Выпуск 44(61) ОПТИМИЗАЦИЯ 1988 г.

Численные методы

УДК 519.8

ДИНАМИЧЕСКОЕ ПРОГРАММИРОВАНИЕ В ОДНОЙ ЗАДАЧЕ АГРЕГИРОВАНИЯ

В.А.Залгаллер

1. Постановка задачи. Имеется N упорядоченных положительных чисел $R_1,\dots,R_N$. Требуется, сохраняя порядок, разбить их на m групп так, чтобы придать минимальное значение сумме

$$S=\sum_{i=1}^{m} s_i x_i = min, \tag{1}$$

где x_i - количество (быть может, нулевое) последовательных чисел R_k, включаемых в m-ю группу;

$$\sum_{i=1}^{m} x_i = N; \tag{2}$$

коэффициенты s_i заданы, причем

$$s_i \geq s_2 \geq \dots \geq s_m > 0. \tag{3}$$

Одновременно требуется удовлетворить условию

$$\underbrace{(R_1+\dots+R_{x_1})}_{\substack{x_1 \text{ слагаемых,} \\ \text{возможно, } x_1=0}}\mu_1+\underbrace{(R_{x_1+1}+\dots+R_{x_1+x_2})}_{\substack{x_2 \text{ слагаемых,} \\ \text{возможно, } x_2=0}}\mu_2+\dots+\underbrace{(R_{N-x_m+1}+\dots+R_N)}_{\substack{x_m \text{ слагаемых,} \\ \text{возможно, } x_m=0}}\mu_m \leq M \tag{4}$$

при заданных коэффициентах

$$0<\mu_1<\mu_2<\dots<\mu_m. \tag{5}$$

Мы дали пример оптимизационной задачи целочисленного программирования со своеобразным комбинаторным ограничением (4). Задача имеет инженерное происхождение. Она поставлена В.И. Трушкиной. Задачу в области ее возникновения требуется решать при порядках величин $N \approx 50$, $m \approx 10$.

Ниже предлагается алгоритм численного решения задачи методом динамического программирования с использованием метода ветвей и границ для сокращения вычислений. Алгоритм ведет к одному из точных решений, не давая перечня всех точных решений, если их несколько.

2. <u>Геометрическая интерпретация</u>. На оси Ox декартовой плоскости отложим последовательно отрезки длиной $R_1, \dots, R_N$ (рис.I), т.е. отметим на оси Ox точки с координатами

$$c_i = \sum_{j=1}^{i} R_j \ . \qquad (6)$$

Каждому набору неотрицательных целых $x_1, \dots, x_m$ с $\sum_{i=1}^{m} x_i = N$ сопоставим выходящую из начала O ломаную линию, i-е звено которой при $x_i \neq 0$ имеет угловой коэффициент ("наклон") μ_i, а проекцией этого звена на ось Ox служит отрезок $[c_{n_i}, c_{n_i + x_i}]$, где $n_i = x_i + \dots + x_{i-1}$. Если $x_i = 0$, то звено ломаной вырождается в точку. Условие (5) означает выпуклость этой ломаной, а условие (4) требование, чтобы правый конец ломаной имел координату $y(c_N) \leq M$. Задача состоит в минимизации суммы (I) на классе таких допустимых ломаных.

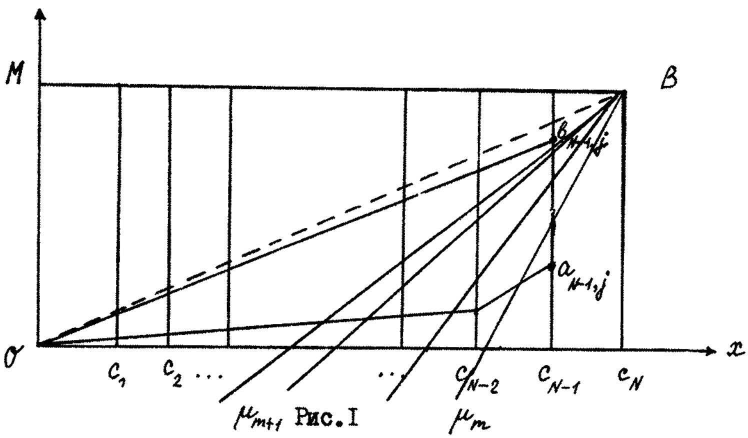

Рис.I

Ввиду выпуклости ломаных все они проходят не выше имеющего наклон $\xi = Mc_N^{-1}$ отрезка OB, где B - точка с координатами (c_N, M) (см. рис.I).

