KYBERNETIKA ČÍSLO 2, ROČNÍK 4/1968

Poznámka o programování v procesech markovského typu

KAREL SLADKÝ

V této práci je ukázáno, že algoritmy navržené Jewellem a Schweitzerem pro nalezení optimálního řízení semimarkovského procesu můžeme obdržet aplikací speciálně upravené revidované simplexové metody na jistým způsobem transformovanou úlohu lineárního lomeného programování, jež popisuje vyšetřovaný semimarkovský proces.

ÚVOD

Pro řízení semimarkovských procesů s konečným počtem stavů pro případ kritéria optimálního průměrného zisku jsou v literatuře popsány dva základní typy algoritmů:

(1) Algoritmy založené na metodě dynamického programování, vzniklé zobecněním Howardova iteračního postupu. Howardův algoritmus je odvozen v práci [1], zobecnění tohoto algoritmu pro případ semimarkovského procesu provedená Jewellem a Schweitzerem jsou uvedena v [2].

(2) Algoritmy odvozené různými úpravami klasické simplexové metody lineárního programování. Algoritmy tohoto typu byly vypracovány pro případ obecné úlohy lineárního programování s lineární lomenou účelovou funkcí (pro stručnost budeme mluvit o lineárním lomeném programování). Tyto postupy jsou popsány v [3], [4], [5], [6], [7], [8]. Na speciální typ úlohy lineárního lomeného programování je též možno převést úlohu o optimalizaci semimarkovského procesu.

V práci [9] Klein navrhl jistou transformaci proměnných, pomocí které úlohu lineárního lomeného programování vyjadřující optimalizaci semimarkovského procesu je možno převést na úlohu lineárního programování. (Na obdobné transformaci je rovněž založena obecná metoda pro řešení úlohy lineárního lomeného programování vypracovaná Charnesem a Cooperem – viz [3].)

V práci [10] bylo ukázáno, že klasický Howardův algoritmus pro nalezení optimálního řízení diskretního markovského procesu je speciálním rozšířením simplexové metody, u které změna přípustného řešení se provádí současně u více proměnných.

Cílem tohoto příspěvku je objasnit souvislost mezi oběma typy algoritmů pro řízení semimarkovských procesů. V práci je ukázáno, že oba algoritmy pro nalezení optimálního řízení semimarkovského procesu uvedené v práci [2] můžeme obdržet aplikací speciálně upravené revidované simplexové metody na vhodně transformovanou úlohu lineárního lomeného programování, která popisuje vyšetřovaný semimarkovský proces. Tento způsob odvození se zdá být jednodušší než postup použitý v práci [2].

FORMULACE PROBLÉMU

Hledání optimální strategie řízeného semimarkovského procesu můžeme formulovat přesně takto:

Mějme semimarkovský proces s n stavy. V i-tém stavu lze použít jedné z s_i strategií. Při použití strategie k v i-tém stavu přejdeme s pravděpodobností p_{ij}^k z i-tého do j-tého stavu. Použití strategie k v i-tém stavu přinese průměrný zisk c_{ik} a uvažovaný proces setrvává v i-tém stavu po náhodnou dobu, jejíž střední hodnota činí $d_{ik} > 0$.

V každém stavu hledáme takovou strategii, aby průměrný zisk za jednotku času g uvažovaného semimarkovského procesu byl co nejvyšší v čase $T \to \infty$. (Pro jednoduchost se omezíme na ergodické procesy.)

Abychom mohli formulovat vyšetřovanou úlohu jako úlohu lineárního lomeného programování, připustíme znáhodněný způsob volby strategií v jednotlivých stavech. Posloupnost jednotlivých stavů vyšetřovaného semimarkovského procesu zřejmě tvoří markovský řetězec. Budeme se zajímat o finální pravděpodobnosti tohoto markovského řetězce v souvislosti s použitými strategiemi v jednotlivých stavech.

Budiž π_{ik} limita pravděpodobnosti, že uvažovaný markovský řetězec bude ve stavu i a že v tomto stavu bude v následujícím kroku použito strategie k, roste-li počet kroků do nekonečna. Lehce se pak přesvědčíme o platnosti vztahu $\pi_i = \sum_{k=1}^{s_i} \pi_{ik}$, kde π_i jsou obvyklé finální pravděpodobnosti.

