

Использование корреляционных связей для предсказания свойств геологических объектов

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

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



Основные вопросы лекции

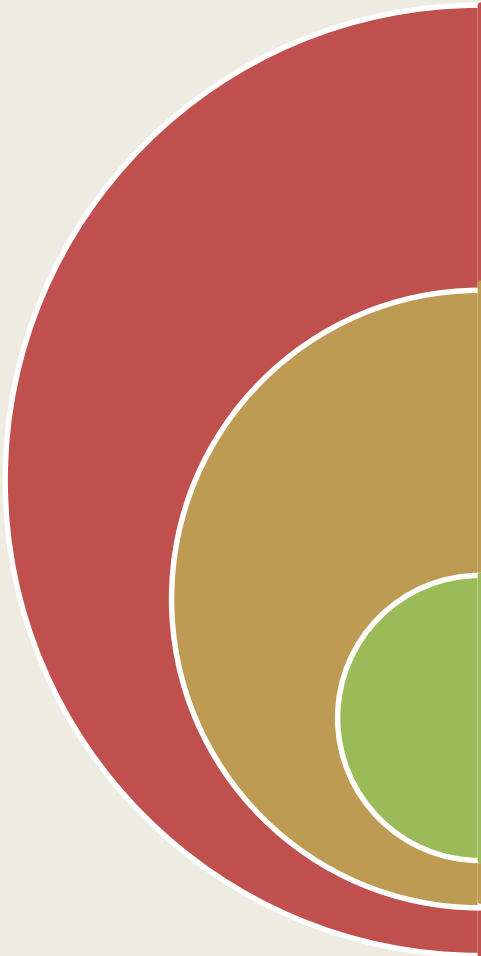
Сущность и условия применения двумерных статистических моделей


Простейшие графические способы преобразования двумерных случайных величин

Корреляционный анализ в геологических исследованиях (проверка гипотезы о наличии корреляционной связи)

Регрессионный анализ

Двумерные статистические модели



При моделировании геологических образований возникает необходимость совместного рассмотрения двух свойств 

Эти свойства могут быть независимы друг от друга

Между свойствами могут быть выявлены взаимосвязи 

ДВУМЕРНАЯ ВЫБОРКА

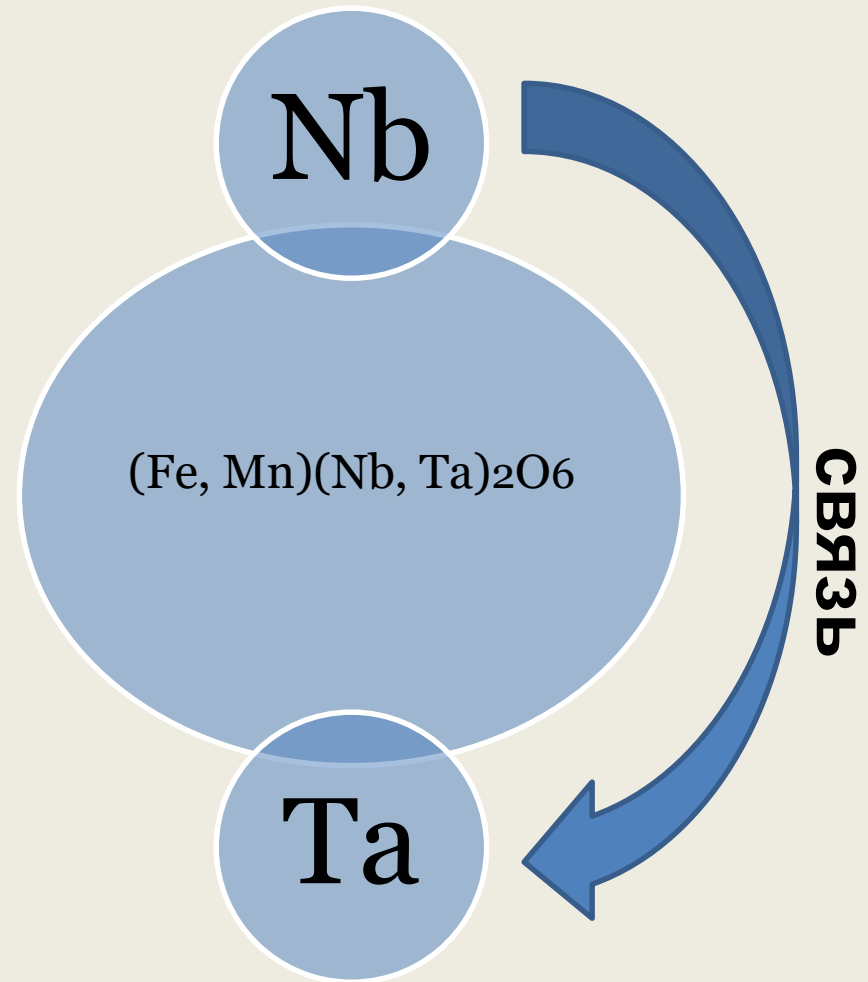
Содержание железа общего и железа магнитного в железных рудах

