ՀԱՅԿԱԿԱՆ ՍՍՀ ԳԻՏՈՒԹՅՈՒՆՆԵՐԻ ԱԿԱԴԵՄԻԱՅԻ ԶԵԿՈՒՅՑՆԵՐ
ДОКЛАДЫ АКАДЕМИИ НАУК АРМЯНСКОЙ ССР
LXIX 1979 4

УДК 519.9

ПРИКЛАДНАЯ МАТЕМАТИКА

А. А. Бабаджанян

Условное укрупнение марковских процессов принятия решений и алгоритм нахождения оптимальной стратегии

(Представлено чл.-корр. АН Армянской ССР Р. Р. Варшамовым 8/V 1979)

Нахождение оптимальной стратегии марковских и полумарковских процессов принятия решений при бесконечном времени планирования основано на двух подходах—итерационном алгоритме Ховарда ([1,2]) и линейном программировании ([2,3]).

Практическое применение методов ([1-3]) затруднительно, в связи с чем в работах ([4-6]) предложены подходы к сокращению возникающих вычислительных трудностей. В ([4]) дан приближенный метод, основанный на группировке состояний процесса. Схема улучшения стратегии по частям в марковских процессах принятия решений была использована в ([5]). В случае сведения к задаче линейного программирования в ([6]) дана схема декомпозиции, в основе которой лежит метод Данцига—Вульфа.

В настоящей работе предлагается метод условного укрупнения марковских процессов принятия решений без переоценки, алгебраической основой которого является результат, полученный в ([7]). На основе условного укрупнения приведен алгоритм нахождения оптимальной стратегии путем фиксирования стратегий по частям. Предложенный метод в отличие от ([4]) является точным. Следует отметить, что нижеизложенные результаты можно распространить на полумарковские процессы принятия решений как с переоценкой, так и без нее.

1. Пусть имеется система, пространство состояний которой есть конечное множество $S=\{1, 2, \ldots, N\}$ (здесь следуем обозначениям ([2])). Каждому состоянию $i\in S$ соответствует конечное множество K_i решений, элементы которого обозначим $k=1, 2, \ldots, K_i$, $K=K_1\times\ldots\times K_N$. Если система находится в состоянии $i\in S$ и принимается решение $k\in K_i$, то система получает доход r_i^k и ее состояние в следующий момент времени определяется вероятностным законом p_{ij}^k $(j\in S)$, где p_{ij}^k — вероятность перехода из состояния i при выборе решения k в

состояние j. Задание отображения (стратегии) $f: S \to K$, где $f(i)=k \in K_i$ ($i \in S$) определяет марковскую цепь, которая полагается эргодической.

Нахождение стационарной оптимальной стратегии, максимизирующей норму дохода $g(f)$ [2], основано на алгоритме Ховарда, состоящем из:

а) процедуры определения весов.

Для произвольной стационарной стратегии f решается система уравнений

$$g+v_i=r_i^k+\sum_{j=1}^{N-1} p_{ij}^k v_j$$

относительно $g, v_1, \ldots, v_{N-1}$ (полагая $v_N=0$);

б) процедуры улучшения стратегии.

По найденным v_i ($i=1, \ldots, N$) и g для каждого $i \in S$ выберем такой элемент множества $G(i, f)$, что

$$r_i^k+\sum_{j=1}^{N-1} p_{ij}^k v_j > g+v_i$$

при любых $k \in K_i$. Если $G(i, f)=\varnothing$ при всех $i \in S$, то f оптимальна.

Если существует i, для которого $G(i, f) \neq \varnothing$, то улучшенная стратегия f' строится так: $f'(i) \in G(i, f)$, если $G(i, f) \neq \varnothing$, и $f'(i)=f(i)$, если $G(i, f)=\varnothing$. После этого следует обратиться к пункту б).

Пусть пространство состояний разбито на части $S=S_1 \cup S_2$, $S_1 \cap S_2=\varnothing$, примем $S_1=\{1, \ldots, h\}$, а $S_2=\{h+1, \ldots N\}$. Предположим, что для состояний $i \in S_1$ множество K_i состоит из единственного элемента, т. е. нет альтернативы в состояниях из S_1. Тогда имеет место следующая основная

Теорема. *Существует марковский процесс решения с числом состояний m ($m=|S_2|+1$) такой, что для любой стратегии f:*

$a)$ $\bar{g}(\bar{f})=g(f)$,

$б)$ $\bar{v}_1=\sum_{i \in S_1} v_i/h$,

$\bar{v}_i=v_{i+h-1} \quad i=2, \ldots, m$.