Jestliže budeme předpokládat pouze stacionárnost použitých strategií, vyšetřovanou úlohu lze s použitím známých vlastností finálních pravděpodobností přepsat jeho úlohu lineárního lomeného programování následovně:

ÚLOHA 1. Za podmínek

$$(1) \qquad \sum_{i=1}^{n} \sum_{k=1}^{s_i} (p_{ij}^k - \delta_{ij})\, \pi_{ik} = 0 \quad \text{pro} \quad j = 1, 2, \ldots, n\,,$$

$$\sum_{i=1}^{n} \sum_{k=1}^{s_i} \pi_{ik} = 1\,, \quad \pi_{ik} \geqq 0\,,$$

 nalezněte maximum účelové funkce:

$$g = \frac{\sum_{i=1}^{n} \sum_{k=1}^{s_i} c_{ik}\pi_{ik}}{\sum_{i=1}^{n} \sum_{k=1}^{s_i} d_{ik}\pi_{ik}} .$$

Symbol δ_{ij} značí v celém textu Kroneckerovo delta.

Jak je známo (viz např. [1]), soustava lineárních omezení ÚLOHY 1 je lineárně závislá. Proto např. n-tá rovnice je lineární kombinací zbývajících rovnic a proto ji nebudeme dále v soustavě lineárních omezení uvažovat.

V práci [9] Klein zavedl následující transformaci, pomocí které je možno ÚLOHU 1 převést na ekvivalentní úlohu lineárního programování.

Jestliže zavedeme

$$t^2 x_{ik} = d_{ik}\pi_{ik} \Rightarrow \pi_{ik} = t^2 \frac{x_{ik}}{d_{ik}} ,$$

kde platí

$$\sum_{i=1}^{n} \sum_{k=1}^{s_i} x_{ik} = 1 , \quad t^2 \sum_{i=1}^{n} \sum_{k=1}^{s_i} \frac{x_{ik}}{d_{ik}} = 1 ,$$

potom ÚLOHU 1 lze přepsat na ÚLOHU 2.

ÚLOHA 2. Za podmínek

$$(2) \qquad \sum_{i=1}^{n} \sum_{k=1}^{s_i} \frac{p_{ij}^k - \delta_{ij}}{d_{ik}} x_{ik} = 0 \quad \text{pro} \quad j = 1, 2, \ldots, n - 1 ,$$

$$\sum_{i=1}^{n} \sum_{k=1}^{s_i} x_{ik} = 1 , \quad x_{ik} \geqq 0 ,$$

nalezněte maximum účelové funkce:

$$g = \sum_{i=1}^{n} \sum_{k=1}^{s_i} \frac{c_{ik}}{d_{ik}} x_{ik} .$$

Zavedeme-li ještě $\bar{x}_{ik} = x_{ik}/d_{ik}$, ÚLOHA 2 přejde na ÚLOHU 3.

ÚLOHA 3. Za podmínek

$$(3) \qquad \sum_{i=1}^{n} \sum_{k=1}^{s_i} (p_{ij}^k - \delta_{ij}) \bar{x}_{ik} = 0 \quad \text{pro} \quad j = 1, 2, \ldots, n - 1 ,$$

$$\sum_{i=1}^{n} \sum_{k=1}^{s_i} d_{ik}\bar{x}_{ik} = 1 , \quad \bar{x}_{ik} \geqq 0 ,$$

nalezněte maximum účelové funkce:

$$g = \sum_{i=1}^{n} \sum_{k=1}^{s_i} c_{ik} \bar{x}_{ik} .$$

Je třeba si všimnout, že při volbě určité strategie v každém stavu hodnota účelové funkce bude v ÚLOHÁCH 1, 2, 3 stejná; rozdílné (se zachováním znaménka) budou pouze hodnoty těch neznámých x_{ik}, $\bar{x}_{ik}$, π_{ik}, které jsou různé od nuly.

Z literatury [2], [9] je známo, že optimální strategii stačí hledat ve třídě „čistých strategií" (tj. při optimálním řízení je v každém stavu používána pouze jediná strategie). O této skutečnosti se lehce můžeme přesvědčit např. z vlastností optimálního řešení úlohy lineárního programování.

Na ÚLOHU 2 a na ÚLOHU 3 budeme aplikovat revidovanou simplexovou metodu.

