РЕШЕНИЕ ЗАДАЧ РАСПРЕДЕЛИТЕЛЬНОГО ТИПА МЕТОДОМ ОБОБЩЕННОГО ГРАДИЕНТНОГО СПУСКА

Г.И.Горбач, Н.З.Шор
(г. Киев)

В Институте кибернетики АН УССР широко используется метод обобщенного градиентного спуска для решения распределительных задач линейного программирования с матрицами больших размеров. Его основные особенности:

1) метод удобен для программирования на ЭВМ;

2) рабочие массивы имеют небольшие размеры, что позволяет размещать их в оперативной памяти ЭВМ вместе с основной программой;

3) число итераций, необходимых для получения решения с заданной точностью, слабо зависит от объема задачи. Это свойство делает данный метод конкурентноспособным по времени счета для больших задач по сравнению с методами, приводящими к решению за конечное число шагов.

Описанный ниже метод для решения задач распределительного типа может рассматриваться как частный случай метода для решения задач блочного программирования, описанного в [1] .

Рассмотрим задачу следующего вида.
Найти

$$\min \sum_{i,j} c_{ij}\, x_{ij} \tag{1}$$

при ограничениях

$$\sum_{i} x_{ij} \geqslant B_j \quad ; \tag{2}$$

$$\sum_{j} \kappa_{ij}^{(\alpha)} x_{ij} \leqslant T_i^{(\alpha)} ; \tag{3}$$

$$x_{ij} \geqslant 0 \quad . \tag{4}$$

Перейдем к вспомогательной задаче.
Найти

$$\min \Big(\sum_{i,j} c_{ij} x_{ij} + \sum_{i,\alpha} u_i^{(\alpha)} \big(T_i^{(\alpha)} - \sum \kappa_{ij}^{(\alpha)} x_{ij} \big) \Big) \tag{5}$$

при ограничениях

$$\sum_{i} x_{ij} \geqslant B_j \quad , \tag{6}$$

$$x_{ij} \geqslant 0 \quad . \tag{7}$$

Эта задача имеет тривиальное решение

$$x_{i^*_{(j)},j} = B_j \quad , \tag{8}$$

$$x_{ij} = 0 \quad , \; i \neq i^* \quad ,$$

где $i^*_{(j)}$ – один из индексов, на котором достигается

$$\min_i \Big[c_{ij} + \sum_{\alpha} \kappa_{ij}^{(\alpha)} u_i^{(\alpha)} \Big] .$$

В соответствии с работой [1] получаем следующий алгоритм определения оптимальных значений $\{u_i^{*(\alpha)}\}$ двойственной задачи:

1) рассматриваются произвольные начальные значения

$$u^\circ = \{ u_{i\alpha}^{(0)} \} \; ; \; u_{i\alpha}^{(0)} = 0 \; ;$$

2) последовательность $u^{(r)} = \{ u_{i\alpha}^{(r)} \}$ вычисляется по следующим рекуррентным формулам:

$$u_{i\alpha}^{(r+1)} = \begin{cases} u_{i\alpha}^{(r)} + h r \Delta_{i\alpha}, \text{ если } u_{i\alpha}^{(r)} + h r \Delta_{i\alpha} \geqslant 0 , & (9) \\ 0 \text{ – в противном случае,} \end{cases}$$

$$\Delta_{i\alpha} = T_i^{(\alpha)} - \sum k_{ij}^{(\alpha)} x_{ij}^{(r)} \qquad (10)$$

получается как решение вспомогательной задачи (5–7) при $u_{i\alpha} = u_{i\alpha}^{(r)}$.

Из работы [2] вытекает, что последовательности $\{u_{i\alpha}^{(r)}\}$ сходятся к оптимальным значениям $\{u_{i\alpha}^{*}\}$ со скоростью геометрической прогрессии, если h_0 выбирается достаточно большим и последовательность $\{h_r\}$ вычисляется по формулам

$$h_{r+1} = h_r \, (1 - \delta \,) \quad ,$$

где δ – достаточно малое положительное число.

