Кібернетичний центр України ім. В. М. Глушкова
Центр інформаційних технологій та систем (ІТС)

Аналіз виборів у невеликій групі

© Юрій Владиславович Дзядик

Відділ віртуальних систем (ІТС/190)
Київ, проспект академіка Глушкова, 40, корпус 7

Київ, 10 червня 2013

План доповіді

1. Постановка задачі. Інформація для дослідження

2. Класи тотожних голосувань

3. Сусіди та антиподи. Графи та кластери

4. Проекції на площини. Факторний аналіз

5. Вибори ревізійної комісії

6. Порівняння ...

7. Інваріанти. Характери. Ментальність.

Дякую за увагу. TEX logo

1. Інформація для дослідження

У невеликій групі відбулись вибори: $M < 100$ виборців, K кандидатів на P посад. Нижче наведені протоколи для двох конкретних голосувань: вибори у деякій організації правління та ревізійної комісії.

1	1	1	1	1	0	1	1	0	1	0	0	0	1	1	1	1	0	0	1	1	1	1	0	1	0	0	1	1	1	1	1	0	1	1	1	1
2	1	1	1	0	1	0	0	1	1	1	1	1	0	0	0	0	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	0
3	1	0	0	1	0	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0	0	1	1	1	1	1	1	1	1
4	0	0	1	1	0	1	1	0	1	0	0	0	1	1	1	1	0	0	1	0	1	1	0	1	1	0	0	0	0	0	1	0	1	1	1	1
5	1	1	1	0	1	0	0	1	0	1	1	1	0	0	0	1	1	1	1	1	0	1	1	0	0	1	1	0	1	1	0	1	0	0	0	0
6	0	0	0	1	0	1	1	0	1	0	0	1	1	0	1	0	0	0	0	0	1	0	1	1	1	0	0	1	1	1	1	0	1	1	1	1
7	0	0	0	1	0	1	1	0	0	0	0	0	1	1	1	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1	1	1	0
8	0	0	1	1	0	1	1	0	1	1	1	0	1	0	1	1	0	0	1	0	1	1	0	0	1	1	0	0	0	0	0	1	1	1	1	1
9	1	1	0	0	1	0	0	1	0	1	1	1	0	0	0	0	1	1	0	1	0	0	1	0	0	1	1	1	0	0	0	1	0	0	0	0
10	1	1	0	0	1	0	0	1	0	1	1	0	0	0	0	0	1	1	0	1	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0
11	0	0	0	0	1	1	1	0	0	0	0	1	1	1	1	1	0	0	1	0	1	1	1	1	1	0	0	1	0	0	1	0	0	1	1	1
12	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	1	0	0	0
13	0	0	1	1	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	1	1	1	0	0	0	0	1
14	1	1	1	0	1	0	0	1	0	1	1	1	0	1	0	0	1	1	0	1	0	0	1	1	0	1	1	1	1	1	0	1	0	0	0	0
15	0	1	0	0	1	0	0	1	0	1	1	1	0	0	0	0	1	1	0	0	0	0	1	0	0	1	1	0	0	0	0	1	0	0	0	0
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36

Рис. 1: Протоколи виборів правління організації: $M = 36$ виборців, $K = 15$ кандидатів на $P = 7$ посад

Вибори таємні, тому нумерація виборців при кожному голосуванні випадкова та незалежна, одна з $M! = 36!$ можливих перестановок.

1	1	1	0	0	1	1	1	0	1	0	1	1	1	0	0	1	0	1	0	1	0	1	1	0	0	0	1	0	1	0	1	1	1	1	1	1
2	1	0	1	1	0	0	0	1	1	1	1	1	1	1	0	0	1	0	1	1	1	0	0	1	1	1	1	0	0	1	0	0	0	0	0	1
3	1	1	0	0	1	1	1	1	1	0	0	0	0	1	1	1	0	1	0	0	0	1	1	0	0	0	0	0	1	0	1	1	1	1	1	1
4	0	0	1	1	0	0	0	1	0	0	1	1	1	1	1	0	1	0	0	1	0	0	0	0	0	0	1	1	0	0	1	0	0	0	0	0
5	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	1	0	0	1	1	1	0	1	0	1	0	1	0	1	1	0
6	0	0	1	1	0	0	1	0	0	1	0	0	0	0	1	0	1	0	1	0	1	0	0	1	1	1	0	0	0	1	0	0	0	0	0	0
7	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	1	1	0	0	0	0	1	1	0	0	0	1	0	0	0
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36

Рис. 2: Протоколи виборів ревізійної комісії організації: $M = 36$ виборців, $K = 7$ кандидатів на $P = 3$ посад

Які висновки з цієї інформації ми можемо зробити про структуру групи? Застосуємо кілька математичних методів.

