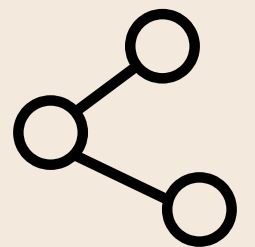


Методы численной оптимизации



**В данной презентации будут
рассмотрены методы
оптимизации, которые
используются в
машинном обучении**





Градиентный спуск

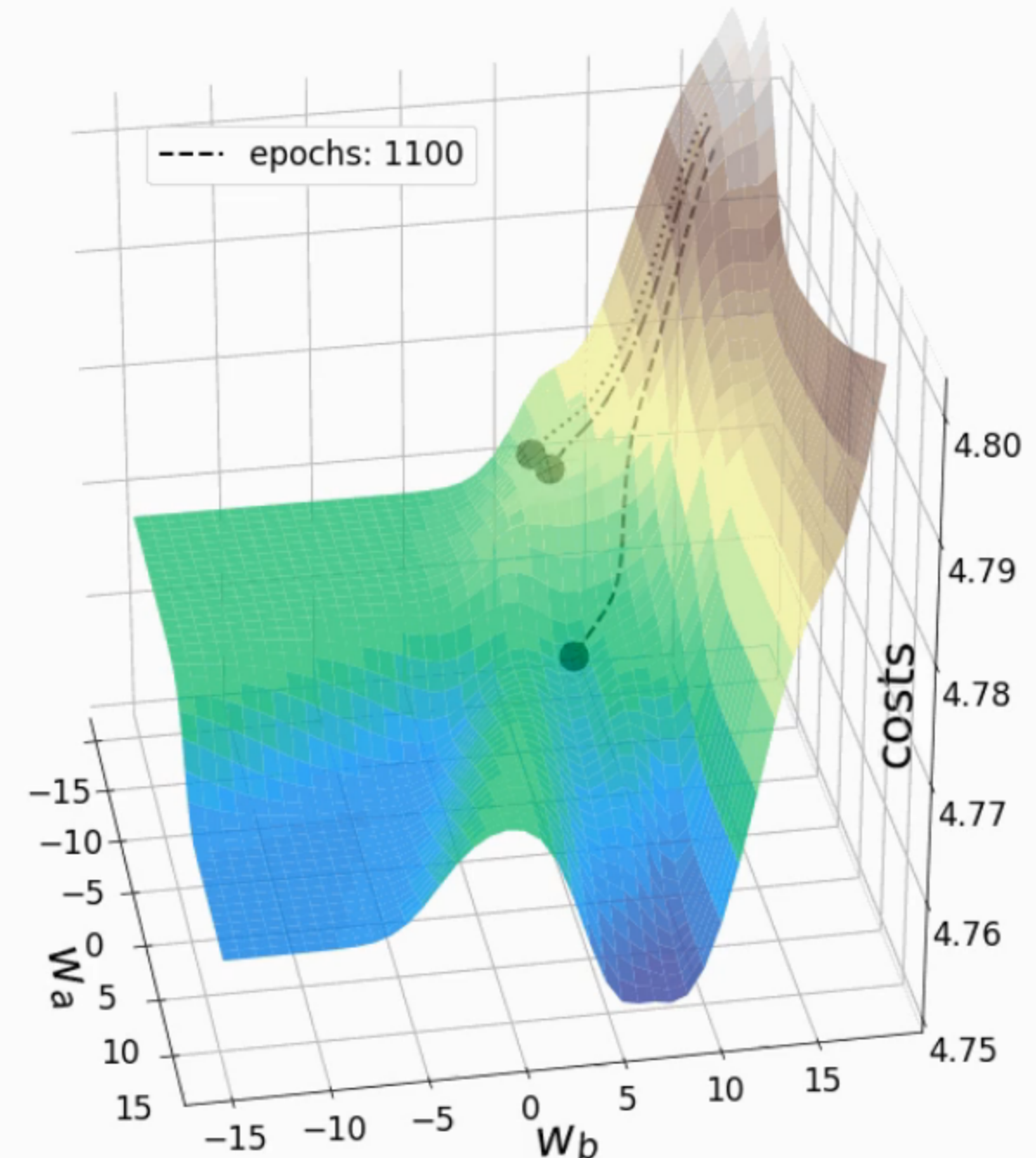
Простой метод оптимизации, позволяющий найти минимум целевой функции. Это **жадный алгоритм**, который делает шаг в сторону максимальной скорости убывания функции.

