

Отчет по первому практическому заданию

Царькова Анастасия

Формулы для градиента и гессиана функции логистической регрессии:

$$f(x) = \frac{1}{m} \sum_{i=1}^m \ln(1 + \exp(-b_i a_i^T x)) + \frac{\lambda}{2} \|x\|_2^2 = \frac{1}{m} \ln(1 + \exp(-b * Ax)) + \frac{\lambda}{2} x^T x$$

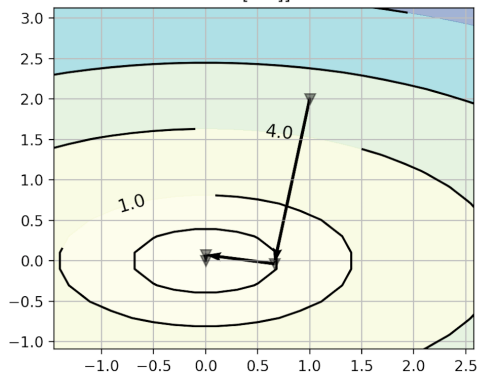
$$\nabla f(x) = \frac{1}{m} A^T \left(b * \left(\frac{1}{1 - \exp(b * Ax)} \right) \right) + \lambda x$$

$$\nabla^2 f(x) = \frac{1}{m} A^T \text{Diag} \left(\left(1 - \left(\frac{1}{1 - \exp(b * Ax)} \right) \right) * \left(\frac{1}{1 - \exp(b * Ax)} \right) \right) A + \lambda I$$

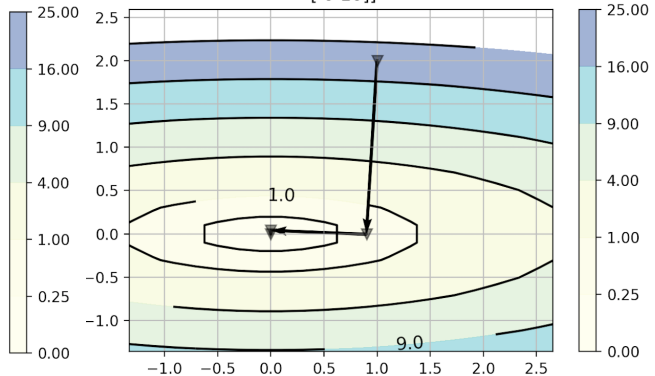
Первый эксперимент.

Зависимость поведения метода от обусловленности функции.

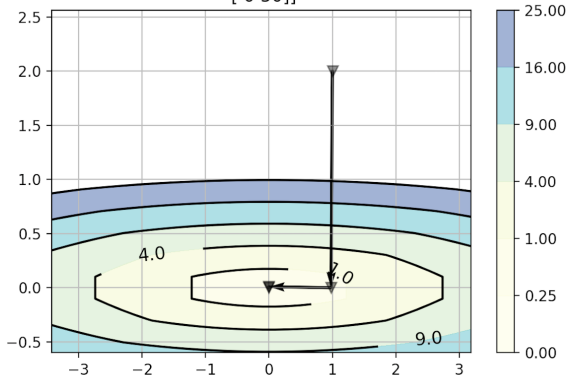
gradient descent with Wolfe linear search, steps: 3
from point [1. 2.]
for Quadratic function with
A = [[1 0]
[0 3]]



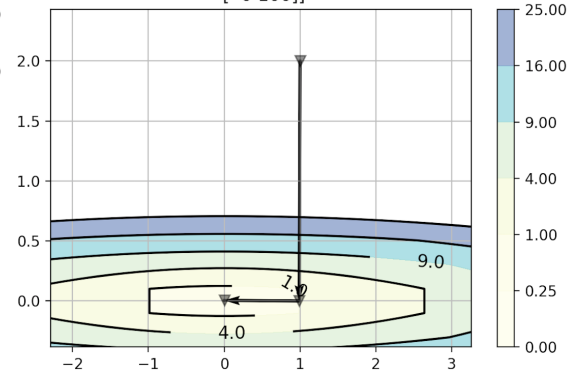
gradient descent with Wolfe linear search, steps: 3
from point [1. 2.]
for Quadratic function with
A = [[1 0]
[0 10]]



