***С. В. Чебаков**, к. ф.-м. н., старший научный сотрудник Объединенного института проблем информатики Национальной академии наук Беларуси,*
***Л. В. Серебряная**, к. т. н., доцент кафедры программного обеспечения информационных технологий Белорусского государственного университета информатики и радиоэлектроники*

Алгоритм нахождения множества Парето на конечном наборе начальных данных

Предлагаемый в работе метод нахождения множества Парето T на конечном наборе начальных данных N относится к двухэтапным алгоритмам решения ряда комбинаторных задач.

На первом этапе без применения переборных операций осуществляется нахождение элементов из множества начальных данных, которые по своей внутренней структуре не могут войти в оптимальное подмножество T. Математическая модель задачи представляет собой двухкритериальные подпространства, для каждого из которых построены паретовские слои на множестве N. Сформулированы достаточные условия существования подмножества, исключаемого из рассмотрения вследствие его удаления из множества T. Достигаемое уменьшение множества начальных данных, из которых формируется оптимальное множество T, ведет к уменьшению времени, необходимого для решения данной комбинаторной задачи.

На втором этапе на основе алгоритма частичного перебора элементов отдельных паретовских слоев и с использованием структур данных, полученных на первом этапе, выполняется поиск всех недоминируемых элементов на множестве N. Разработанный алгоритм позволяет рассматривать оба этапа нахождения множества Парето как единый процесс, делая его эффективнее за счет существенного уменьшения времени решения поставленной задачи.

Введение

Многокритериальная задача нахождения множества Парето T на конечном множестве начальных данных N относится к классическим комбинаторным оптимизационным задачам. Для ее решения требуется найти все недоминируемые элементы множества N в заданном многокритериальном пространстве предпочтений. Методы решения этой комбинаторной задачи представлены, например, здесь [1] и основаны на различных способах перебора элементов множества N. С увеличением числа элементов в N количество требуемых попарных сравнений его элементов будет достаточно большим. Следовательно, разработка алгоритмов, уменьшающих общее время решения комбинаторной задачи, является актуальной проблемой. Основные усилия в этом направлении — реализация алгоритмов параллельных вычислений [2]. Предложенный в работе метод нахождения множества Парето продолжает разработку алгоритмов в рамках предлагаемой авторами двухэтапной схемы решения ряда комбинаторных задач [3]. На первом этапе на основе алгоритмов поиска в упорядоченных структурах данных осуществляется построение решения комбинаторной задачи или нахождение подмножества J начального множества N, элементы которого по своей структуре не могут войти в множество T. Если достаточные условия существования множества Парето выполняются, то решение задачи заканчивается и оценка сложности алгоритма решения совпадает с оценкой операций в упорядоченных структурах данных. В случае выполнения условий существования подмножества J все его элементы исключаются из рассмотрения при построении на втором этапе, на основе разработанного в работе специального варианта переборного алгоритма, множества Парето T. В результате образуется новая задача с множеством начальных данных $N^{/}$, где число элементов в $N^{/}$ может быть существенно меньше чем в N. Оценки сложности алгоритмов поиска в упорядоченных структурах существенно превосходят по эффективности подобные оценки для алгоритмов переборного типа. Следовательно, использование в качестве начальных данных элементов множества $N^{/}$ может привести к уменьшению времени, необходимого для решения комбинаторной задачи.

В статье приводится алгоритм построения множества Парето, основанный на структурах данных, сформированных на первом этапе. Разработка такого алгоритма дает возможность рассматривать двухэтапную схему нахождения множества Парето как единое целое, где отдельные этапы связаны между собой.

При отсутствии такого алгоритма при невыполнении на первом этапе условий существования множества Парето или подмножества J пришлось бы начинать решение комбинаторной задачи с «чистого листа», и выполнение операций первого этапа не имело бы смысла.