| Номер пробы n | Содержание железа, % | |
|--------------------|----------------------|----------------------|
| | общего x | магнетитового y |
| 1 | 52,0 | 45,7 |
| 2 | 49,4 | 45,4 |
| 3 | 34,5 | 28,4 |
| 4 | 41,5 | 36,6 |
| 5 | 36,5 | 22,1 |
| 6 | 22,7 | 10,9 |
| 7 | 42,3 | 27,5 |
| 8 | 20,0 | 10,3 |
| 9 | 23,9 | 17,3 |
| 10 | 23,8 | 16,0 |
| 11 | 33,2 | 23,8 |
| 12 | 61,8 | 55,8 |
| 13 | 63,7 | 57,3 |
| 14 | 22,1 | 15,2 |

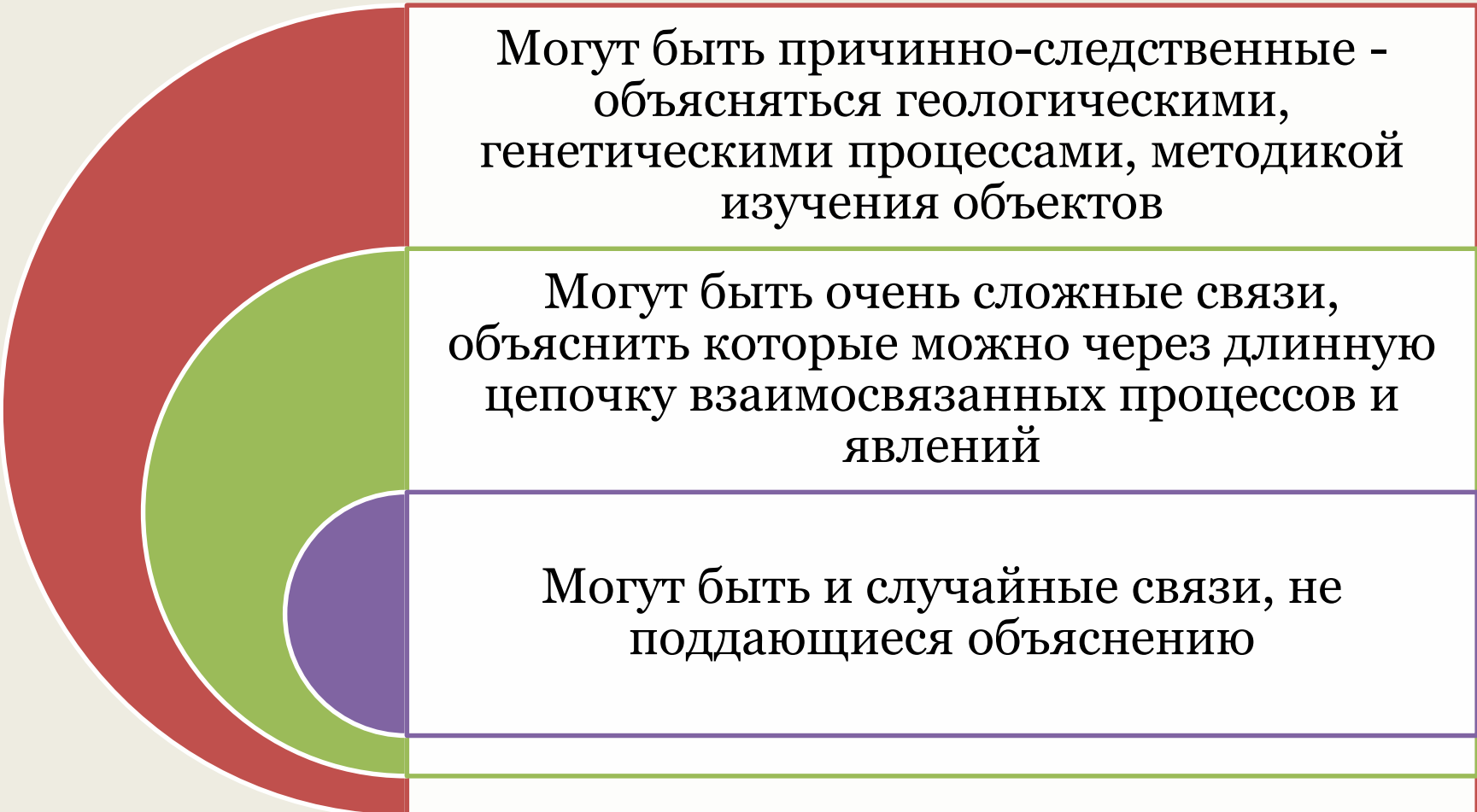


Пример