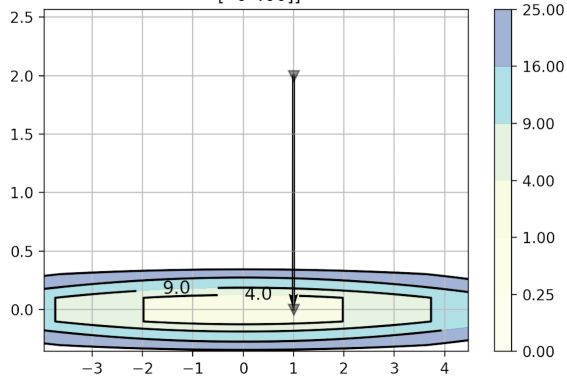
gradient descent with Wolfe linear search, steps: 3
from point [1. 2.]
for Quadratic function with
A = [[1 0]
[0 50]]



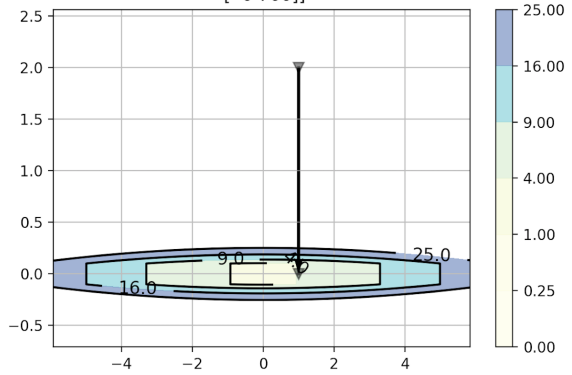
gradient descent with Wolfe linear search, steps: 2
from point [1. 2.]
for Quadratic function with
A = [[1 0]
[0 100]]



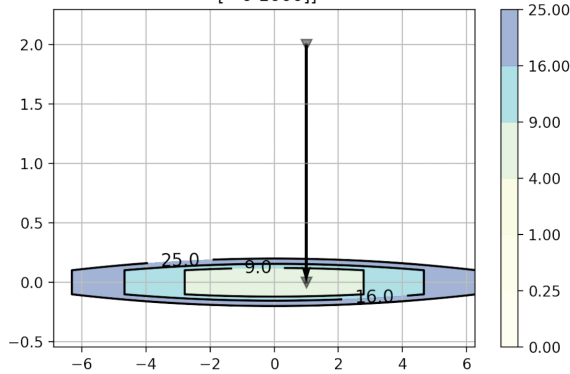
gradient descent with Wolfe linear search, steps: 1
 from point [1. 2.]
 for Quadratic function with
 $A = \begin{bmatrix} 1 & 0 \\ 0 & 400 \end{bmatrix}$



gradient descent with Wolfe linear search, steps: 1
 from point [1. 2.]
 for Quadratic function with
 $A = \begin{bmatrix} 1 & 0 \\ 0 & 700 \end{bmatrix}$

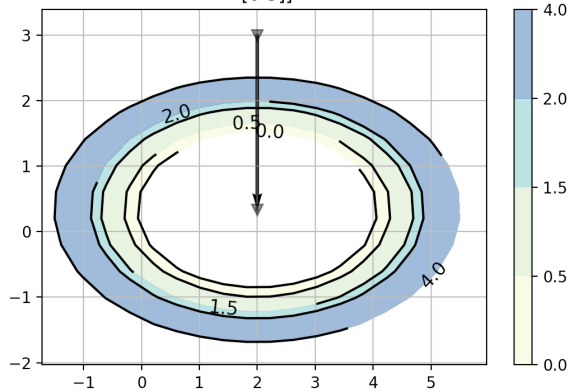


gradient descent with Wolfe linear search, steps: 1
 from point [1. 2.]
 for Quadratic function with
 $A = \begin{bmatrix} 1 & 0 \\ 0 & 1000 \end{bmatrix}$

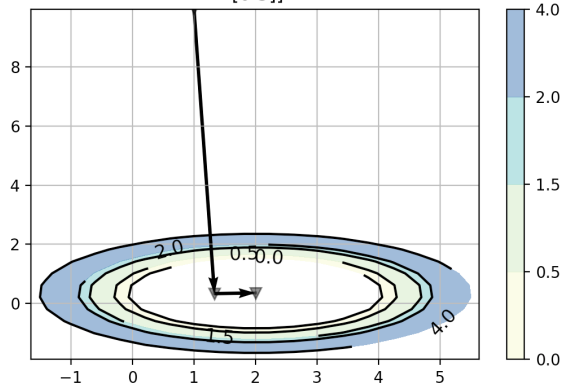


Зависимость поведения метода от начальной точки

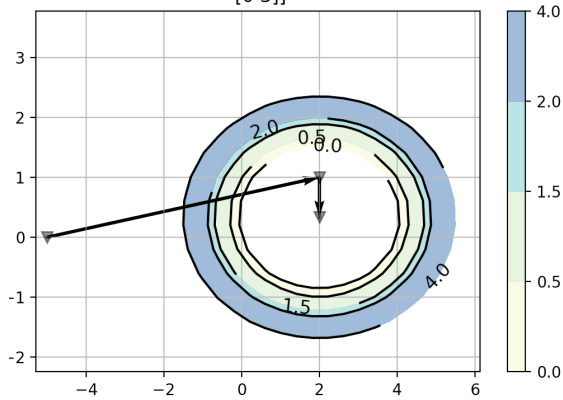
gradient descent with Wolfe linear search, steps: 1
 from point [2. 3.]
 for Quadratic function with
 $A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$