Градиент

$$\Delta w_i = -\eta * \overbrace{f'_w(w_i)}^{\text{Градиент}} + \alpha w_{i-1}$$

η - скорость обучения (шаг сходимости)

α - момент (смещение)





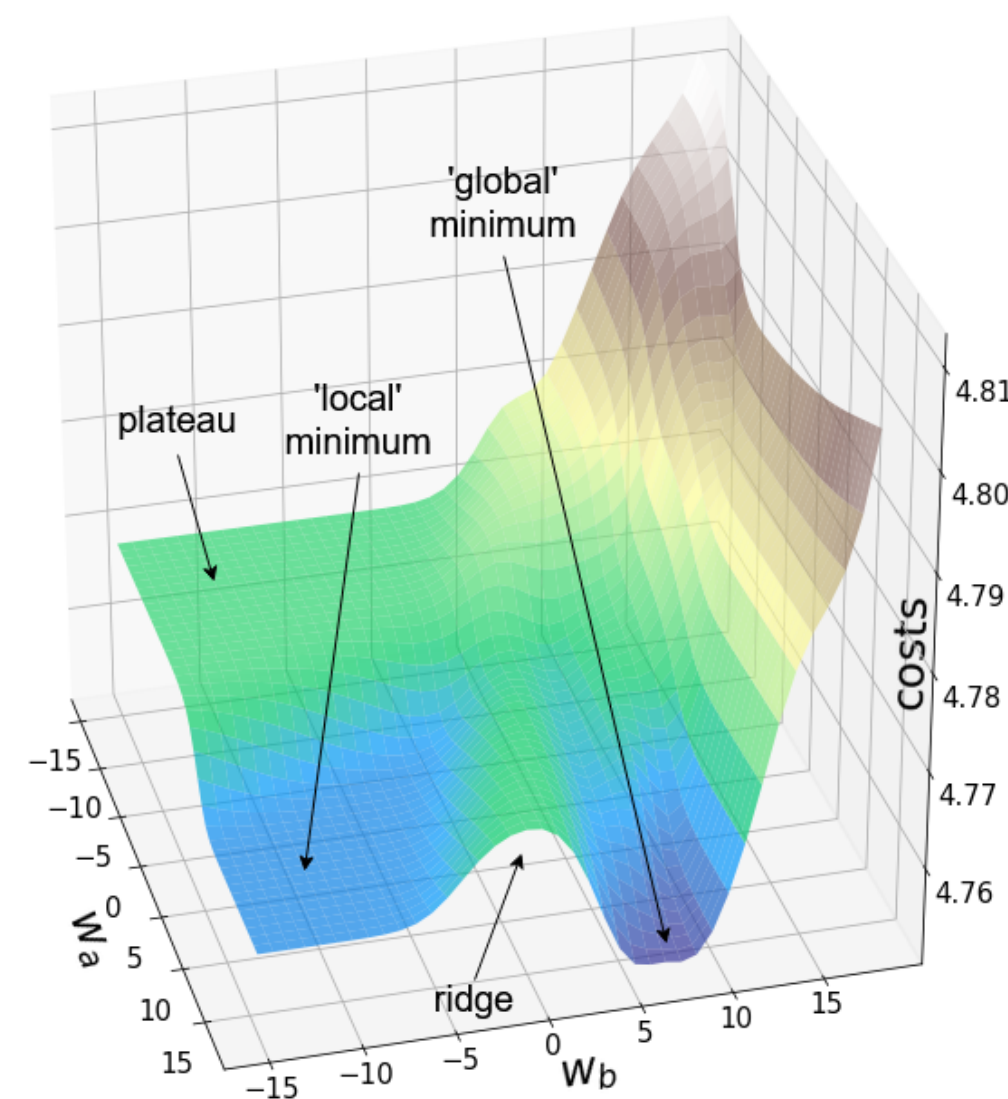
Проблемы

Ловушка локального минимума

Попадание в ловушку локального минимума, что является прямым следствием жадности этого алгоритма.

Осцилляция (плато)

Явление, которое возникает, когда значение функции не изменяется существенно независимо от направления, в котором оно движется (так называемое плато).



2 Метод Ньютона

Метод Ньютона итеративный алгоритм. В многомерном случае используется и градиент и матрица Гессе.

