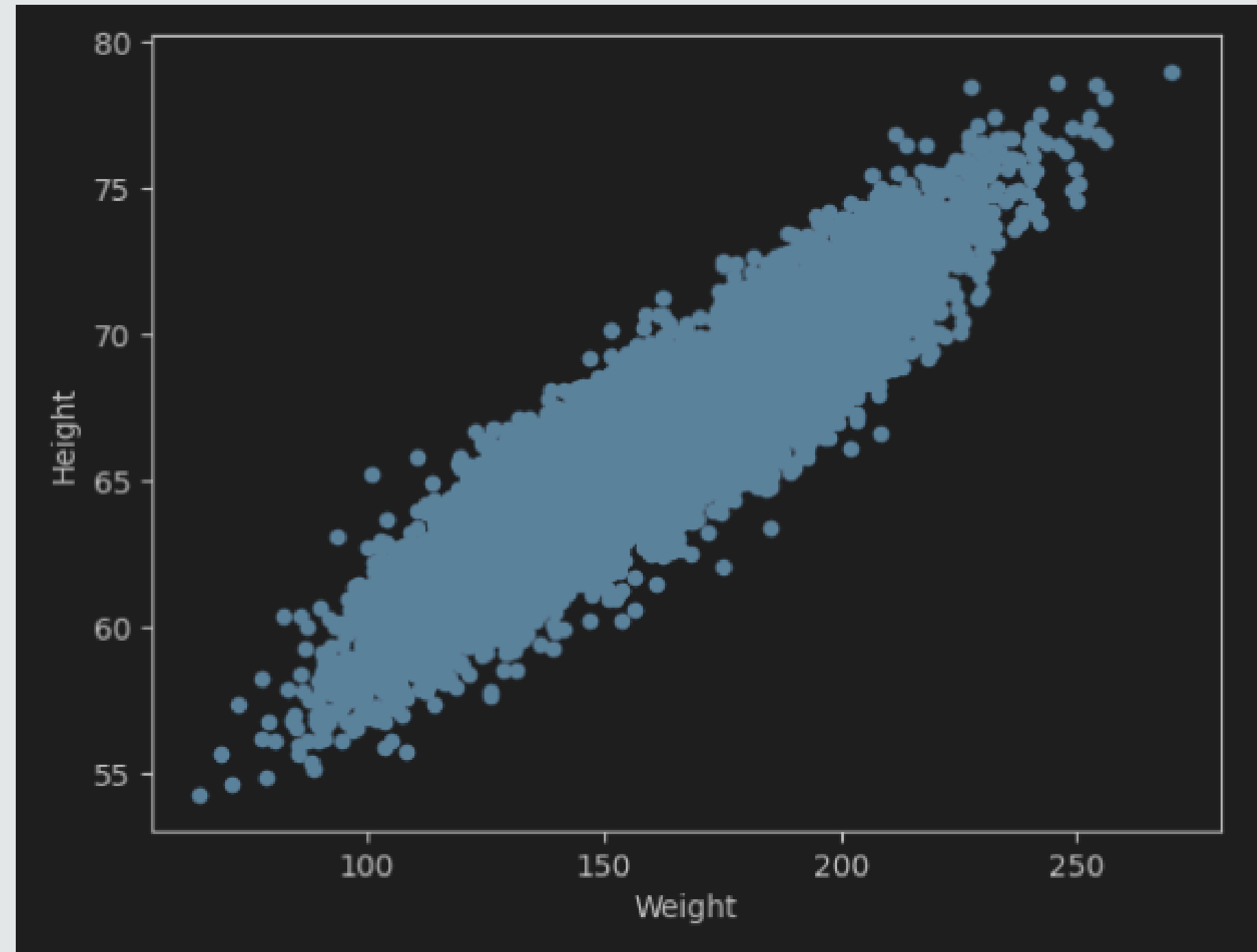


Линейная регрессия

Лебедев Марк Алексеевич 23221

