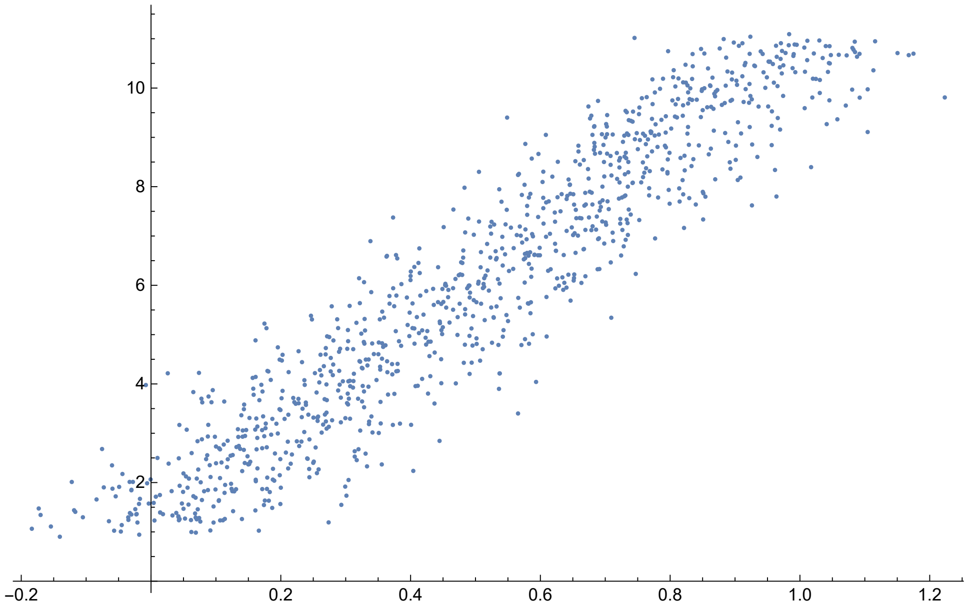
Реализация метода градиентного спуска

```
n = 1000;
data =
  Transpose[{Range[n] / n, 10 Range[n] / n + 1}] + 0.1 RandomVariate[NormalDistribution[], {n, 2}];
ListPlot[data, ImageSize → 500]
```



Порядок действий:

1. Модель линейной регрессии имеет вид:

$$y = \theta_0 + \theta_1 x.$$

Создайте две переменные X и y, где X – матрица «объекты-признаки», а y – вектор ответов.

2. Поскольку выбор из нормального закона, то наилучшей оценкой параметров θ является оценка наименьших квадратов. То есть, функционал качества имеет следующий вид:

$$Q(a,\, X) = \frac{1}{2\,l} \sum_{i=1}^{l} (y(x_i) - a(x_i))^2 \to \min_{\theta},$$

где l – число объектов обучающей выборки. Запишите в отдельную переменную Q функционал качества для выборки $(X,\, y)$. Для удобства можно записать Q в виде функции двух переменных θ_0 и θ_1.

3. Решите задачу методом градиентного спуска. Возьмите начальное приближение $\theta_0 = 0$, $\theta_1 = 0$, а размер шага градиентного спуска $\eta = 0.1$. Максимальное число итераций установите равным 1000. Точность, при которой наступает сходимость, укажите равной 10^{-5}.

4*. Отобразите на графике значение функционала качества на каждой итерации обучения. Как менялось качество предсказаний?