3. Условие разрешимости и случаи тривиальных решений. Если $\mu_1 > \xi$, то класс допустимых ломаных пуст, поскольку даже при минимальном наклоне μ_1 всех звеньев ломаной на ее правом конце будет $y(c_N) > M$.

Если $\mu_1 = \xi$, решением задачи служит $x_1 = N, x_k = 0$ при $k \neq 1$. Если $\mu_m \leq \xi$, решением задачи служит $x_m = N, x_k = 0$ при $k \neq m$. В дальнейшем исключаем эти тривиальные случаи и предполагаем, что $\mu_1 < \xi < \mu_m$.

Полезно разбить значения $\mu_1, \dots, \mu_m$ на небольшие и большие ξ:

$$0 < \mu_1 < \dots < \mu_p \leq \xi < \mu_{p+1} < \dots < \mu_m. \qquad (7)$$

4. Случаи понижения размерности задачи. Если $p = 1$, то над $[0, c_1]$ участок допустимой ломаной обязан идти с наклоном μ_1. В этом случае задачу можно переформулировать, уменьшив на единицу значение N. Далее считаем $p \geq 2$.

Если значение μ_m настолько велико, что $c_{N-1}\mu_1 + R_N\mu_m > M$, то ни одна допустимая ломаная не может иметь звено с наклоном μ_m. В этом случае можно отбросить сведения о μ_m, s_m и переформулировать задачу, уменьшив на единицу число m. Поэтому далее считаем, что

$$c_{N-1}\mu_1 + R_N\mu_m \leq M. \qquad (8)$$

5. Отрезки $[z_{ij}]$ и их разбиения. Над каждой точкой c_i оси $0x$ рассмотрим вертикальный отрезок $[z_i]$ длины $c_i\xi$. Вдоль него координата y лежит в пределах $[0, c_i\xi]$. Если вышедшая из точки 0 допустимая ломаная достигла отрезка $[z_i]$, причем на своем участке при подходе слева к $[z_i]$ имела наклон μ_j, то эта ломаная заведомо могла достичь отрезка $[z_i]$ не на всем его протяжении $[0, c_i\xi]$, а только в пределах более короткой его части $[a_{ij}, b_{ij}]$, которую нетрудно ограничить.

Ломаная могла вплоть до отрезка $[z_{i-1}]$ идти с наименьшим возможным наклоном μ_1 и только на участке от c_{i-1} до c_i с наклоном μ_j. Это позволяет положить

$$a_{ij} := c_{i-1}\mu_1 + R_i\mu_j. \qquad (9)$$

При выборе b_{ij} различаем два случая:

I) если $j \leq p$, то до $[z_i]$ ломаная шла с наклоном, не большим μ_j, поэтому полагаем

$$b_{ij} := c_i \mu_j ; \tag{10}$$

2) если $j > p$, то на оставшемся пути $c_N - c_i$ допустимая ломаная идет с наклоном, не меньшим μ_j, и подымается не выше M, поэтому полагаем

$$b_{ij} := M - (c_N - c_i)\mu_j . \tag{11}$$

Если в последнем случае окажется, что $a_{ij} > b_{ij}$, то подход слева к $[z_i]$ с наклоном μ_j для допустимой ломаной невозможен.

Обозначим через $[z_{ij}]$ части $[a_{ij}, b_{ij}]$ ($i = 1,\dots,N-1$, $j = 1,\dots,m$) отрезков $[z_i]$. В дальнейшем рассматриваем самостоятельно каждый из отрезков $[z_{ij}]$, отбросив те из них, для которых $a_{ij} > b_{ij}$.

В силу (9), (10) все отрезки $[z_{i1}]$ одноточечные: $a_{i1} = b_{i1}$. Кроме того, отрезки $[z_{im}]$ не потребуют в дальнейшем особого разбиения, так как за ними ломаная может идти только с наклоном μ_m. Поэтому общее число подвергаемых далее специальным разбиениям отрезков $[z_{ij}]$ не превосходит $(m-2)(N-1)$.

Для каждой точки $y \in [z_{ij}]$ существует конечное число возможных продолжений допустимой ломаной, если она пришла слева в точку y с наклоном μ_j. Поэтому существует минимальное значение $S(y)$ вклада, который внесет это продолжение в сумму (I) и номер $k(y)$ наклона $\mu_{k(y)}$ первого звена того продолжения, которое реализует $S(y)$. Если значение $S(y)$ достигается при нескольких разных k, то через $k(y)$ обозначим наибольшее из этих k.

