Оптимално разпределение на мандати за 40. Народно събрание по избирателни райони

[18.12.2005]

Задача за разпределение на мандатите по партии

Дадени са p партии (p=7), които събират повече от 4% от действителните гласове на изборите за 40. Народно събрание. Означаваме партиите с 3@, 6@, 8@, 12@, 14@, 17@ и 19@. Мандатите за тези партии се определят пропорционално на получените гласове (Метод на Д'Ондт). Нека i-тата партия е получила v_i гласа и n_i мандати ($i=1,2,\ldots,p$). Броят на мандатите е решение на следната оптимизационна задача:

$$\min_i \frac{v_i}{n_i} \rightarrow \max_i \left(\max_i \frac{n_i}{v_i} \rightarrow \min_i \right)$$

$$\sum_{i=1}^{p} n_i = M, \quad (M = 240),$$

където n_i са цели неотрицателни числа и решението е:

No	партия	гласове	мандати
i		v_i	$\mid n_i \mid$
1	3@	1129196	82
2	6@	725314	53
3	8@	234778	17
4	12@	189268	13
5	14@	296848	21
6	17@	467400	34
7	19@	280323	20
	Общо		240

Задача за разпределение на мандатите на партиите по райони

Дадени са r избирателни района (r=31), като за всеки район е определен предварително броят на мандатите m_j в този район (по закон пропорционално на населението му).

No	Избирателен	мандати
$\mid j \mid$	район	m_j
01:	Благоевград	10
02:	Бургас	13
03:	Варна	14
04:	Велико Търново	9
05:	Враца	7
07:	Габрово	4
08:	Добрич	7
09:	Кърджали	5
10:	Кюстендил	5
11:	Ловеч	5
12:	Монтана	6
13:	Пазарджик	9
14:	Перник	5
15:	Плевен	10
16:	Пловдив град	10

No	Избирателен	мандати
$\mid j \mid$	район	m_{j}
17:	Пловдив окръг	11
18:	Разград	5
19:	Pyce	8
20:	Силистра	4
21:	Сливен	7
22:	Смолян	4
23:	София 1	13
24:	София 2	11
25:	София окръг	8
27:	Стара Загора	11
28:	Търговище	4
29:	Хасково	8
30:	Шумен	6
31:	Ямбол	5
	Общо:	240

Задачата е да се разпределят получените мандати на партиите по избирателните райони.

Нека $X = \left\{x_{ij}\right\}_{i,j}$ са разпределените мандати на i-тата партия в j-тия район. Тогава трябва да бъдат изпълнени равенствата:

$$\sum_{j=1}^{r} x_{ij} = n_i, \quad i = 1, 2, \dots, p, \qquad \sum_{i=1}^{p} x_{ij} = m_j, \quad j = 1, 2, \dots, r,$$
 (1)

като x_{ij} са *цели неотрицателни* числа. Тези ограничения са същите както при целочислена транспортна задача. Известна ни е и матрицата на вота с елементи v_{ij} – броят на гласовете на i-тата партия в j-тия район, $i=1,2,\ldots,p;\ j=1,2,\ldots,r.$ Ще решаваме задачата при следните две предположения за разглежданите партии:

- регистрация във всички избирателни райони с достатъчно дълга партийна листа;
- $-v_{ij}>0$, т.е. всяка партия получава поне един глас във всеки избирателен район.