2. Класи тотожних голосувань

Насамперед знайдемо класи виборців, які голосували тотожно.

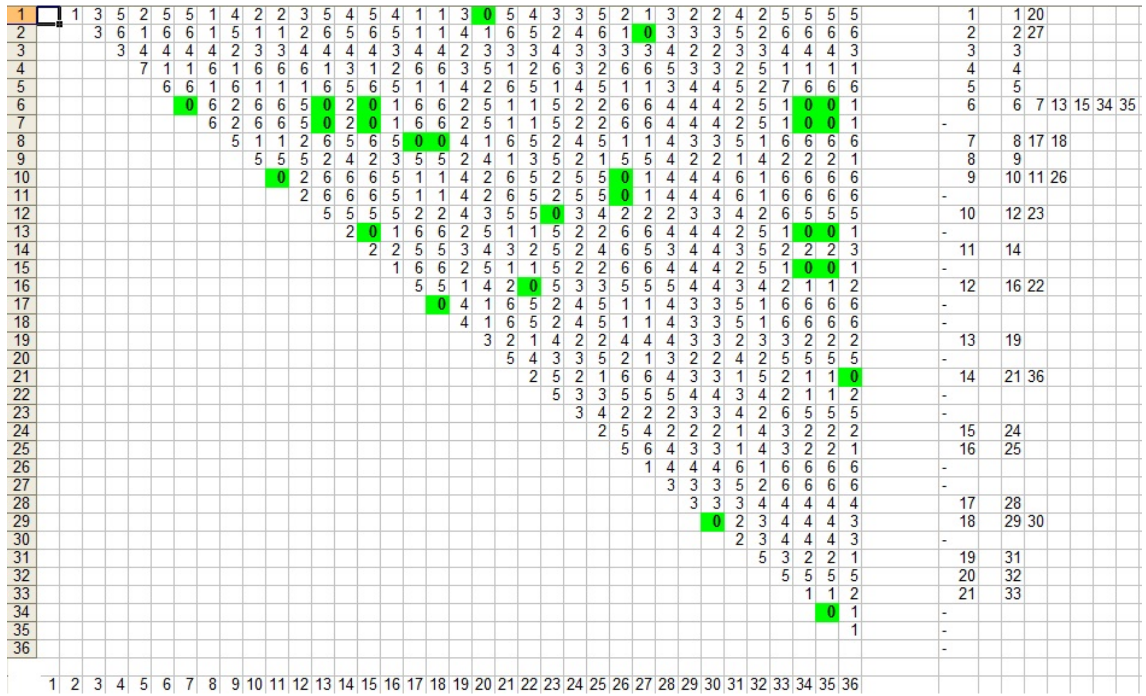
Спочатку проаналізуємо вибори правління. Всього існує $\binom{K}{P} = \binom{15}{7} = \frac{15!}{7!8!} = 6435$ **можливих варіантів голосувань. Але ці варіанти мають рівну імовірність лише за умови рівності кандидатів в очах усіх виборців, що нереально.**

Деякі з виборців є однодумцями, тобто користуються тотожними системами оцінок. Інші виборці можуть домовлятися між собою. Тому цілком природньо, що навіть у невеликій групі деякі виборці голосують тотожно.

Кожному з M **виборців** $\{a, b, \ldots\}$ **відповідає** K**-вимірний бінарний вектор (**$M = 36, K = 15$**). Щоб визначити класи тотожних виборців, насамперед підрахуємо відстані між векторами виборців:**

