 233–246. (Pozn. překl.)

časem). Nesmíme zapomenout ani na to, že simplexová metoda má za sebou třicetiletý vývoj, měli bychom tedy dát stejnou příležitost i elipsoidové metodě. Mohlo by se dokonce ukázat, že „tím pravým“ bude vhodná kombinace obou metod.

Tuto část ukončíme jednou negativní poznámkou. Zajímavou (ale negativní) vlastností elipsoidové metody je to, že počet potřebných aritmetických operací závisí na tom, kolik míst mají koeficienty. Simplexová metoda se v tomto směru chová jinak, u ní je počet iterací omezen funkcí n a m (mají-li koeficienty mnoho míst, ovlivní to samozřejmě i simplexovou metodu, protože se zvětší čas potřebný k provedení jedné aritmetické operace). Chceme-li tento rozdíl formulovat přesněji, můžeme říci, že simplexová metoda pracuje nad libovolně uspořádaným tělesem v tom smyslu, že máme-li k dispozici způsoby, jak ukládat, sčítat, odčítat, násobit, dělit a porovnávat prvky tohoto tělesa, pak nad tímto tělesem můžeme maximalizovat lineární účelovou funkci, která splňuje lineární omezení. Elipsoidová metoda však pracuje jen tehdy, je-li příslušné uspořádané těleso archimedovské (protože se vzrůstajícím k se objem E_k musí blížit k 0). Neznám sice žádný dobrý příklad nearchimedovského uspořádaného tělesa, nad kterým bychom případně mohli chtít řešit lineární programy, kdyby se však takový problém objevil, toto teoretické omezení se elipsoidové metodě stane osudným.

Není to tak úplně lineární programování

Není těžké si všimnout, že elipsoidová metoda téměř nikde nevyužívá tu skutečnost, že přípustná oblast P je mnohostěn. Důležité je pouze to, že P je konvexní a omezená a že pro každý bod, který není přípustný, lze najít nadrovinu, která jej odděluje od P . Efektivní podprogram, který to dokáže, může existovat dokonce i v případě, že P není mnohostěn. Jako příklad můžeme uvést množinu všech pozitivně semidefinitních symetrických matic řádu n s jednotkovou stopou v n^2 dimenzionálním eukleidovském prostoru všech matic řádu n . (Rozpoznat pozitivně semidefinitní matice a „oddělit“ od nich matice, která pozitivně semidefinitní není, je jednoduché cvičení z klasické lineární algebry.) Toto jednoduché pozorování umožňuje použít elipsoidovou metodu k řešení známého a (dříve) neřešeného problému kombinatorické optimalizace, totiž k určení maximálního počtu nezávislých bodů v perfektním grafu v polynomiálním čase. (O perfektních grafech viz Berge [2], funkci \mathfrak{g} , kterou je třeba spočítat, zavedl Lovász [12].)

Elipsoidová metoda, jak se zdá, přinese některé revoluční změny v celé teorii kombinatorické optimalizace. Většinu problémů kombinatorické optimalizace lze formulovat takto: je-li dána konečná množina S vektorů v R^n a lineární účelová funkce cx , najděte maximum cx pro $x \in S$. Množina S je většinou bohatě strukturována, protože představuje nějaký kombinatorický objekt. Podle povahy tohoto kombinatorického objektu dostáváme různé známé problémy operačního výzkumu. Uvažme například pro daný graf G s hranami e_1, \dots, e_n množinu S všech n -tíc $(x_1, \dots, x_n) \in \{0, 1\}^n$ takových, že kdykoliv $x_i = x_j = 1$, $i \neq j$, pak hrany e_i, e_j nemají společný vrchol. Je-li navíc $c = (1, \dots, 1)$, pak $\max \{cx : x \in S\}$ je párovací číslo daného grafu (viz výše). Tímto způsobem lze formulovat i některé další známé problémy, např. problém obchodního cestu-

jícího, přiřazovací problém, problém pakování atd. Stanovení algoritmu, který by dokázal určit maximum cx je samozřejmě triviální: vypočítáme cx pro každé $x \in S$ (S je konečná). S však může být ve srovnání s n exponenciálně velká a my bychom chtěli, aby algoritmus byl v n polynomiální.

Obecný přístup, zkoumaný především Fordem, Fulkersonem, Hoffmanem a Edmondsem, je tento (podrobněji viz Chvátal [3]): označíme-li P konvexní obal množiny S , můžeme účelovou funkci cx maximalizovat také na P , protože cx nabývá optima v nějakém vrcholu a tento vrchol leží samozřejmě v S . Tím však dostáváme problém lineárního programování. Umíme-li nějak určit množinu nerovností, které definují S , pak nám zbývá jen problém lineárního programování, který umíme efektivně vyřešit.

Takový přístup má však jeden závažný nedostatek: počet stěn mnohostěnu P (tj. minimální počet nerovností, které definují P) je až na několik výjimek exponenciálně velký vzhledem k dimenzi n . To tedy znamená, že i když problém lineárního programování umíme řešit efektivně, velikost lineárního programu je exponenciální. Elipsoidová metoda však naštěstí nevyžaduje explicitní výčet celého systému nerovností. Stačí, máme-li k dispozici podprogram, který pro daný vektor, jenž není přípustný, dokáže určit tu z předpokládaných nerovností, kterou tento vektor porušuje. Při správném zápisu takto dostaneme nový problém kombinatorické optimalizace, jehož řešení může být jednodušší než řešení původního problému. Tento druhý problém nazveme „polarně sdruženým“ k prvému problému (při formalizaci postupu skutečně dostaneme druhý problém z prvního pomocí polarity ve smyslu projektivní geometrie). Není bez zajímavosti, že dvojnám užitím polarity dostaneme opět původní problém.