Одномерный случай

$$w_{i+1} = w_i - \frac{f'(w_i)}{f''(w_i)}$$

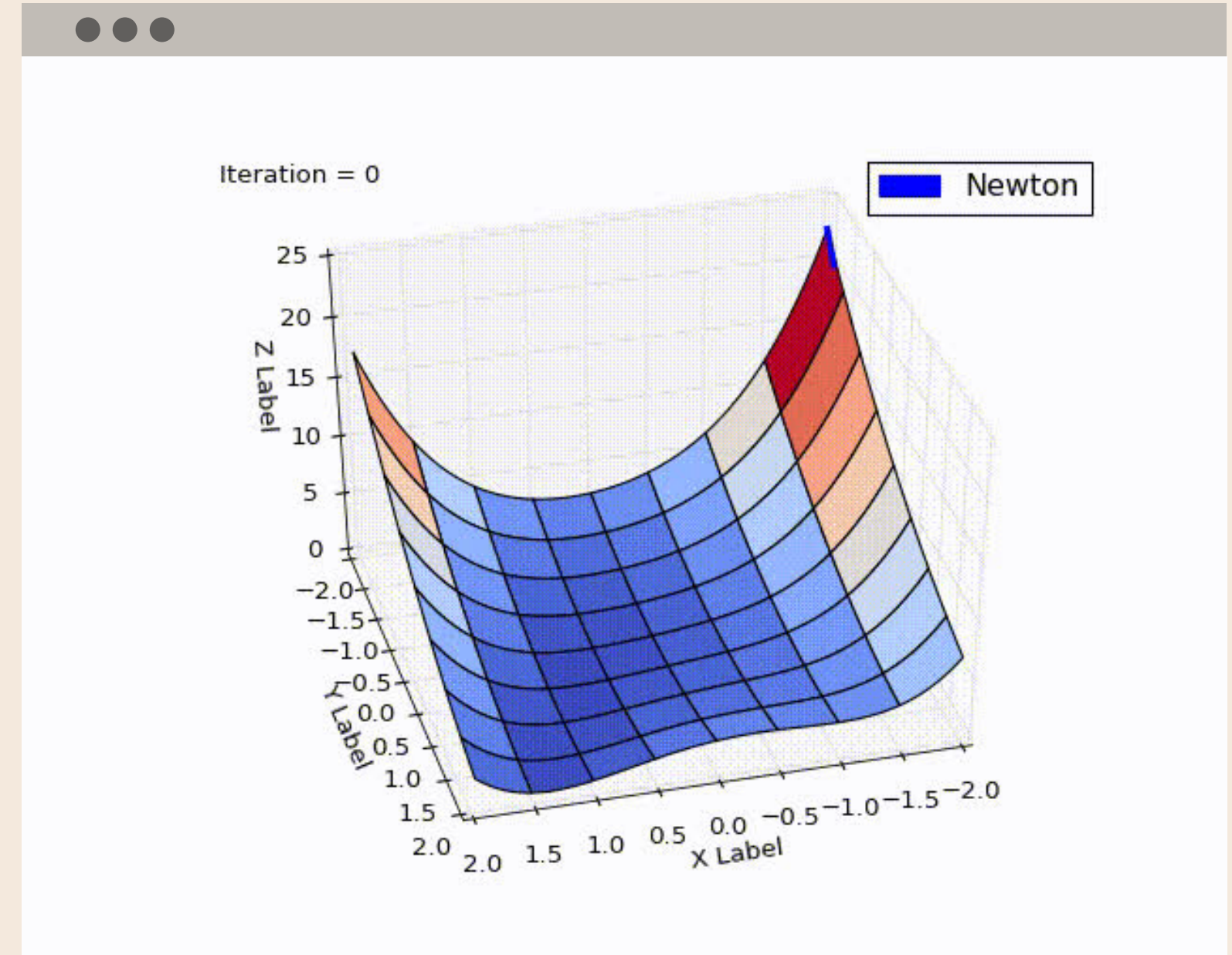
Многомерный случай

$$f'(w_i) = \nabla f(w_i) = g_i$$

$$H_i = \nabla^2 f(w_i)$$

$$H_i g_i = -g_i$$

$$w_{i+1} = w_i - H_i^{-1} g_i$$