	3@	6@	8@	12@	14@	17@	19@
j	v_{1j}	v_{2j}	v_{3j}	v_{4j}	v_{5j}	v_{6j}	v_{7j}
01:	34771	28448	7376	7497	6142	25472	14899
02:	56050	35201	8683	7468	20679	28557	12697
03:	55942	60248 27521	13130 7247	11707	17920	18916	15427
04:	45205 24545	10869	7247 2594	7357 2795	15272 2608	6121 8026	10226 4327
06:	38996	19679	2394 3685	6515	2008 6808	1213	6771
00.	19817	16417	4025	3622	6303	4041	6459
08:	35478	21457	4002	5955	7748	14058	3926
09:	17746	4870	1064	540	2229	63570	1975
10:	27685	16955	3762	2818	3065	978	4844
11:	25531	16765	4538	3793	6485	5658	5867
12:	30826	15824	2461	3138	4418	3047	6514
13:	39100	23978	5560	7089	9163	14829	12725
14:	23444	15455	3839	2889	4029	261	7058
15:	52499	26851	6470	8609	14165	10238	11920
16:	42434	42996	14105	7678	12946	11797	21187
17:	54575	33498	6027	7474	12314	15331	15060
18: 19:	15975 33400	9549 25945	1430 7050	1772 7396	5133 15125	36435 14483	2290 10558
20:	23262	10162	1360	1805	4250	27514	2130
21:	28006	14878	3363	4363	7941	8414	5745
22:	18308	9813	4097	5365	2461	9060	5624
23:	62206	43311	36536	11129	16992	899	20148
24:	47318	36011	29869	13133	15437	4097	15336
25:	53267	35644	20693	13636	17554	954	14911
26:	40210	20860	5607	5612	11590	4712	7296
27:	56809	28980	9435	9848	15325	8929	12046
28:	22528	6262	2493	2759	5897	28610	2003
29:	42585	23826	4452	8505	10033	21156	6535
30:	26972	16018	2468	2050	10412	27455	6065
31:	27581	13591	2324	3472	4837	1943	4207

Преди да формулираме оптимизационни задачи, ще дадем разпределението на мандатите според ЦИК, т.е. според методиката на ЦИК. Ще отбележим, че по тази методика са разпределяни мандатите на всички парламентарни избори от 1991 г. до сега.

No	3@	6@	8@	12@	14@	17@	19@	Сума:
	x_{1j}	x_{2j}	xз j	x_{4j}	x_{5j}	x_{6j}	$\frac{x_{7j}}{2}$	n_j
01: 02: 03: 04: 05:	Ž 3 3 3	Ž 2 4 2 1	Í 1 1 1 0	1 1 1 1	0 2 2 1 0	$egin{array}{c} x_{6j} \\ 2 \\ 3 \\ 1 \\ 0 \\ 0 \\ \end{array}$	1 2 1 0	10 13 14 9 4 7
04: 05: 06: 07: 08: 09: 10:	$\begin{array}{ c c c c c c c c c c c c c c c c c c c$	x_{2j} 242122222222230202122222	0 0 0 0	0 0 0 0 0	0 0 1 0 0	0 0 1 5 0	0 0 0 0	7 4 7 5 5 5 6
11: 12: 13: 14: 15:	3 4 3 4	2 2 2 2	0 0 0 0 0	0 0 0 0 0 0 1 1 1 0 0 0	0 0 1 0 1 1	0 0 1 0 1	0 0 1 0 1	6 9 5 10 10
15: 16: 17: 18: 19: 20:	4 2 3 1 2 1 3 2 1	3 0 2 0	0 0 0 0	1 0 1 0	1 0 1 0	1 4 1 3 1	2 2 0 1 0	10 11 5 8 4 7
18: 19: 20: 21: 22: 23: 24: 25: 26: 27:	3 2 1 1 3 5 5	2 1 2 2 2	0 0 5 4 2 0	1 1	1 0 2 1 2 1	1 0 0 0	0 0 2 2 2 0	4 13 11
26: 27: 28: 29: 30:	5 5 1 2 1	2 2 0 2 1	0 1 0 0	0 1 0 1 0	1 1 0 1 1	0 0 3 2 3	0 1 0 0	12 8 11 4 8 6
31:	4	1	0	0	0	0	0	5
Сума:	82	53	17	13	21	34	20	240

Накратко, методиката е следната: най-напред мандатите на всяка партия се разпределят пропорционално на получените гласове по районите (по Д'Ондт). Тъй като сумата от мандатите на различните партии в един район (изобщо) не е равна на определените мандати за този район, то се прави преразпределение на мандатите по районите в рамките на всяка партия. Това разпределение ще означаваме с X_0 .

Защо това разпределение не ни харесва и кое е "хубаво" разпределение?