$$\operatorname{dist}(a, b) = \frac{1}{2} \sum_{i=1}^{K} |a_i - b_i|$$

Матриця відстаней симетрична і має нулі на головній діагоналі. Тому достатньо відобразити та проаналізувати лише її верхню половину.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
1		1	3	5	2	5	5	1	4	2	2	3	5	4	5	4	1	1	3	0	5	4	3	3	5	2	1	3	2	2	4	2	5	5	5	5
2			3	6	1	6	6	1	5	1	1	2	6	5	6	5	1	1	4	1	6	5	2	4	6	1	0	3	3	3	5	2	6	6	6	6
3				3	4	4	4	4	2	3	3	4	4	4	4	3	4	4	2	3	3	3	4	3	3	3	3	4	2	2	3	3	4	4	4	3
4					7	1	1	6	1	6	6	6	1	3	1	2	6	6	3	5	1	2	6	3	2	6	6	5	3	3	2	5	1	1	1	1
5						6	6	1	6	1	1	1	6	5	6	5	1	1	4	2	6	5	1	4	5	1	1	3	4	4	5	2	7	6	6	6
6							0	6	2	6	6	5	0	2	0	1	6	6	2	5	1	1	5	2	2	6	6	4	4	4	2	5	1	0	0	1
7								6	2	6	6	5	0	2	0	1	6	6	2	5	1	1	5	2	2	6	6	4	4	4	2	5	1	0	0	1
8									5	1	1	2	6	5	6	5	0	0	4	1	6	5	2	4	5	1	1	4	3	3	5	1	6	6	6	6
9										5	5	5	2	4	2	3	5	5	2	4	1	3	5	2	1	5	5	4	2	2	1	4	2	2	2	1
10											0	2	6	6	6	5	1	1	4	2	6	5	2	5	5	0	1	4	4	4	6	1	6	6	6	6
11												2	6	6	6	5	1	1	4	2	6	5	2	5	5	0	1	4	4	4	6	1	6	6	6	6
12													5	5	5	5	2	2	4	3	5	5	0	3	4	2	2	2	3	3	4	2	6	5	5	5
13														2	0	1	6	6	2	5	1	1	5	2	2	6	6	4	4	4	2	5	1	0	0	1
14															2	2	5	5	3	4	3	2	5	2	4	6	5	3	4	4	3	5	2	2	2	3
15																1	6	6	2	5	1	1	5	2	2	6	6	4	4	4	2	5	1	0	0	1
16																	5	5	1	4	2	0	5	3	3	5	5	5	4	4	3	4	2	1	1	2
17																		0	4	1	6	5	2	4	5	1	1	4	3	3	5	1	6	6	6	6
18																			4	1	6	5	2	4	5	1	1	4	3	3	5	1	6	6	6	6
19																				3	2	1	4	2	2	4	4	4	3	3	2	3	3	2	2	2
20																					5	4	3	3	5	2	1	3	2	2	4	2	5	5	5	5
21																						2	5	2	1	6	6	4	3	3	1	5	2	1	1	0
22																							5	3	3	5	5	5	4	4	3	4	2	1	1	2
23																								3	4	2	2	2	3	3	4	2	6	5	5	5
24																									2	5	4	2	2	2	1	4	3	2	2	2
25																										5	6	4	3	3	1	4	3	2	2	1
26																											1	4	4	4	6	1	6	6	6	6
27																												3	3	3	5	2	6	6	6	6
28																													3	3	3	4	4	4	4	4
29																														0	2	3	4	4	4	3
30																															2	3	4	4	4	3
31																																5	3	2	2	1
32																																	5	5	5	5
33																																		1	1	2
34																																			0	1
35																																				1
36																																				

1	1	1	20				
2	2	2	27				
3	3	3					
4	4	4					
5	5	5					
6	6	6	7	13	15	34	35
7	-						
8	7	8	17	18			
9	8	9					
10	9	10	11	26			
11	-						
12	10	12	23				
13	-						
14	11	14					
15	-						
16	12	16	22				
17	-						
18	-						
19	13	19					
20	-						
21	14	21	36				
22	-						
23	-						
24	15	24					
25	16	25					
26	-						
27	-						
28	17	28					
29	18	29	30				
30	-						
31	19	31					
32	20	32					
33	21	33					
34	-						
35	-						
36	-						

Рис. 3: Частина матриці відстаней вище головної діагоналі. Яскраво-зеленим кольором відзначені пари виборців, які мали тотожне голосування

Ми бачимо, що при виборах правління усі $M = 36$ виборців розбиваються на 21 клас: в одному з класів 6 виборців, у 2-х по 3, у 6-ти по 2, у 12-ти по 1.

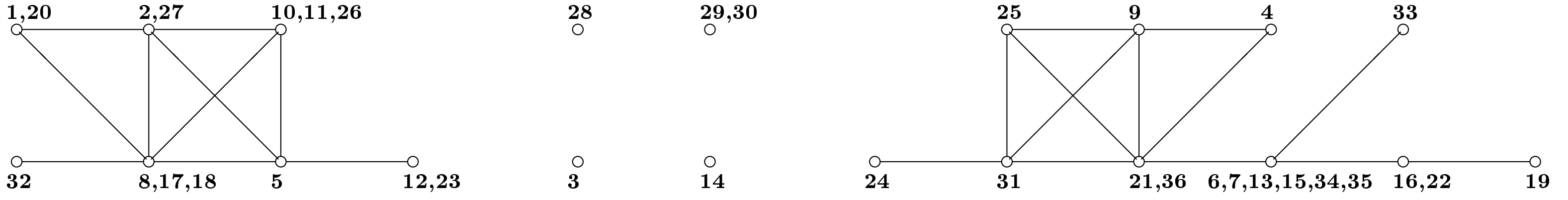
3. Сусіди та антиподи. Графи та кластери

Пару виборців (a, b) будемо називати сусідами, якщо $\mathrm{dist}\ (a, b) = 1$, та антиподами, якщо $\mathrm{dist}\ (a, b) \geqslant P - 1$, де P – кількість посад, $P \geqslant 3$.