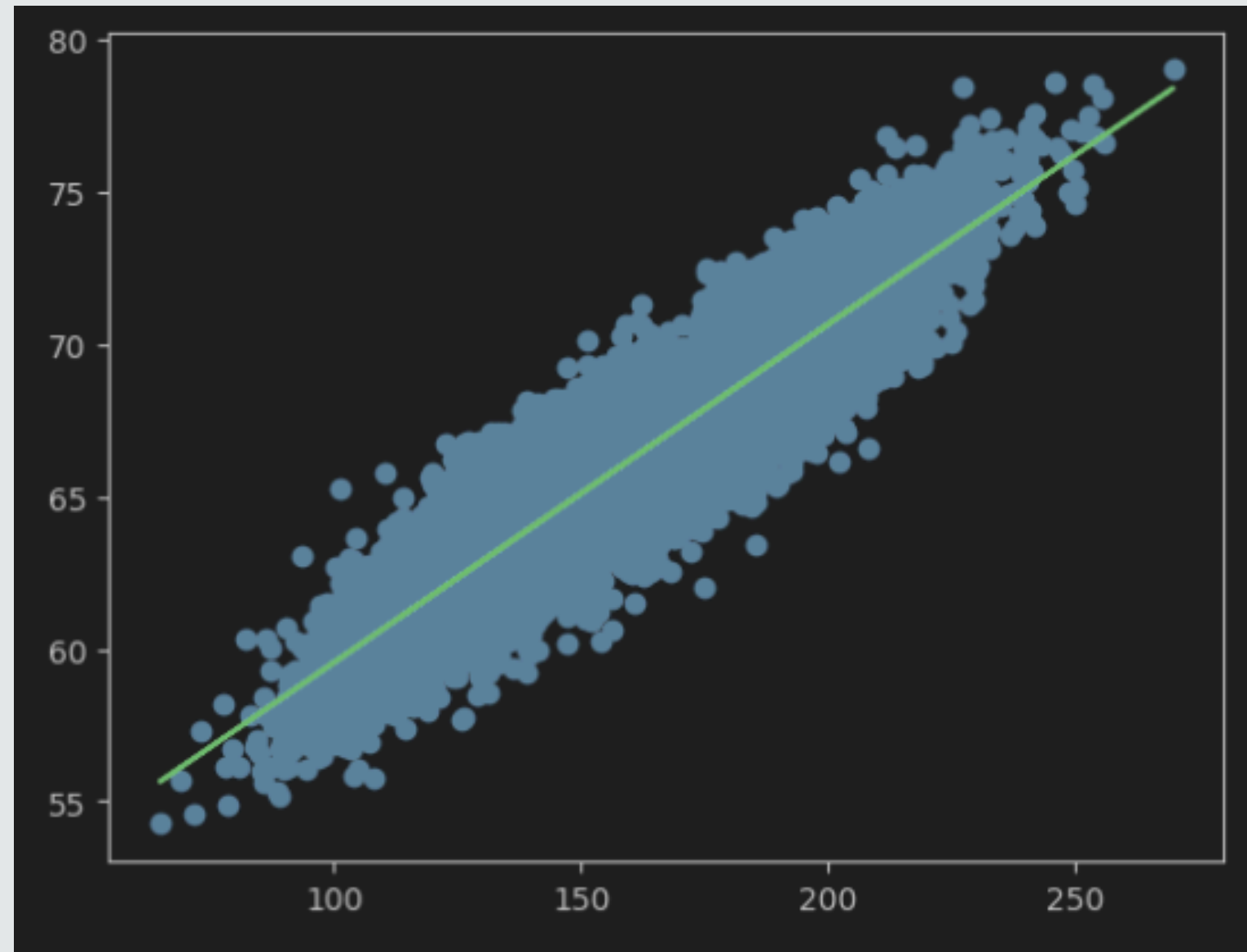
Задача

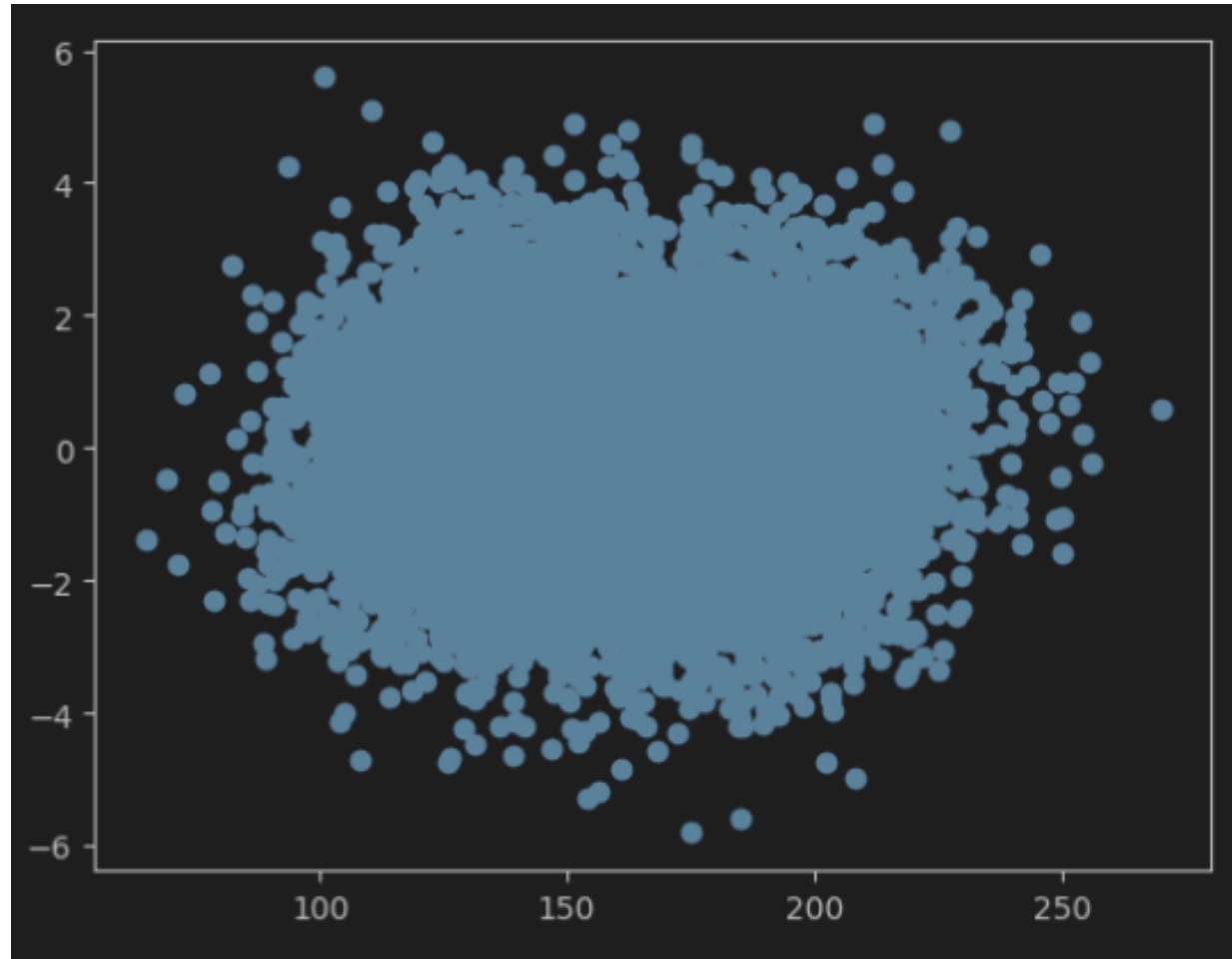
Зачем нужна линейная
регрессия?



Решение

Аппроксимируем прямой





Ошибки

Как подбирать параметры?

