МИНОБРНАУКИ РОССИИ

Федеральное государственное бюджетное образовательное учреждение высшего образования
«САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ ЭКОНОМИЧЕСКИЙ УНИВЕРСИТЕТ»

Факультет информатики и прикладной математики
Кафедра прикладной математики и экономико-математических методов

ОТЧЁТ

по дисциплине:

«Модели комбинаторной оптимизации»

на тему:

«Задание №1»

Направление: 01.03.02
Обучающийся: Бронников Егор Игоревич
Группа: ПМ-1901

Санкт-Петербург
2022

Задача №1

Сформулировать точную постановку задачи разбиения множества векторов на k групп (MDMWNPP) как задачи целочисленного программирования. Каждый вектор имеет длину l, при этом l_1 количественных характеристик (положительные вещественные числа), а l_2 качественных ($l_1 + l_2 = l$). Для каждой из качественных дано множество допустимых значений. По исходным данные рассчитываются «идеальные» суммарные характеристики для групп по количественным характеристикам и по количеству представителей в группах различных значений качественных характеристик. Целевая функция: минимизация суммы абсолютных отклонений суммарных характеристик групп от идеальных значений.

Решение

k – количество групп
n – количество векторов
l – количество элементов в векторе
$S = \{s_i = (s_{i1}, s_{i2}, \dots, s_{il}) \mid \forall i \in \{1, \dots, n\}\}$ – мультимножество
l_1 – количество количественных признаков в векторе
l_2 – количество качественных признаков в векторе
Должно выполняться следующее соотношение: $l = l_1 + l_2$
$D_i \quad \forall i \in \{l_1 + 1, \dots, l\}$ – множество допустимых значений для i-ой качественной характеристики
$s_{ij} \in \mathbb{R} \quad \forall i \in \{1, \dots, n\}, \forall j \in \{1, \dots, l_1\}$ – множество допустимых значений для количественных характеристик

Разобьём понятие «идеальных» суммарных характеристик для количественных признаков и для качественных признаков.

Пусть $\widetilde{p}_1$ – это «идеальная» сумма для количественных признаков, а $\widetilde{p}_2$ – это «идеальная» сумма для качественных признаков.

Тогда «идеальная» сумма для количественных признаков может быть найдена следующим образом:

$$\widetilde{p}_1 = \frac{\sum\limits_{i=1}^{n} \sum\limits_{j=1}^{l_1} s_{ij}}{k}$$

Теперь переходим к качественным признакам.

Пусть $r = \{|D_{l_1+1}|, \ldots, |D_l|\}$ – это вектор мощностей множеств D_i $\forall i \in \{l_1 + 1, \ldots, l\}$

Введём матрицы $I_i = \{y_{ijk}\}$ – элементами которых являются значения 0 или 1. (бинарная матрица)

$$y_{ijk} = \begin{cases} 1, & \text{if в } k\text{-том векторе присутствует } j\text{-тый признак} \\ 0, & \text{в противном случае} \end{cases}$$

$\forall i \in \{l_1 + 1, \ldots, l\},\ \forall j \in r,\ \forall k \in \{1, \ldots, n\}$

$\dim(I_i) = (n \times m)$, где n – это количество векторов, а m – это мощность D_i $\forall i \in \{l_1 + 1, \ldots, l\}$

Пример

Рассмотрим следующее мультимножество:
$S = \{(1, 2, 3, \text{red}, \text{big}, \text{triangle}), (1, 2, 3, \text{green}, \text{big}, \text{rectangle}), (2, 4, 5, \text{green}, \text{small}, \text{circle})\}$

Пусть даны следующие множества допустимых значений качественных признаков:
$D_1 = \{\text{red}, \text{green}, \text{yellow}\}$
$D_2 = \{\text{big}, \text{small}\}$
$D_3 = \{\text{triangle}, \text{rectangle}, \text{circle}\}$

Составим матрицу I_1 для первого качественного признака:

$$I_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

Составим матрицу I_2 для второго качественного признака:

$$I_2 = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Составим матрицу I_3 для третьего качественного признака:

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Тогда можно рассчитать «идеальную» сумму для каждого i-го качественного признака:

$$\widetilde{p}_{2_i} = \frac{\sum\limits_{j=1}^{r_i} I_{ij}}{k}$$

где I_{ij} – это j-тый столбец матрицы I_i

Пример

Продолжаем предыдущий пример.