Проблемы

Вторая производная

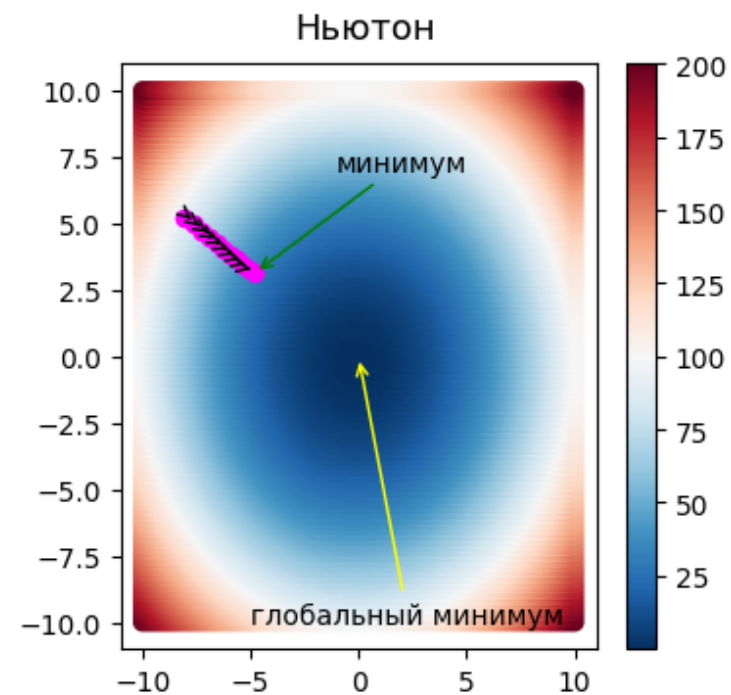
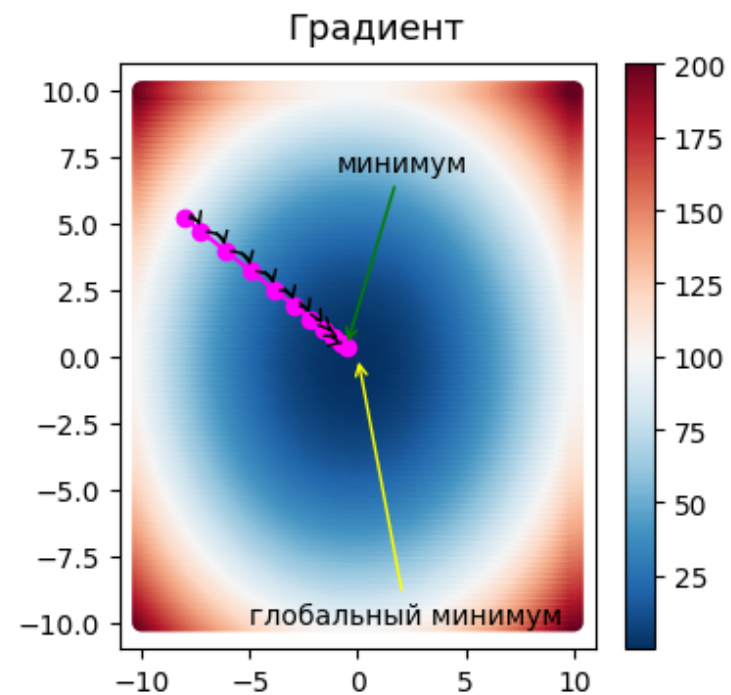
Для алгоритма требуется не только первая производная, но и вторая. Для некоторых функций это может стать проблемой.