- В редкометальных пегматитах ниобий и тантал ВХОДЯТ В СОСТАВ минерала – танталит-колумбита



Связи между признаками



Могут быть причинно-следственные -
объясняться геологическими,
генетическими процессами, методикой
изучения объектов

Могут быть очень сложные связи,
объяснить которые можно через длинную
цепочку взаимосвязанных процессов и
явлений

Могут быть и случайные связи, не
поддающиеся объяснению

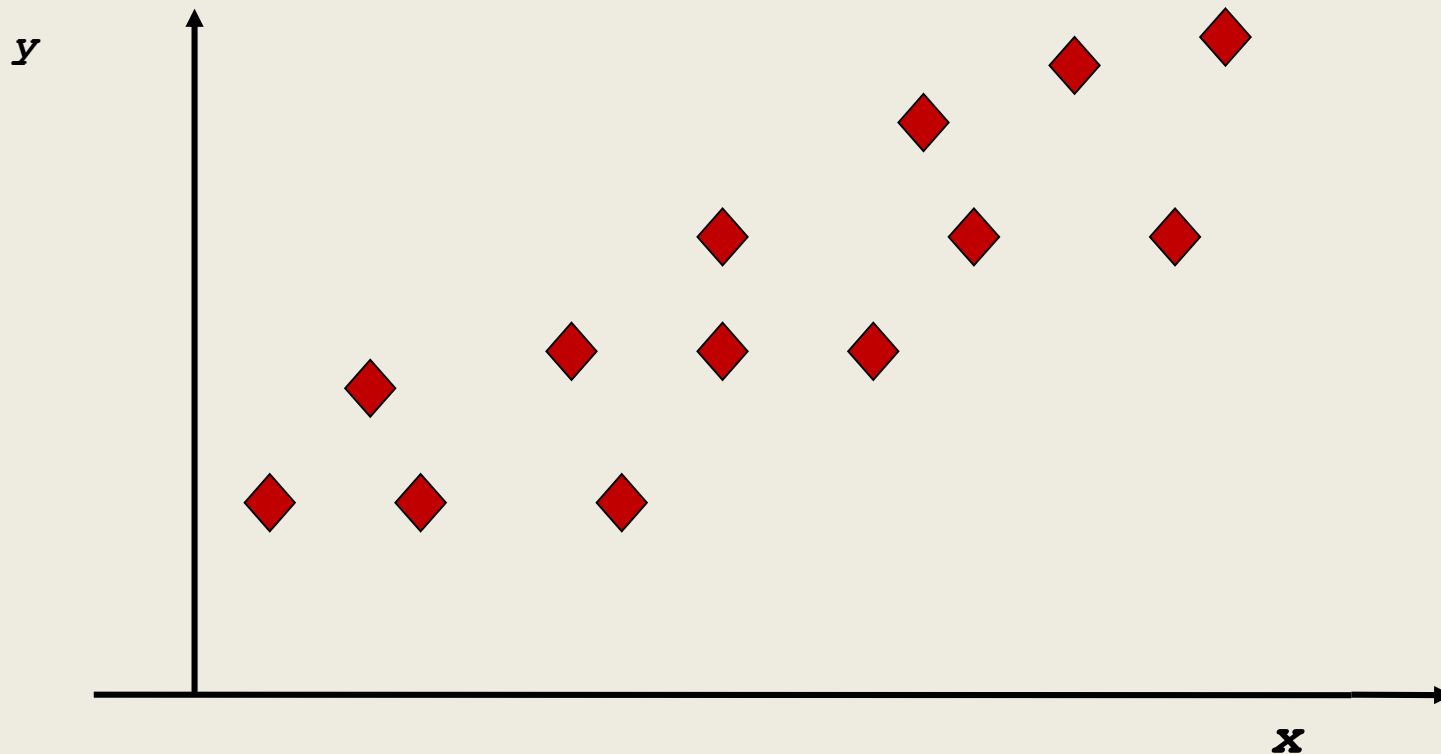
↓ Основная характеристика двумерной модели