$$\widetilde{p}_{2_1} = \left(\frac{1}{k}, \frac{2}{k}, 0\right)$$
$$\widetilde{p}_{2_2} = \left(\frac{2}{k}, \frac{1}{k}\right)$$
$$\widetilde{p}_{2_3} = \left(\frac{1}{k}, \frac{1}{k}, \frac{1}{k}\right)$$

Введём переменные x_{ij}.

$$x_{ij} = \begin{cases} 1, & \text{if } j\text{-ый вектор попал в } i\text{-ую группу} \\ 0, & \text{в противном случае} \end{cases}$$

$\forall i \in \{1, \dots, k\}$, $\forall j \in \{1, \dots, n\}$, где n – это количество векторов, а k – количество групп.

Введём p_1 – «фактическое» значение суммы для количественных признаков.

$$p_1 = \sum_{i=1}^{n} \sum_{j=1}^{l_1} x_{ij} \cdot s_{ij}$$

Введём p_{2_j} – «фактическое» значение суммы для j-го качественного признака.

$$p_{2_j} = x_i \cdot I_j$$

$\forall i \in \{1, \ldots, k\}, \ \forall j \in \{l_1 + 1, \ldots, l\}$, где k – количество групп.

Тогда можно составить целевую функцию.

Целевая функция:

$$|\widetilde{p}_1 - p_1| + \sum_{i=l_1+1}^{l} \sum_{j=1}^{r_i} |\widetilde{p}_{2_{ij}} - p_{2_{ij}}| \longrightarrow \min$$

НО, получившаяся целевая функция не удовлетворяет линейности, поэтому введём дополнительные переменные z_1 и $z_{2_{ij}}$.

Целевая функция:

$$z_1 + \sum_{i=l_1+1}^{l} \sum_{j=1}^{r_i} z_{2_{ij}} \longrightarrow \min$$

Ограничения:

$$\begin{cases} z_1 \geq p_1 - \widetilde{p}_1 \\ z_1 \geq \widetilde{p}_1 - p_1 \\ z_{2_{ij}} \geq \widetilde{p}_{2_{ij}} - p_{2_{ij}} \\ z_{2_{ij}} \geq p_{2_{ij}} - \widetilde{p}_{2_{ij}} \end{cases}$$

$$\sum_{i=1}^{k} x_{ij} = 1 \quad \forall j \in \{1, \ldots, n\}$$

$$x_{ij} \in \{0; 1\} \quad \forall i \in \{1, \ldots, k\}, \ \forall j \in \{1, \ldots, n\}$$

$$p_1 \in \mathbb{R}, \quad z_1 \in \mathbb{R}$$

$$p_{2_{ij}} \in \mathbb{R}, \ z_{2_{ij}} \in \mathbb{R} \quad \forall i \in \{l_1 + 1, \ldots, l\}, \ \forall j \in \{1, \ldots, r_i\}$$

Задача №2

Сформулировать точную постановку задачи разбиения мультимножества чисел на k групп (multiway NPP) как задачи целочисленного программирования. Дано множество пар позиций чисел из мультимножества. Если пара чисел на этих позициях оказывается в одной группе, то из суммы чисел в группе вычитается среднее от пары этих чисел. Целевая функция: минимизация разности между наибольшей и наименьшей суммой чисел среди групп, при этом суммы чисел в группах находятся с учетом «штрафных баллов».

Решение

k – количество групп
n – количество векторов
$S = \{s_1, s_2, \ldots, s_n\}$ – мультимножество
$D = \{(a_1, b_1), (a_2, b_2), \ldots, (a_m, b_m)\}$ - множество пар позиций чисел из мультимножества
m – мощность множества пар позиций чисел

Пример

Мультимножество:
$S = \{10, 20, 30, 40, 50, 60, 70, 80, 90, 100\}$

Множество пар позиций чисел из этого мультимножества:
$D = \{(1, 2), (2, 3), (4, 5), (4, 6), (7, 8)\}$

Пусть p_i – это сумма чисел в i-ой группе. $\forall i \in \{1, \ldots, k\}$
Тогда пусть $p_{\max}$ – это максимальная сумма среди всех групп, а $p_{\min}$ – это минимальная сумма среди всех групп.