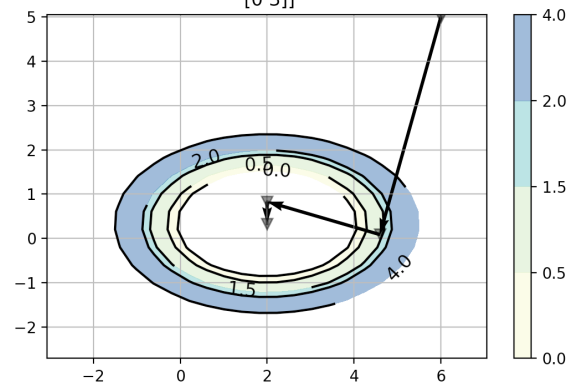
gradient descent with Wolfe linear search, steps: 2
 from point [1. 10.]
 for Quadratic function with
 $A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$



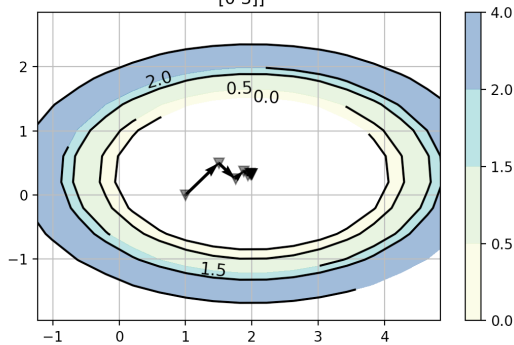
gradient descent with Wolfe linear search, steps: 2
from point [-5. 0.]
for Quadratic function with
 $A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$



gradient descent with Wolfe linear search, steps: 3
from point [6. 5.]
for Quadratic function with
 $A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$

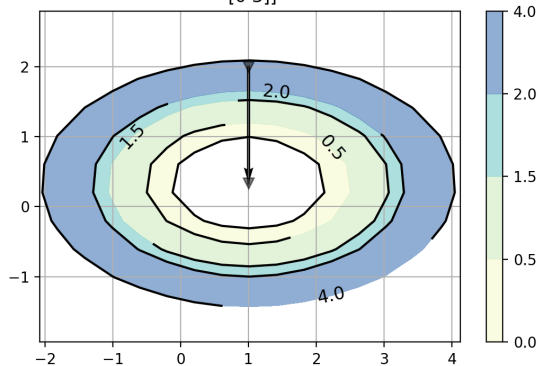


gradient descent with Wolfe linear search, steps: 9
from point [1. 0.]
for Quadratic function with
 $A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$

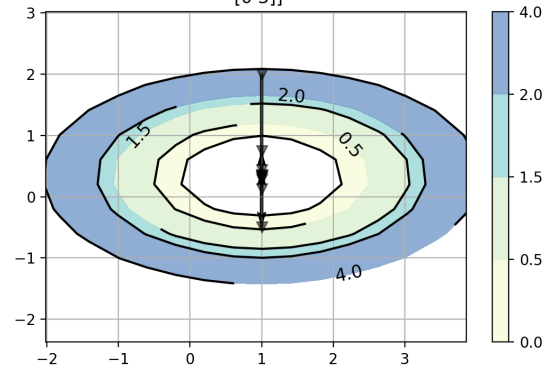


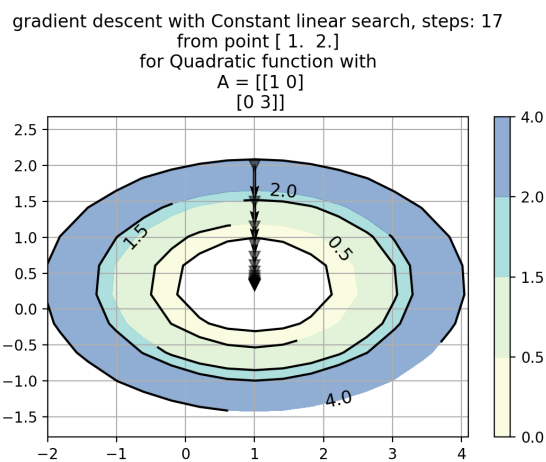
Зависимость поведения метода от стратегии выбора шага.