Однако этот метод в чистом виде не дает нам одновременно решения прямой задачи (1–4). Для получения приближенного решения этой задачи применяется так называемая процедура ε – сглаживания (ε – заданное положительное число), которая сводится к тому, что вместо формул (8) применяются следующие формулы:

$$x_{ij} = P_{ij} \, B_j \quad , \qquad (11)$$

где $$P_{ij} = \frac{\alpha_{ij}}{\sum \alpha_{ij}} \quad ;$$

$$(12)$$

$$\alpha_{ij} = \begin{cases} 0 \text{ , если } (c_{ij} + \sum\limits_{\alpha} k_{ij}^{(\alpha)} u_{i\alpha}) - v_j > \varepsilon , \\ \varepsilon - [(c_{ij} + \sum\limits_{\alpha} k_{ij}^{(\alpha)} u_{i\alpha}) - v_j] ; \end{cases} \quad (13)$$

в противном случае:

$$v_j = \min_i [c_{ij} + \sum_{\alpha} k_{ij}^{(\alpha)} u_{i\alpha}] . \quad (14)$$

Метод опробован на задачах по оптимальной загрузке прокатных станов. Эта работа выполнена в отделе экономической кибернетики ИК АН УССР в течение 1963–1967 гг. в содружестве с ВНИИОЧЕРМЕТом г. Харькова.

Для формировки задачи приняты такие обозначения.

Имеется m прокатных станов (индекс стана $i = 1, \ldots, m$). На всех станах производится n видов прокатной продукции ($k = 1, .., n$). На каждом из них производится часть всех видов продукции, в соответствии со специализацией каждого стана. Вся готовая продукция отправляется r потребителям (индекс потребителя $j = 1,\ldots, r$).

$P_{ik} \frac{\text{тонн}}{\text{час}}$ – часовая производительность i-го стана при прокатке k-го вида продукции;

$t_{ik} = \frac{1}{P_{ik}}$ – удельное время, т.е. время, необходимое для изготовления 1 тонны k-г вида продукции на i-м стане;

B_i – план производства в тоннах всех видов продукции на i-м стане;

T_i – рабочее время в часах для планируемого периода;

$E_{i\alpha}$ – ограничение к плану производства некоторых видов продукции на i-м стане;

M_{jk} - заявки j-го района на k-й вид продукции,

c_{ij} - расстояние (себестоимость перевозок) от i-го стана к j-му району потребления;

c'_{ik} - расходы по пределу, приходящиеся на 1 тонну k-го вида продукции на i-м стане.

Задачу оптимального распределения заказов между прокатными станами можно ставить различными способами в зависимости от выбранного критерия и типа ограничений.

а). Критерий - минимизация объема (себестоимости) транспортных перевозок при ограничениях на план производства в тоннах B_i и горячее время T_i на каждом стане. Необходимо разместить заказы районов-потребителей между прокатными станами, учитывая удельное время (t_{ik}) обработки k-го вида продукции на i-м стане.

$$\sum_j \sum_k t_{ik} x_{ijk} \leqslant T_i \quad , i = 1, \ldots, m ; \tag{15}$$

$$\sum_j \sum_k x_{ijk} \leqslant B_i \quad , \quad i = 1, \ldots, m ; \tag{16}$$

$$\sum_i x_{ijk} = M_{jk} \quad , \quad j = 1, \ldots, r ; \; k = 1, \ldots, n ; \tag{17}$$

$$x_{ijk} \geqslant 0 , \; i = 1, \ldots, m , \; j = 1, \ldots, r , \; k = 1, \ldots, n ; \tag{18}$$

найти

$$\min \sum_i \sum_j \sum_k c_{ij} x_{ijk} \; . \tag{19}$$

б). Критерий - учитывает, кроме транспортных издержек, себестоимость производства на станах. Ограничения - на горячее время на станах.

$$\sum_j \sum_k t_{ik} x_{ijk} \leqslant T_i \quad ; \tag{20}$$

$$\sum_i x_{ijk} = M_{jk} \quad ; \tag{21}$$

$$x_{ijk} \geq 0 \qquad (22)$$

найти

$$\min\Big[\sum_i c_i' \Big(\sum_j \sum_k t_{ik} x_{ijk}\Big) + \sum_i \sum_j \sum_k c_{ij} x_{ijk}\Big]. \qquad (23)$$

в). Критерий – как и в задаче а). Даны ограничения на план каждого стана в тоннах и ограничения на план производства для некоторых выделенных сортов проката.

$$x_{ijk} \geq 0 \; ; \qquad (24)$$

$$\sum_j \sum_k x_{ijk} \leq B_i \;, \; i = 1, \ldots, m \; ; \qquad (25)$$

$$\sum_i x_{ijk} = M_{jk} \;, \; j = 1, \ldots, r \;, \; k = 1, \ldots, n \; ; \qquad (26)$$

$$\sum_j \sum_k r_{ik\alpha} x_{ijk} \leq E_{i\alpha} \;, \qquad (27)$$

где

$r_{ik\alpha} = \begin{cases} 0 \\ 1 \end{cases}$ – технологический коэффициент;

α – тип ограничений.