Введём переменные x_{ij}.

$$x_{ij} = \begin{cases} 1, & \text{if } j\text{-ое число попало в } i\text{-ую группу} \\ 0, & \text{в противном случае} \end{cases}$$

$$\forall i \in \{1, \ldots, k\}, \quad \forall j \in \{1, \ldots, n\}$$

Введём параметры y_{ij}.

$$y_{lj} = \begin{cases} 1, & \text{if в паре } i \text{ есть } j\text{-ое число} \\ 0, & \text{в противном случае} \end{cases}$$
$$\forall l \in \{1, \ldots, m\}, \quad \forall j \in \{1, \ldots, n\}$$

Если $\sum\limits_{j=1}^{n} x_{ij} \cdot y_{lj} \geq 2$, то значит в группе есть числа, позиции которых находятся в одной паре. В противном случае в группе нет чисел, позиции которых находятся в одной паре.

Хотелось бы сделать так:

if $\sum\limits_{j=1}^{n} x_{ij} \cdot y_{lj} \geq 2$

then:

$\sum\limits_{j=1}^{n} x_{ij} \cdot s_j - \sum\limits_{j=1}^{n} \frac{a_{lj} \cdot s_j}{2} \leq p_{\max}$

$\sum\limits_{j=1}^{n} x_{ij} \cdot s_j - \sum\limits_{j=1}^{n} \frac{a_{lj} \cdot s_j}{2} \geq p_{\min}$

else:

$\sum\limits_{j=1}^{n} x_{ij} \cdot s_j \leq p_{\max}$

$\sum\limits_{j=1}^{n} x_{ij} \cdot s_j \geq p_{\min}$

$\forall i \in \{1, \ldots, k\}, \quad \forall l \in \{1, \ldots, m\}, \quad \forall j \in \{1, \ldots, n\}$

Импликацию можно заменить следующим образом:

$$A_1 \Rightarrow A_2 \longrightarrow \text{NOT } A_1 \text{ OR } A_2.$$

Если рассуждать на языке алгебры логики, то имеются следующие события A и B:

A: $\sum\limits_{j=1}^{n} x_{ij} \cdot y_{lj} < 2$ OR $(\sum\limits_{j=1}^{n} x_{ij} \cdot s_j - \sum\limits_{j=1}^{n} \frac{a_{lj} \cdot s_j}{2} \leq p_{\max}$ AND $\sum\limits_{j=1}^{n} x_{ij} \cdot s_j - \sum\limits_{j=1}^{n} \frac{a_{lj} \cdot s_j}{2} \geq p_{\min})$

B: $\sum_{j=1}^{n} x_{ij} \cdot y_{lj} \geq 2$ OR $(\sum_{j=1}^{n} x_{ij} \cdot s_j \leq p_{\max}$ AND $\sum_{j=1}^{n} x_{ij} \cdot s_j \geq p_{\min})$

Нам бы хотелось в конечном итоге, чтобы выполнялись условия: A OR B.

Рассмотрим более подробно событие A.
Пусть M_A – большое число и введём параметр $z_A \in \{0; 1\}$, тогда событие A можно переписать следующим образом:

$$\begin{cases} \sum_{j=1}^{n} x_{ij} \cdot y_{lj} < 2 + M_A \cdot z_A \\ \sum_{j=1}^{n} x_{ij} \cdot s_j - \sum_{j=1}^{n} \frac{a_{lj} \cdot s_j}{2} \leq p_{\max} + M_A \cdot (1 - z_A) \\ \sum_{j=1}^{n} x_{ij} \cdot s_j - \sum_{j=1}^{n} \frac{a_{lj} \cdot s_j}{2} \geq p_{\min} + M_A \cdot (1 - z_A) \end{cases}$$
$$\forall i \in \{1, \ldots, k\}, \quad \forall l \in \{1, \ldots, m\}, \quad \forall j \in \{1, \ldots, n\}$$

