Теория оптимизации

Максим Сахаров
Кандидат технических наук, старший консультант по Data Science, BasisSoft

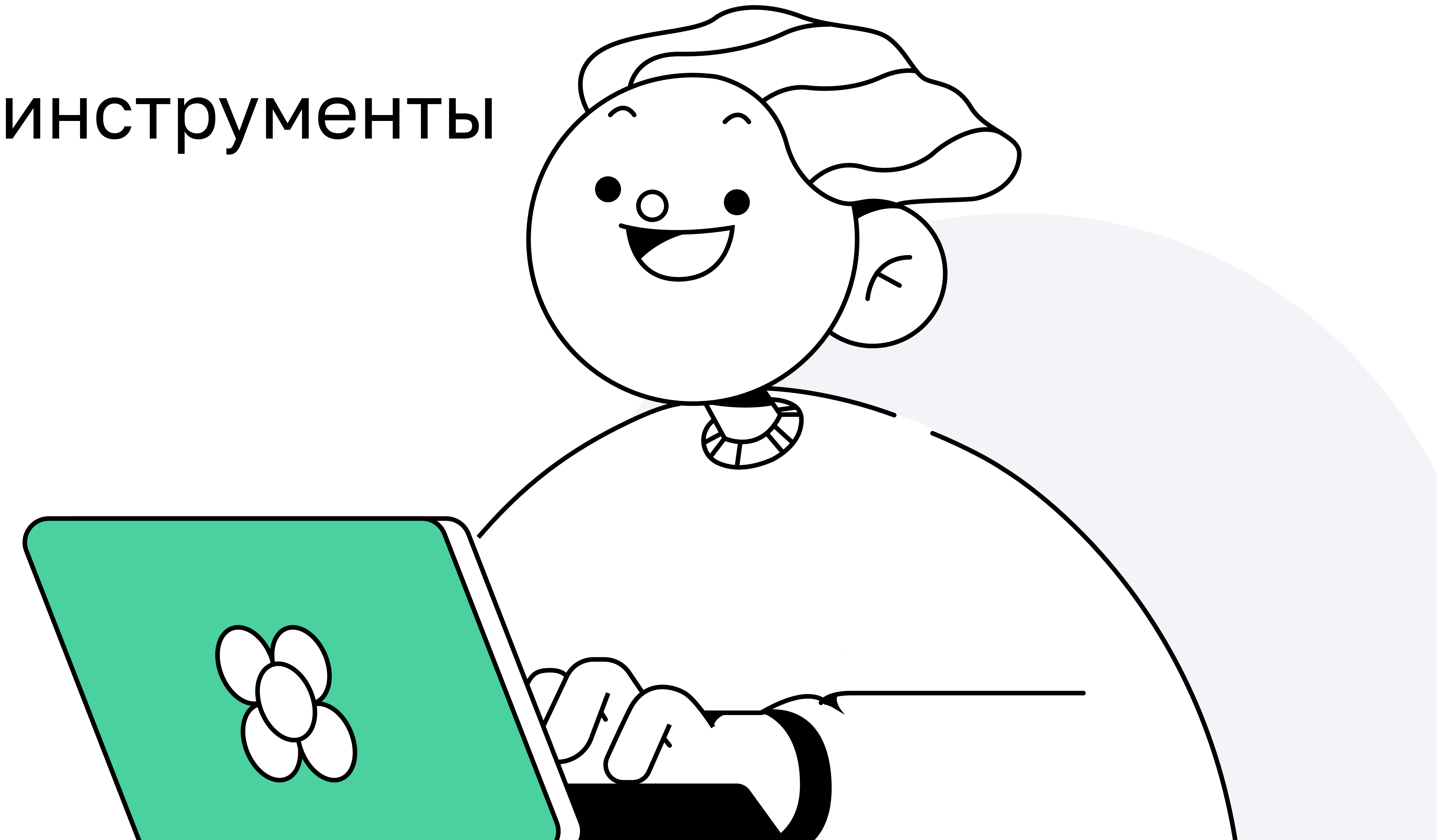
Максим Сахаров

О спикере:

- Старший консультант по Data Science, BasisSoft
- к. т. н., доцент МГТУ им. Н. Э. Баумана
- Автор более 40 научных работ
- Области интересов:
  — анализ данных и машинное обучение
  — математическая оптимизация
  — статистическое управление процессами

План занятия

1. Постановка задачи оптимизации
2. Обучение как оптимизация
3. Методы оптимизации
4. Практика в Python: встроенные инструменты оптимизации SciPy