двумерная функция распределения случайных
величин
 x и y

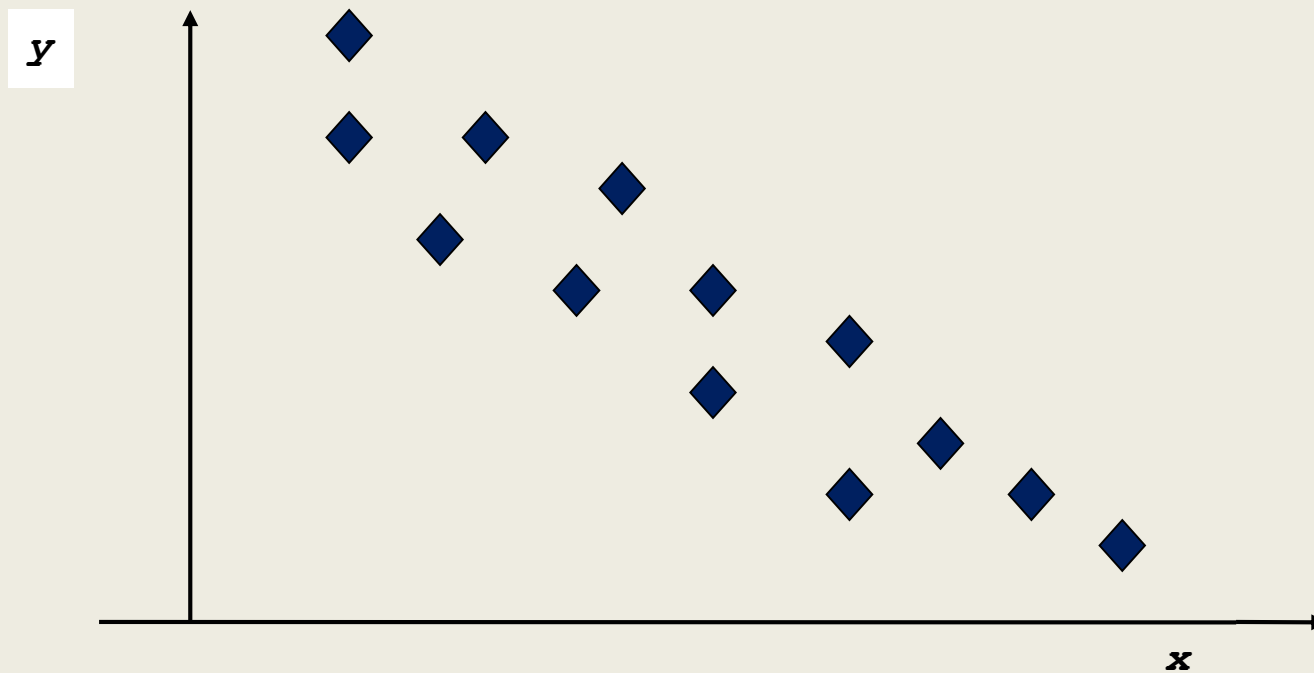
Заданному значению x *соответствует*
определенное значение y

