

Основы теории принятия решений

Лекция 2

Содержание

1. Алгоритм Forel 2
2. Выделение устойчивых таксонов
3. Таксономия динамически описываемых объектов.
4. Выбор алгоритмов таксономии.
5. Пример 1.
6. Примеры прикладных задач таксономии:
 - прогнозирование успеваемости;
 - ранжирование объектов.

Назначение и свойства алгоритма Forel 2

- Предназначен для группировки объектов в таких условиях, когда:
- все характеристики объектов однородны,
- число таксонов **N** задано,
- обладает таксонами, которые имеют форму гиперсферы.

Алгоритм FOREL-2 (шаги 1-5)

- Шаг 1. Все признаки объектов нормируются так, чтобы их значения были в диапазоне 0 - 1.
- Шаг 2. $R_0 = +\infty$
- Шаг 3. Все точки считаем непомеченными.
- Шаг 4. На множестве непомеченных точек выбирается произвольная $\mathbf{x}_i$, после чего осуществляется переход к шагу 5. Если таковых точек нет, то перейти к шагу 8.
- Шаг 5. Ищется максимальное расстояние R от $\mathbf{x}_i$ до остальных точек.

Алгоритм FOREL-2 (шаги 5 - 12)

- Шаг 6. $R0 = \min \{R0; R\}$.
- Шаг 7. Точка x_i помечается. Если помечены все точки, то перейти к шагу **8**, нет - к шагу **4**.
- Шаг 8. $R = R0 - \varepsilon$
- Шаг 9. Если множество точек пусто, то перейти к шагу **16**, нет - к шагу **10**.
- Шаг 10. Все точки считаем непомеченными.
- Шаг 11. На множестве непомеченных точек выбирается произвольная точка x_i .
- Шаг 12. Определяется число точек $P(x_i)$, расстояние которых до x_i не превышает R .

Алгоритм FOREL-2 (шаги 13 - 18)

- Шаг 13. Точку x_i считаем помеченной. Если помечены все точки, то перейти к шагу **14**, нет - к шагу **11**.
- Шаг 14. Выбирается j -я точка, для которой справедливо: $P(x_i) \leq P(x_j)$, $i=1,2,3,\dots,n$.
- Шаг 15. Все точки, расстояние от которых до не превышает **R**, удаляются. Перейти к шагу 9.
- Шаг 16. Если число таксонов меньше **N**, то перейти к шагу 17, иначе к шагу 19.
- Шаг 17. $R=R-\varepsilon$.
- Шаг 18. Все точки возвращаются на «свои места», перейти к шагу 10.

Алгоритм FOREL-2 (шаги 19 - 22)

- Шаг 19. Если число таксонов равно **N**, то перейти к шагу 22, нет - к шагу 20.
- Шаг 20. $\varepsilon = \varepsilon / 2$.
- Шаг 21. $R = R + \varepsilon$, перейти к шагу 18.
- Шаг 22. Конец алгоритма.

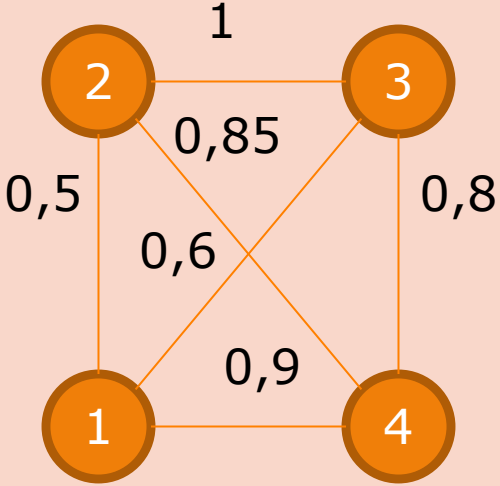
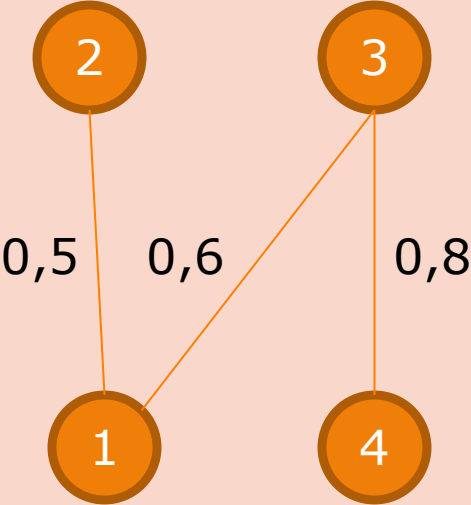
Пример 1