Polarita představuje velmi silnou metodu pro redukci problémů kombinatorické optimalizace. Zdá se, že skoro všechny problémy kombinatorické optimalizace, které jsou řešitelné pomocí efektivního algoritmu, lze řešit kombinací elipsoidové metody a *jednoho* dalšího poměrně jednoduchého algoritmu (tzv. hladového algoritmu). Uvědomíme-li si, jak jsou některé algoritmy kombinatorické optimalizace důmyslné a komplikované, je tento fakt více než překvapivý.

Připomíná to trochu detektivku; největší bankovní podvod na světě byl objeven jenom proto, že se svědomitý úředník anglické banky snažil zjistit, proč se mu o tři pence rozcházejí účty. Pokus opravit špatné chování simplexové metody pro několik uměle zkonstruovaných případů povede možná k převratným změnám kombinatorické optimalizace.

Literatura

- [1] C. BERGE: *Sur le couplage maximum d'un graphe*. C. R. Acad. Sci. 247 (1958), 258—259.
- [2] C. BERGE: *Graphs and hypergraphs*. North-Holland, 1973.
- [3] V. CHVÁTAL: *Edmonds polyhedra and a hierarchy of combinatorial problems*. Discrete Math. 4 (1973), 305—337.
- [4] G. B. DANTZIG: *Maximization of a linear function of variables subject to linear inequalities*, in: T. C. KOOPMANS, ed.: *Activity Analysis of Production and Allocation*. New York, Wiley (1951), 339—347.
- [5] J. EDMONDS: *Paths, trees and flowers* Canad. J. Math. 17 (1965), 449—467.

- [6] J. EDMONDS: *Maximum matching and a polyhedron with 0,1-vertices*. J. Res. Nat. Bur. Stand. B1-2 (1965), 125—130.
- [7] J. FARKAS: *Über die theorie der einfachen Ungleichungen*. J. Reine Angew. Math. 124 (1902), 1—24.
- [8] D. B. JUDIN a A. S. NEMIROVSKIJ: *Informational complexity and effective methods for the solution of convex extremal problems* (rusky). Ekon. i Matem. Metody 12, 2 (1976), 357—369.
- [9] L. V. KANTOROVIČ: *Mathematical methods in the organization and planning of production*. Izd. Leningrad. Gos. Univ., 1939.
- [10] L. G. CHAČIJAN: *A polynomial algorithm for linear programming*. Doklady Akad. Nauk SSSR 244 (1979), 1093—1096.
- [11] V. KLEE a G. MINTY: *How good is simplex algorithm?* In: O. SHIBA, ed.: *Inequalities III*. New York, Academic Press (1972), 159—175.
- [12] L. LOVÁSZ: *On the Shannon capacity of a graph*. IEEE Trans. Inf. Th. IT — 25 (1979), 1—7.
- [13] G. L. MILLER: *Riemann's hypothesis and tests for primality*. J. Comput. System Sci. 13 (1976), 300—317.
- [14] J. VON NEUMANN: *On a maximization problem* (rukopis), Institute for Advanced Study, Princeton, 1947.
- [15] N. Z. ŠOR: *Convergence rate of the gradient descent method with dilation of the space*. Kibernetika 2 (1969), 80—85.
- [16] W. T. TUTTE: *The factorization of linear graphs*. J. London Math. Soc. 22 (1947), 107—111

vyučování

Před více než 12 lety byla na sovětských středních školách zahájena výuka matematiky podle nových osnov a učebnic. To znamená, že podle těchto osnov již dva ročníky absolvovaly celou střední školu. V současné době probíhá mezi sovětskými matematiky diskuse o úspěšnosti, resp. neúspěšnosti této reformy. Tato diskuse se odrazila i v tisku. Prvním kritickým rozbořem byl článek akademiků L. S. Pontrjagina, A. N. Tichonova a V. S. Vladimirova v časopise „Matematika v škole“ (roč. 1979, č. 3, str. 12—14). V následujícím čísle tohoto časopisu byl otištěn článek akademiků L. V. Kantoroviče a S. L. Soboleva, kteří obhajují reformu jako celek, i když s konkrétní-

mi kritickými připomínkami souhlasí. Na podzim roku 1980 pak vyšel v časopise ÚV KSSS „Kommunist“ (č. 14) obsáhlý článek L. S. Pontrjagina „Matematika a kvalita její výuky“, který uvedenou problematiku rozebírá v širších souvislostech. Redakční komentář k tomuto článku a také redakční článek v č. 18 (prosinec 1980) téhož časopisu (shrnující došlé ohlasy na Pontrjaginův článek) ukazují, že kritické stanovisko v diskusi převládá.

V tomto čísle Pokroků otiskujeme dva ze zmíněných článků, z nichž první vyjadřuje názor kritiků a druhý názor stoupenců současné reformy. Umožní čtenáři získat jistou představu jednak o reformě výuky matematiky na sovětských středních školách, zahájené v 60. letech, jednak o charakteru současné diskuse k této reformě. Obě strany se v podstatě shodují v kritice všech konkrétních uváděných nedostatků;