

**Многомерные статистические модели.
Модели классификации
геологических объектов и
признаков**

Для магистрантов направления 7М07202
– Геология и разведка месторождений
полезных ископаемых

Портнов Василий Сергеевич
Доктор технических наук,
Профессор Кафедры ГРМПИ
НАО «КарТУ»



Основные темы

Сущность и условия применения
многомерных статистических моделей

Множественная корреляция

Множественная регрессия

Классификационные модели в геологии

Многомерная статистическая модель

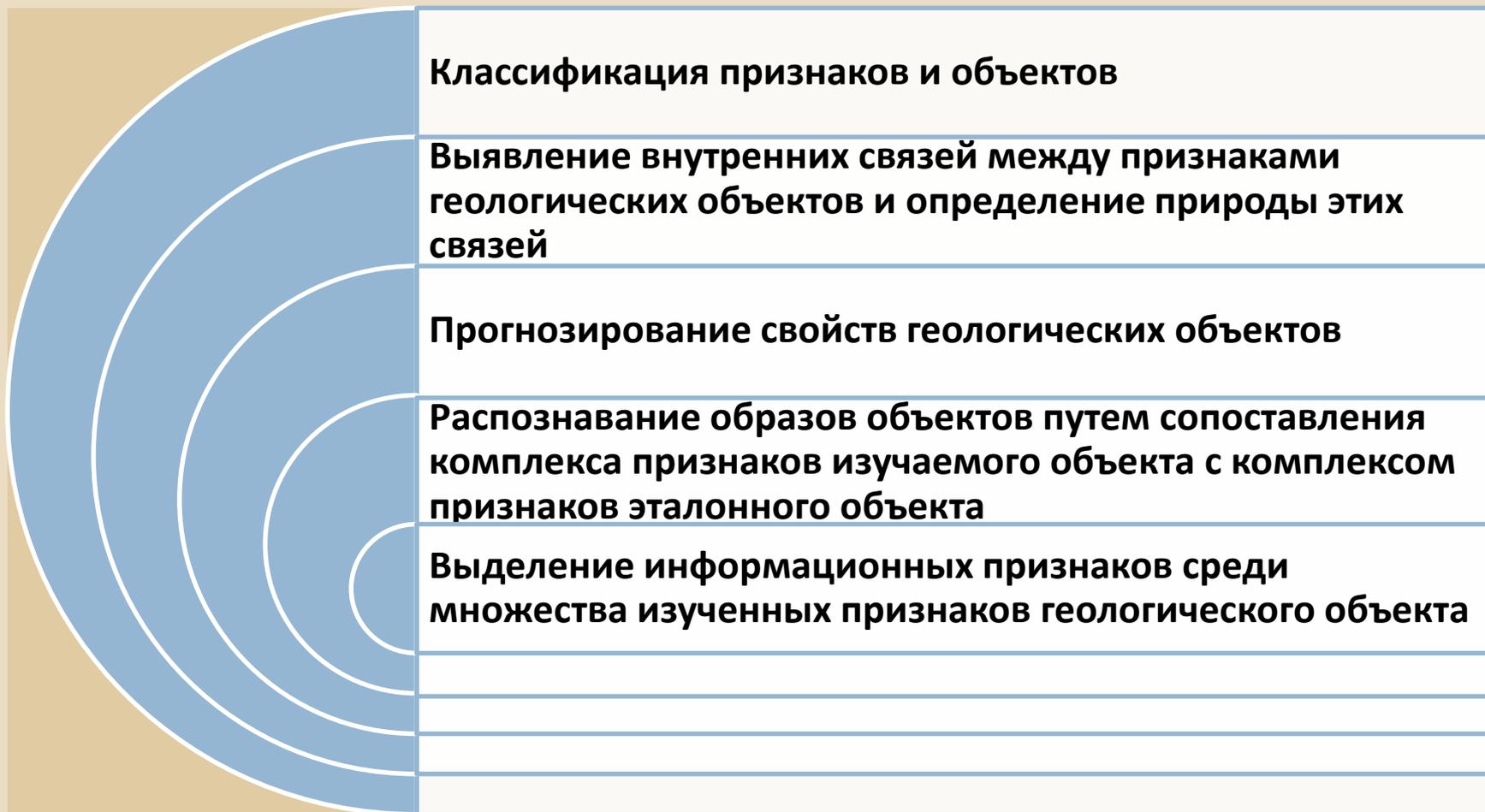
- В практике геологических исследований часто возникает необходимость совместного рассмотрения комплекса признаков
- Многомерная статистическая модель состоит из совокупности множества сопряженных случайных величин