По райони:

Район	3@		6@		8@		17@	
09:	17746	0	-	-	-	-	63570	5
23:	62206	1	43311	2	36536	5	-	-
24:	47318	1	36011	2	29869	4	-	-

По партии:

Партия	6:		17:	17:		
3@	38996	5	-	-	62206	1
6@	_	-	33498	3	43311	2

Модел 1.

Как да определим "цените на превозите", за да минимизираме транспортните разходи? Тъй като връзката между гласове и мандати трябва да е правопропорционална, т.е. за повече гласа се получават не по-малко мандати, разумно е "транспортните разходи" в нашия модел да са обратно пропорционални на броя гласове. По този начин целевата функция (общите "транспортни разходи") е:

$$F_1(X) = \sum_{i=1}^p \sum_{j=1}^r \frac{x_{ij}}{v_{ij}}.$$
 (2)

Задачата е: $\min F_1(X)$ при ограничения (1) и полученото оптимално

решение е

```
4
                                                                                      6000000050000000050400000004460
                                         00004707055605100000074000850405
                                                               001490040000090001000000000060000
                                                                                                            000000000000000000000012000000
                                                                                                                                 X_1 =
```

със стойност на целевата функция $F_1(X_1)=0.0090816706$. За решението на ЦИК имаме: $F_1(X_0)=0.0116575$.

Модел 2.

Максимизираме най-малката цена на мандата. Цената на мандата е за i-тата партия в j-тия район е $\frac{v_{ij}}{x_{ij}}$.

$$\min_{i,j} \frac{v_{ij}}{x_{ij}} \rightarrow \max.$$
 (3)

За да избегнем безкрайни стойности и за да линеаризираме модела, вземаме реципрочна стойност на целевата функция.

$$F_2(X) = \max_{i,j} \frac{x_{ij}}{v_{ij}} \quad \to \quad \min. \tag{4}$$

Едно оптималното решение е

със стойност на целевата функция $F_2(X_2) = 0.0001450273739$. Оптималното решение се достига в точката (3@, 31:), т.е.

$$\max_{i,j} \frac{x_{ij}}{v_{ij}} = \frac{x_{1}}{v_{1}} \frac{31}{31} = \frac{4}{27581} = 0.0001450273739.$$

За решението на ЦИК имаме $F(X_0) = F(X_2)$, т.е. и то е оптимално решение в този модел.

Модел 3.

Минимална разлика между най-голяма и най-малка цени на мандати.

$$\max_{i,j} \frac{x_{ij}}{v_{ij}} - \min_{i,j: \ x_{ij} > 0} \frac{x_{ij}}{v_{ij}} \rightarrow \min, \tag{5}$$

$$F_3(X) = \max_{i,j} \frac{x_{ij}}{v_{ij}} - \min_{i,j} \frac{x_{ij} + 1}{v_{ij}} \to \min.$$
 (6)

Намереното оптимално решение е:

със стойност на целевата функция $F_3(X_3) = 7.35377009 \times 10^{-5}$. Тази стойност се достига за (19@, 06:)

$$\max_{i,j} \frac{x_{ij}}{v_{ij}} = \frac{x_{76}}{v_{76}} = \frac{1}{6771} = 14.7689 \times 10^{-5}$$

и (3@, 30:)

$$\min_{i,j} \frac{x_{ij} + 1}{v_{ij}} = \frac{x_{130} + 1}{v_{130}} = \frac{2}{26972} = 7.4151 \times 10^{-5}.$$

Стойността на целевата функция за решението на ЦИК съвпада с оптималната и се достига в същите райони при същите партии.

Модел 4.

Да се опитаме да решим задачата за монотонност едновременно по партии и по райони. Това означава, че когато една партия е получила повече гласове в един район в сравнение с друг, то получава и не помалко мандати в първия район в сравнение с втория. Аналогично когато една партия е получила повече гласове от друга партия в даден район, тя получава не по-малко мандати в този район.

