Третьякова Д. А.

Лабораторная работа №3

Метод барьеров

1. Вспомогательная функция $f_t(x, u)$

Для задачи LASSO, минимизируемой методом барьеров, вспомогательная функция имеет вид:

$$f_t(x,u) = \frac{1}{2}|Ax - b|^2 + \lambda\langle \mathbf{1_n}, u\rangle - \ln u + x - \ln(u - x)$$

где A — матрица коэффициентов, b — вектор наблюдений, λ — коэффициент регуляризации, $\mathbf{1_n}$ — вектор единиц.

1.1 Система линейных уравнений

Ньютоновское направление $d_k = (d_k^x, d_k^u)$ задается следующей системой линейных уравнений:

$$\nabla^2 f_t(x_k, u_k) d_k = -\nabla f_t(x_k, u_k)$$

где градиенты и Гессиан вычисляются следующим образом:

Градиенты:

$$\nabla_x f_t(x,u) = A^T(Ax - b) - \frac{1}{u+x} + \frac{1}{u-x}$$
$$\nabla_u f_t(x,u) = \lambda \mathbf{1_n} - \frac{1}{u+x} - \frac{1}{u-x}$$

Гессиан:

$$\nabla^2 f_t(x,u) = \begin{pmatrix} \nabla^2_{xx} f_t(x,u) & \nabla^2_{xu} f_t(x,u) \\ \nabla^2_{ux} f_t(x,u) & \nabla^2_{uu} f_t(x,u) \end{pmatrix}$$

где

$$\nabla^2_{xx} f_t(x,u) = A^T A + \mathrm{Diag}\left(\frac{1}{(u+x)^2} + \frac{1}{(u-x)^2}\right)$$
$$\nabla^2_{uu} f_t(x,u) = \mathrm{Diag}\left(\frac{1}{(u+x)^2} + \frac{1}{(u-x)^2}\right)$$

$$\nabla^2_{xu} f_t(x,u) = \mathrm{Diag}\left(\frac{1}{(u+x)^2} - \frac{1}{(u-x)^2}\right)$$

1.2 Максимальная допустимая длина шага α

Максимальная допустимая длина шага определяется с учетом ограничений задачи. Для текущей точки (x_k, u_k) и направления d_k ограничения могут быть заданы как $|x| \leq u$.

Максимальная длина шага α^{max} определяется следующим образом:

1. Если $x_k \geq 0$:

$$\alpha \leq \frac{u_k - x_k}{d_k^x - d_k^u}, \quad \text{при } d_k^x - d_k^u > 0$$

2. Если $x_k < 0$:

$$\alpha \leq -\frac{u_k + x_k}{d_k^x + d_k^u}, \quad \text{при } d_k^x + d_k^u < 0$$

В результате, выбираем $\alpha < \alpha^{max}$ и начинаем с $\alpha^0 = \min(1{,}0.99 \cdot \alpha^{max})$.

1.3 Начальная точка (x_0, u_0)

Для метода барьеров можно выбрать начальную точку (x_0, u_0), удовлетворяющую ограничениям:

$$Q = \{(x,u) \in R^n : |x| - u \leq 0\}$$

Рекомендуется выбирать (x_0, u_0) в пределах области допустимых решений, например, с маленькими положительными значениями для u_0 и x_0, чтобы обеспечить соблюдение ограничений.

2. Эксперимент *a*. Варьирование γ и ε_{inner}

В этом эксперименте исследуется влияние параметров γ и ε_{inner} на эффективность метода.

Мы используем следующие параметры:

$$n = 300$$
$$m = 600$$
$$\lambda = 1 \times 10^{-4}$$

Элементы матрицы A и вектора b создаются на основе стандартного нормального распределения.