Пространством состояний здесь принято множество $\bar{S}=\{1, \ldots m\}$, здесь и ниже черта сверху означает соответствие укрупненному процессу.

Доказательство теоремы проводится при помощи формального приема „банка“ из [8] с использованием результата [7], где дан и сам алгоритм построения такого укрупнения.

Назовем построенный марковский процесс решения условно—укрупненным, а соответствующую оптимальную стратегию условно—оптимальной.

Замечание 1. Аналогичная теорема имеет место и при произвольном разбиении пространства состояний $S=\bigcup_e S_e$.

2. Ниже схематично излагается алгоритм нахождения оптимальной стратегии.

Пусть f — произвольная стратегия. Разобьем пространство состояний на части $S = S_1 \cup S_2$, $S_1 \cap S_2 = \varnothing$.

Предположим зафиксированным f на S_1. Перейдя к условно-укрупненному процессу, найдем с помощью алгоритма Ховарда $\bar{f}$-условно-оптимальное. Ясно, что $g(\bar{f}) > g(f)$, хотя $\bar{f}(i) = f(i)$, $i \in S_1$.

Теперь улучшим стратегию $\bar{f}$, оставляя ее неизменной на S_2. Снова, перейдя к условно-укрупненному процессу, найдем $\bar{\bar{f}}$-условно-оптимальное, где $g(\bar{\bar{f}}) > g(\bar{f})$. Принимаем $\bar{\bar{f}}$ вновь за произвольную и поступаем с ней аналогично f.

Таким образом, процесс нахождения оптимальной стратегии сводится к определению конечной последовательности условно-оптимальных стратегий укрупненных процессов решения.

З а м е ч а н и е 2. Следует отметить, что метод, предложенный в ([4]), приближенный, так как не удовлетворяет условию совместимости (см. ([7])), тем самым не имеет место б).

З а м е ч а н и е 3. Идея фиксации части стратегий широко применяется в динамических задачах; в марковских процессах решения она применена в ([5]) с использованием блочного алгоритма Гаусса (схема улучшения стратегий по частям).

Отметим, что могут быть предложены другие разновидности схемы улучшения стратегий по частям с использованием условно-укрупненных процессов.

3. Связь с линейным программированием и условное укрупнение полумарковских процессов принятия решений выделены в самостоятельную работу.

В заключение выражаю благодарность К. А. Абгаряну за внимание к работе.

Армянский научно-исследовательский
институт энергетики

Ա. Ա. ԲԱՐԱՋԱՆՅԱՆ

Որոշումների ընդունման մարկովյան պրոցեսների պայմանական խոշորացումը և օպտիմալ ստրատեգիան գտնելու ալգորիթմը

Աշխատանքում մտցվում է որոշման ընդունման մարկովյան պրոցեսի պայմանական խոշորացումը և նրան համապատասխանող պայմանական-օպտիմալ ստրատեգիայի հասկացությունը։

Այդպիսի խոշորացումը հնարավոր է դառնում հեղինակի կողմից՝ գծային բալանսային հավասարումների ագրեգացման տեսության մեջ ստացված արդյունքի շնորհիվ։

Դրա հիման վրա առաջարկվում է օպտիմալ ստրատեգիան որոշելու ալգորիթմը՝ ստրատեգիան մաս-մաս բարելավելու սխեմայի ճանապարհով։

Վերոհիշյալ մեթոդը թույլ է տալիս որոշումների ընդունման տրված մարկովյան պրոցեսը բերելու պայմանական խոշորացված պրոցեսների վերջավոր հաջորդականությանը։

ЛИТЕРАТУРА — ԳՐԱԿԱՆՈՒԹՅՈՒՆ

[1] *Р. А. Ховард,* Динамическое программирование и марковские процессы, «Советское радио», М., 1964. [2] *Х. Майн, С. Осаки,* Марковские процессы принятия решений, «Наука», М., 1977. [3] *Ф. Вольф, Д. Данциг,* Кибернетический сборник, вып. 4 (1967). [4] *М. Г. Теплицкий,* Автоматика и телемеханика, № 10 (1969). [5] *Буй Куанг Зиеу,* Вестник ЛГУ, № 19 (1978). [6] *В. К. Демин, А. А. Осадченко,* Известия АН СССР, Техническая кибернетика, № 3 (1979). [7] *А. А. Бабаджанян,* ДАН Арм. ССР, LXVIII № 4 (1979). [8] *Дж. Кемени, Дж. Снелл,* Конечные цепи Маркова, «Наука», М., 1970.