Пример многомерной выборки

n	Ag г/т	Cu%	Zn%	Pb%
1	2,0	0,5	2,0	3,0
2	4,0	0,3	1,1	4,1
3	15,0	2,4	4,1	7,2
4	3,0	2,0	3,4	3,6
5	8,0	2,4	2,5	5,4
6	1,0	2,1	3,8	2,0
7	0.2	3,5	6,1	1,1
8	0.6	2,1	3,5	1,2
9	1,0	2,9	4,2	2,0
10	2,0	2,7	4,1	2,8

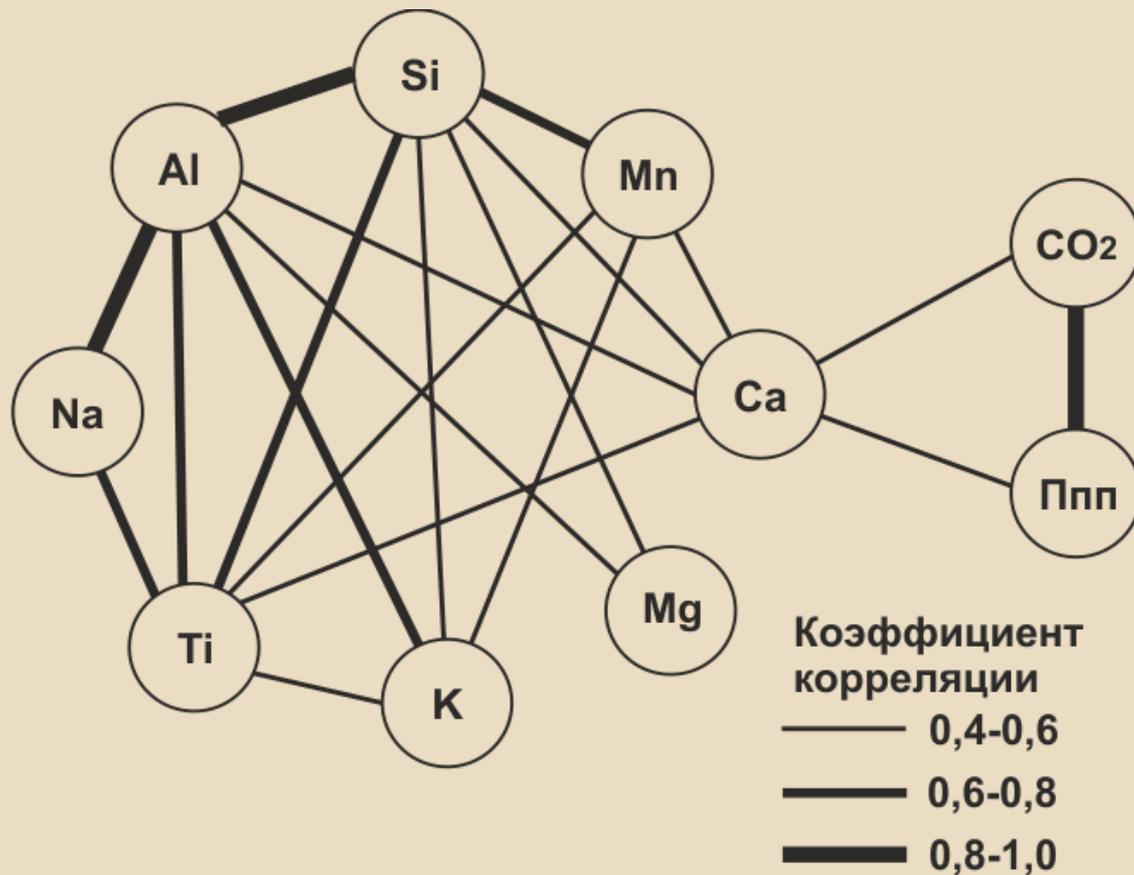
ПРИМЕНЕНИЕ МНОГОМЕРНОЙ СТАТИСТИЧЕСКОЙ МОДЕЛИ В ГЕОЛОГИИ



Примеры математических методов, использующихся для построения многомерных моделей

- Корреляционный анализ (графы связей, многомерная корреляция)
- Многомерный регрессионный анализ
- Кластерный анализ
- Дискриминантный анализ
- Факторный анализ
- И др.

Пример построение графа связей на основе корреляционной матрицы, который наглядно иллюстрирует характер взаимосвязей между свойствами и несет определенную геологическую информацию



Граф связей

Коэффициент множественной корреляции

- Это совокупный коэффициент корреляции, который рассчитывается если установлены корреляционные связи между несколькими парами признаков (x-y, x-z, y-z)

$$r_{x,y,z} = \sqrt{\frac{r_{xy}^2 + r_{yz}^2 - 2r_{xy}r_{xz}r_{yz}}{1 - r_{xy}^2}}$$

Множественная линейная регрессия

Во многих случаях возникает необходимость изучить зависимость одной случайной величины от множества других случайных величин.

Многофакторная зависимость обычно выражается уравнением множественной линейной регрессии

$$y = a_1x_1 + a_2x_2 + \dots + a_nx_n + b$$

где x_1, x_2, \dots, x_n – свойства; a_1, a_2, \dots, a_n, b – постоянные коэффициенты

Процедура множественного регрессионного анализа

1. Постановка задачи с определением зависимой переменной y (необходимого нам признака, который может зависеть от других признаков)
2. Проведение корреляционного анализа с выбором информативных признаков