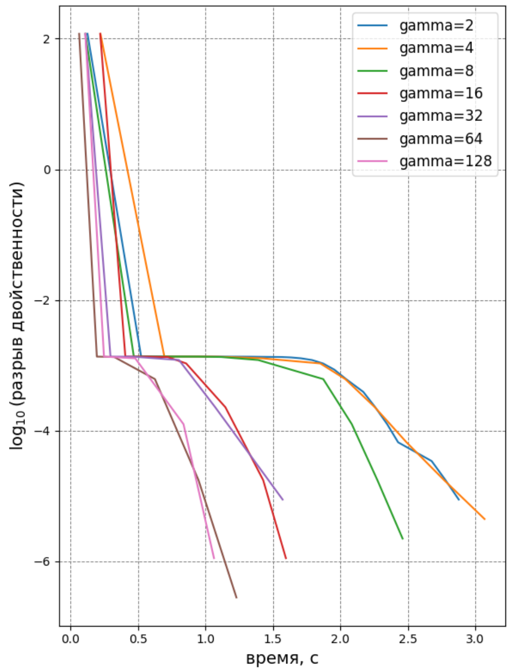
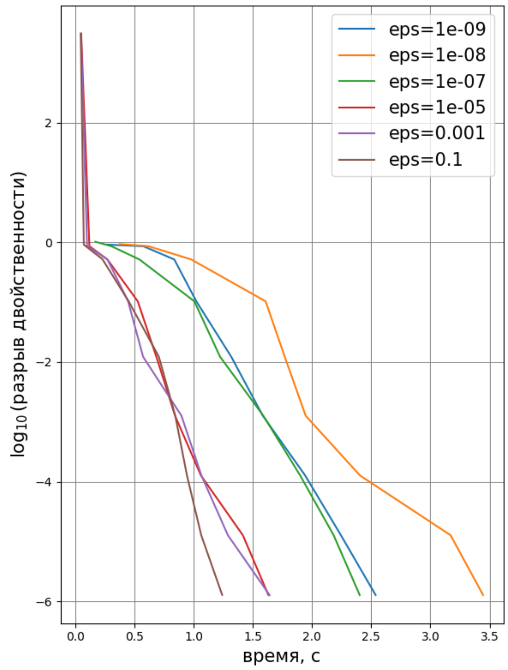
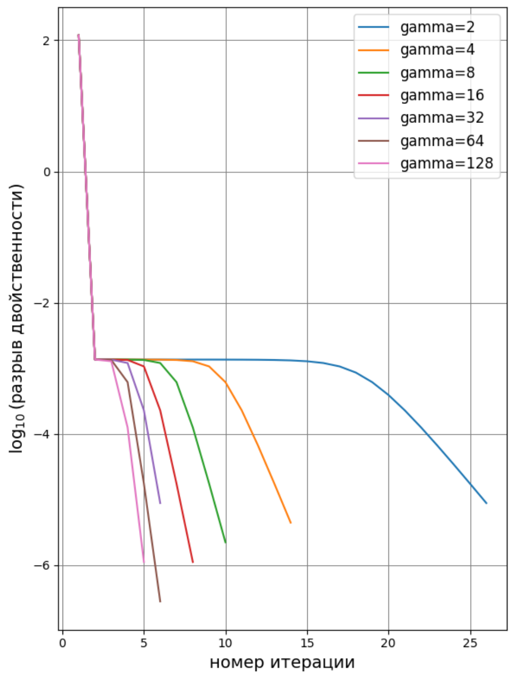
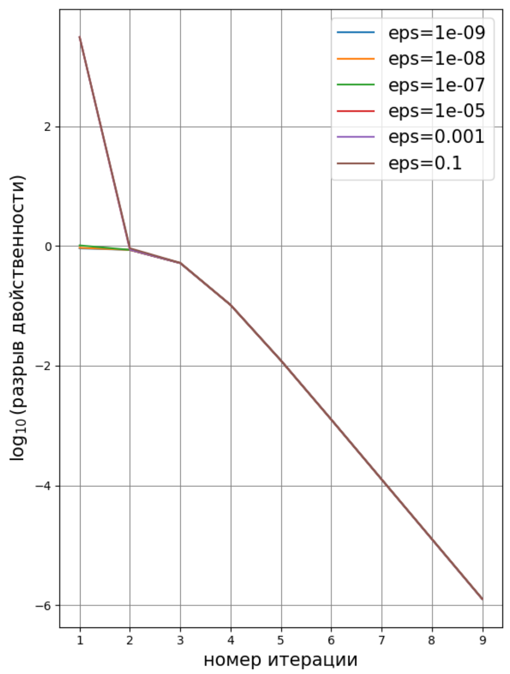
В качестве метода линейного поиска для метода Ньютона применяется метод Армихо с бектрекингом, где параметр равен $\frac{1}{2}$ и коэффициент $c_1 = 1 \times 10^{-4}$.

Для эксперимента выбираем значения

$$\gamma \in (2, 4, 8, 16, 32, 64, 128)$$

и

$$\varepsilon_{inner} \in (1 \times 10^{-9}, 1 \times 10^{-8}, 1 \times 10^{-7}, 1 \times 10^{-6}, 1 \times 10^{-5}, 1 \times 10^{-4}, 1 \times 10^{-3}, 1 \times 10^{-2}, 1 \times 10^{-1})$$

<table>
<tr><td rowspan="2"></td><td colspan="2">Зазор двойственности</td></tr>
<tr><td>Параметр γ</td><td>Параметр ε_{inner}</td></tr>
<tr><td>Зависимость от времени</td><td>
</td><td>
</td></tr>
<tr><td>Зависимость от итерации</td><td>
</td><td>
</td></tr>
</table>

Увеличение параметра γ значительно ускоряет сходимость метода, как в плане реального времени, так и по количеству итераций. Это свидетельствует о том, что более высокие значения γ способствуют более быстрому достижению результата.