Графические способы преобразования двумерных случайных величин

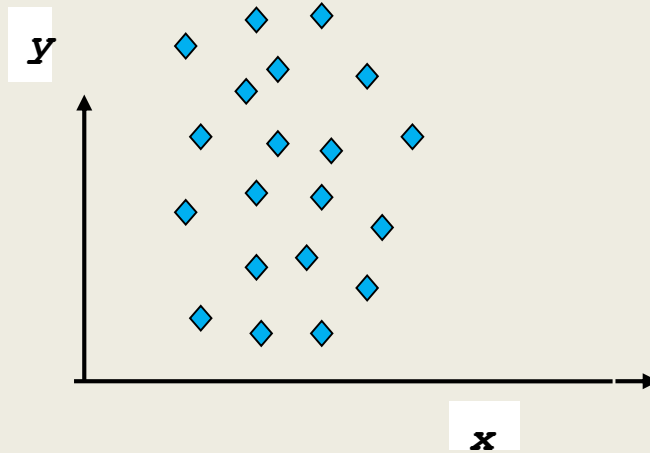
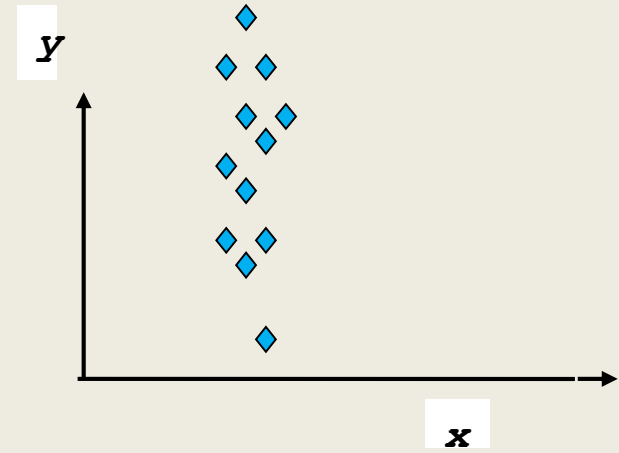
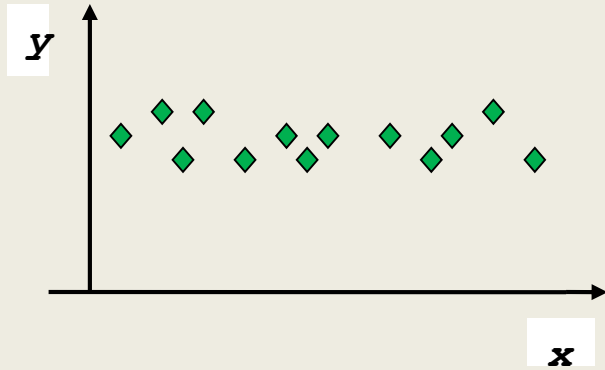
Связь прямая – зависимость положительная