Ако за i-тата партия в j-тия и k-тия райони е в сила $v_{ij}>v_{ik}$, то трябва да бъде изпълнено $x_{ij}\geq x_{ik}$. И ако в j-тия район за i-тата и k-тата партии е изпълнено $v_{ij}>v_{kj}$, то е в сила и $x_{ij}\geq x_{kj}$.

$$sign(v_{ij} - v_{ik})(x_{ij} - x_{ik}) \ge 0,
i = 1, 2, ..., p, k = 1, 2, ..., r, j = 1, 2, ..., k;
sign(v_{ij} - v_{kj})(x_{ij} - x_{kj}) \ge 0,
j = 1, 2, ..., r, k = 1, 2, ..., p, j = 1, 2, ..., k.$$
(7)

За съжаление областта, определена от транспортните равенства (1) и неравенствата (7) е празна. За да елиминираме малки разлики в гласовете, модифицираме модела, като разликите делим целочислено на 1000 и след това определяме знака. Но и в този случий допустимата област на задачата е празна. Ако образуваме релаксирана задача по следния начин:

$$F_4(X) = z \rightarrow \min$$

при условия (1) и

$$sign(v_{ij} - v_{ik})(x_{ij} - x_{ik}) + z \ge 0,$$

$$i = 1, 2, \dots, p, \quad k = 1, 2, \dots, r, \quad j = 1, 2, \dots, k;$$

$$sign(v_{ij} - v_{kj})(x_{ij} - x_{kj}) + z \ge 0,$$

$$j = 1, 2, \dots, r, \quad k = 1, 2, \dots, p, \quad j = 1, 2, \dots, k.$$

$$(8)$$

то тази задача има решение с оптимална стойност z=1.

Модел 5.

Идеята на този модел е да минимизираме отклонението от средното за всеки район и за всяка партия. Например ако една партия е получила $\frac{1}{k}$ част от гласовете в даден район, то тази партия би трябвало да получи приблизително $\frac{1}{k}$ част от определените мандати за този район, т.е. $\frac{v_{ij}}{w_j} \approx \frac{x_{ij}}{m_j}$, където w_j е сумата от гласовете на участващите в разпределението партии в този район.

$$w_j = \sum_{i=1}^p v_{ij}, \quad j = 1, 2, \dots, r.$$

Аналогично се реализира тази идея и за разпределението на мандатите на една партия по райони — искаме $\frac{v_{ij}}{v_i} \approx \frac{x_{ij}}{m_i}$, където v_i е сумата от гласовете на партия i във всички райони.

$$v_i = \sum_{j=1}^r v_{ij}, \quad i = 1, 2, \dots, p.$$

Прецизирането на понятието близост ни дава този и следващите два модела.

Най-напред да формулираме задачата с равномерна норма:

$$F_5(X) = \max_{i,j} \left| \frac{v_{ij}}{v_i} - \frac{x_{ij}}{n_i} \right| + \max_{i,j} \left| \frac{v_{ij}}{w_j} - \frac{x_{ij}}{m_j} \right| \rightarrow \min$$
 (9)

Решението е:

```
26622312132322252512131526331212
X_5 =
                                                      0
0
1
```

със стойност на целевата функция

$$F_5(X_5) = 0.161243 + 0.0377498 = 0.1989926$$

при (12@, 13:), (6@, 28:) и за решението на ЦИК

$$F_5(X_0) = 0.135089 + 0.359636 = 0.494726$$

при (8@, 23:),(17@, 20:).

Партия	3@	6@	8@	12@	14@:	17@	19@
09:	17746 1	4870 0	1064 0	540 0	2229 0	63570 4	1975 0
13:	39100 2	23978 3	5560 1	7089 0	9163 0	14829 2	12725 1
20:	23262 1	10162 1	1360 0	1805 0	4250 0	27514 2	2130 0
23:	62206 5	43311 2	36536 3	11129 1	16992 1	899 0	20148 1
24:	47318 2	36011 4	29869 2	13133 1	15437 1	4097 0	15336 1
28:	22528 1	6262 1	2493 0	2759 0	5897 0	28610 2	2003 0

Модел 6.