Влияние ε_{inner} на время выполнения метода заметно, однако это значение не влияет на количество итераций, необходимых для достижения требуемой точности.

3. Эксперимент *b*. Варьирование n, m, λ

В этом разделе мы исследуем чувствительность метода к размерности задачи n, количеству наблюдений m и коэффициенту регуляризации λ.

Для анализа фиксируем

$$\varepsilon = 1 \times 10^{-5}$$
$$\varepsilon_{inner} = 1 \times 10^{-8}$$
$$\gamma = 10$$

Линейный поиск выполняется методом Армихо.

Анализ влияния размерности n:

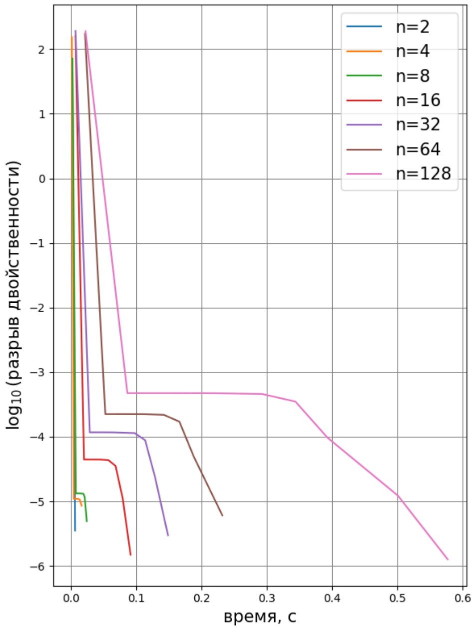
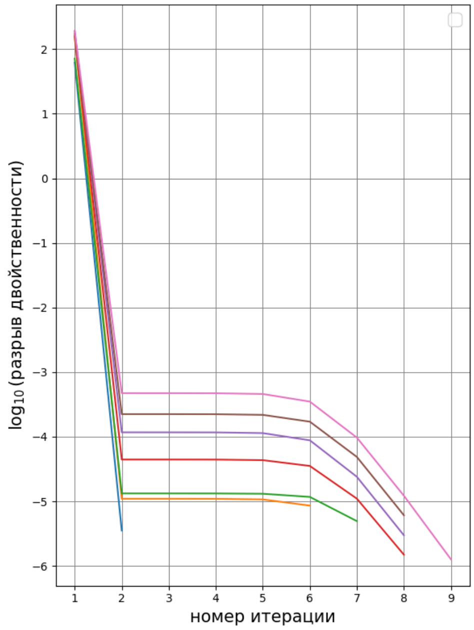
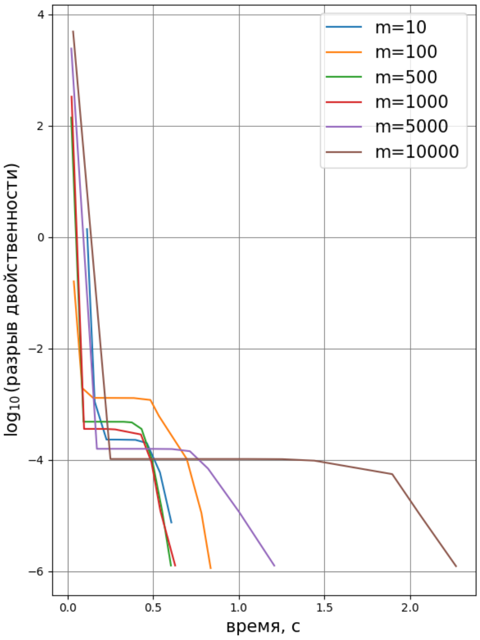
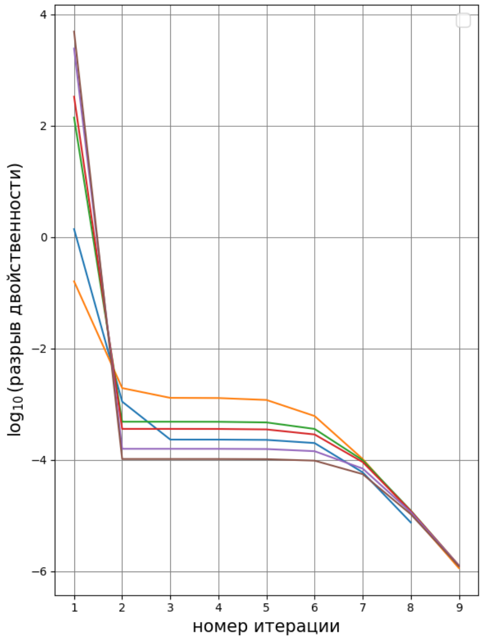
$$m = 600$$
$$\lambda = 1 \times 10^{-4}$$
$$n \in (2, 4, 8, 16, 32, 64, 128)$$