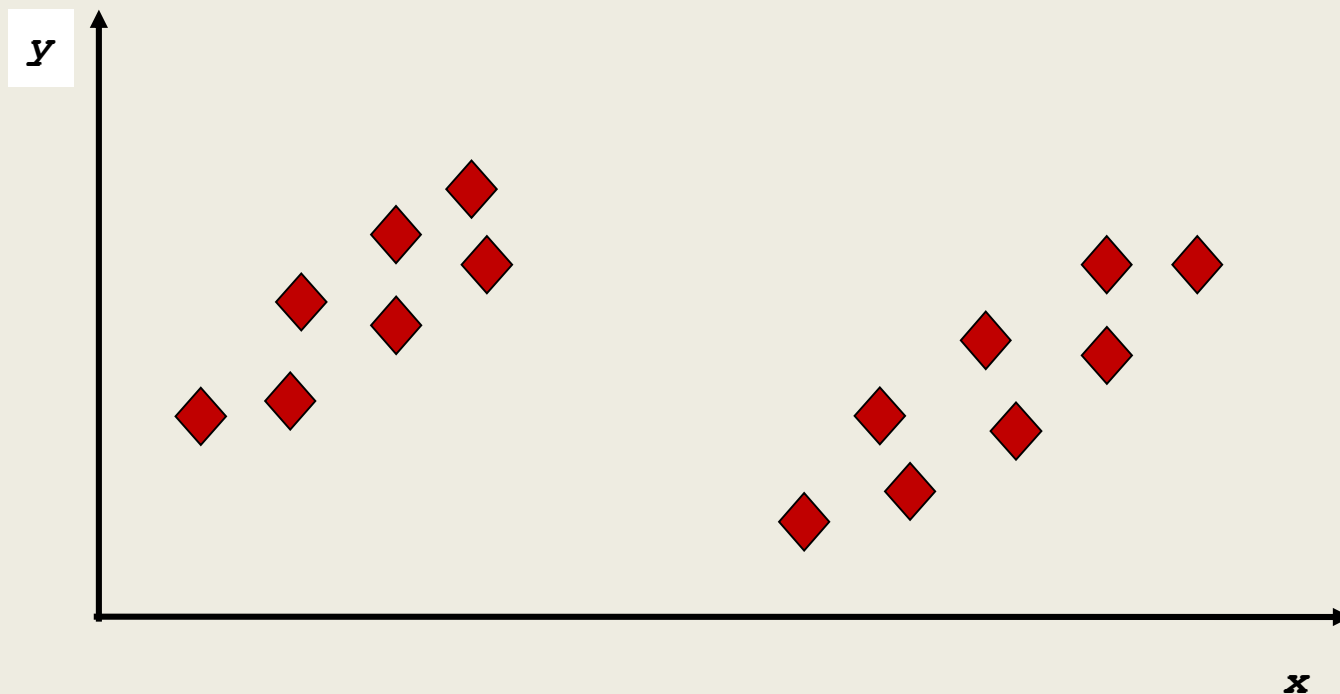
Отрицательная связь



Отсутствие связи



Две однородные совокупности, которые надо изучать отдельно



Корреляционный анализ

$$r = \frac{\sum (x - \bar{x})(y - \bar{y})}{nS_x S_y}$$

Где

x – значения $x_1, x_2, x_3, x_4, \dots, x_n$

y – значения $y_1, y_2, y_3, y_4, \dots, y_n$

n - количество точек наблюдения

S – среднеквадратичное отклонение

Коэффициент корреляции

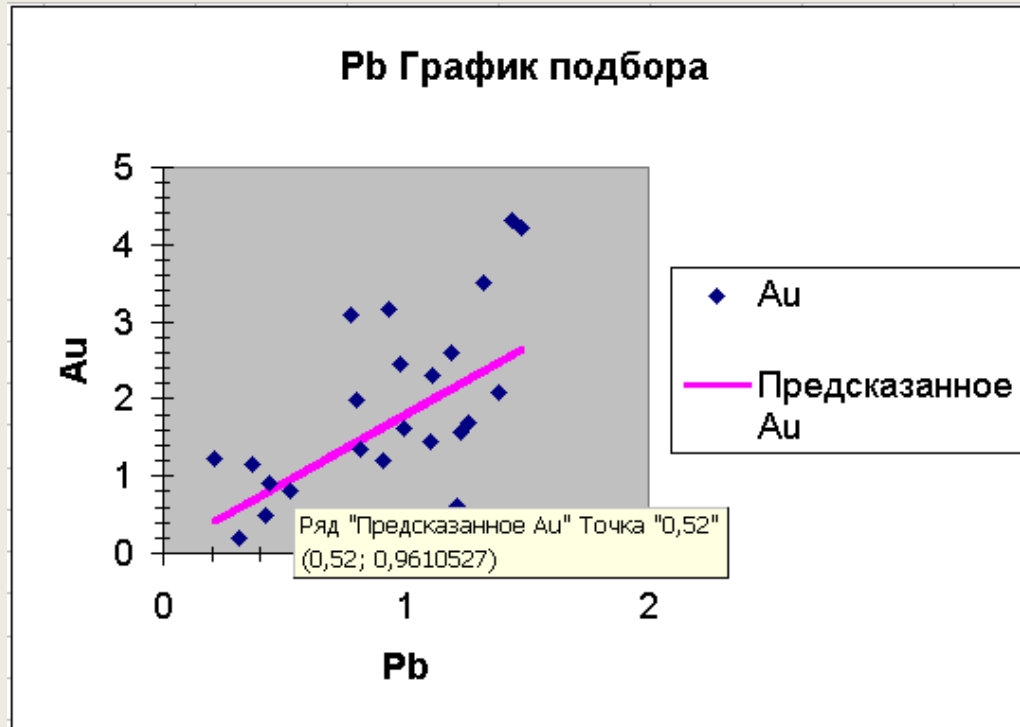
- $r \leq 1$
- Может быть + и –
- Значимый и незначимый
 - $|r| > r_{кр}$
- $r_{кр}$ зависит от количества точек наблюдения
- Значимый положительный
- Значимый отрицательный