Ці відношення є симетричними, але не транзитивними.

Відношення сусідства дозволяє встановити на множині виборців структуру неорієнтованого графу. А саме, вершини $(a,\ b)$ з'єднані ребром тоді і лише тоді, коли вони сусіди, тобто $\mathrm{dist}\ (a, b) = 1$.

Для виборів правління цей граф складається із шести зв'язних компонент: 4 ізольовані вершини (5 виборців) та два кластери $\{\mathbf{A}, \mathbf{B}\}$. Кластер A містить 7 класів (14 виборців), кластер B – 10 класів (17 виборців). Граф має вигляд:



Зауважимо, що класи $\{2, 5, 8, 10\}$ та $\{9, 21, 25, 31\}$ поєднані між собою кожен з кожним, тобто складають дві правильні 3-вимірні піраміди з одиничними ребрами.

Тепер використаємо відношення антипо́дії. Будемо говорити, що кластер є спільнотою, якщо серед його елементів відсутні антиподи.

1	0	1	3	5	2	5	1	4	2	3	4	4	3	5	3	5	3	2	4	2	5
2	1	0	3	6	1	6	1	5	1	2	5	5	4	6	4	6	3	3	5	2	6
3	3	3	0	3	4	4	4	2	3	4	4	3	2	3	3	3	4	2	3	3	4
4	5	6	3	0	7	1	6	1	6	6	3	2	3	1	3	2	5	3	2	5	1
5	2	1	4	7	0	6	1	6	1	1	5	5	4	6	4	5	3	4	5	2	7
6	5	6	4	1	6	0	6	2	6	5	2	1	2	1	2	2	4	4	2	5	1
7	1	1	4	6	1	6	0	5	1	2	5	5	4	6	4	5	4	3	5	1	6
8	4	5	2	1	6	2	5	0	5	5	4	3	2	1	2	1	4	2	1	4	2
9	2	1	3	6	1	6	1	5	0	2	6	5	4	6	5	5	4	4	6	1	6
10	3	2	4	6	1	5	2	5	2	0	5	5	4	5	3	4	2	3	4	2	6
11	4	5	4	3	5	2	5	4	6	5	0	2	3	3	2	4	3	4	3	5	2
12	4	5	3	2	5	1	5	3	5	5	2	0	1	2	3	3	5	4	3	4	2
13	3	4	2	3	4	2	4	2	4	4	3	1	0	2	2	2	4	3	2	3	3
14	5	6	3	1	6	1	6	1	6	5	3	2	2	0	2	1	4	3	1	5	2
15	3	4	3	3	4	2	4	2	5	3	2	3	2	2	0	2	2	2	1	4	3
16	5	6	3	2	5	2	5	1	5	4	4	3	2	1	2	0	4	3	1	4	3
17	3	3	4	5	3	4	4	4	4	2	3	5	4	4	2	4	0	3	3	4	4
18	2	3	2	3	4	4	3	2	4	3	4	4	3	3	2	3	3	0	2	3	4
19	4	5	3	2	5	2	5	1	6	4	3	3	2	1	1	1	3	2	0	5	3
20	2	2	3	5	2	5	1	4	1	2	5	4	3	5	4	4	4	3	5	0	5
21	5	6	4	1	7	1	6	2	6	6	2	2	3	2	3	3	4	4	3	5	0
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21

1	2	7				A						
2	1	5	7	9		A	4	6	14	16	21	
3												
4	6	8	14	21		B	2	7	9	10		
5	2	7	9	10		A	6	8	14			
6	4	12	14	21		B	2	5	7	9		
7	1	2	5	9	20	A	4	6	14	21		
8	4	14	16	19		B	5					
9	2	5	7	20		A	4	6	11	14	19	21
10	5					A	4	21				
11							9					
12	6	13				B						
13	12					B						
14	4	6	8	16	19	B	2	5	7	9		
15	19					B						
16	8	14	19			B	2					
17												
18												
19	8	14	15	16		B	9					
20	7	9				A						
21	4	6				B	2	7	9	10		
	однодумці						антагоністи					

Рис. 4: Сусіди та антиподи для кожного класу виборців. Літерами **A**, **B** позначені два великих кластери. Увага: на цій сторінці кожен кластер позначений своїм порядковим номером, а не номером першого виборця у цьому класі!