Положительно определённая матрица

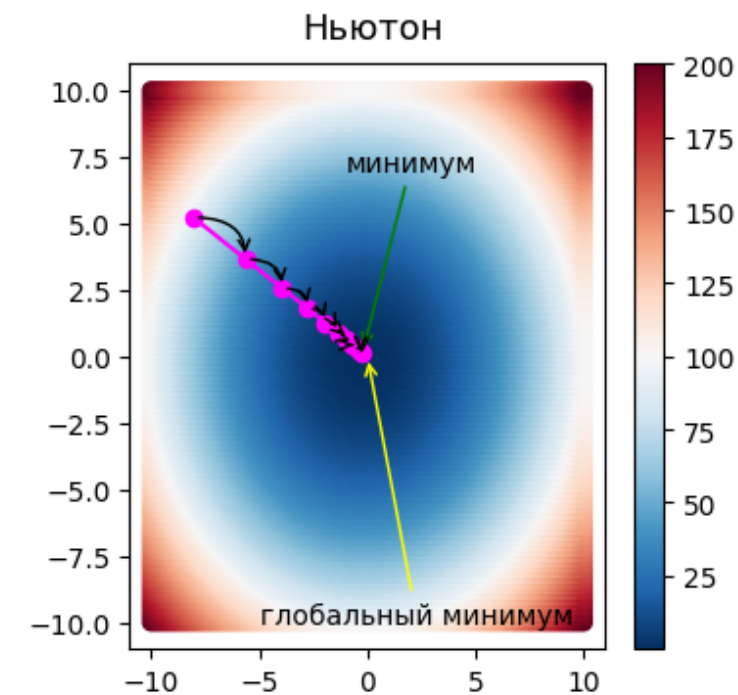
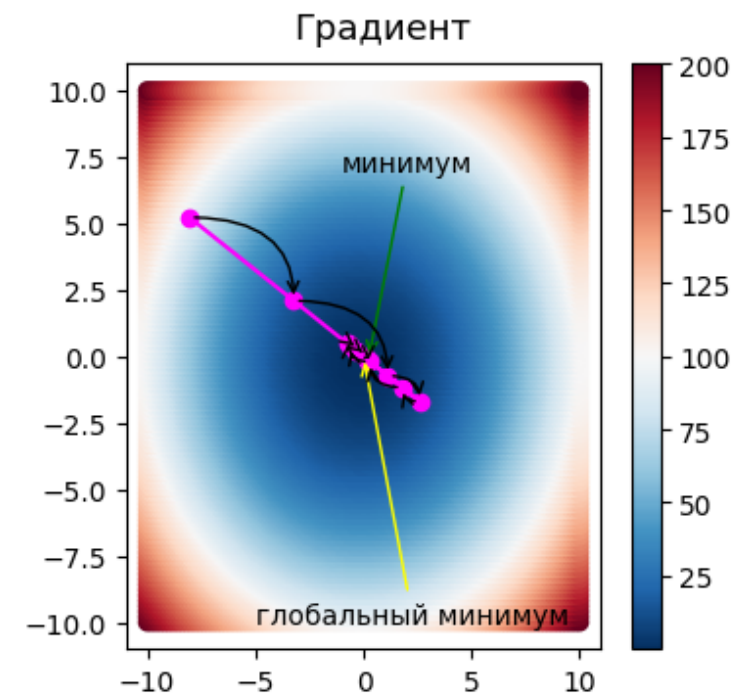
Матрица Гессе должна быть **положительно определена**. Это не обязательно, но если это условие выполнено, то метод Ньютона является методом спуска и обладает свойством глобальной, а не только локальной сходимости.