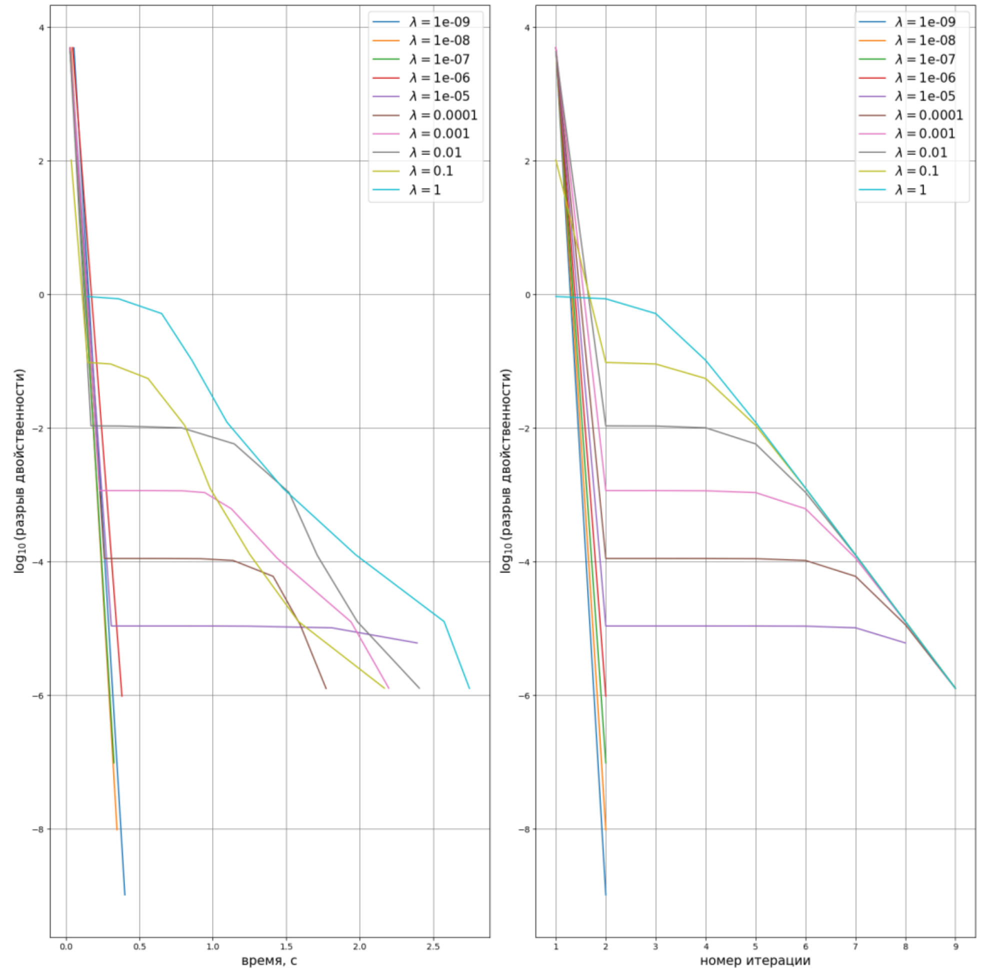
Влияние количества наблюдений m:

$$n = 300$$
$$\lambda = 1 \times 10^{-4}$$
$$m \in (10, 100, 500, 1000, 5000, 10000)$$

Влияние коэффициента регуляризации λ:

$$n = 300$$
$$m = 600$$
$$\lambda \in (1 \times 10^{-9}, 1 \times 10^{-8}, 1 \times 10^{-7}, 1 \times 10^{-6}, 1 \times 10^{-5}, 1 \times 10^{-4}, 1 \times 10^{-3}, 1 \times 10^{-2}, 1 \times 10^{-1}, 1)$$

Параметр	Зависимость зазора двойственности	
	От времени	От итерации
n		
m		

λ	

С увеличением размерности n наблюдается резкий рост количества необходимых итераций, однако при больших n вариация итераций становится незначительной (от 7 до 9). Размерность также значительно влияет на время выполнения: чем больше n, тем больше время работы метода.

Количество наблюдений m не влияет на количество итераций, необходимых для сходимости, но время работы метода возрастает с увеличением m, аналогично влиянию размерности.

При очень малом λ метод сходится быстро (за две итерации) и требует минимального времени. С увеличением λ количество итераций увеличивается, но при этом оно становится независимым от значения λ.