Ми бачимо, що елементи кожного кластеру або зовсім не мають антиподів, або мають їх з протилежного кластеру, за одним винятком: антиподом до класу №9 з кластеру A є ізольований клас №11. Отже, кластери A, B є спільнотами.

4. Проекції на площини. Факторний аналіз

Поставимо задачу: спроектувати M векторів виборців $\{a, b, \dots\}$ з K-вимірного простору оператором проекції $\mathbf{P}$ на таку двомірну площину, щоб втратити найменше інформації. Тоді ми наочно побачимо на площині конфігурацію, максимально подібну до K-вимірної.

Справедлива теорема, що оптимальною є проекція на площину головних компонент, отриманих в результаті факторного аналізу. А саме, має місце більш загальна

Теорема. Позначимо через $\mathbf{P}_s$ оператор проекції на підпростір перших s факторів $\{\mathbf{z}_1, \ \dots, \ \mathbf{z}_s\}$, $G(\mathbb{R}; k, s)$ – грасманів многовид усіх s-вимірних підпросторів дійсного M-вимірного простору, $\{\mathbf{P}\}$ – відповідні проектори. Тоді

$$\sum_{i=1}^{M}\sum_{j=1}^{i-1}\left|\mathbf{P}_s\left(\mathbf{x}_i-\mathbf{x}_j\right)\right|^2=\max_{\mathbf{P}\in G(\mathbb{R};k,s)}\sum_{i=1}^{M}\sum_{j=1}^{i-1}\left|\mathbf{P}\left(\mathbf{x}_i-\mathbf{x}_j\right)\right|^2.$$

Перший та другий фактори містять лише 65,4% інформації про відстані (54,1% та 11,3% відповідно), третій – 8,5%. Тому представимо дві діаграми.

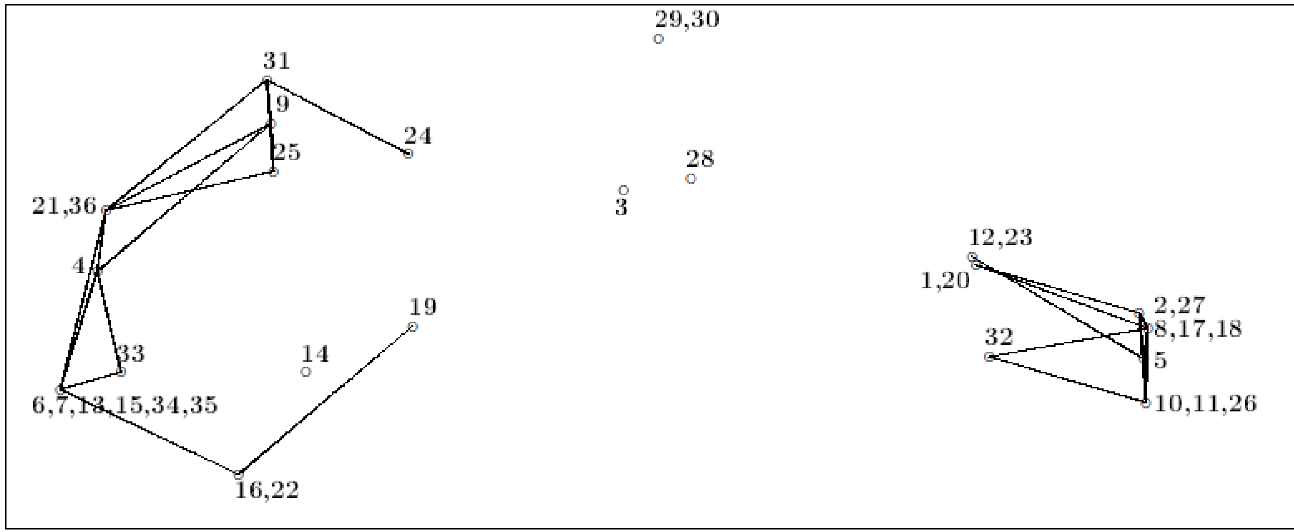


Рис. 5: Проекція на головні компоненти (тобто на 1 і 2 фактори)

Проекція на 1 та 2 фактори. Зліва 10 елементів кластера B, справа 7 елементів кластера A, між ними 3 ізольовані класи 3 (№3), 28 (№17), 29 (№18), четвертий ізольований клас 14 (№11) неначе знаходиться всередині кластера B.