Значения $S(y)$, $k(y)$ дискретны. Поэтому отрезок $[z_{ij}] = [a_{ij}, b_{ij}]$ разбивается на следующие друг за другом отрезки $q_0^{ij}, q_1^{ij}, \dots, q_{r(i,j)}^{ij}$, на каждом из которых значения $S(y)$, $k(y)$ постоянны. Отрезок q_0^{ij} замкнут и может состоять всего из одной точки a_{ij}. Все последующие отрезки открыты слева и замкнуты справа. Вместо $S(y), k(y)$ мы вправе писать $S(q_\ell^{ij}), k(q_\ell^{ij})$. При этом

$$S(q_{\ell+1}^{ij}) \geq S(q_\ell^{ij}), \tag{12}$$

и если $S(q_{\ell+1}^{ij}) = S(q_\ell^{ij})$, то $k(q_{\ell+1}^{ij}) < k(q_\ell^{ij})$. На отрезках $[z_{im}]$ есть всего одна часть

$$q_0^{im}=[z_{im}] \text{ с } S(q_0^{im})=(N-i)s_m, \; K(q_0^{im})=m.$$

6. Построение начального разбиения отрезков $[z_{N-1,j}]$. Проведем через точку B с координатами (c_N, M) прямые с наклонами $\mu_{p-1}, \dots, \mu_m$. На рис.I эти прямые обозначены указанием их наклонов. Ввиду (8), (II) хоть одна из этих прямых достигает отрезков $[z_{N-1,j}]$. Они разобьют каждый отрезок $[a_{N-1,j}, b_{N-1,j}]$ на части $q_0^{N-1,j}, \dots, q_{r(N-1,j)}^{N-1,j}$, причем $r(N-1,j) \leq m-p+1$. Для каждой из этих частей очевидны значения $S(q_l^{N-1,j})$, $K(q_l^{N-1,j})$.

7. Динамическое программирование. Допустим, что разбиения всех отрезков $[z_{i+1,k}]$ при $K=1,\dots,m$ уже построены. Это позволяет строить разбиение отрезка $[z_{ij}]$ следующим образом. Для каждого $K=j, j+1, \dots, m$ проектируем "с наклоном μ_K" разбиение отрезка $[z_{i+1,k}]$ на отрезок $[z_{ij}]$ (рис.2). Для каждого K получаем таким проектированием из каждой части $q_l^{i+1,k}$ на $[z_{i+1,k}]$ некоторый участок $\tilde{q}^{ij}$ на $[z_{ij}]$. На участке $\tilde{q}^{ij}$ заведомо достижимо значение $\tilde{S}(\tilde{q}^{ij})=s_k+S(q_l^{i+1,k})$ при $K(\tilde{q}^{ij})=K$. Наложение таких разбиений отрезка $[z_{ij}]$ при разных $K=j, j+1, \dots, m$, позволяет построить окончательное разбиение $[z_{ij}]$ на части q_l^{ij} с указанием $S(q_l^{ij})$ и $K(q_l^{ij})$. (В зоне перекрытий участков $\tilde{q}$, построенных для разных K, отдается предпочтение меньшему $S(y)$, а при равных $S(y)$ – большему K. Полученные участки с общим $S(y)$, $K(y)$ объединяются при формировании q_l^{ij}.)

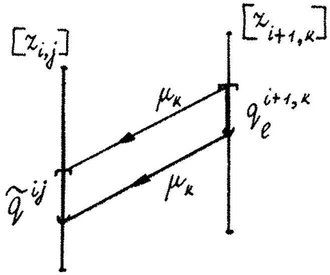

Рис.2

После того, как таким образом будут индуктивно построены разбиения всех отрезков $[z_{ij}]$, легко строится точное решение – оптимальная ломаная. Из точки $\mathcal{O}$ ее первое звено на участке $[0,c_1]$ идет с тем наклоном μ_i, для которого минимально (по всем $i=1,\dots,p$) значение $s_i+S(y_i)$, где y_i – точка с координатой $\mu_i R_1$ на отрезке $[z_{1i}]$. При равенстве значений $s_i+S(y_i)$ для разных i выбирается то i, для которого больше $k(y_i)$. Значение $k(y_i)$ при выбранном i укажет ломаной на следующем участке (до отрезка $[z_2]$) и т.д. Ломаная построится однозначно. Значения $x_1,\dots,x_m$ определяются ее строением.