Сравним методы

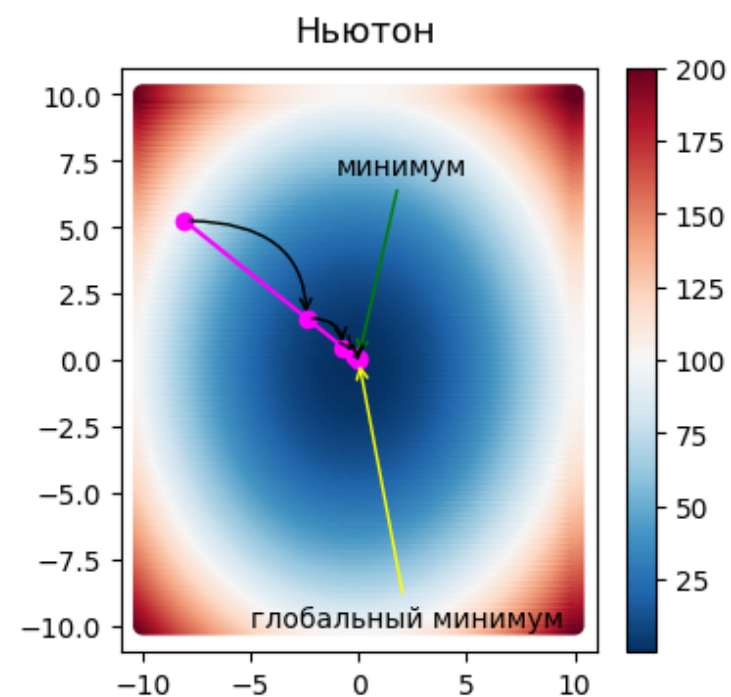
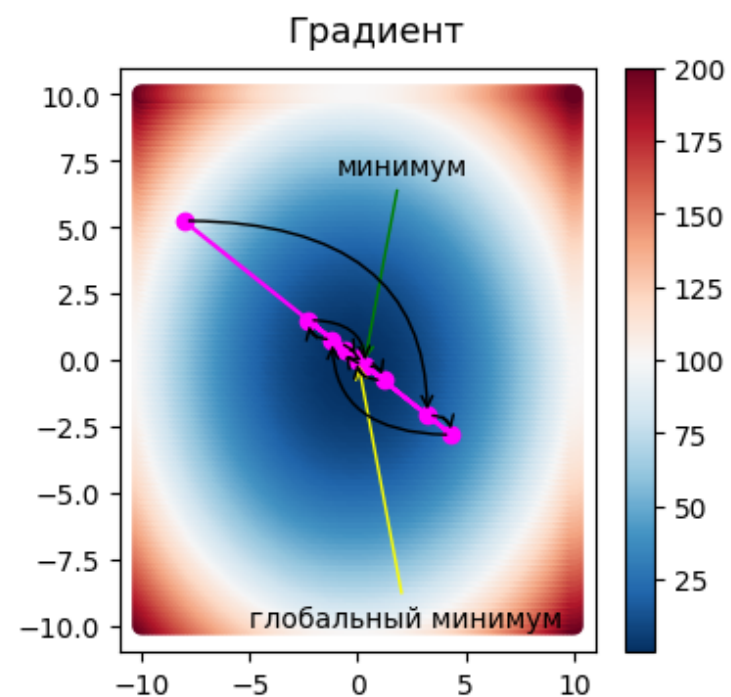
Скорость = 0.05