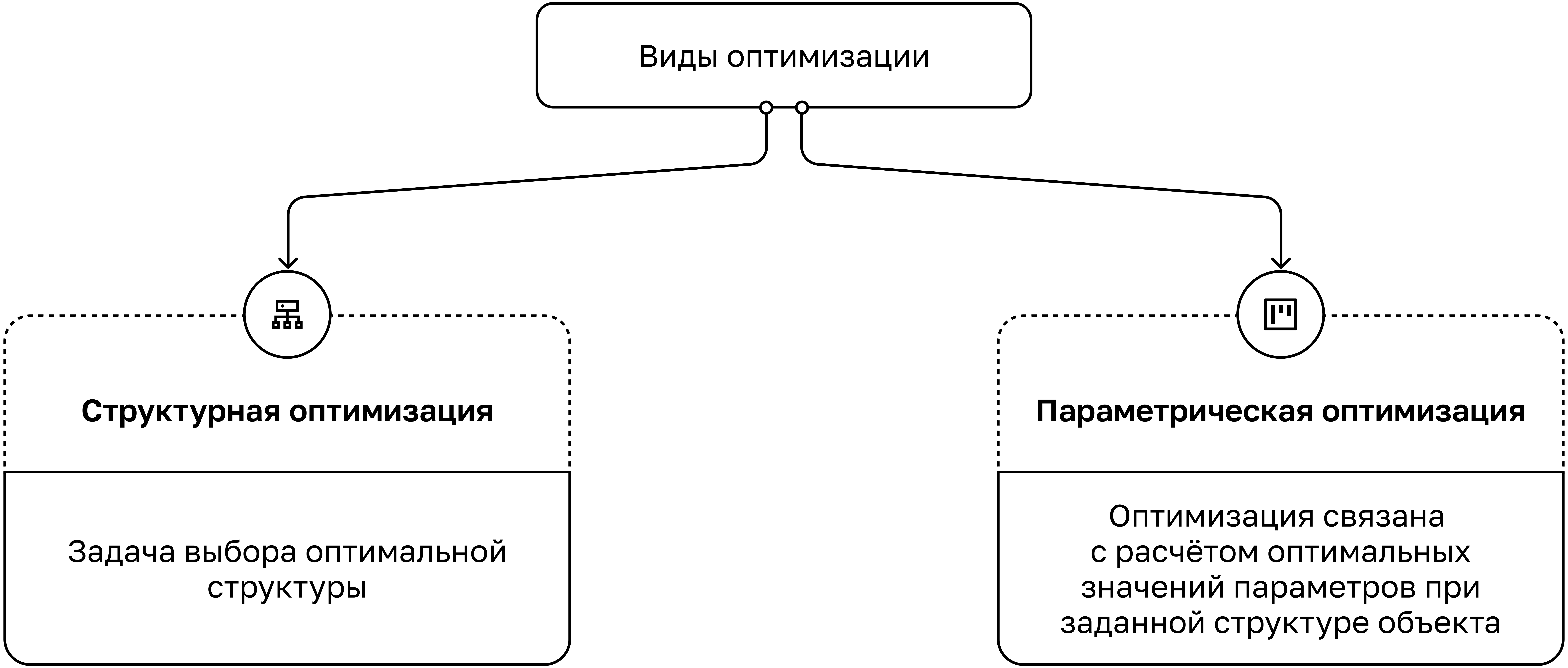
Постановка задачи оптимизации

1

Примеры задач оптимизации

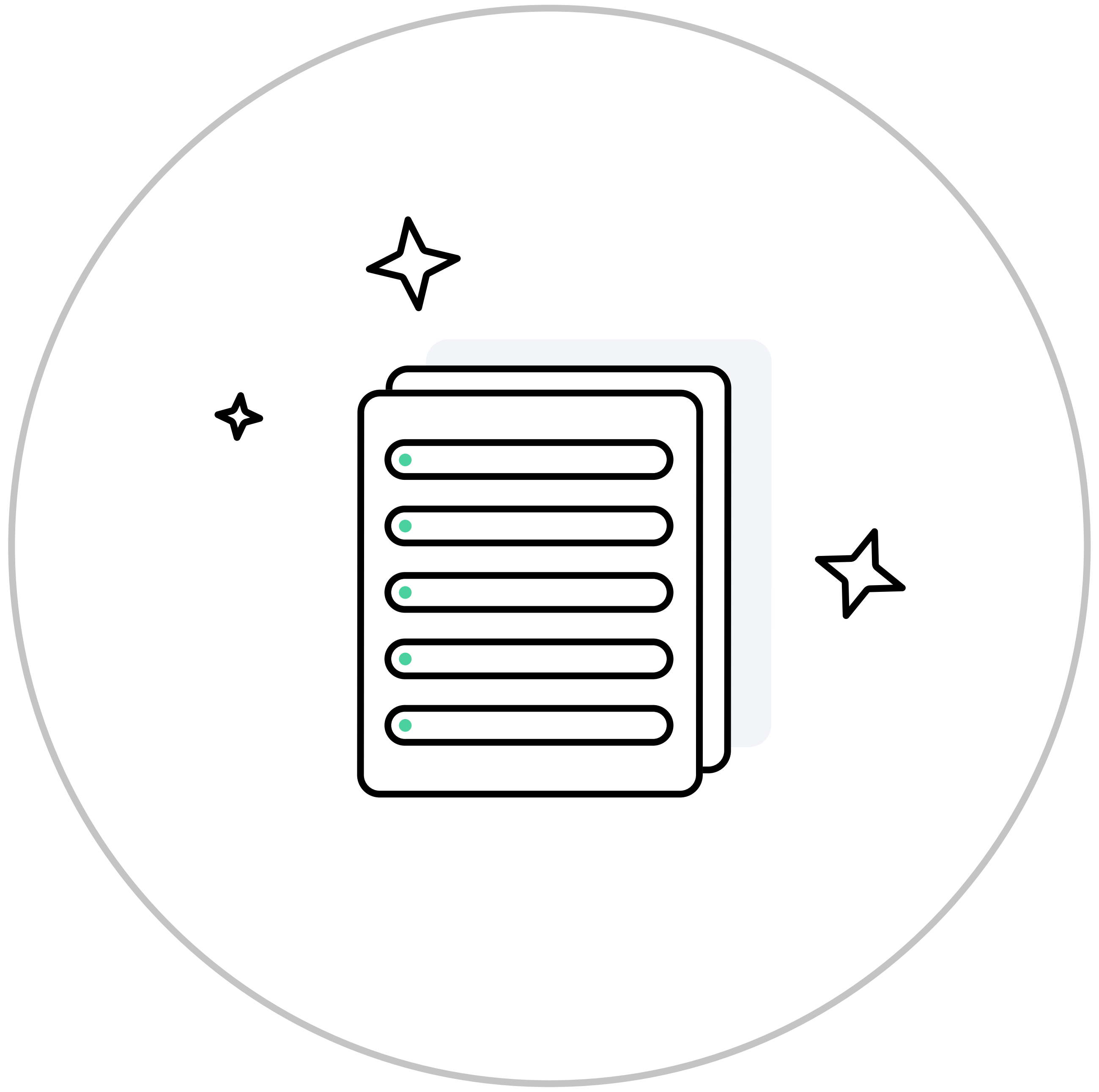
- Минимизация: как выполнить работу за максимально короткий срок
- Максимизация: как максимально увеличить прибыль портфеля акций
- Оптимизация с целевым значением, заданным требованием: как получить изделие определённой массы, не более n кг и не менее m кг

Оптимизационная задача — это задача определения наилучшей структуры или значений параметров объектов


Виды оптимизации
Структурная оптимизация
Задача выбора оптимальной структуры
Параметрическая оптимизация
Оптимизация связана с расчётом оптимальных значений параметров при заданной структуре объекта

Классификация задач оптимизации

- Локальная и глобальная оптимизация
- Условная и безусловная оптимизация
- Однокритериальная и многокритериальная оптимизация
- Одномерная и многомерная оптимизация
- Другие классификации



Локальная и глобальная оптимизация

Глобальная оптимизация	Локальная оптимизация
Ищет самый большой максимум или самый маленький минимум на всей области определения функции	Ищет хоть какой-нибудь максимум и хоть какой-нибудь минимум в некоторой окрестности