1. Построение паретовских множеств на двухкритериальных подпространствах

Предположим, что размерность n заданного критериального пространства представляет собой четное число. Пусть заданное отношение предпочтения между элементами множества N транзитивно, и значения критериев представляют собой действительные числа. Будем считать, что число элементов в N представляет собой достаточно большую величину. Произвольным образом разбиваем весь набор критериев на отдельные пары. Разбиение проводим, например, следующим образом. Критерии нумеруются от 1 до n, затем формируется первая пара C_1 из первых двух критериев, вторая пара C_2 — из следующих двух критериев и т. д. Количество таких пар будет равно величине $n/2$. На каждом сформированном двухкритериальном подпространстве на множестве N определяем паретовские множества P_i, $i = 1, n/2$.

Утверждение 1. Любой элемент множества N, который входит в какое-либо паретовское множество P_i, $i = 1, n/2$, является недоминируемым на n-мерном критериальном пространстве и принадлежит множеству Парето T.

Доказательство. На каждой отдельной паре критериев C_i при построении паретовских множеств P_i, $i = 1, n/2$, рассматривается одно и то же множество начальных данных N. Пусть некоторый элемент A из N принадлежит паретовскому множеству P_1 на паре критериев C_1. Следовательно, элемент A доминирует любой элемент из N хотя бы по одному из двух критериев из подпространства C_1. Вместе с тем, для того чтобы элемент A был доминируем в n-мерном критериальном пространстве, требуется, чтобы имелся элемент F из N такой, что не существует ни одного критерия, по которому F уступал бы A. Следовательно, элемент A является недоминируемым в n-мерном критериальном пространстве и должен быть включен в множество T.

Очевидно, что подобные рассуждения справедливы для всех элементов, составляющих паретовские множества P_i, $i = 1, n/2$. Утверждение доказано.

Пусть имеется вычислительная система, состоящая из m однородных процессоров. Для построения паретовских множеств P_i, $i = 1, n/2$, рассматриваются элементы одного и того же множества N и, следовательно, их формирование является независимым друг от друга. Тогда построение множеств P_i может быть реализовано в режиме параллельных вычислений.

Объединение всех элементов, составляющих паретовские множества P_i, $i = 1, n/2$ обозначим через Z. Далее естественно предположить, что могут существовать условия, при выполнении которых подмножество Z совпадает с требуемым множеством Парето T. Один и тот же элемент из множества N может входить в множества P_i на нескольких парах критериев C. Далее будем использовать в наших рассуждениях понятие паретовского слоя. Множество N следующим образом представимо в виде объединения отдельных паретовских слоев [1]. Множество Парето включает в себя все недоминируемые элементы и является первым паретовским слоем. Паретовский слой с номером m представляет собой совокупность паретовских элементов на той части множества начальных данных, которая остается после удаления элементов, принадлежащих всем предыдущим слоям. Следовательно, для каждого элемента, входящего во второй и последующие паретовские слои, существует хотя бы один элемент из предыдущего слоя, который его доминирует.

Определение 1. Верхней критериальной границей некоторого паретовского слоя V_m является вектор L^+_m чьи координаты представляют собой максимум по предпочтению по каждой координате среди всех элементов, образующих этот паретовский слой.

Определение 2. Нижней критериальной границей паретовского слоя V_m является вектор L^-_m чьи координаты представляют собой минимум по предпочтению на каждой координате среди элементов данного паретовского слоя.

Из способа построения векторов L^+_m, L^-_m следует, что вектор L^+_m доминирует все альтернативы паретовского слоя V_m, а вектор L^-_m доминируется каждым его элементом.

Утверждение 2. Верхняя критериальная граница любого паретовского слоя V_m доминирует верхние критериальные границы всех последующих паретовских слоев V_{i+1}, V_{i+2},..., V_v

Доказательство. Пусть V_i и V_{i+1} представляют собой соседние паретовские слои, а векторы L^+_i и L^+_{i+1} — их верхние критериальные границы. Покажем, что вектор L^+_i доминирует вектор L^+_{i+1}. Предположим противоположное, что эти два вектора либо находятся между собой в отношении Парето, либо L^+_{i+1} доминирует L^+_i. По определению верхних критериальных границ существует элемент в слое V_{i+1}, который хотя бы по одной координате доминирует все элементы предыдущего слоя V_i. Это противоречит тому факту, что для каждого элемента из V_{i+1} в предыдущем слое V_i существует хотя бы один доминирующий его элемент. Из транзитивности отношения предпочтения следует, что вектор L^+_i доминирует вектор L^+_{i+2} который доминирует вектор L^+_{i+3}. Тогда L^+_i доминирует вектор L^+_{i+3}. Аналогичным образом можно доказать, что вектор L^+_i доминирует верхние критериальные оценки L^+_{i+4},...., L^+_v последующих паретовских слоев. Утверждение доказано.