8. <u>Отсечение ветвей.</u> Пусть нам известно, что существует допустимая ломаная с довольно малым "рекордным" для нескольких просмотренных ломаных значением $S_{рек}$ суммы (I). "Ветвь", состоящая из тех допустимых ломаных, которые проходят отрезок q^{ij}_{ℓ} (с подходом к нему слева с наклоном μ_j) дает сумму (I), оцениваемую снизу:

$$S \geq i s_j + S(q^{ij}_{\ell}). \tag{I3}$$

Если правая часть неравенства (I3) превосходит $S_{рек}$, то весь участок q^{ij}_{ℓ} можно удалить из рассмотрения при разбиении отрезка $[z_{ij}]$. При этом отрезок $[z_{ij}]$ лишь укорачивается, что сокращает дальнейшую реализацию описанного выше динамического программирования.

9. <u>Начальный выбор рекорда $S_{рек}$</u>. I) Для каждой фиксированной пары i,j номеров с $1\leq i\leq p$, $p+1\leq j\leq m$ в классе не более чем двух двузвенных отрезок с наклонами звеньев μ_i, μ_j легко обнаруживается ломаная с наименьшим значением суммы (I): проведем через точку $\mathcal{O}$ прямую с наклоном μ_i, а через B – прямую с наклоном μ_j. Пусть (α,β) – координаты точки пересечения этих прямых и

$$c_{t-1} \leq \alpha < c_t . \tag{I4}$$

Тогда лучшей в указанном классе ломаной будет та, для которой $x_i=t$; $x_j=N-t$; $x_k=0$ при $k\neq i,j$. Ей отвечает значение

$$S=s_i t + s_j (N-t). \tag{I5}$$

Просмотрев $p(m-p)$ пар номеров i,j, можно выбрать в качестве $S_{рек}$ меньшее из значений (I5).

2) Полезно найти значение S для ломаной, получаемой наивным стремлением идти все время с наибольши наклоном, при котором еще продолжение отрезка ломаной не идет выше точки B. Участок от O до $[z_1]$ идет с наклоном μ_p. Дойдя до $[z_1]$, выбираем наибольший допустимый наклон (он будет не меньше μ_p) и идем с ним до $[z_2]$ и т.д. Подсчет значения S для этой ломаной может иногда улучшить ранее выбранное значение $S_{рек}$.

Недостаток последнего варианта в том, что выбираемая ломаная не зависела от конкретных значений $S_1 \geq S_2 \geq \ldots \geq S_m$. Улучшить выбор ломаной может следующий вариант.

3) Подсчитаем таблицу разностей $S_{i-1} - S_i$ и выделим только те значения i, для которых

$$S_{i-1} - S_i \geq d, \tag{16}$$

назначив $d > 0$ таким, чтобы число выделенных значений i было не менее трех-четырех.

После этого, сохранив только μ_1 и μ_i с выделенными номерами. осуществляем построение ломаной, как в п.2). Это дает существенную дополнительную возможность улучшить "начальный рекорд" $S_{рек}$.

Последний способ позволяет, варьируя d, легко получать приближенные решения задачи.

Знание удачного начального рекорда $S_{рек}$ позволит сократить лежащий в начале индуктивной схемы динамического программирования перечень разбиений отрезков $[z_{N-1,j}]$, отрезки с малыми значениями $j = 1,2,3$ могут оказаться полностью отброшенными или существенно укороченными снизу из-за того, что сумма $S_j(N-1) + S(y)$ при всех $y \in [z_{N-1,j}]$ или при $y \in q_0^{N-1,j}$, $y \in q_1^{N-1,j}, \ldots, y \in q_k^{N-1,j}$ превысит $S_{рек}$.

10. <u>Текущий рекорд.</u> После того, как при фиксированном i построены все разбиения отрезков $[z_{ij}]$, для каждого из разбиений $[z_{ij}] = \bigcup_{\ell=0}^{r(i,j)} q_\ell^{ij}$ легко найти при данных j, ℓ наименьшее значение $S_{j\ell}$ суммы (1), которую могут давать ломаные, имеющие на участке до $[z_i]$ не более чем два звена, последнее из которых имеет наклон μ_j, и пересекающие отрезок $[z_{ij}]$ в точке с координатой y, не большей $\{\max y \mid y \in q_\ell^{ij}\}$.

Значение $S = \min\limits_{j,\ell} S_{j\ell}$ способно улучшить $S_{рек}$, а знание лучшего рекорда может позволить отбросить некоторые из

$[x_{ij}]$ или укоротить их, отбросив некоторые q_e^{ij}. Это сократит построение разбиений следующих отрезков $[x_{i-1,k}]$.

Кроме того, такой подход позволяет найти окончательное решение задачи по разбиениям отрезков $[x_{2,j}]$ без построения всех разбиений $[x_{1,j}]$.

Поступила в ред.-изд. отдел
14.07.1988 г.