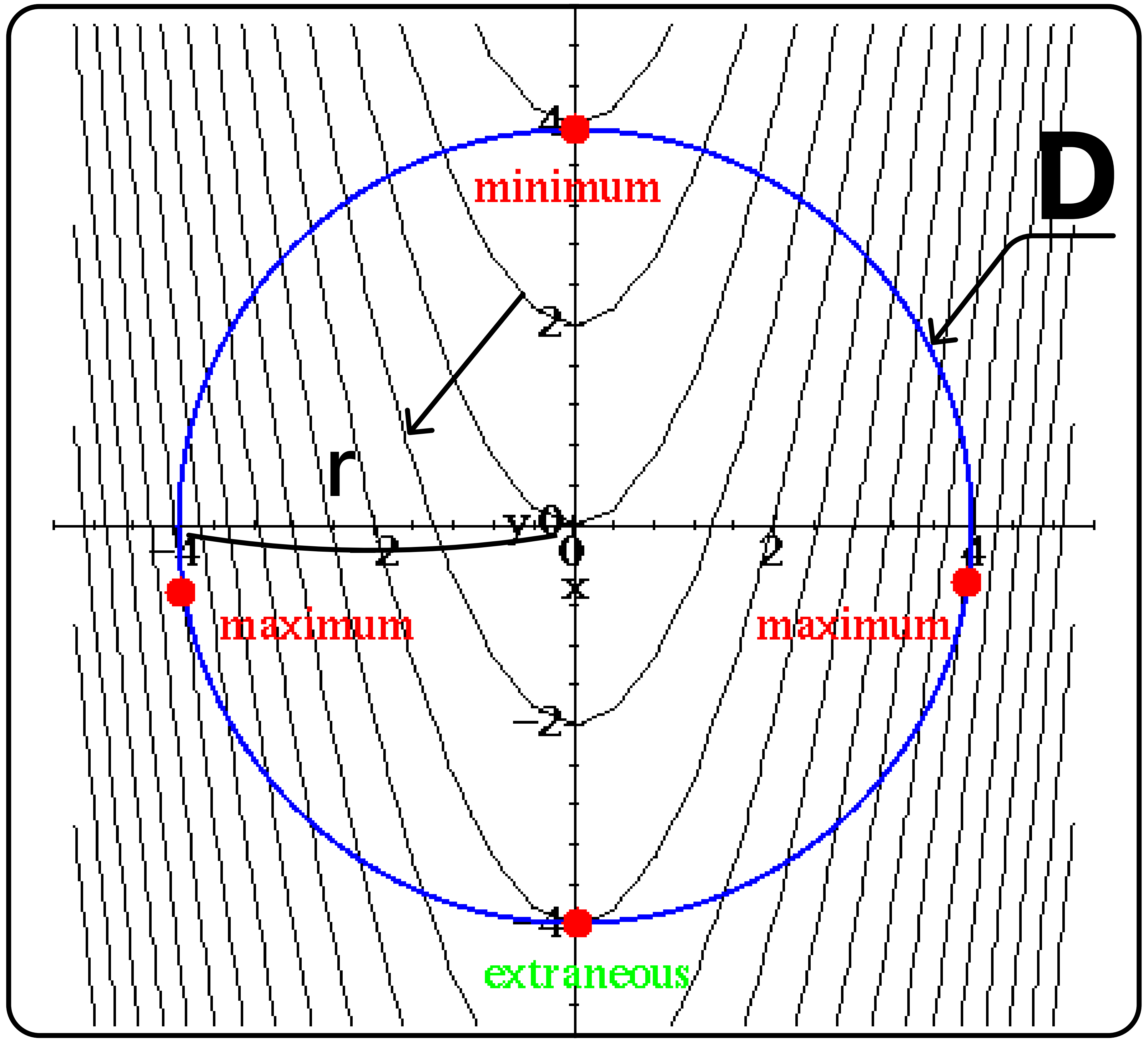
Условная и безусловная оптимизация

Безусловная оптимизация	**Условная оптимизация**
Нет ограничений на входные и выходные параметры	Есть ограничения на входные и выходные параметры
Чаще в теоретических задачах	Чаще в практических задачах

Условная и безусловная оптимизация

- Ограничения на входные параметры x
- Ограничения на выходные параметры f

$$x_1^2 + x_2^2 \leq r^2$$



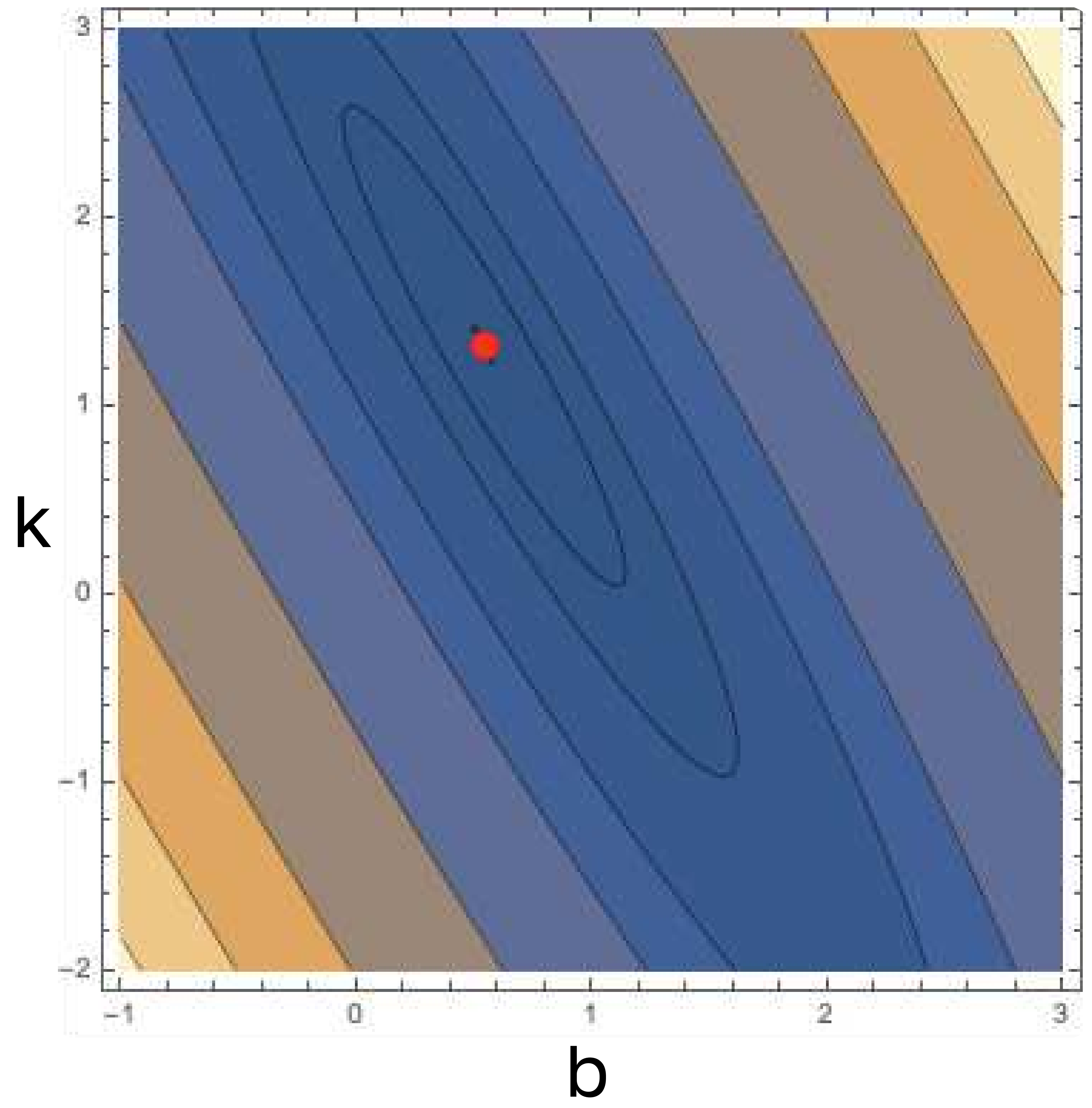
Однокритериальная и многокритериальная оптимизация