Корреляционная матрица

| | FE | TI | MN | CU | NI | V | MO | SN | PB |
|----|---------------------|---------------------|---------------------|--------------------|---------------------|---------------------|--------------------|-------|---------------------|
| FE | 1 | | | | | | | | |
| TI | 0,07 | 1,00 | | | | | | | |
| MN | 0,17 | <u>0,59</u> | 1,00 | | | | | | |
| CU | 0,00 | <u>0,89</u> | <u>0,33</u> | 1,00 | | | | | |
| NI | <u>0,45</u> | <u>0,67</u> | 0,27 | <u>0,65</u> | 1,00 | | | | |
| V | <u>0,63</u> | <u>0,61</u> | <u>0,34</u> | <u>0,58</u> | <u>0,91</u> | 1,00 | | | |
| MO | <u>0,45</u> | 0,14 | 0,24 | 0,10 | 0,25 | <u>0,33</u> | 1,00 | | |
| SN | <u>-0,33</u> | <u>0,62</u> | <u>0,47</u> | <u>0,60</u> | 0,17 | 0,09 | 0,09 | 1,00 | |
| PB | <u>0,81</u> | <u>0,33</u> | 0,23 | <u>0,36</u> | <u>0,71</u> | <u>0,84</u> | <u>0,36</u> | -0,11 | 1,00 |
| CO | -0,31 | <u>-0,35</u> | <u>-0,54</u> | -0,21 | <u>-0,47</u> | <u>-0,50</u> | -0,05 | -0,03 | <u>-0,45</u> |