Пусть подмножество $W = Z/N$. Элементы этого подмножества не входят ни в одно паретовское множество P_i, i= 1, $n/2$. Если W — пустое множество, т. е. Z совпадает с N, то, по утверждению 1, требуемое множество Парето T определено. Оно совпадает с множеством начальных данных N, и решение задачи заканчивается. Пусть W — непустое множество. Будем предполагать, что на каждом двухкритериальном подпространстве C_i после определения паретовского множества P_i существуют еще элементы, не входящие в него. Тогда на каждой паре C_i определяем второй паретовский слой. Покажем, что и в этом случае существуют условия, когда подмножество Z представляет собой решение рассматриваемой комбинаторной задачи.

Утверждение 3. Если выполняются следующие два условия:

1. Нижняя критериальная граница каждого из паретовских множеств P_i, $i = 1$, $n/2$ доминирует верхнюю критериальную границу второго паретовского слоя на двухкритериальном подпространстве C_i.

2. Существует хотя бы один элемент K из Z, который принадлежит всем паретовским множествам P_i, $i = 1$, $n/2$, то подмножество Z представляет собой требуемое множество Парето T.

Доказательство. Покажем, что каждый элемент из W доминируется хотя бы одним элементом из Z, что и докажет данное утверждение. По первому условию нижние критериальные границы паретовских множеств P_i, $i = 1, n/2$, доминируют верхние критериальные границы вторых паретовских слоев на соответствующих двухкритириальных пространствах C_i. Тогда, исходя из транзитивности отношения предпочтения, свойств критериальных границ паретовских множеств и справедливости утверждения 2, получаем, что на каждом отдельном двухкритериальном подпространстве C_i любой элемент множества P_i, $i = 1, n/2$ будет доминировать все элементы из множества N, которые принадлежат второму и последующим паретовским слоям. По второму условию существует элемент K из Z, который входит во все P_i, $i = 1, n/2$. Следовательно, K будет доминировать на каждом двухкритириальном пространстве C_i все элементы из N, которые принадлежат только второму и последующим паретовским слоям. По определению подмножества W все его элементы на каждой паре критериев и принадлежат именно второму и последующим паретовским слоям. Тогда каждый из них будет доминироваться элементом K на любой паре критериев, а, следовательно, и на всем n-мерном критериальном пространстве. Таким образом, получаем, что подмножество Z содержит все недоминируемые элементы на множестве N и представляет собой требуемое множество Парето T. Утверждение доказано.

Алгоритм построения Z применим при любом способе разбиения критериев на отдельные пары. Если размерность критериального пространства нечетна, то последняя группа критериев имеет размерность 3. Тогда требуется найти паретовское множество P_i на трехкритериальном подпространстве. Суть утверждений 1 и 3 и алгоритм их доказательств не изменятся.

Покажем, что операции построения паретовских слоев в двухкритериальных подпространствах, на которых основываются вышеприведенные рассуждения, представляют действия с упорядоченными структурами данных. Формирование множеств P_i и паретовских слоев в двухкритериальных пространствах C_i отвечает следующему факту из теории многокритериальной оптимизации [4]. Если элементы множества Парето в двухкритериальном пространстве упорядочить по возрастанию значения предпочтения одного из критериев, то по второму критерию эти же элементы будут следовать друг за другом в порядке убывания их предпочтения. Значения критериев представляют собой действительные числа. Тогда при построении паретовских множеств P_i и всех последующих паретовских слоев достаточно использовать алгоритм двоичного поиска в упорядоченных структурах данных, чья эффективность по сравнению с алгоритмами перебора достаточно высока. При выполнении условий утверждения 3 для нахождения решения рассматриваемой комбинаторной задачи вообще не применялись алгоритмы перебора элементов множества N.