Jestliže ve stavu i zvolíme strategii k_i, potom simplexoví násobitelé pro ÚLOHU 2, které označíme $v_1, v_2, \ldots, v_n$ jsou definovány vztahem (4) (využíváme přitom vztahu (D.1) z dodatku):

$$\sum_{j=1}^{n-1} \frac{p_{ij}^{k_i} - \delta_{ij}}{d_{ik_i}} v_j + v_n = \frac{c_{ik_i}}{d_{ik_i}} \quad \text{pro} \quad i = 1, 2, \ldots, n .$$

Úpravou vztahu (4) dostaneme stejný tvar rovností, jako když vypočítáváme simplexové násobitele pro ÚLOHU 3. Tento tvar po malé úpravě je následující:

$$\sum_{j=1}^{n-1} p_{ij}^{k_i} v_j + d_{ik_i} v_n = v_i + c_{ik_i} \quad \text{pro} \quad i = 1, 2, \ldots, n-1 , \tag{5}$$

$$\sum_{j=1}^{n-1} p_{nj}^{k_n} v_j + d_{nk_n} v_n = c_{nk_n} .$$

Existující řešení je možno zlepšit použitím té proměnné (využíváme vztahu (D.2) z dodatku), pro kterou jsou následující výrazy kladné:

pro ÚLOHU 2 výraz

$$\frac{c_{ik}}{d_{ik}} - \sum_{j=1}^{n-1} \frac{p_{ij}^{k} - \delta_{ij}}{d_{ik}} - v_n , \tag{6}$$

pro ÚLOHU 3 výraz

$$c_{ik} - \sum_{j=1}^{n-1} (p_{ij}^{k} - \delta_{ij}) v_j - d_{ik} v_n . \tag{7}$$

V *simplexové metodě* se v každém *iteračním* kroku zavádí pouze jediná nová proměnná, která maximalizuje výše uvedené výrazy, neboť je třeba ještě vyšetřit, kterou z proměnných báze vyloučíme.

Hodnotu účelové funkce g můžeme vyčíslit pomocí simplexových násobitelů (viz vztah (D.3) v dodatku). V našem případě dostaneme $g = v_n$.

Zavedeme-li ještě $w_i = -v_i$ pro $i = 1, 2, \ldots, n - 1$; $w_n = 0$ rovnice (5) přejde na tvar:

$$w_i + d_{ik_i}g = c_{ik_i} + \sum_{j=1}^{n} p_{ij}^{k_i}w_j \tag{8}$$

$$\text{pro} \quad i = 1, 2, \ldots, n\,, \quad \text{když klademe} \quad w_n = 0\,.$$

Tvaru (8) je použito pro určení „váhových koeficientů" w_i a průměrného zisku g v Jewellově i Schweitzerově algoritmu v práci [2].

Zavedeme-li ještě hodnoty w_i, g místo hodnot v_i do kritérií (6) a (7), kritérium (6) nabude potom tvaru

$$\frac{c_{ik}}{d_{ik}} + \sum_{j=1}^{n} \frac{p_{ij}^k}{d_{ik}} w_j - \frac{w_i}{d_{ik}} - g \tag{6*}$$

a kritérium (7) bude:

$$c_{ik} + \sum_{j=1}^{n} p_{ij}^k w_j - w_i - g d_{ik}\,, \tag{7*}$$

jestliže v obou případech klademe $w_n = 0$.

Optimální strategii stačí hledat ve třídě „čistých strategií", jak již bylo připomenuto. Tím odpadne nutnost vyšetřit, kterou bázickou proměnnou máme vyloučit při zavedení nové proměnné do báze. Je proto možné při známých hodnotách simplexových násobitelů (neboli „váhových koeficientů") v každém stavu použít tu strategii, která maximalizuje pro daný stavu (tedy pro jisté i) kritéria (6*) a (7*). Současné zavedení více nových proměnných do báze má za následek, že nemůžeme obdržet nové simplexové násobitele jednoduchým způsobem, kterého se využívá v revidované simplexové metodě. Proto je obvykle výhodnější vypočítávat „nové" simplexové násobitele řešením soustavy lineárních rovnic podle jejich definičního vztahu (4). Lehce se přesvědčíme, že při pevném i, jestliže hledáme maximum přes k, místo kritéria (6*) lze použít poněkud jednoduššího kritéria ve tvaru

$$\frac{1}{d_{ik}} \left[c_{ik} + \sum_{j=1}^{n} p_{ij}^k w_j - w_i\right], \tag{6**}$$

což je kritérium použité v algoritmu navrženém Jewellem; místo kritéria (7*) stačí uvažovat pouze maximalizaci výrazu

$$c_{ik} + \sum_{j=1}^{n} p_{ij}^k w_j - d_{ik} g\,. \tag{7**}$$

Tohoto kritéria je použito ve Schweitzerově algoritmu.