*Выделены значимые коэффициенты корреляции

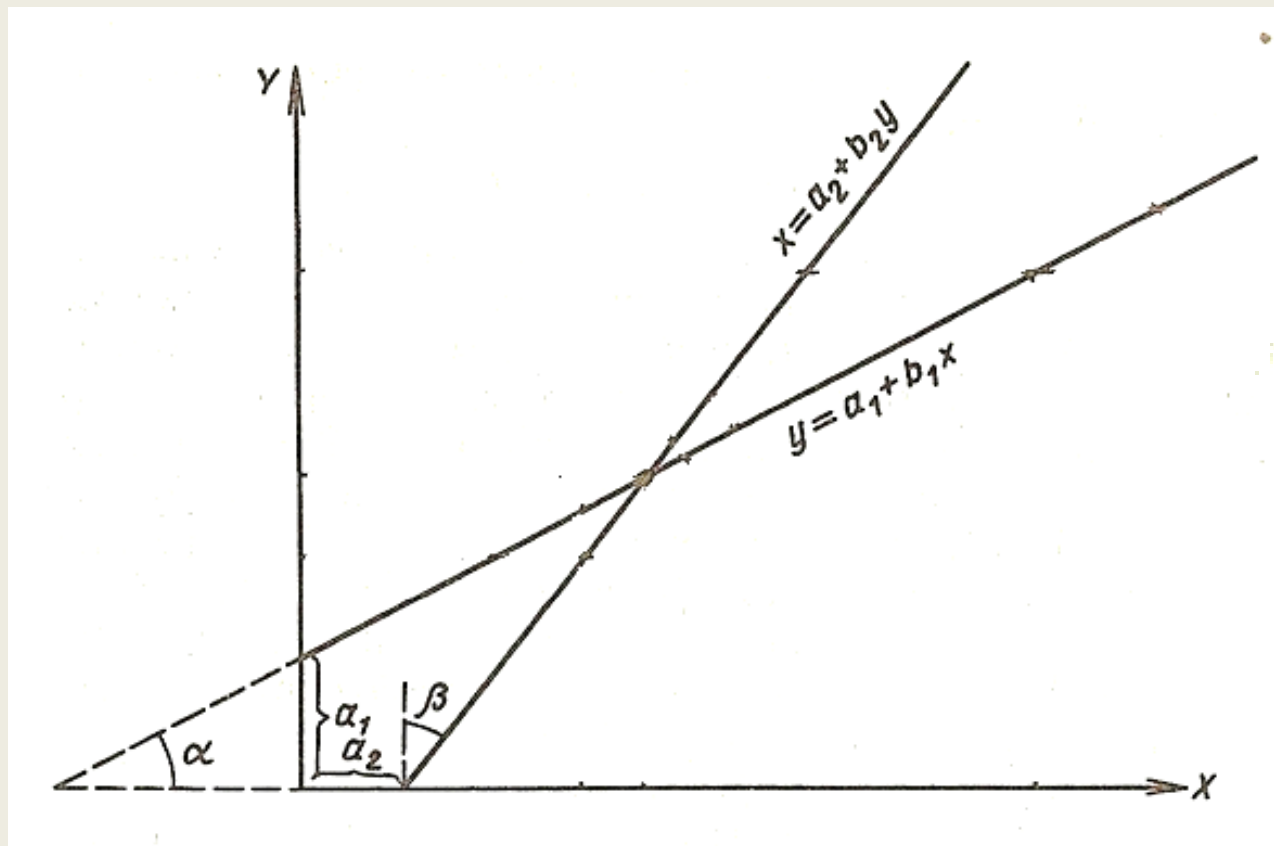
Предсказание свойств



Регрессионный анализ. Уравнение линейной регрессии

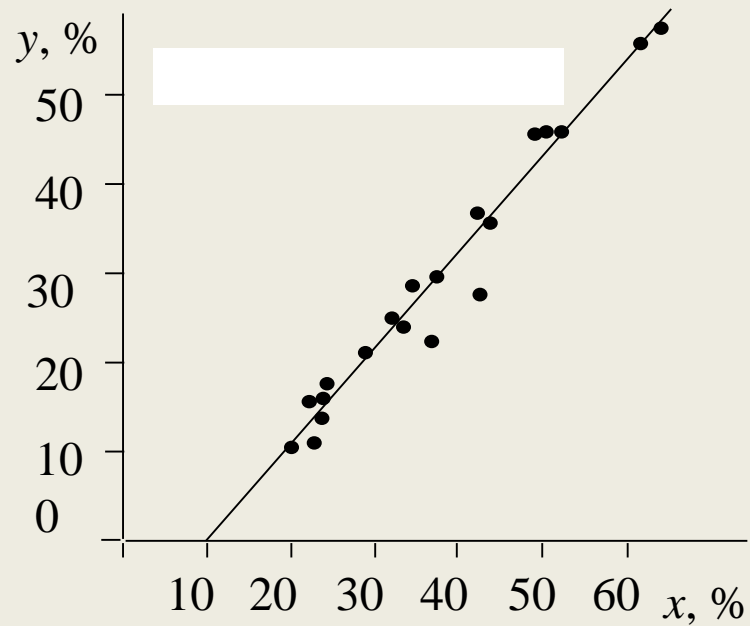
- Если между величинами x и y установлена линейная статистическая зависимость, то представляет интерес найти ее выражение в виде уравнения прямой линии
- $y = ax + b$ (где a и b – коэффициенты)
- Такое уравнение называется *уравнением регрессии*
- *Заданному значению x соответствует определенное значение y*
- *X – независимая переменная*
- *Y – зависимая переменная*

Линейная зависимость между признаками x и y



$$b_1 = \operatorname{tg} \alpha; \quad b_2 = \operatorname{tg} \beta$$

Зависимость содержания
магнетитового железа x от содержания
общего железа y



$$y = 1,08x - 11,0$$

Примеры геологических задач

выявление парагенетических ассоциаций минералов и химических элементов в горных породах и рудах

изучение поведения химических элементов в процессе гипергенного и метасоматического изменения пород

палеогеографический анализ условий формирования осадочных пород

выяснение источников рудного вещества

СПАСИБО ЗА ВНИМАНИЕ