Однокритериальная	Многокритериальная
Оптимизация по одному критерию, например доходность или время	Оптимизация по нескольким критериям, например и доходность и время

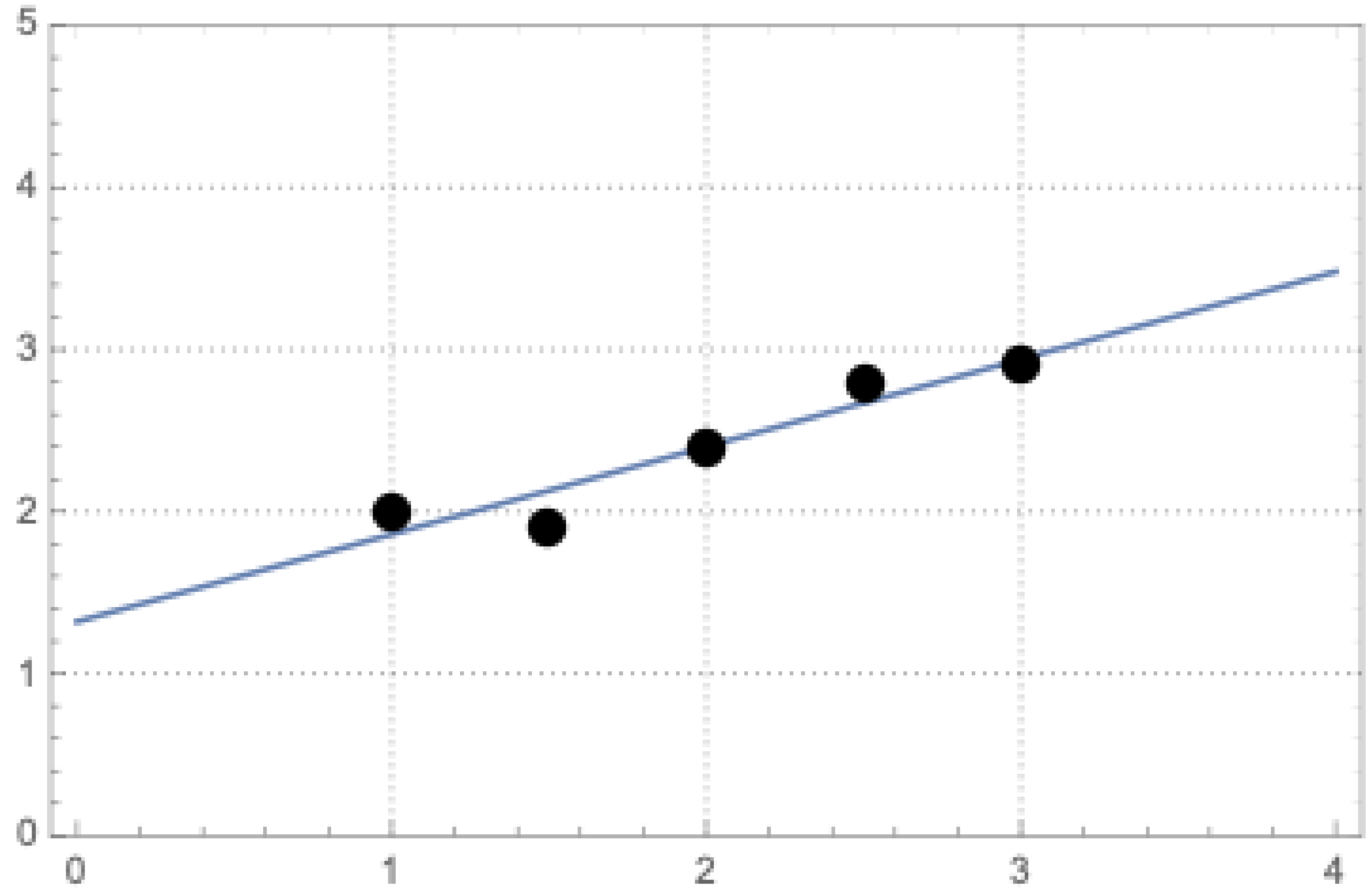
Обучение как оптимизация

2