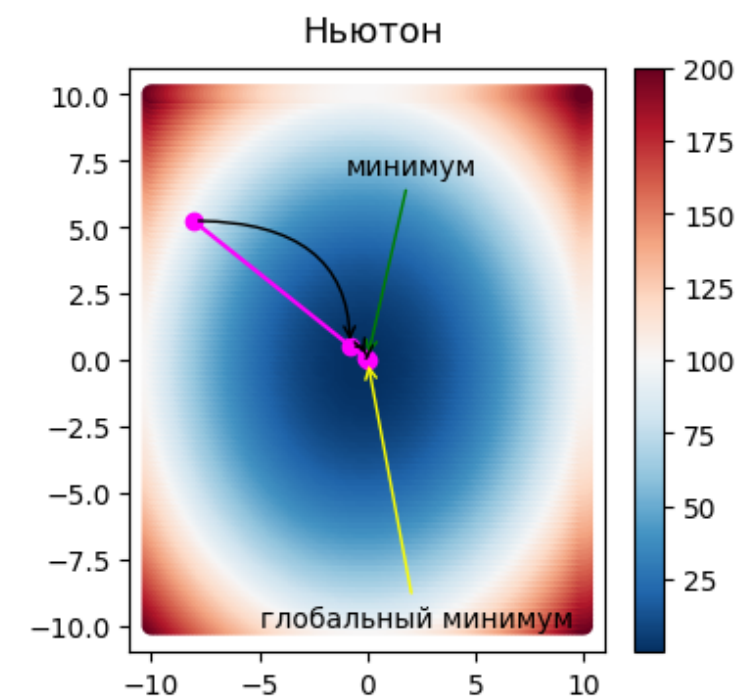
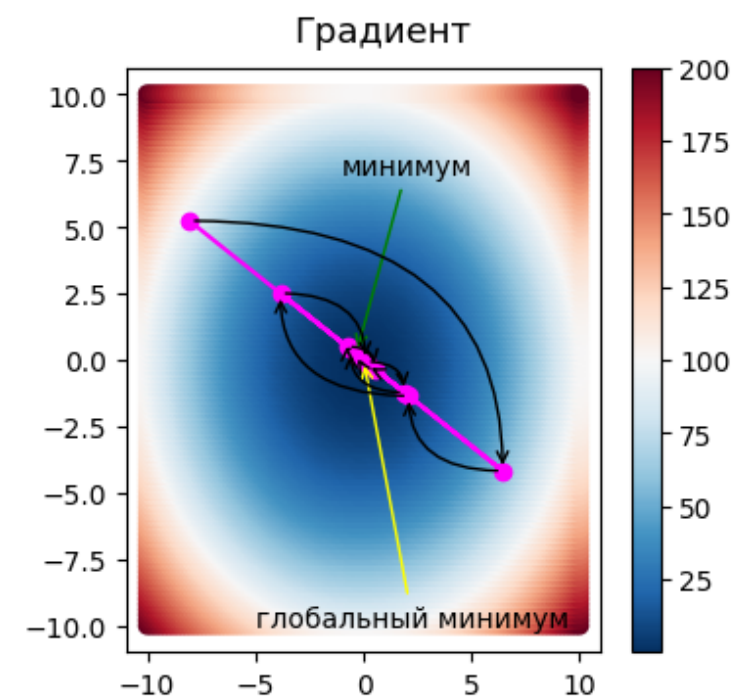
Скорость = 0.3



