ИЗВЕСТИЯ
ТОМСКОГО ОРДЕНА ОКТЯБРЬСКОЙ РЕВОЛЮЦИИ
И ОРДЕНА ТРУДОВОГО КРАСНОГО ЗНАМЕНИ
ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА им. С. М. КИРОВА

Том 290 1974

ОБ ИСПОЛЬЗОВАНИИ ЭЛЕМЕНТОВ МЕТОДА ВЕТВЕЙ И ГРАНИЦ В СХЕМАХ ДИНАМИЧЕСКОГО ПРОГРАММИРОВАНИЯ

Э. И. ГЕРМАН, И. П. МАКАРОВ

(Представлена научно-техническим семинаром кафедры АСУ и лаборатории управления)

В настоящее время в связи с бурным развитием средств вычислительной техники среди методов оптимизации все большее распространение получают различного рода схемы упорядоченного частичного перебора вариантов. Обширной областью их применения являются задачи определения экстремума функций конечного числа переменных, заданных на дискретных множествах, например, задачи дискретного программирования [1], экстремальные задачи на графах [2], дискретные динамические задачи распределения ресурсов [3] и т. п.

Значительное число задач оптимизации функций и функционалов с непрерывными областями определения переменных не допускают аналитического решения. К их числу следует отнести обширный класс задач по определению различного рода оптимальных траекторий. При их решении численными методами на ЭВМ используется их конечномерная аппроксимация и само решение происходит, как правило, в рамках одной из схем упорядоченного перебора [3, 4].

К числу наиболее распространенных схем упорядоченного перебора относятся метод ветвей и границ [1], метод динамического программирования [3, 5], его модификация — последовательный анализ вариантов и ряд других [6].

Существенное различие идей, на которых основаны эти методы, наличие определенных достоинств и недостатков у каждого из них, неоднозначность в построении алгоритмов и вычислительных схем при реализации методов для решения конкретных задач, существенная зависимость эффективности алгоритмов от того, насколько полно учтена специфика решаемой задачи, — все это делает, в частности, актуальными попытки совместного использования элементов различных схем в рамках одного алгоритма.

В данной работе предлагается один из вариантов использования идей метода ветвей и границ при решении одного класса условных конечномерных экстремальных задач с дискретными областями определения функций методом динамического программирования.

Примером рассматриваемого класса задач является задача определения экстремума нелинейной сепарабельной функции

$$f(\vec{x}) = \sum_{j=1}^{N} f_j(x_j) \quad (1)$$

при ограничениях

$$x_j \in \Omega_j, \; j = \overline{1, N}, \tag{2}$$

где Ω_j — дискретные конечные множества.

Другим примером являются конечномерно-аппроксимированные задачи определения оптимальных траекторий следующего вида:

найти экстремум

$$f(\vec{x}) = \sum_{j=1}^{N} f_j(x_j, \; x_{j+1}) \tag{3}$$

при ограничениях вида (2).

Задачи с целевыми функциями вида (3) допускают простую геометрическую интерпретацию. Вводится в рассмотрение N гиперплоскостей и каждому допустимому значению $x_j \in \Omega_j$ сопоставляется взаимно-однозначно точка в Ω_j-й гиперплоскости.

Не теряя общности, можно использовать одномерные гиперплоскости — прямые. Соединим попарно ребрами точки гиперплоскостей, имеющих соседние индексы. В качестве длины ребра, соединяющего очки $x_j \in \Omega_j$ и $x_{j+1} \in \Omega_{j+1}$, примем значение функции $f_j(x_j, \; x_{j+1})$.

Совокупности точек $x_j^* \in \Omega_j$, $j = \overline{1, N}$, выбранных по одной для каждой гиперплоскости, может быть сопоставлен путь от первой до N-й гиперплоскости, которому может быть поставлено в соответствие значение функции (3)

$$f(\vec{x}^*) = \sum_{j=1}^{N} f_j(x_j^*, \; x_{j+1}^*). \tag{4}$$

Тогда задача определения экстремума функции (3) может быть интерпретирована как задача определения экстремального пути в сформированном графе.

Заметим, что задача (1) — (2) может рассматриваться как частный случай задачи (2) — (3). Поэтому в дальнейшем описание достаточно вести для задачи вида (2) — (3), считая для определенности, что ищется минимум (3).

Прежде чем перейти к описанию предлагаемой модификации схемы динамического программирования, рассмотрим вкратце стандартную схему решения задачи (2) — (3) по методу динамического программирования.

Решение задачи (2) — (3) методом динамического программирования осуществляется путем последовательного решения уравнений Беллмана следующего вида:

$$l(x_{j+1}) = \min_{x_j \in \Omega_j} \{l(x_j) + f_j(x_j, \; x_{j+1})\}, \; j = \overline{2, N-1}, \tag{5}$$

$$l(x_2) = \min_{x_1 \in \Omega_1} f_1(x_1, \; x_2). \tag{6}$$

Функции Беллмана $l(x_j)$ в терминах рассмотренной геометрической интерпретации соответствуют длине кратчайшего пути от первой гиперплоскости до вершины x_j j-й гиперплоскости.

Тогда оптимальное значение функции (3) может быть определено как

$$\min_{\vec{x}} f(\vec{x}) = \min_{x_N \in \Omega_N} l(x_N). \tag{7}$$

Решение задачи с использованием уравнений (5)—(6) осуществляется путем последовательного вычисления функций $l(x_{j+1})$ для всех $x_{j+1} \in \Omega_{j+1}$ по известным значениям $l(x_j)$ для $x_j \in \Omega_j$.