Регрессионный анализ

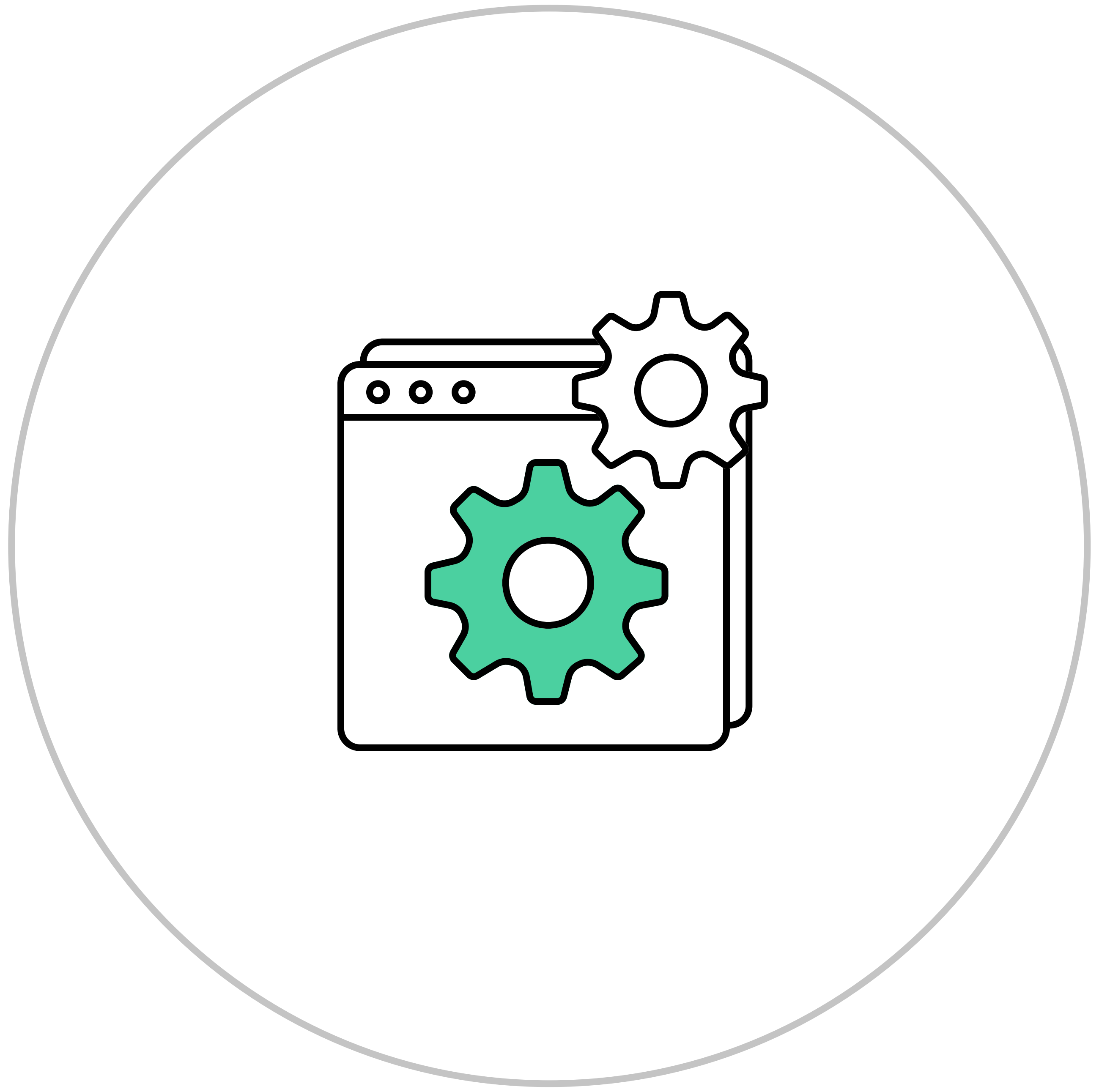


$y = 0.54 * x + 1.32$



Другие примеры

- Матричные разложения
- Обучение нейронных сетей

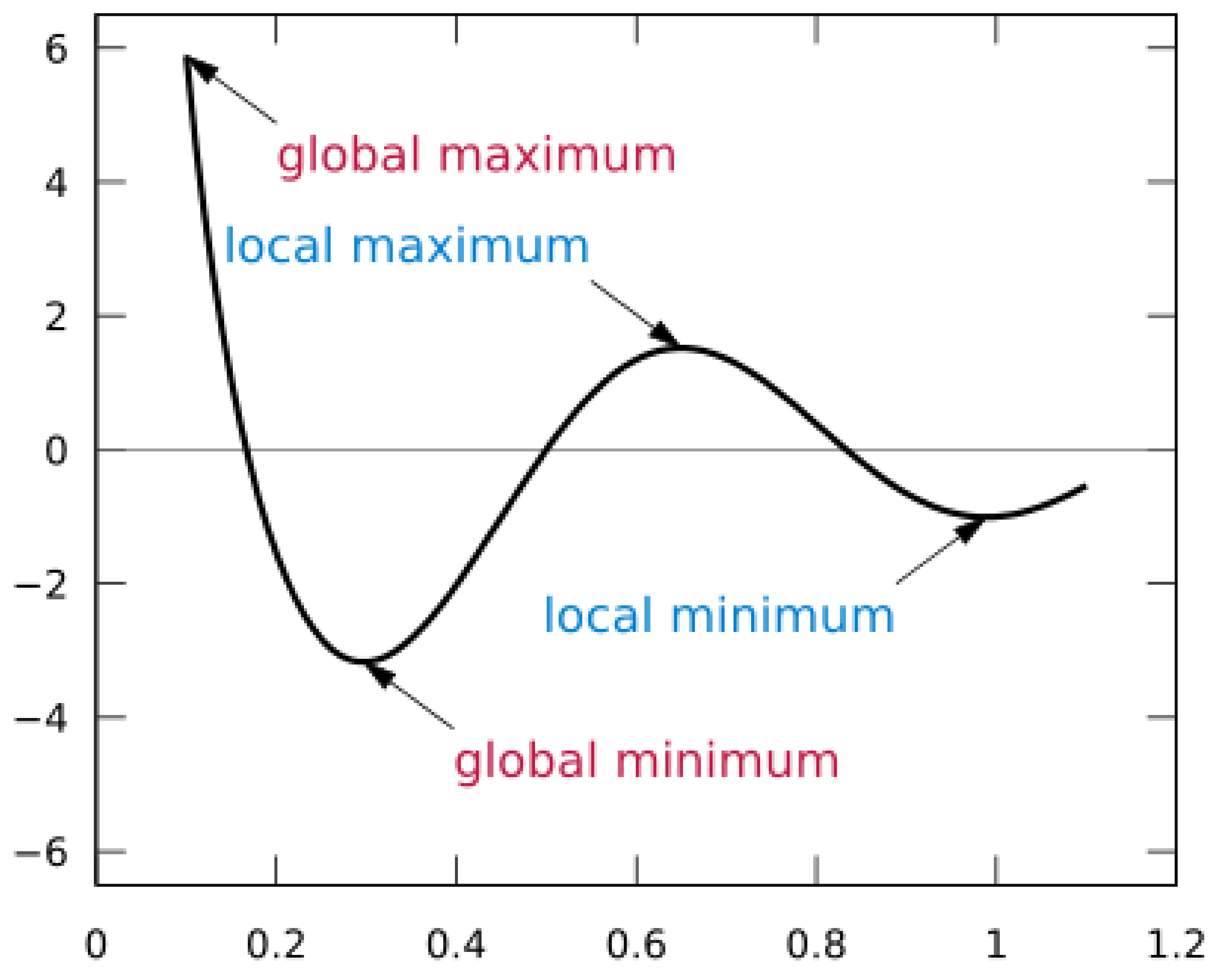


Методы оптимизации

3

А что же с экстремумами?