- Разбить 4 объекта на три таксона, если λ -расстояния между объектами приведены ниже в таблице M , $\varepsilon = 0,1$.

● $M =$

0	0,5	0,6	0,9
0,5	0	1	0,85
0,6	1	0	0,8
0,9	0,85	0,8	0

Решение задачи примера 1

Исходный граф	Остов графа	Число вершин, принадлежащих каждому таксону
		$P(x_1)=3; R=0,6.$ $P(x_2)=2; R=0,5.$ $P(x_3)=3; R=0,8.$ $P(x_4)=2; R=0,8.$ $\text{Max } P(x_i) = P(x_1)=3;$ $R=0,6$

Итерация №1



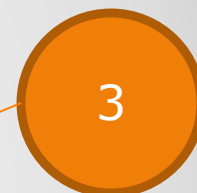
Таксон № 1;
 $R=0,6$.



0,5



0,6



0,8

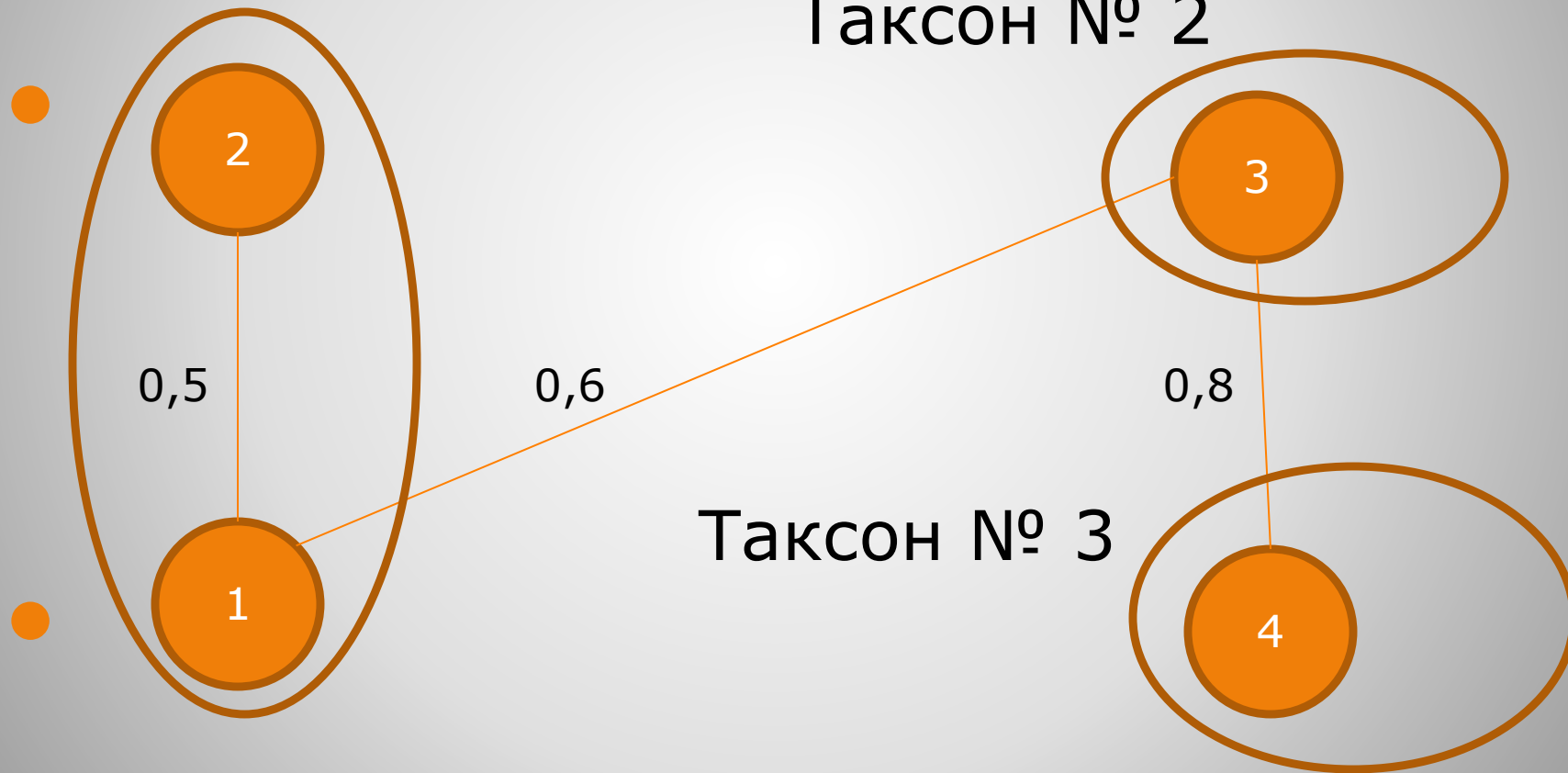


Таксон № 2



Итерация №2

Таксон № 1; $R=0,5$.



Самостоятельно

1. Заполнить матрицу M , имеющую 4 строки и два столбца: $M(i, j) = (k - i)^2 + (k - j)^2$, $i = 1, 2, 3, 4$; $j = 1, 2$.
где k – порядковый номер студента.
2. Полагая, что строки отвечают объектам, а столбцы – критериям, выделить, пользуясь Forel 1, таксоны при условии, что $\xi = 0,5$.
3. Разбить все объекты на 3 таксона, пользуясь Forel 2.

САМОСТОЯТЕЛЬНО

**● Определить
достоинства и
недостатки
алгоритма Forel 2.**

Алгоритм Skat

- Шаг 1. Определяется таксономия S для m объектов с помощью FOREL-1.