Предлагаемая модификация рассмотренной стандартной схемы позволяет исключать из рассмотрения неперспективные значения переменных x_j, сужая тем самым области Ω_j, что способствует сокращению необходимого числа вычислений при определении $l(x_{j+1})$ по (5). При этом, учитывая что $l(x_{j+1})$ вычисляется для всех x_{j+1}, получаем, что исключение одного значения x_j уменьшает число вычислений на $|\Omega_{j+1}|$, где $|\Omega_{j+1}|$ — мощность множества Ω_{j+1}.

Рассмотрим подробнее предлагаемую процедуру сужения областей Ω_j. Пусть известно одно из значений $l(x_N)$, иными словами, значение функции (3) на одном из путей между первой и N-й гиперплоскостями. Тогда, если для произвольного значения $l(x_j)$ выполняется соотношение

$$l(x_j) \geqslant l(x_N), \tag{8}$$

то соответствующее значение x_j может быть исключено из множества Ω_j как неперспективное. В обоснование этого заметим, что при любом продолжении пути из данной вершины j-й гиперплоскости до N-й гиперплоскости значение целевой функции может только возрасти по сравнению с $l(x_j)$, т. е. любой путь, содержащий данную вершину x_j даст значение целевой функции (3) не меньше, чем на известном уже нам пути.

Используя терминологию метода ветвей и границ, можно рассматривать $l(x_N)$ как рекорд, т. е. значение целевой функции на лучшем из просмотренных решений, а величины $l(x_j)$ для всех $x_j \in \Omega_j$ и $j = \overline{1, N-1}$ как оценки, т. е. нижние границы целевой функции на множестве решений. И тогда соотношение (8) аналогично принятому правилу в методе ветвей и границ исключения из рассмотрения тех вершин дерева, которые имеют оценку худшую, чем рекорд.

Известно, что схема метода ветвей и границ позволяет получить решение с любой априорно заданной близостью к оптимальному. Аналогично, используя для оценки перспективности вместо соотношения (8) соотношение вида

$$\frac{l(x_N - l(x_j)}{l(x_N)} \cdot 100 \leqslant \varkappa, \tag{9}$$

где $\varkappa$ — априорно заданная точность решения в процентах, можно исключать из рассмотрения те x_j, которые „обещают“ незначительное, с точки зрения $\varkappa$, улучшение значения (3).

Таким образом, в предлагаемой схеме каждое вновь вычисленное по (5) значение $l(x_j)$ сравнивается с $l(x_N)$ по (8) (или с использованием (9)) и в случае, если соответствующее значение x_j признано неперспективным, то оно исключается из Ω_j.

Описание рассматриваемой модификации схемы динамического программирования можно считать оконченным, если дополнительно оговорить способ получения и корректировки $l(x_N)$.

Прежде всего заметим, что при решении многих практических задач в качестве $l(x_N)$ может быть взято значение (3) на лучшем из известных решений, полученных из физических или иных соображений.

Другим способом получения $l(x_N)$ является использование вспомогательных, сравнительно простых эвристических алгоритмов, позволяющих получить локально оптимальные решения. Например, если речь идет о классической задаче определения кратчайшего маршрута в сети, то в качестве первого приближения для $l(x_N)$ может быть взята длина маршрута, построенного с использованием следующего правила. Из множества дуг, позволяющих перейти из заданной вершины слоя j в слой $j+1$, каждый раз выбирается дуга с наименьшей длиной.

Процедура определения $l(x_N)$, которую можно условно назвать процедурой зондирования, не обязательно должна предшествовать процессу решения уравнений (5) — (6). Зондирование может проводиться, начиная с любого этапа решения уравнений (5) — (6), иными словами, поиск локально оптимальных путей до N-й гиперплоскости может быть начат в принципе с любой вершины любой гиперплоскости. Естественно, что при неоднократном зондировании в (8) используется меньшее из полученных значений $l(x_N)$.

Эффективность предлагаемой схемы решения определяется тем, сколь значительно удается сузить области Ω_j. Последнее, в свою очередь, зависит от того, насколько существен разброс в значениях $f_j(x_j, x_{j+1})$ при фиксированном j и различных $x_j \in \Omega_j$, $x_{j+1} \in \Omega_{j+1}$. В терминах геометрической интерпретации это означает, что чем больше различие в длине дуг между парами гиперплоскостей, тем более вероятно появление частичных путей (путей от 1-й до j-й гиперплоскости, где $j < N$), имеющих длину не меньшую, чем длина известного пути от 1-й до N-й гиперплоскости.

Приведенные соображения позволяют определять, насколько целесообразно использовать предлагаемую процедуру сужения областей Ω_j при решении конкретных задач и, в том числе, с какого этапа j следует ожидать выполнения соотношения (8) для достаточно большого числа значений x_j. Иными словами, при решении конкретных задач всегда следует соизмерять эффект от сужения областей Ω_j с затратами на зондирование.

Описанный способ использования элементов метода ветвей и границ в стандартной схеме динамического программирования применен при решении задачи [7].

ЛИТЕРАТУРА

1. А. А. Корбут, Ю. Ю. Финкельштейн. Дискретное программирование. М., «Наука», 1969.
2. С. Цой, С. М. Цхай. Прикладная теория графов. Алма-Ата, «Наука», 1971.
3. Р. Беллман, С. Дрейфус. Прикладные задачи динамического программирования. М., «Наука», 1965.
4. Н. Н. Моисеев. Численные методы в теории оптимальных систем. М., «Наука», 1971.
5. Р. Беллман. Динамическое программирование. ИЛ, 1960.
6. В. С. Михалевич. Последовательные алгоритмы оптимизации и их применение. «Кибернетика», 1965. № 1, 2.
7. И. П. Макаров, В. Г. Ротарь. Модель и алгоритм планирования процессов создания АСУ и наращивания мощности ВЦ. В сб.: «Кибернетика и вуз», вып. 5. Томск, ТПИ, 1972.