Если в формулировках задач (**а, б, в**) вместо пары индексов (j,k) применить один индекс j, то мы приходим к задачам типа (1-4). Отсюда получаются следующие алгоритмы решения вышеприведенных задач по методу обобщенного градиентного спуска со сглаживанием.

Описание алгоритма решения задачи типа (а)

Переходим к двойственной задаче:

найти

$$\max\Big(\sum_j \sum_k M_{jk} \lambda_{jk} - \sum_i T_i v_i - \sum_i B_i u_i\Big) \qquad (28)$$

при условии

$$\lambda_{jk} - t_{ik} v_i - u_i \leqslant c_{ij} \quad . \tag{29}$$

Отсюда в оптимальном плане:

$$\lambda_{jk} = \min_i \left(c_{ij} + t_{ik} v_i + u_i \right) \quad ; \tag{30}$$

u_i, v_i, λ_{jk} – новые переменные двойственной задачи. Начальные значения переменных u_i, v_i выбираются равными 0. ε – выбирается равной

$$\approx 1\% \max_{i,j} c_{ij} \, .$$

Начальное значение коэффициента h рекомендуется выбирать в зависимости от исходных данных по формуле, полученной эмпирическим путем,

$$h_0 = \frac{20 \cdot \max c_{ij} \cdot m}{S \cdot \sum B_i} \quad , \tag{31}$$

S – количество итераций при постоянном шаге.

Вычисляем

$$q_{ijk} = c_{ij} + t_{ik} v_i + u_i \quad . \tag{32}$$

Определяем

$$\lambda_{jk} = \min_i \left(c_{ij} + t_{ik} v_i + u_i \right) \, . \tag{33}$$

Находим отклонения от минимального элемента

$$q_{ijk} - \lambda_{jk} = a_{ijk} \; , \quad i = 1, \ldots, m \, . \tag{34}$$

Если полученное отклонение больше выбранного отклонения ε, тогда засылаем 0 вместо полученного результата.

Проверка этого условия:

$$\alpha_{ijk} = \varepsilon - a_{ijk} = \begin{cases} 0 \, , & \text{если} \quad \alpha_{ijk} \leqslant 0 \quad ; \\ \alpha_{ijk} \, , & \text{если} \quad \alpha_{ijk} > 0 \quad . \end{cases} \tag{35}$$

Определяем P_{ijk} по формуле

$$P_{ijk} = \frac{\alpha_{ijk}}{\sum\limits_i \alpha_{ijk}} . \tag{36}$$

Величина P_{ijk} представляет ту часть заказа, которая будет прикрепляться к i -му стану, а величина заказа равна $M_{jk} \cdot P_{ijk}$. Это и есть текущие значения x_{ijk}, т.е. значения неизвестных в прямой задаче, полученные на данном шаге вычислений.

Определяем величину разбалансов:

$$B_i - \sum_j M_{jk} P_{ijk} = B_i' \quad ; \tag{37}$$

$$T_i - \sum_{j,k} M_{jk} P_{ijk} t_{ik} = T_i' . \tag{38}$$

В конце итерации по станам получаем разбалансы в тоннах B_i' и по времени T_i'. Если станы перегружены, соответствующие значения разбалансов имеют отрицательный знак; недогруженные – положительные значения разбалансов.

Потенциалы перегруженных станов повышаем пропорционально полученным разбалансам; потенциалы недогруженных – понижаем по каждому из типов ограничений.