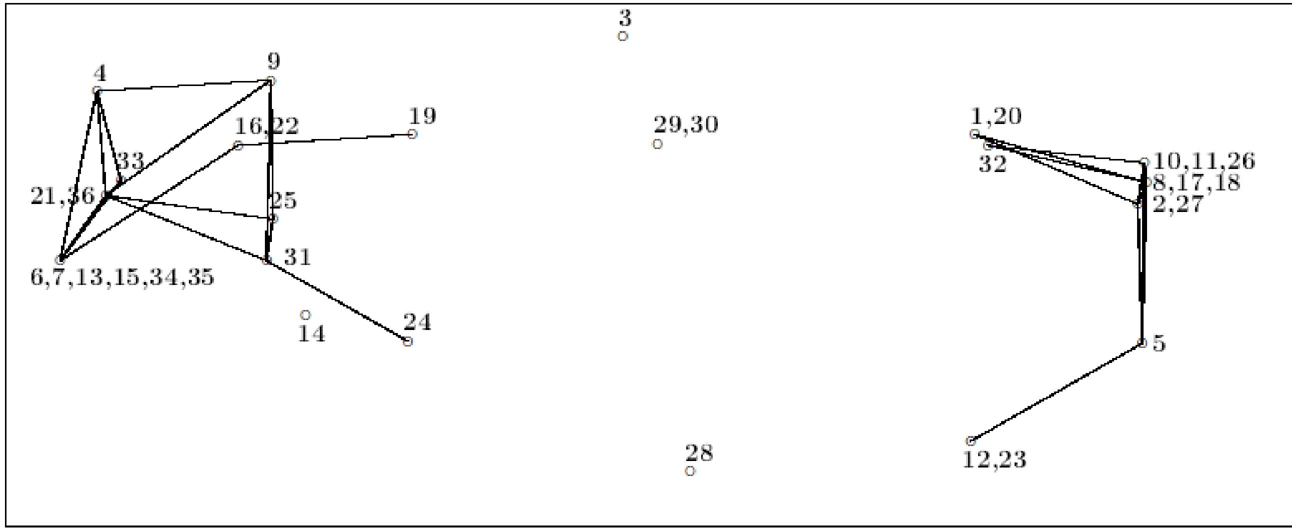


Рис. 6: Проекція на 1 та 3 фактори

Проекція на 1 та 3 фактори доповнює інформацію, яку дає проекція на головні компоненти. Зокрема, ізольований клас 14 вже знаходиться зовні кластеру B.

5. Вибори ревізійної комісії

Тепер проаналізуємо вибори ревізійної комісії. Ми відразу застосуємо метод проекції на площину двох головних компонент.