Скорость = 0.7

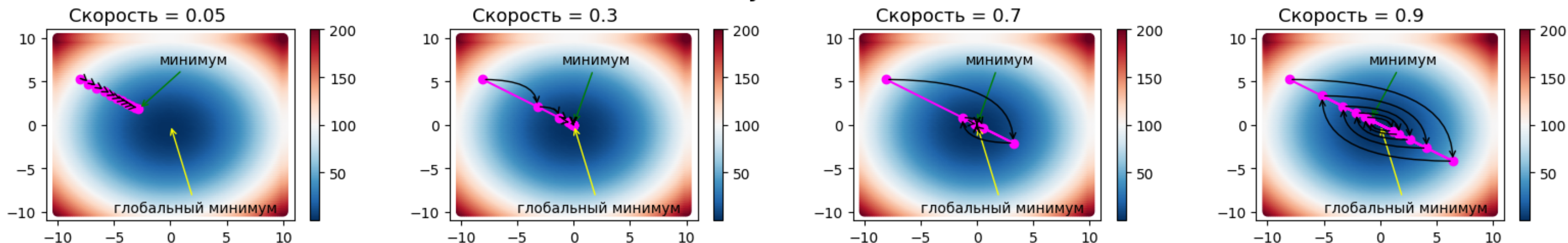


Скорость = 0.9

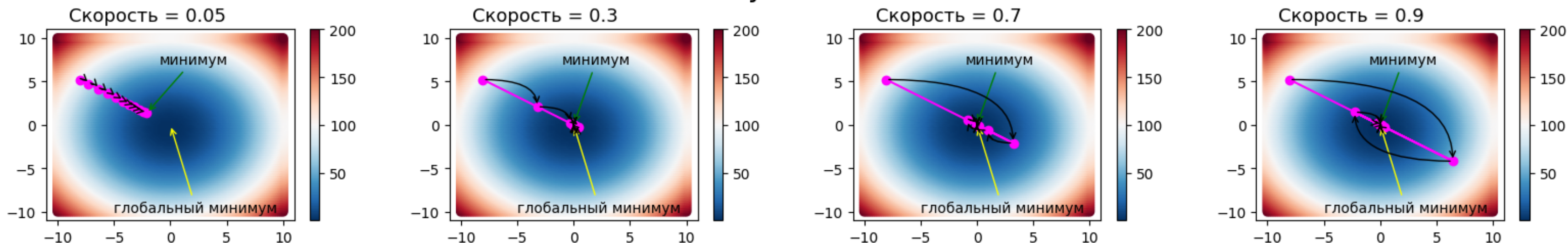


Зависимость от параметров

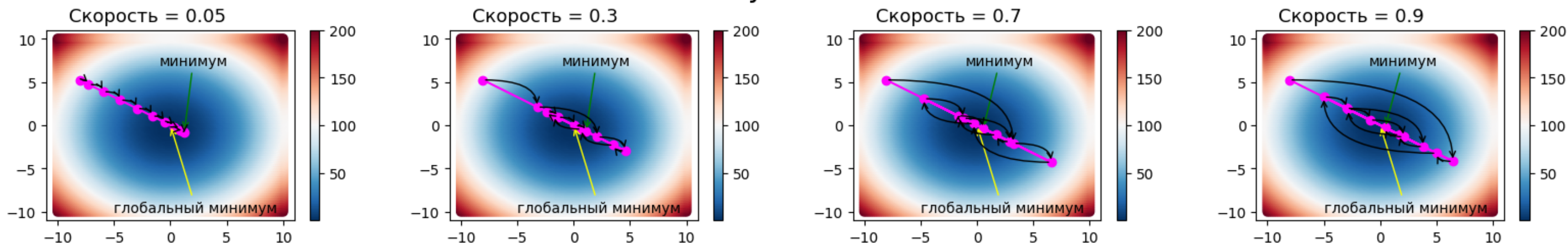
Импульс = 0



Импульс = 0.2



Импульс = 0.7



Итог

Алгоритмы, которые стоит рассмотреть

Методы линий

Одномерная минимизация

Метод наискорейшего спуска

Метод Бroyдена Флетчера Гольдфарба Шанно (BFGS)

Нелинейные сопряжённые градиенты (NCG)

Усечённый метод Ньютона (TNM)

Отсканируйте QR-код, чтобы
перейти к репозиторию