При невыполнении условий утверждения 3 для построения множества Парето разработан итерационный алгоритм, описанный в третьем параграфе данной работы, использующий подмножество Z как свою первую итерацию.

2. Определение избыточности множества начальных данных *N*

Пусть условия утверждения 3 не выполняются. Можно предположить, что даже их частичное выполнение позволяет определить некоторое подмножество доминируемых элементов из множества N, используя только операции в упорядоченных структурах данных. Пусть первое условие утверждения 3 либо не выполняется, либо выполняется на некоторой части паретовских множеств P_i. Тогда существует группа H пар критериев C_j, $j \leq n$, таких, что нижняя критериальная граница соответствующего множества P_j не доминирует верхнюю критериальную границу второго паретовского слоя. Предположим, что второе условие утверждения 3 выполняется, т. е. существует элемент A, который входит во все множества $P_i = 1, n/2$. На парах критериев C_j из группы H продолжаем процесс построения паретовских слоев, начиная с третьего. Пусть на каждой паре критериев C_j определены паретовские слои с номерами S_j, для которых выполняется условие доминируемости их верхней критериальной границы нижней критериальной границей соответствующего паретовского множества P_j. Возможность существование таких паретовских слоев следует из справедливости утверждения 2. Любой элемент из W после представления множества N на каждой паре критериев $C_i = 1, n$ в виде объединения отдельных паретовских слоев принадлежит некоторому паретовскому слою с номером больше 1. Пусть существует подмножество элементов из W, которые на каждой паре критериев C_j из H входят только в паретовские слои с номерами больше или равными S_j. Обозначим это подмножество через J. Справедливо следующее утверждение.

Утверждение 4. Все элементы подмножества J являются доминируемыми на всем n-мерном критериальном пространстве.

Доказательство. Пусть $j < n$. Любой элемент из подмножества J на всех парах критериев C_q, не входящих в H, включается в паретовские слои, начиная со второго. Для пар критериев C_q элемент A входит во все паретовские множества P_q. По транзитивности отношения предпочтения и утверждению 2 каждый элемент из J будет доминируемым элементом A на всех парах критериев C_q. На каждой паре критериев из H элементы J принадлежат слоям с доминируемыми верхними критериальными границами. Тогда и на этих парах они также будут доминируемыми A. Таким образом, любой элемент из J доминируется A на n-мерном критериальном пространстве. Если $j = n$, то, как следует из приведенных рассуждений, все элементы из J и в этом случае будут доминироваться элементом A. Утверждение доказано.

Элементы подмножества J не могут войти в требуемое множество T. Алгоритм построения J представляет собой операции в упорядоченных структурах данных, не использующих операции перебора. Предположим, что номера паретовских слоев S_j, чьи верхние критериальные границы доминируются нижними границами множеств P_j, значительно меньше числа паретовских слоев на подпространствах C_j. Тогда подмножество J может содержать количество элементов достаточно близкое к их числу в множестве N.

3. Поиск недоминируемых элементов на отдельных паретовских слоях

Пусть проведено разбиение множества N на паретовские слои в подпространствах C_j, входящих в H и в соответствии с утверждением 4 сформировано подмножество J. Сформулируем подзадачу разработки алгоритма построения множества Парето T на основе построенных паретовских слоев. По утверждению 1 все элементы подмножества Z входят в T. Возьмем пару $C_к$ с максимальным числом построенных паретовских слоев и сформируем все последующие слои на новом множестве начальных данных N^l, исключив из N элементы подмножества J. Новые элементы множества T могут принадлежать на подпространстве $C_к$ любому паретовскому слою с номером больше 1.