Рассмотрим более подробно событие B.
Пусть M_B – большое число и введём параметр $z_B \in \{0; 1\}$, тогда событие B можно переписать следующим образом:

$$\begin{cases} \sum_{j=1}^{n} x_{ij} \cdot y_{lj} \geq 2 + M_B \cdot z_B \\ \sum_{j=1}^{n} x_{ij} \cdot s_j \leq p_{\max} + M_B \cdot (1 - z_B) \\ \sum_{j=1}^{n} x_{ij} \cdot s_j \geq p_{\min} + M_B \cdot (1 - z_B) \end{cases}$$
$$\forall i \in \{1, \ldots, k\}, \quad \forall l \in \{1, \ldots, m\}, \quad \forall j \in \{1, \ldots, n\}$$

Осталось всё собрать вместе и рассмотреть событие: A OR B.
Пусть M – большое число, а ε – маленькое число и введём параметр $z \in \{0; 1\}$, тогда событие A OR B можно переписать следующим образом:

$$\begin{cases} \sum\limits_{j=1}^{n} x_{ij} \cdot y_{lj} \leq 2 + M_A \cdot z_A + M \cdot z - \varepsilon \\ \sum\limits_{j=1}^{n} x_{ij} \cdot s_j - \sum\limits_{j=1}^{n} \dfrac{a_{lj} \cdot s_j}{2} \leq p_{\max} + M_A \cdot (1 - z_A) + M \cdot z \\ \sum\limits_{j=1}^{n} x_{ij} \cdot s_j - \sum\limits_{j=1}^{n} \dfrac{a_{lj} \cdot s_j}{2} \geq p_{\min} + M_A \cdot (1 - z_A) + M \cdot z \\ \sum\limits_{j=1}^{n} x_{ij} \cdot y_{lj} \geq 2 + M_B \cdot z_B + M \cdot (1 - z) \\ \sum\limits_{j=1}^{n} x_{ij} \cdot s_j \leq p_{\max} + M_B \cdot (1 - z_B) + M \cdot (1 - z) \\ \sum\limits_{j=1}^{n} x_{ij} \cdot s_j \geq p_{\min} + M_B \cdot (1 - z_B) + M \cdot (1 - z) \end{cases}$$

$$\forall i \in \{1, \ldots, k\}, \quad \forall l \in \{1, \ldots, m\}, \quad \forall j \in \{1, \ldots, n\}$$

Целевая функция:

$$p_{\max} - p_{\min} \longrightarrow \min$$

Ограничения:

$$\begin{cases}
\sum\limits_{j=1}^{n} x_{ij} \cdot y_{lj} \le 2 + M_A \cdot z_A + M \cdot z - \varepsilon \\
\sum\limits_{j=1}^{n} x_{ij} \cdot s_j - \sum\limits_{j=1}^{n} \dfrac{a_{lj} \cdot s_j}{2} \le p_{\max} + M_A \cdot (1 - z_A) + M \cdot z \\
\sum\limits_{j=1}^{n} x_{ij} \cdot s_j - \sum\limits_{j=1}^{n} \dfrac{a_{lj} \cdot s_j}{2} \ge p_{\min} + M_A \cdot (1 - z_A) + M \cdot z \\
\sum\limits_{j=1}^{n} x_{ij} \cdot y_{lj} \ge 2 + M_B \cdot z_B + M \cdot (1 - z) \\
\sum\limits_{j=1}^{n} x_{ij} \cdot s_j \le p_{\max} + M_B \cdot (1 - z_B) + M \cdot (1 - z) \\
\sum\limits_{j=1}^{n} x_{ij} \cdot s_j \ge p_{\min} + M_B \cdot (1 - z_B) + M \cdot (1 - z) \\
\sum\limits_{i=1}^{k} x_{ij} = 1
\end{cases}$$

$$\forall i \in \{1, \dots, k\}, \quad \forall l \in \{1, \dots, m\}, \quad \forall j \in \{1, \dots, n\}$$

$$x_{ij} \in \{0; 1\} \quad \forall i \in \{1, \dots, k\}, \; \forall j \in \{1, \dots, n\}$$

$$y_{lj} \in \{0; 1\} \quad \forall l \in \{1, \dots, m\}, \quad \forall j \in \{1, \dots, n\}$$

$$z_A \in \{0; 1\} \quad z_B \in \{0; 1\} \quad z \in \{0; 1\}$$

$$p_{min} \in \mathbb{R}, \; p_{max} \in \mathbb{R}, \; M_A \in \mathbb{R}, \; M_B \in \mathbb{R}, \; M \in \mathbb{R}, \; \varepsilon \in \mathbb{R}$$