Изменение потенциалов проводим по следующей формуле:

$$u_i^0 - hB_i' = \begin{cases} 0, & \text{если} \quad u_i \leqslant 0 , \\ u_i', & \text{если} \quad u_i > 0 ; \end{cases} \tag{39}$$

$$v_i^0 - hT_i' \left(\frac{B_i}{T_i}\right)^2 = \begin{cases} 0, & \text{если} \quad v_i^1 \leqslant 0 , \\ v_i^1, & \text{если} \quad v_i > 0 . \end{cases} \tag{40}$$

После каждой итерации h умножается на (1–б) либо после определенного числа итераций с постоянным h – коэффициент измельчается (уменьшается в два раза).

Задача **б** является упрощенной модификацией задачи **а**, поэтому алгоритм решения задачи **б** получается из алгоритма решения задачи **а** путем очевидных упрощений.

Описание алгоритма решения задачи(в)

Запишем двойственную задачу:
найти

$$\max\Big[\sum_j\sum_k M_{jk}\lambda_{jk}-\sum_i B_i u_i-\sum_i\sum_\alpha E_{i\alpha}v_{i\alpha}\Big] \quad (41)$$

при условии

$$\lambda_{jk}-u_i-\sum_\alpha r_{ik\alpha}v_{i\alpha}\leqslant c_{ij}\ ,\quad u_i\geqslant 0\ . \quad (42)$$

В оптимальном случае:

$$\lambda_{jk}=\min_i\,(c_{ij}+u_i+\sum_\alpha r_{ik\alpha}v_{i\alpha})\ . \quad (43)$$

Краткое описание алгоритма:

1. $q_{ijk}=c_{ij}+u_i+\sum_\alpha r_{ik\alpha}v_{i\alpha}$.
2. $\lambda_{jk}=\min_i(c_{ij}+u_i+\sum_\alpha r_{ik\alpha}v_{i\alpha})$.
3. $a_{ijk}=q_{ijk}-\lambda_{jk}$.
4. $\max(0,\ \varepsilon-a_{ijk})=\alpha_{ijk}$.
5. $p_{ijk}=\alpha_{ijk}/\sum_i\alpha_{ijk}$.
6. $B_i-\sum_j M_{jk}p_{ijk}=B_i'\,;\ E_{i\alpha}-\sum_{j,k}M_{jk}p_{ijk}r_{ik\alpha}=E_{i\alpha}'$
7. $$u_i^{s+1}=u_i^s-hB_i'=\begin{cases}0\ , & \text{если } u_i^{s+1}\leqslant 0\ ,\\ u_i^{s+1}\ , & \text{если } u_i^{s+1}>0\ .\end{cases}$$
$$v_{i\alpha}^{s+1}=v_{i\alpha}^s-h_1E_{i\alpha}'=\begin{cases}0\ , & \text{если } v_{i\alpha}^{s+1}\leqslant 0\ ,\\ v_{i\alpha}^{s+1} & \text{если } v_{i\alpha}^{s+1}>0\ .\end{cases}$$

Кроме вышеперечисленных задач, решались задачи на максимум выпуска продукции в заданном соотношении или на минимум времени, необходимого для производства заданного количества продукции. Вторая из этих задач формулируется следующим образом:

$$\sum_i x_{ik}\geqslant M_k\ ,\quad k=1,\dots,n\ ; \quad (44)$$

$$\sum_j t_{ik}x_{ik}=\theta T_i\ ,\quad i=1,\dots,m\ ; \quad (45)$$

$$x_{ik}\geqslant 0\ ,\quad i=1,\dots,m\ ,k=1,\dots,n\ . \quad (46)$$

Найти

$$Z=\min\theta\ . \quad (47)$$

Обозначим через

$$y_{ik} = \frac{x_{ik}}{M_k} \quad . \tag{49}$$

Тогда задача (47) примет вид:

$$y_{ik} \geqslant 0 \quad ; \tag{49}$$

$$\sum_i y_{ik} = 1 \ , \quad k = 1, \ldots, n . \tag{50}$$

$$\sum_k \frac{t_{ik} M_k}{T_i} y_{ik} \geqslant \theta \quad , \tag{51}$$

$$Z = \min \theta .$$

Построим двойственную задачу:
найти

$$L = \max \left(\sum_k v_k \right) \tag{52}$$

при ограничениях

$$v_k - \frac{t_{ik} M_k}{T_i} u_i \leqslant 0 \ ; \tag{53}$$

$$\sum_i u_i = 1 \quad . \tag{54}$$

В двойственной задаче переменные u_i и v_k называются потенциалами. Для решения такой задачи применяется метод обобщенного градиентного спуска в пространстве потенциалов u_i и v_k . Из неравенства (53) получаем:

$$v_k \leqslant \frac{t_{ik} M_k}{T_i} u_i \quad ,$$

отсюда

$$v_k = \min_i \frac{t_{ik} M_k}{T_i} u_i \quad . \tag{55}$$

Начальные значения потенциалов u_i^o выбираем равными между собой так, чтобы выполнялось условие (54). Опыт вычислений показал, что лучшими являются

$$u_1^o = u_2^o = \ldots = u_m^o = \frac{1}{m} \quad .$$

При фиксированном κ определяем

$$\lambda_{i\kappa} = \frac{M_\kappa t_{i\kappa}}{T_i} u_i \quad , \quad i = 1,\ldots,m \; .$$

Находим

$$v_\kappa = \min_i \frac{M_\kappa t_{i\kappa}}{T_i} u_i \quad .$$

Определяем область сглаживания :

$$\ell_\kappa = 0{,}03\, v_\kappa \quad ,$$

$$\ell_\kappa - \lambda_{i\kappa} = \Delta\varepsilon_{i\kappa} \quad .$$

Вычисляем

$$P_{i\kappa} = \frac{\Delta\varepsilon_{i\kappa}}{\sum_\kappa \Delta\varepsilon_{i\kappa}} \quad .$$

Находим

$$r_\kappa = \frac{M_\kappa t_{i\kappa}}{T_i} P_{i\kappa} \quad .$$

Распространяя суммирование по всем κ, получаем значение

$$\alpha_i = \sum_\kappa \frac{M_\kappa t_{i\kappa}}{T_i} P_{i\kappa} \; .$$

Находим среднее арифметическое по всем станам:

$$R^m = \frac{1}{m} \sum_i \sum_\kappa \frac{M_\kappa t_{i\kappa}}{T_i} P_{i\kappa} \; .$$

Определяем величину отклонения на каждом стане от среднего арифметического:

$$\Delta\alpha_i = \sum_\kappa \frac{M_\kappa t_{i\kappa}}{T_i} P_{i\kappa} - R^m \; .$$

Изменение потенциалов проводим по формуле

$$\Delta u_i = h \Delta \alpha_i \ ,$$

$$u_i^{s+1} = u_i^s (1 + \Delta u_i) \ ,$$

где s – число итераций. После определенного числа итераций h уменьшается в два раза. Значение целевой функции θ равно обратному значению среднего арифметического;

а) $u_0 < 1$ – в этом случае можно не только разместить все заказы на станах, но и получить дополнительную продукцию в том же сортаменте;

б) $u_0 = 1$ – в этом случае можно разместить все заказы, но дополнительную продукцию получить нельзя;

в) $u_0 > 1$ – разместить все заказы в этом случае нельзя, но точно известно, на сколько надо уменьшить значения M_k, чтобы можно было разместить на станах все заказы в планируемый период.

Для решения задач **а, б** и **в**, а также задачи (44–47) в отделе экономической кибернетики ИК АН УССР разработаны стандартные программы для ЭВМ М-20. С помощью этих программ проведены многочисленные экспериментальные расчеты по оптимальной загрузке прокатных станов. Максимальный объем решенных задач типа **а)**: 8000 заказов, 116 районов потребления, 50 прокатных станов. Время счета этой задачи – 26 час., количество итераций – 500.

Аналогичные стандартные программы разработаны ВНИИОЧЕРМЕТом для ЭВМ „Минск-22". С помощью этих программ в конце 1967 г. было проведено внедрение планирования оптимальной загрузки для проволочных станов в „Союзглавметалле".

В ИК АН УССР на ЭВМ М-20 осуществлена большая серия расчетов по оптимальному распределению флота по видам перевозок и линиям движения, многие из которых внедрены в практику работы судоходства в Днепровском бассейне.

Таким образом, в настоящее время накоплен значительный опыт применения обобщенного градиентного спуска для решения разнообразных задач распределительного типа.

Литература

1. Н.З. Шор, Применение обобщенного градиентного спуска в блочном программировании, журн. „Кибернетика", №3, 1967.

2. Н.З.Шор, О скорости сходимости обобщенного градиентного спуска, журн. „Кибернетика", №3, 1968 (в печати).

Доложено на семинаре 26.II 1968 г.