Пусть $T_1 = Z$. Предположим, что некоторый элемент d принадлежит второму слою S_2 на подпространстве $C_к$. Тогда из алгоритма построения паретовских слоев и транзитивности отношения предпочтения следует, что элемент d хотя бы по одному из двух критериев подпространства $C_к$ превосходит, в смысле предпочтения, любой элемент как слоя S_2, так и последующих паретовских слоев. Следовательно, d может доминироваться в n-мерном пространстве только элементами, входящими в множество $P_к$, представляющего собой первый паретовский слой. Определим возможную доминируемость всех элементов второго паретовского слоя, сравнивая каждый из них с элементами из $P_к$. Пусть существуют элементы второго слоя, обозначим это подмножество через X_1, которые оказались недоминируемыми после сравнения их с элементами подмножества $P_к$. Тогда все они являются недоминируемыми на всем множестве начальных данных N_l. Включаем их в множество T и получаем его новую итерацию $T_2 = T_1 \cup X_1$. Все доминируемые элементы второго слоя исключаем из рассмотрения.

Рассмотрим третий паретовский слой. Примем без доказательства тот факт, что если некоторые элементы третьего слоя, обозначим их через X_2, не доминируются в n-мерном пространстве элементами из T_2, то они являются недоминируемыми на множестве N^l и включаются в множество Парето T. Тогда для определения доминируемости элементов третьего паретовского слоя достаточно сравнить их только с элементами множества T_2. Полагаем $T_3 = T_2 \cup X_2$ и доминируемые элементы третьего паретовского слоя исключаются из рассмотрения. Все элементы четвертого паретовского слоя проверяем на доминируемость элементами из T_3. Недоминируемые элементы слоя образуют подмножество X_3 и включаются в множество T. Формируем множество $T_4 = T_3 \cup X_3$ и переходим к элементам следующего паретовского слоя. Множество T_r, полученное после рассмотрения элементов последнего паретовского слоя представляет собой решение данной комбинаторной задачи.

Общая схема построения множества Парето на множестве начальных данных N может быть описана следующим образом:

1. Первый этап. Разбиение всего критериального пространства на двухкритериальные подпространства.

2. Нахождение паретовских множеств P_i. При выполнении условий утверждения 3 решение задачи заканчивается.

3. Построение паретовских слоев на двухкритериальных подпространствах C_i, $i = 1, n/2$. Если условия утверждения 4 выполняются, то получаем новое сокращенное множество начальных данных N^I.

4. *Второй этап.* Нахождение требуемого множества Парето T последовательным рассмотрением элементов отдельных паретовских слоев на выбранном двухкритериальном подпространстве $C_к$.

Заключение

Предложен двухэтапный алгоритм для решения оптимизационной задачи нахождения множества Парето на заданном конечном множестве начальных данных N. Разработан способ нахождения подмножества доминируемых элементов начального множества N путем построения паретовских слоев на двухкритериальных подпространствах не использующий переборные алгоритмы. Предложен итерационный алгоритм построения множества Парето.

Литература

1. Дубов, Ю. А. Многокритериальные модели формирования и выбора вариантов систем / Ю. А. Дубов, С. И. Травкин, В. Н. Якимец. – М.: Наука, 1986. – 296 с.

2. Посыпкин, М. А. Комбинированный параллельный алгоритм решения задачи о ранце / М. А. Посыпкин // Труды четвертой международной конференции «Параллельные вычисления и задачи управления» (Москва, 27–29 октября 2008 г.). – С. 177–189.

3. Чебаков, С. В. Алгоритм решения заданных комбинаторных задач на основе модели многокритериальной оптимизации / С. В. Чебаков, Л. В. Серебряная // Доклады БГУИР. – 2015. – № 4 (90). – С 16–22.

4. Kung, H. F. On Finding the Maxima of a set of Vectors / H. F. Kung, F. P. Preparata // Journal of the Association for Computing Machinery. – 1975. – Vol. 22. – P. 469–476.

S. V. Chebakov, L. V. Serebryanaya

ALGORITHM OF FINDING A SET OF PARETO ON A FINAL SET OF INITIAL DATA

A two-stage algorithm for solving the optimization task for finding the Pareto set on a given finite set of initial data *N* is proposed. A method is developed for finding a subset of dominated elements of the initial set *N* by constructing the Pareto layers on two-criterial subspaces that does not use backtracking algorithms. An iterative algorithm for constructing the Pareto set is proposed.

Статья поступила 24.05.2017