Нахождение экстремумов функции в практически значимых задачах возможно далеко не всегда!



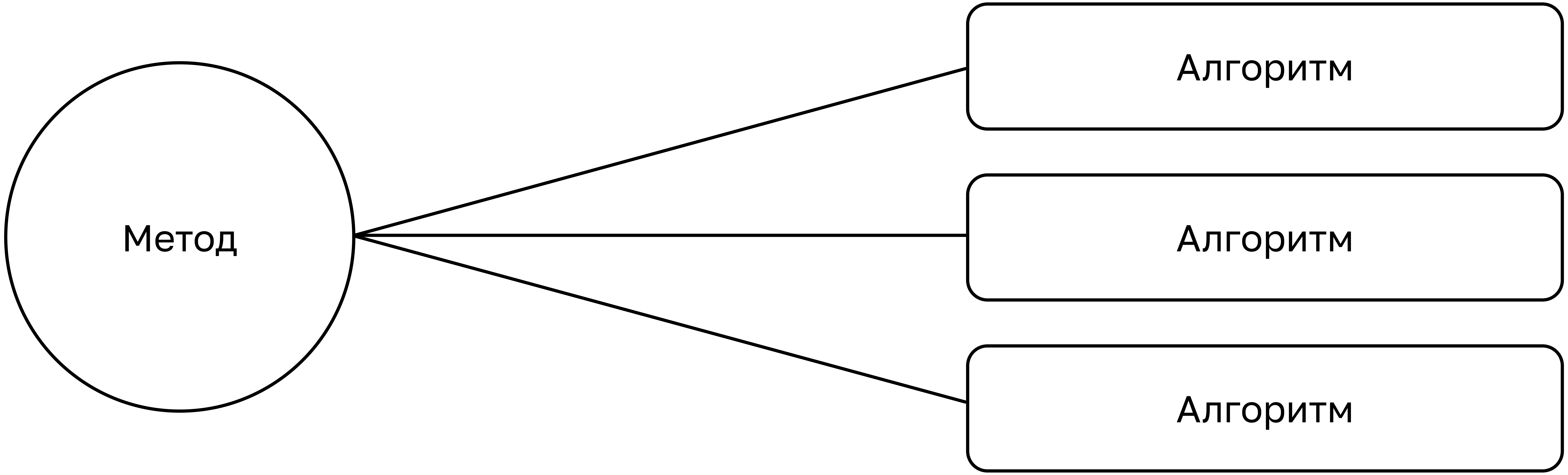
Почему?

- Мы работаем не вручную, а на компьютере: представление функции всегда дискретно
- Подходы, которые мы используем вручную (аналитическое решение), не являются для компьютера оптимальными, эффективными

На компьютерах при решении задач оптимизации используются поисковые методы оптимизации — вычислительные процедуры, последовательности шагов, удобные, в том числе, для написания программ

Метод или алгоритм

Одному методу может соответствовать несколько алгоритмов!

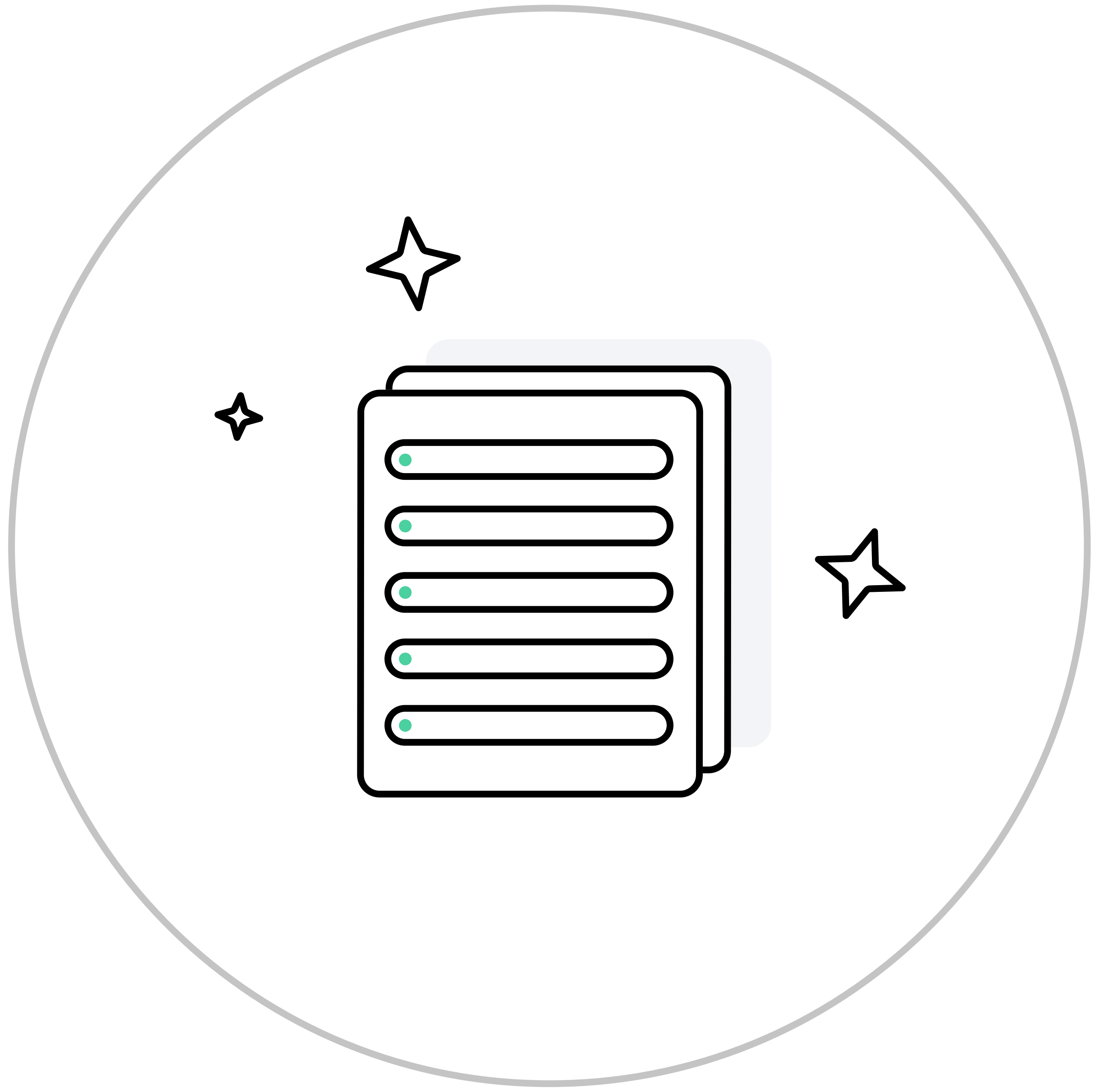


Метод — понятие более высокого уровня, которое содержит несколько последовательных шагов