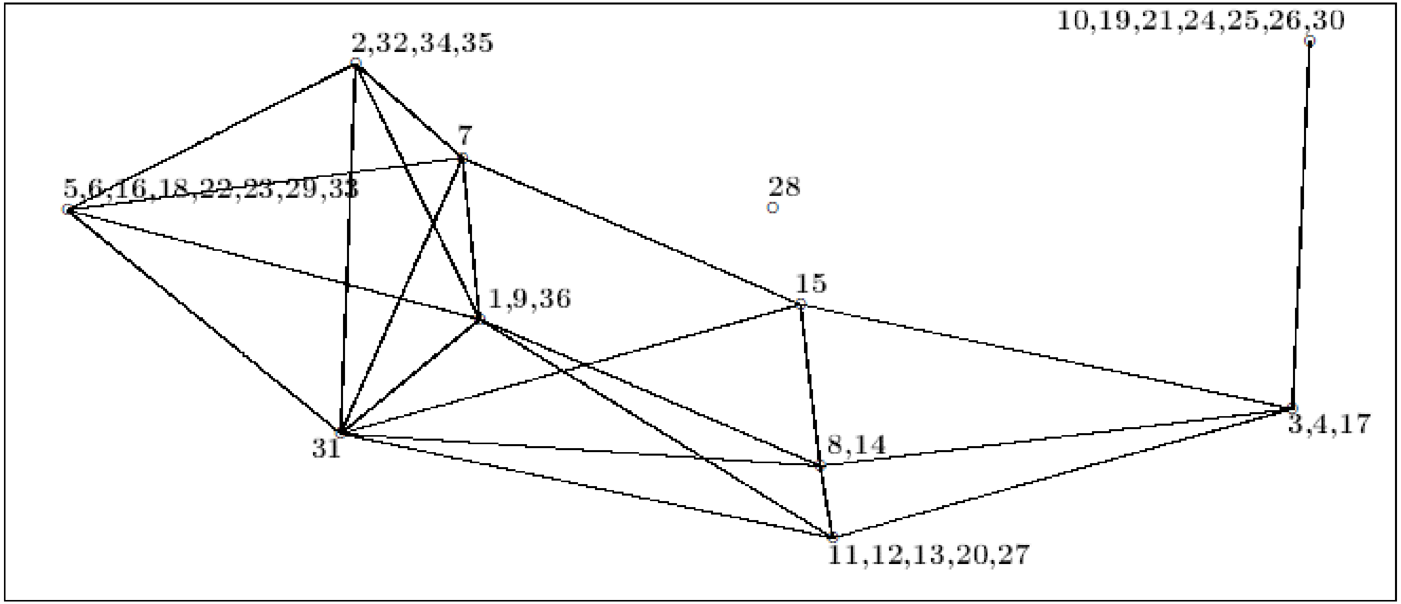


Рис. 7: Проекція матриці виборів ревізійної комісії на головні компоненти. Усі сусіди з'єднані лініями

Як і раніше, позначаємо клас найменшим номером з його виборців.

Ми бачимо, що по одному з класів містять 8, 7, 5, 4-х та 2-х виборців, 4 класи містять по одному виборцю, 2 – по три. Усього є $5+2+4=11$ **класів,** $(4\cdot 1+2)+(2\cdot 3+4)+(5+7+8)=6+10+20=36$ **(виборців).**

Є один ізольований клас 28, решта 10 класів об'єднані сусідствами у один кластер.

Класи 7 та 1 є антиподами ізольованого класу 28, класи 2, 5 та 31 – антиподами класів 3 та 10 (у тому ж кластері).

Зауважимо, що 5 класів 1, 2, 5, 7, 31 поєднані між собою кожен з кожним, тобто складають правильну 4-вимірну піраміду з одиничними ребрами.

6. Порівняння ...

Спробуємо порівняти результати двох виборів в цій групі.

З'єднаємо на діаграмі 5 (вибори правління) штриховими лініями тих сусідів, для яких $\operatorname{dist}(a,b)=2$**, отримаємо діаграму 8. Тепер на діаграмах обох виборів (7 та 8) будуть відображені усі сусіди, які задовольняють єдиній умові:**

$$\operatorname{dist}(a,b) \leqslant \frac{P}{3}, \text{where } \frac{P}{3} \in \left\{\frac{7}{3}=2, \frac{3}{3}=1\right\}.$$

Ми бачимо, що тепер кластер B на рис. 8 є многогранником, і нагадує 4-вимірну піраміду на діаграмі 7, яка складається з класів 7, 1/3, 31, 5/8, 2/4 і так само, як і кластер B, містить 17 виборців.

Покажемо, що (при одному припущенні) *піраміда на діаграмі 7 містить принаймні одного виборця з кластеру* **A**, *отже, не співпадає з кластером* **B.**