3. Составление уравнения множественной регрессии и расчет коэффициентов регрессии



Выявление информативных признаков

Компонент	Cu	Zn	S	Pb
Cu	1,000			
Zn	-0,130	1,000		
S	0,370	0,276	1,000	
Pb	0,132	0,659	0,217	1,000
Ag	<u>0,541</u>	0,451	0,394	<u>0,811</u>

Зависимая переменная (y) - **Ag**

$$\mathbf{Ag = 1,923Pb + 1,074Cu + 2,871}$$

Предсказание свойств объекта с помощью уравнения множественной регрессии

- Уравнение регрессии, полученное для эталонного объекта, используется для аналогичного, но малоизученного объекта, в котором значения признака **y**, отнесенного нами к зависимой переменной, неизвестны, но известны значения признаков **x, z**
- $$Ag = 1,923Pb + 1,074Cu + 2,871$$

y **x** **z**
- **Далее можно подставлять значения независимых переменных x и z в данное уравнение для получения прогнозируемого значения y в любой точке наблюдения**

Классификационные методы математической статистики, используемые в геологии

Примеры классификаций

При классификации признаков геологических объектов можно выявлять парагенетические ассоциации элементов, информативные прогнозные признаки.

Можно выделять группы метасоматических пород, разделяя их на околорудные и безрудные.

Можно классифицировать осадочные породы по набору фациальных признаков

Разделять кимберлиты на алмазоносные и неалмазоносные, граниты – на редкометалльные и безрудные и т.д.

Классифицировать можно как признаки, так и сами объекты

Принципы классификации

**Классификация без учителя
(т.е без аналога)**

- Графы связей
- **Кластерный анализ**
- Факторный анализ

**Классификация с учителем
(с аналогом)**

- **Дискриминантный анализ**

Благодарю за внимание