- Шаг 2. Используя центры таксонов в S , как новые стартовые точки для FOREL-1, определяются таксономии **$S_1, S_2, \dots, S_n$** .
- Шаг 3. Выбор устойчивых таксонов.
- Шаг 4. Конец алгоритма.

САМОСТОЯТЕЛЬНО

- Пользуясь алгоритмом SKAT и, в его рамках, алгоритмом FOREL 1, выделить устойчивые таксоны в таблице:

№	Оценка	Пропуски
1	2	9
2	4	8
3	5	2

Решение задачи

- 1) Таксономия S1 при исходной нумерации:
1-й таксон: 1, 2; 2-й таксон: 3.
- 2) Таксономия S2 при обратной нумерации объектов (строк):
1-й таксон: 2, 3; 2-й таксон: 1.
- 3) Вывод: таксоны неустойчивы.

САМОСТОЯТЕЛЬНО

**Определите
достоинства и
недостатки
алгоритма SKAT**

САМОСТОЯТЕЛЬНО

- Составить блок-схему алгоритма SKAT.
- Реализовать алгоритм программно.
- Исследовать созданную программу, построив график зависимости времени счета от размерности исходной матрицы.

Динамическая таксономия

- Ниже рассматривается ситуация, когда в ходе исследований объекты возникают по одному или группами и на каждой итерации осуществляется таксономия. В этом случае используется алгоритм DINA, описание которого приводится ниже.