Алгоритм — способ, который можно запрограммировать и использовать для решения задачи

Классификация методов оптимизации

- Локальные и глобальные методы
- Детерминированные и стохастические методы
- Методы на основе производных и прямые методы
- Другие классификации



Преимущества метода Нелдера-Мида

Метод Нелдера-Мида, или метод деформируемого многогранника, применяется для нахождения решения задачи оптимизации вещественных функций многих переменных

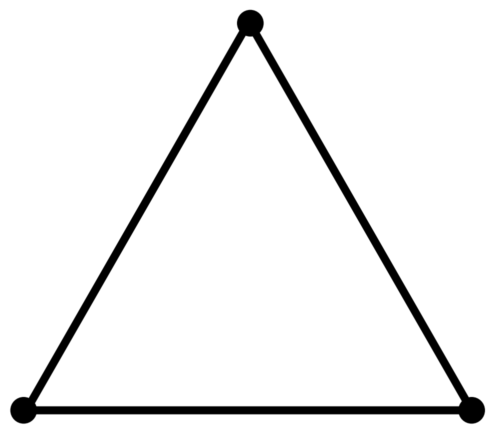
- Метод прост в реализации и полезен на практике
- **Классификация метода:** прямой детерминированный метод локальной оптимизации

Метод Нелдера-Мида

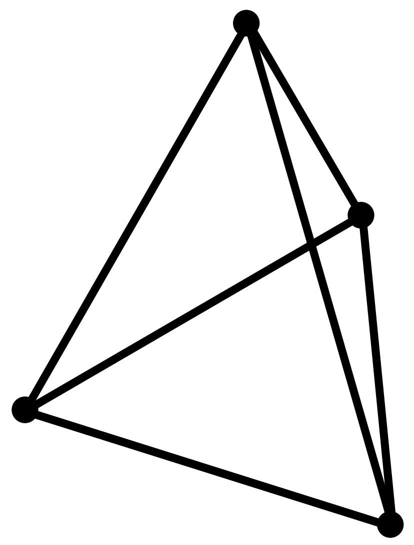
На каждой итерации для функции от **n** переменных требуется вычислить значение функции в **n+1** точке

1. Строим симплекс в зависимости от размерности пространства:

1D симплекс



2D симплекс



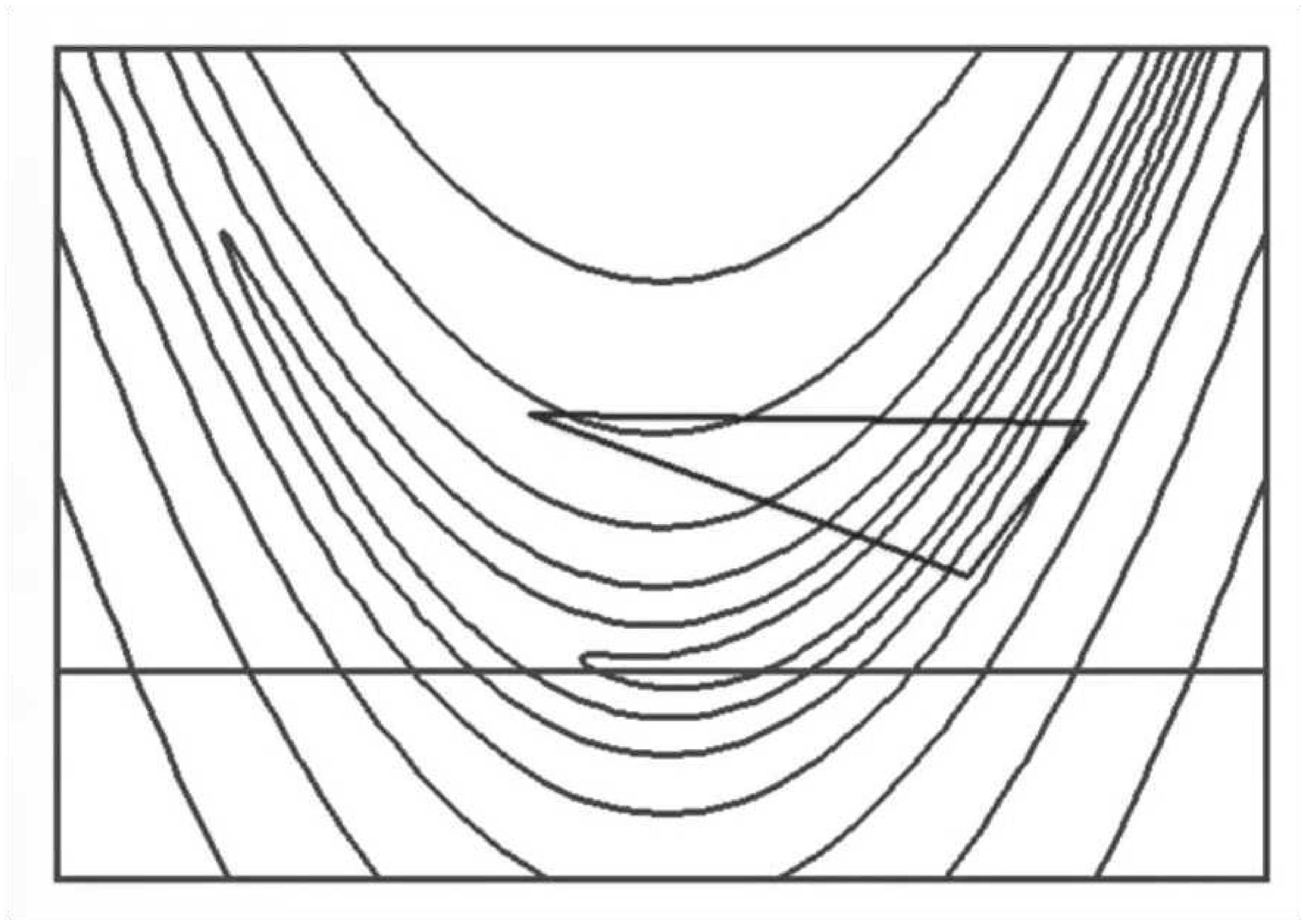
3D симплекс

Метод Нелдера-Мида

2 Сортируем значения функции в вершинах:

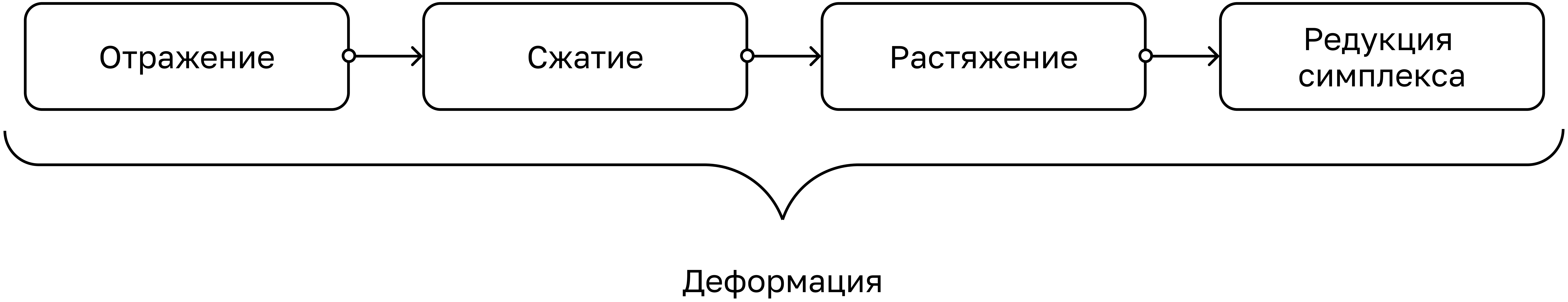
$$f\,(x_1) \le f\,(x_2) \le \ldots \le f\,(x_{n+1})$$

Уменьшить значение функции за один шаг в самой плохой точке $\mathbf{x}_{n+1}$.

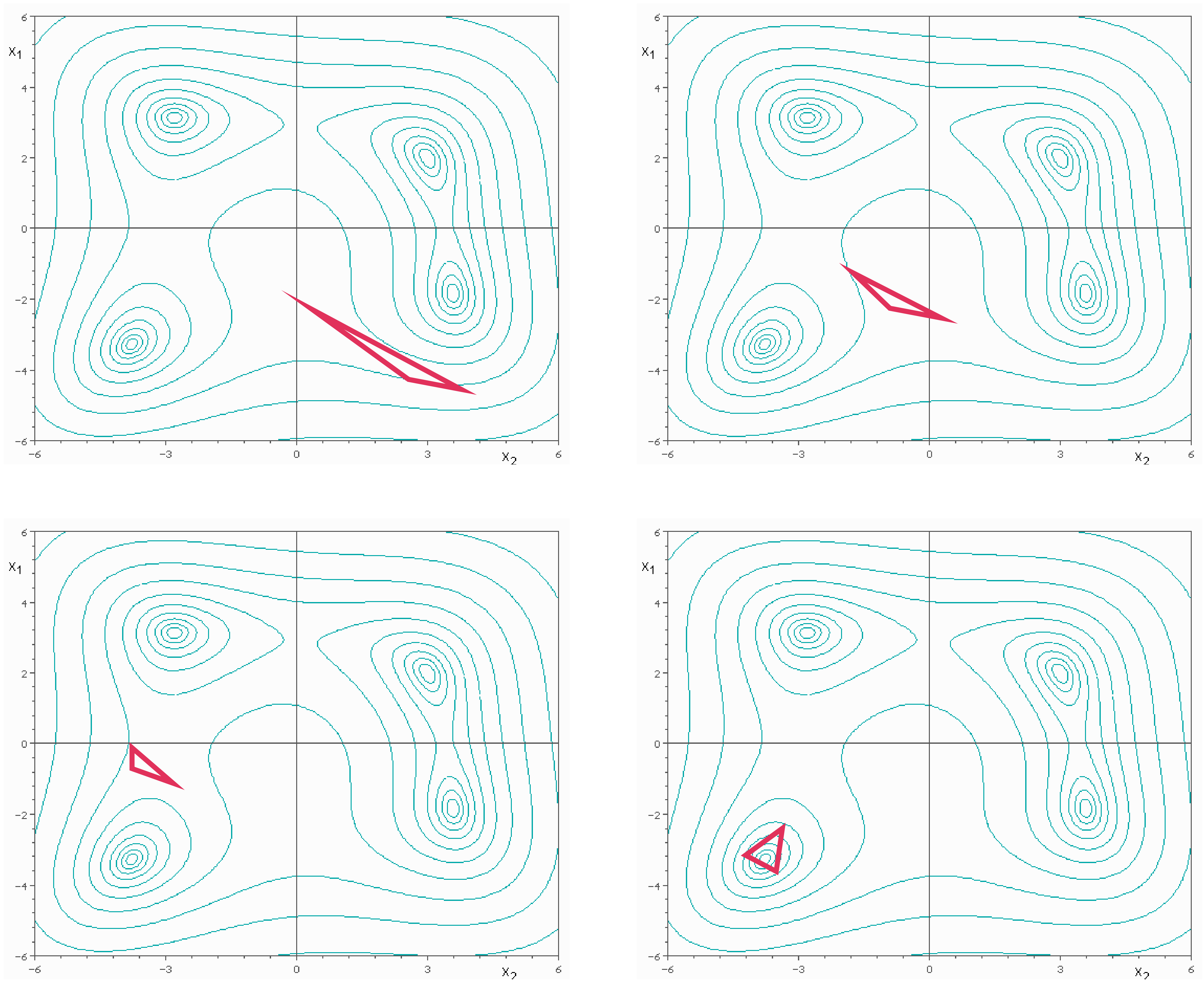


Метод Нелдера-Мида

3. Деформируем симплекс так, чтобы он «подползал» к минимуму функции:



Метод Нелдера-Мида



Метод Нелдера-Мида

4. Проверяем условия останова: симплекс стал достаточно маленьким или др.

Другими словами, максимальная разница между двумя любыми вершинами симплекса меньше заданной точности решения ε_F

$$\boxed{\max_{i,k\in[1:n+1],i\neq k}|F(X_i) - F(X_k)| < \varepsilon_F}$$

Практика в Python

Встроенные инструменты оптимизации SciPy

Итоги занятия

1. Задачи оптимизации возникают повсеместно и, разумеется, в науке о данных и машинном обучении
2. Обучение различных математических моделей сводится к решению задачи оптимизации, где необходимо минимизировать разницу между исходными данными и результатами модели
3. Существует огромное число различных методов оптимизации, предназначенных для решения классов задач
4. Метод Нелдера-Мида — простой и эффективный алгоритм локальной оптимизации, который использует только значения целевой функции
5. Решение задачи оптимизации — это всегда баланс между скоростью и качеством решения

Теория оптимизации

Максим Сахаров
Кандидат технических наук, старший консультант по Data Science, BasisSoft