DODATEK

V dodatku připomeneme některé známé vztahy pro revidovanou simplexovou metodu, kterých využíváme v této práci. Tyto vztahy lze nalézt na př. v [11].

Uvažujme ÚLOHU lineárního programování: Za podmínek

$$\sum_{j=1}^{n} a_{ij}x_j = b_i \,, \quad x_j \geqq 0 \quad \text{pro} \quad i = 1, 2, \ldots, m \,; \quad \text{kde} \quad m < n \,,$$

nalezněte maximum účelové funkce:

$$z = \sum_{j=1}^{n} c_j x_j \,.$$

Předpokládejme, že uvažované bázické řešení ÚLOHY je tvořeno proměnnými $x_1, x_2, \ldots, x_m$. Potom pro simplexové násobitele, které označíme $v_1, v_2, \ldots, v_m$ platí

$$\sum_{i=1}^{m} a_{ij} v_i = c_j \quad \text{pro} \quad j = 1, 2, \ldots, m \,. \tag{D.1}$$

Na základě známých simplexových násobitelů je možno provést:

zlepšení existujícího řešení ÚLOHY zavedením k-té proměnné do báze $(k > m)$, pro kterou výraz

$$\bar{c}_k = c_k - \sum_{i=1}^{m} a_{ik} v_i \tag{D.2}$$

bude kladný;

výpočet hodnoty účelové funkce na základě vztahu

$$z = \sum_{i=1}^{m} b_i v_i \,. \tag{D.3}$$

(Došlo dne 22. září 1967.)

LITERATURA

[1] R. A. Howard: Dynamic Programming and Markov Processes. Technology Press and Wiley Press, New York 1960. Ruský překlad Sovetskoe radio 1965.

[2] W. S. Jewell: Markov Renewal Programming I., II. Operations Research *11* (1963), 6, 938—971.

[3] A. Charnes, W. W. Cooper: Programming with Linear Fractional Functionals. Naval Research Logistics Quarterly *9* (1962), 181—186.

[4] W. Dinkelbach: Die Maximierung eines Quotienten zweier linearen Funktionen unter linearen Nebenbedingungen. Zeitschrift für Wahrscheinlichkeitsteorie *1* (1962), 141—145.

[5] W. S. Dorn: Linear Fractional Programming. IBM Research Report RC-830 (Nov. 17, 1962).

[6] P. C. Gilmore, R. E. Gomory: A Linear Programming Approach to the Cutting Stock Problem — Part II. Operations Research, *11* (1963), 863—888.

[7] J. R. Isbell, W. H. Marlow: Attrition Games. Naval Research Logistics Quarterly *3* (1956), 71—94.

[8] B. Martos: Hyperbolic Programming. Naval Research Logistics Quarterly *11* (1964), 134—155.

[9] M. Klein: Inspection — Maintenance — Problem under Markovian Deterioration. Management Science *9* (1962), 1, 25—32.

[10] de Ghellink, G.: Les Problèmes de Décisions Séquentielles. Cahiers du Centre d' Étude de Recherche Opérationelle *2* (1960), 2, 161—179.

[11] G. Dantzig: Linear Programming and Extensions. Princenton University Press, Princenton 1963. Ruský překlad Progress 1966; slovenský překlad SVTL 1967.

SUMMARY

A Note on the Programming in Markov-Type Processes

KAREL SLADKÝ

In the following paper two types of algorithms dealing with the optimization of semi-Markov processes are discussed:

(1) The algorithms based on dynamic programming approach (the extensions of Howard's algorithm developed by Jewell and Schweitzer).

(2) The algorithms based on linear fractional programming.

(It can be easily shown that the optimization of a semi-Markov process can be also formulated as a linear fractional programming problem).

It is shown in this paper that Jewell's and Schweitzer's algorithms are special extensions of the revised simplex method applied to the transformed linear fractional programming problem.

Ing. Karel Sladký, Ústav teorie informace a automatizace ČSAV, Vyšehradská 49, Praha 2.