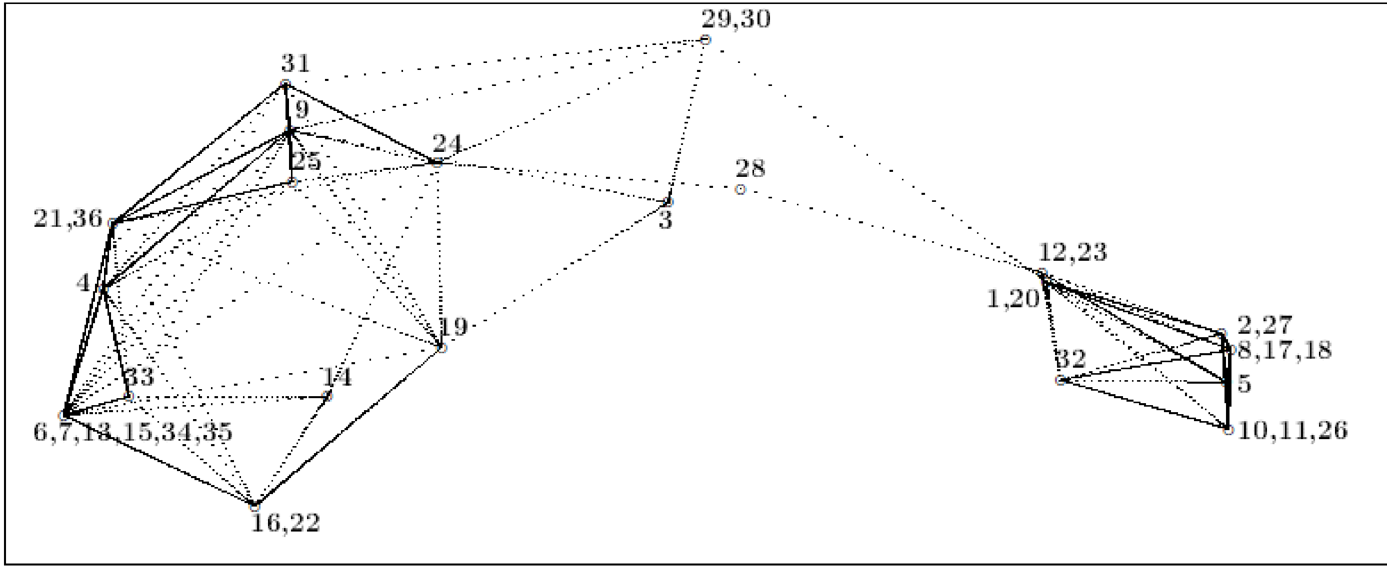


Рис. 8: fact-15-36-12-2

Для перенесення на рис. 7 кластеру A застосуємо булеву логіку. Позначимо $M-$вимірні булеві вектори кандидатів на виборах у правління через $\mathbf{g}_i,\ i = 1, \dots, 15$, булеві вектори кандидатів на виборах у ревізійну комісію – через $\{\mathbf{h}_i\},\ i = 1, \dots, 7$.

Відзначимо, що $\mathbf{A}\&\mathbf{g}_3 = \mathbf{A}\&\mathbf{g}_4 = \mathbf{A}\&\mathbf{g}_6 = \mathbf{A}\&\mathbf{g}_{10} = \mathbf{A},\ \mathbf{A}\&\mathbf{g}_{11} = \mathbf{g}_{11} = \mathbf{A}\setminus \{1, 20\}$.

Нам також відомо, що кандидати

$$\text{name}\ (\mathbf{g}_6) = \text{name}\ (\mathbf{h}_1),\ \ \text{name}\ (\mathbf{g}_9) = \text{name}\ (\mathbf{h}_2),\ \ \text{name}\ (\mathbf{g}_{11}) = \text{name}\ (\mathbf{h}_6),\ \ \text{name}\ (\mathbf{g}_{13}) = \text{name}\ (\mathbf{h}_7)$$

– одні й ті самі особи, причому кандидат $\text{name}\ (\mathbf{g}_{11}) = \text{name}\ (\mathbf{h}_6)$ **отримав однакову кількість (12) голосів:**

$$\text{ord}\ (\text{set}\ (\mathbf{g}_{11})) = \text{ord}\ (\text{set}\ (\mathbf{h}_6)) = \text{ord}\ (\{3, 4, 7, 10, 15, 17, 19, 21, 24, 25, 26, 30\}) = 12.$$

Припустимо, що $\text{set}(\mathbf{g}_{11}) = \text{set}(\mathbf{h}_6)$, *тобто що за нього віддавали голоси одні й ті ж виборці*. **Тоді нам відомі 12 із 14 виборців класу A при виборах ревізійної комісії, а саме:**

$$\text{set}(\mathbf{h}_6) = \{3, 4, 17\} \cup \{10, 19, 21, 24, 25, 26, 30\} \cup 7 \cup 15.$$

Оскільки виборець 7 **належить до 4-вимірної піраміди (рис. 7), і водночас**

$$7 \in \text{set}(\mathbf{h}_6)\ \&\ \text{name}\ (\mathbf{g}_{11}) = \text{name}\ (\mathbf{h}_6)\ \&\ \text{set}(\mathbf{g}_{11}) \subset \mathbf{A},$$

то ми дійшли висновку:

якщо $\text{set}(\mathbf{g}_{11}) = \text{set}(\mathbf{h}_6)$, *то піраміда на діаграмі 7 містить принаймні одного виборця з кластеру* $\mathbf{A}$, *отже, не співпадає з кластером* $\mathbf{B}$,

що й потрібно було довести.

7. Інваріанти. Характери. Ментальність.

Подібність конфігурацій на рис. 7 та рис. 8 говорить про існування деяких інваріантів двох виборів. Водночас, щойно було доведено, що ці інваріанти не є тотожністю кластерів.

Спробуємо знайти якісь математичні форми для твердження, що голосування обох виборів (правління та ревізійної комісії) є відображення ментальності виборців, їх способу прийняття рішень. Тоді ми зможемо уточнювати сенс цього твердження (у різних його математичних формах), в рідкісних крайніх випадках – повністю довести або спростувати.

Будемо міркувати та шукати.

...

...

Дякую за увагу

Рис. 9: TEX logo