В този модел използваме l_1 норма. Линеаризацията на модулите дават задача с 652 променливи, 907 ограничения.

$$\sum_{i=1}^{p} \sum_{j=1}^{r} \left| \frac{v_{ij}}{v_i} - \frac{x_{ij}}{n_i} \right| + \sum_{j=1}^{r} \sum_{i=1}^{p} \left| \frac{v_{ij}}{w_j} - \frac{x_{ij}}{m_j} \right| \to \min$$
 (10)

Едновременно използване на равномерна и l_1 норми, намалява размера на линеаризираната задача (257 променливи, 908 ограничения).

$$F_6(X) = \sum_{i=1}^{p} \max_{j} \left| \frac{v_{ij}}{v_i} - \frac{x_{ij}}{n_i} \right| + \sum_{j=1}^{r} \max_{i} \left| \frac{v_{ij}}{w_j} - \frac{x_{ij}}{m_j} \right| \to \min$$
 (11)

Стойността на целевата функция на непрекъснатата задача е 0.45.

Едно целочислено решение (без доказателство за оптималност) е:

```
3543231312233243412131434441312
X_6 =
```

Партия	3@		6@		8@		12@		14@:		17@		19@	
01:	34771	3	28448	2	7376	1	7497	1	6142	0	25472	2	14899	1
23:	62206	4	43311	3	36536	3	11129	1	16992	1	899	0	20148	1
24:	47318	3	36011	3	29869	2	13133	1	15437	1	4097	0	15336	1
31:	27581	2	13591	1	2324	1	3472	0	4837	1	1943	0	4207	0

Модел 7.

Най-използваният модел за отклонение от средното е минимизация с l_2 норма (най-малки квадрати).

$$F_7(X) = \sum_{i=1}^p \sum_{j=1}^r \left(\frac{v_{ij}}{v_i} - \frac{x_{ij}}{n_i}\right)^2 + \sum_{j=1}^r \sum_{i=1}^p \left(\frac{v_{ij}}{w_j} - \frac{x_{ij}}{m_j}\right)^2 \to \min$$
 (12)

Линеаризация на задачата.

За всяка двойка $(i,j), i=1,\ldots,p; j=1,\ldots,r$ полагаме:

$$x_{ij} = x_{ij1} + 2x_{ij2} + \dots + qx_{ijq} = \sum_{k=1}^{q} x_{ijk}$$

$$\left(\frac{v_{ij}}{v_i} - \frac{k}{n_i}\right)^2 = f_{ijk}^{(1)}, \quad \left(\frac{v_{ij}}{w_j} - \frac{k}{m_j}\right)^2 = f_{ijk}^{(2)}, k = 0, 1, \dots, q.$$

Целевата функция е:

$$F(X) = \sum_{i=1}^{p} \sum_{j=1}^{r} \sum_{k=0}^{q} \left(f_{ijk}^{(1)} + f_{ijk}^{(2)} \right) x_{ijk}.$$

Търсим $\min F(X)$ при ограничения:

$$\sum_{j=1}^{r} \sum_{k=1}^{q} kx_{ijk} = n_i, \quad i = 1, 2, \dots, p; \quad \sum_{i=1}^{p} \sum_{k=1}^{q} kx_{ijk} = m_j, \quad j = 1, 2, \dots, r;$$

$$\sum_{k=0}^q x_{ijk} = 1, \quad i = 1, 2, \dots, p, \quad j = 1, 2, \dots, r; \quad x_{ijk} \in \{0, 1\}, \;$$
 за всяко $i, j, k.$

Размери на задачата: (p = 7, r = 31, q = 6)

- неизвестни: pr(q+1) = 1519;
- събираеми в целевата функция: pr(q+1) = 1519;
- ограничения: p+r=38 от първия вид и pr=217 от втория вид.

Целевата функция е 0.7020591912 за следното решение:

```
3443231312233243412131434341323
```

Софтуер за линейна оптимизация

The GLPK (GNU Linear Programming Kit) package is intended for solving large-scale linear programming (LP), mixed integer programming (MIP), and other related problems. It is a set of routines written in ANSI C and organized in the form of a callable library.

The GLPK package includes the following main components:

- * Revised simplex method.
- * Primal-dual interior point method.
- * Branch-and-bound method.
- * Translator for GNU MathProg.
- * Application program interface (API).
- * Stand-alone LP/MIP solver.