$$MAE = \frac{\sum_i^N |y_i - \hat{y}_i|}{N}$$

$$MSE = \frac{\sum_i^N (y_i - \hat{y}_i)^2}{N}$$

$$SSE = \sum_i^N (y_i - \hat{y}_i)^2$$

Два подхода к подбору параметров

Решение задачи минимизации ошибки



Аналитический



Градиентный спуск

Аналитический ПОДХОД

Метод наименьших квадратов

$$Q(w) = ||Xw - \hat{y}||_2^2$$

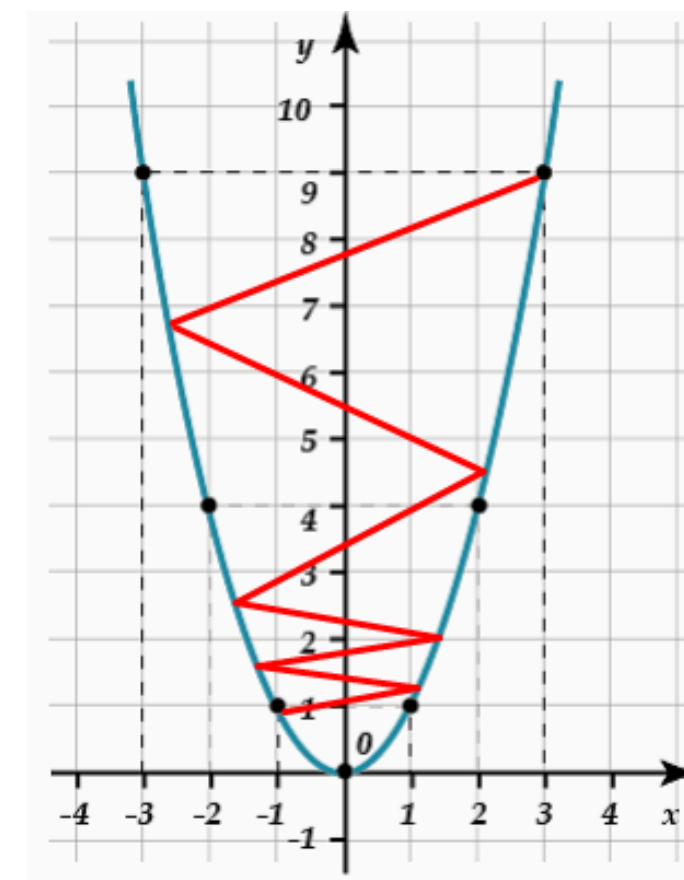
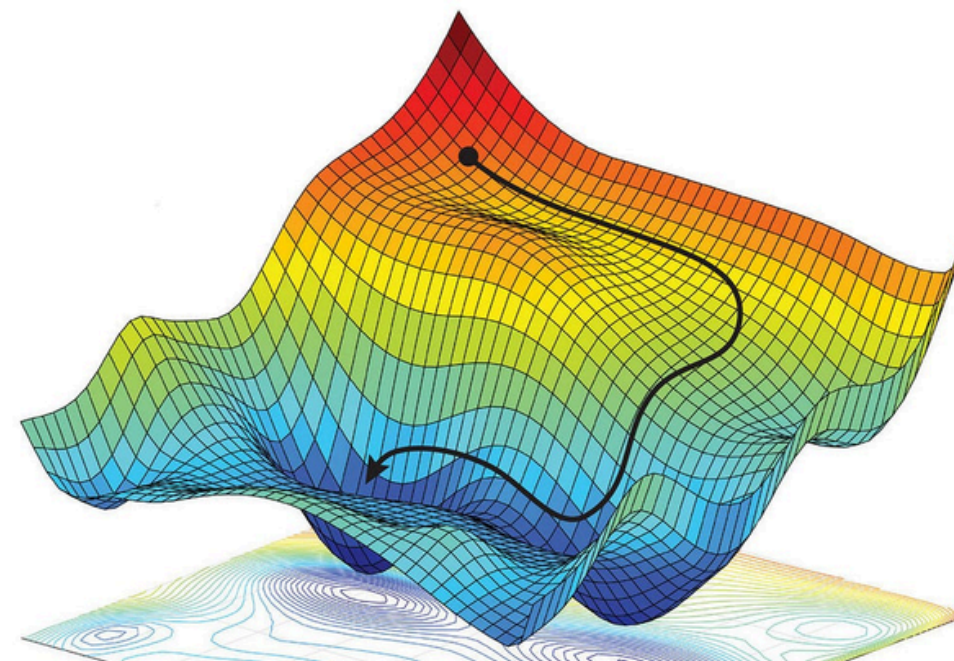


$$\hat{w} = (X^T X)^{-1} X^T \hat{y}$$

Итеративный ПОДХОД

3

Градиентный спуск



Преимущества

- Простота интерпретации
- Простота расчета

Недостатки

- Ограниченная применимость
- Большое количество проблем, которые решаются усложнением модели (коллинеарность, регуляризация, выбросы и т д)