gradient descent with Wolfe linear search, steps: 1
from point [1. 2.]
for Quadratic function with
 $A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$

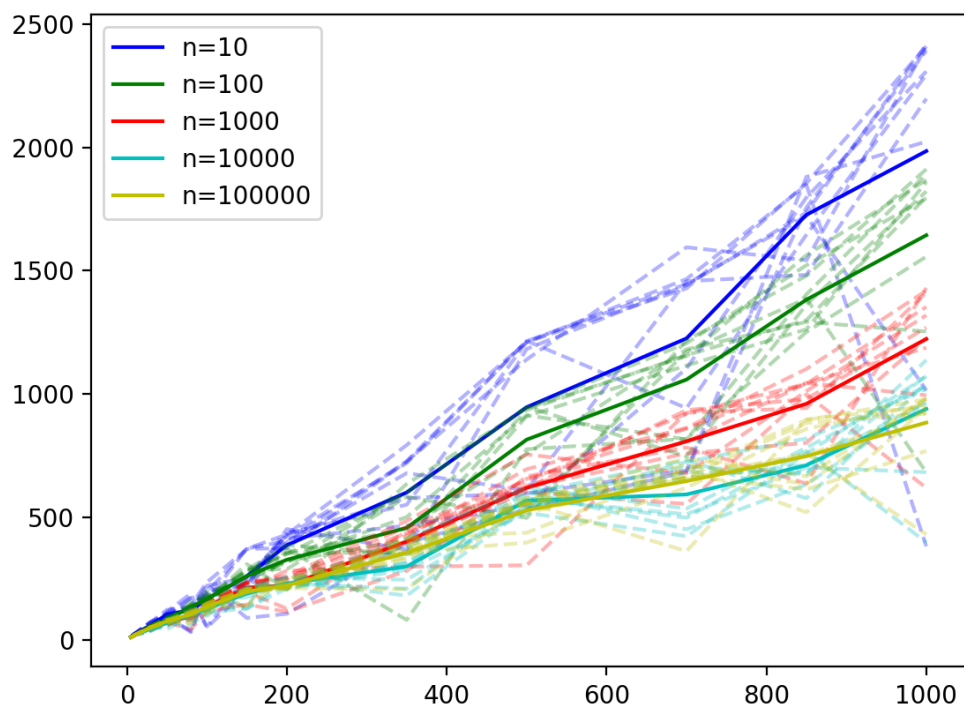


gradient descent with Armijo linear search, steps: 9
from point [1. 2.]
for Quadratic function with
 $A = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$



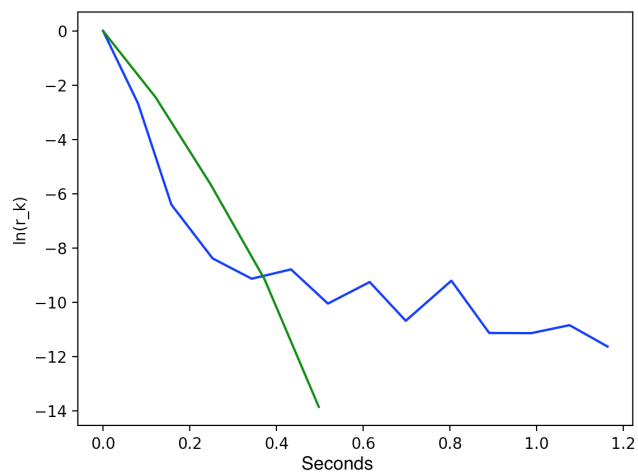
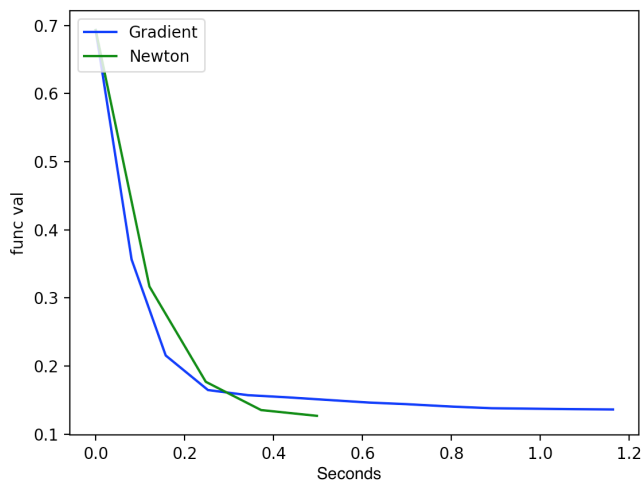


Зависимость числа итераций градиентного спуска от числа обусловленности и размерности пространства.