Алгоритм DINA

- Шаг 1. Ввод **R** - радиуса таксона.
- Шаг 2. **i=1**.
- Шаг 3. Ввод **i**-го объекта.
- Шаг 4. Если **i>1**, то перейти к шагу 6.
- Шаг 5. Введённый объект отвечает центру первого таксона.
- Шаг 6. Определяется расстояние от **i**-й точки до центров всех созданных таксонов.
- Шаг 7. Если кратчайшее из этих расстояний меньше **R**, то **i**-й объект принадлежит соответствующему таксону, в противном случае **i**-й объект - центр нового таксона.
- Шаг 8. **i=i+1**.
- Шаг 9. Если ввод объектов завершён, то перейти к шагу 10, нет - к шагу 3.
- Шаг 10. Конец алгоритма.

Пример 2

Определить, пользуясь DINA, таксономию трёх последовательно возникающих объектов: круга, квадрата и треугольника, характеризующихся Таблицей, представленной на следующем слайде, $R=1$.

Таблица исходных данных

Объект	Число углов	Площадь	Цвет
	0	7	1
	4	3	2
	3	8	3

Таблица нормированных данных

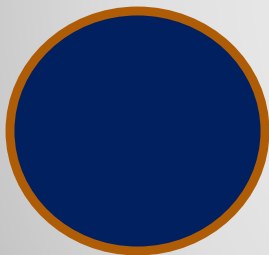
Объект	Число углов	Площадь	Цвет
	0	0,875	0,33
	1	0,375	0,66
	0,75	1	1

Порядок решения

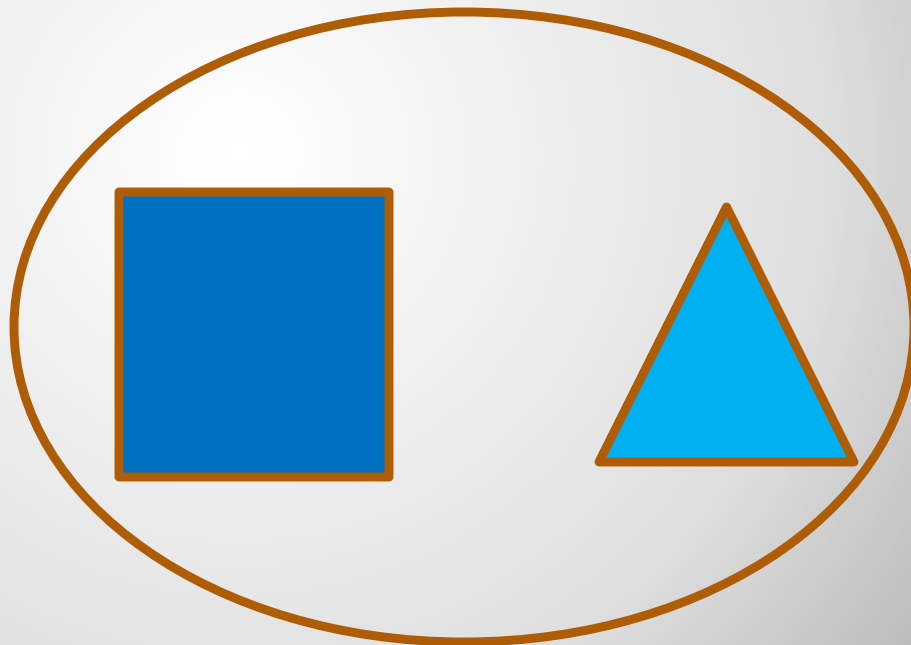
- Круг - центр первого таксона.
Расстояние между квадратом и кругом больше **R** : **$r_1=1.1657$** , следовательно квадрат - центр второго таксона.
- Расстояние между треугольником и квадратом **$r_2=0.7541$** меньше, чем между треугольником и кругом **$r_3=1.013422$** и меньше **R** , следовательно треугольник принадлежит ко второму таксону.

ОТВЕТ

Таксон № 1



Таксон № 2



САМОСТОЯТЕЛЬНО

Пользуясь DINA, определить таксономию четырех объектов, описанных в таблице и появляющихся последовательно:

№	Признак 1	Признак 2
1	2	50
2	3	40
3	4	15
4	5	5