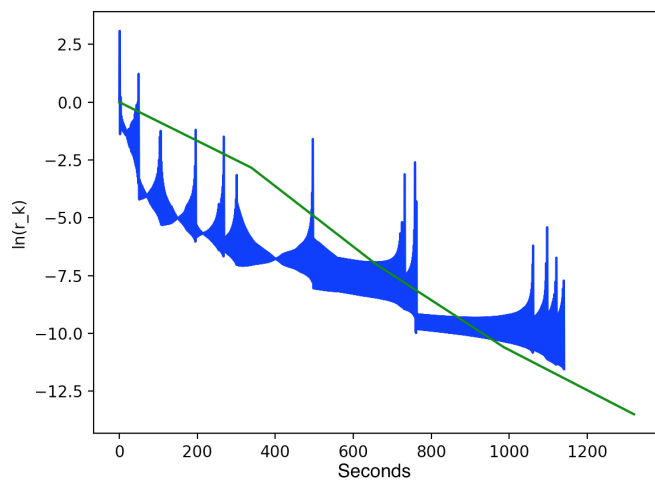
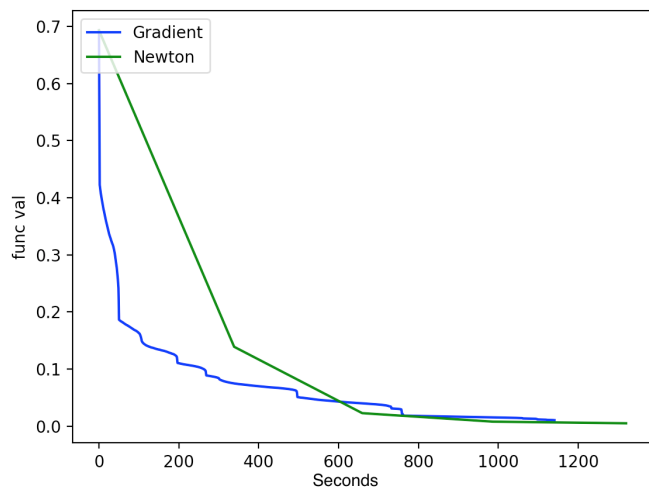


Сравнение методов градиентного спуска и Ньютона на реальной задаче логистической регрессии.

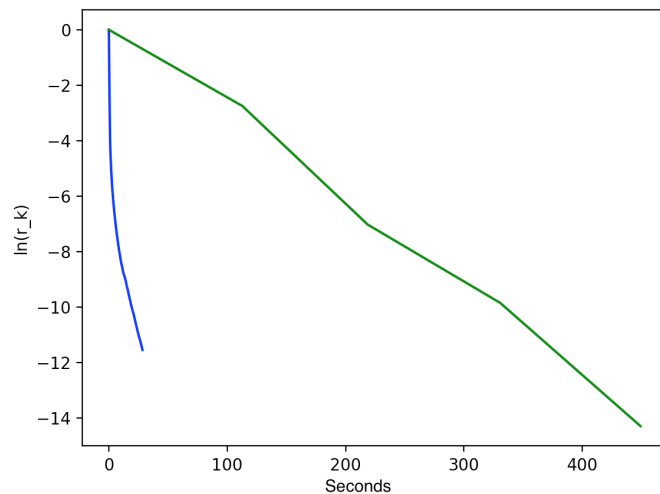
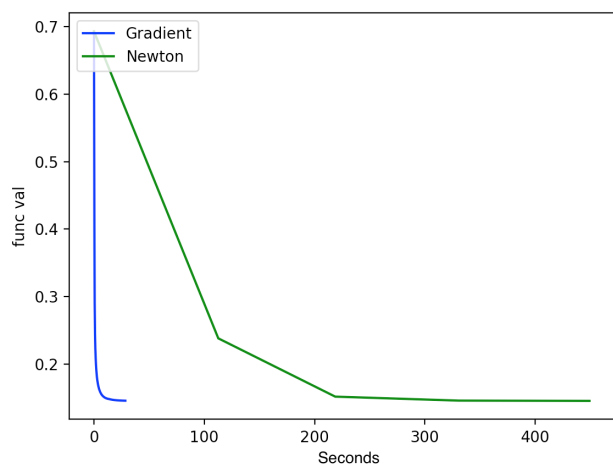
w8a



gisette-scale



real-sim



Оптимизация вычислений в градиентном спуске.

