Ханмурзин Алим Ринатович. МОАД

Лабораторная работа 3: Метод барьеров.

1. Метод Барьеров в LASSO

Вспомогательная функция:

$$f_t(x,u) = t * \left(\frac{1}{2}\|Ax-b\|_2^2 + \lambda(1_n, u)\right) - \sum(\log(u_i + x_i) + \log(u_i - x_i))$$

Найдем градиент по x и u:

$$\nabla_x f_t(x,u) = t * A^T(Ax-b) - \frac{1}{u+x} + \frac{1}{u-x}$$

$$\nabla_u f_t(x,u) = t * \lambda * 1_n - \frac{1}{u+x} - \frac{1}{u-x}$$

Напишем гессиан:

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

$$\nabla_{xx}^2 f_t = t * A^T A + diag\left(\frac{1}{(u+x)^2} + \frac{1}{(u-x)^2}\right)$$

$$\nabla_{uu}^2 f_t = diag\left(\frac{1}{(u+x)^2} + \frac{1}{(u-x)^2}\right)$$

$$\nabla_{xu}^2 f_t = diag\left(\frac{1}{(u+x)^2} - \frac{1}{(u-x)^2}\right)$$

Методом Ньютона решаем систему:

$$\begin{pmatrix} \nabla_{xx}^2 f_t & \nabla_{xu}^2 f_t \\ \nabla_{xu}^2 f_t & \nabla_{uu}^2 f_t \end{pmatrix} \begin{pmatrix} d_k^x \\ d_k^u \end{pmatrix} = \begin{pmatrix} -\nabla_x f_t(x,u) \\ -\nabla_u f_t(x,u) \end{pmatrix}$$

Нахождение максимального допустимого α:

1) Должно выполняться ограничение $-u_i \le x_i \le u_i$
2) Если $d_k^{x_i} > d_k^{u_i}$, то $\alpha \le 0.99 * \frac{(u_i - x_i)}{d_k^{x_i} - d_k^{u_i}}$
3) Если $d_k^{x_i} < -d_k^{u_i}$, то $\alpha \le -0.99 * \frac{(u_i + x_i)}{d_k^{x_i} + d_k^{u_i}}$

Начальная точка для метода барьеров:

1) Для нулевой итерации требуется, чтобы начальная точка была внутренней
2) В остальном же можем брать любые, подходящие под ограничения.

2.1 Эксперимент: Чувствительность к выбору параметров γ и ϵ_{inner}

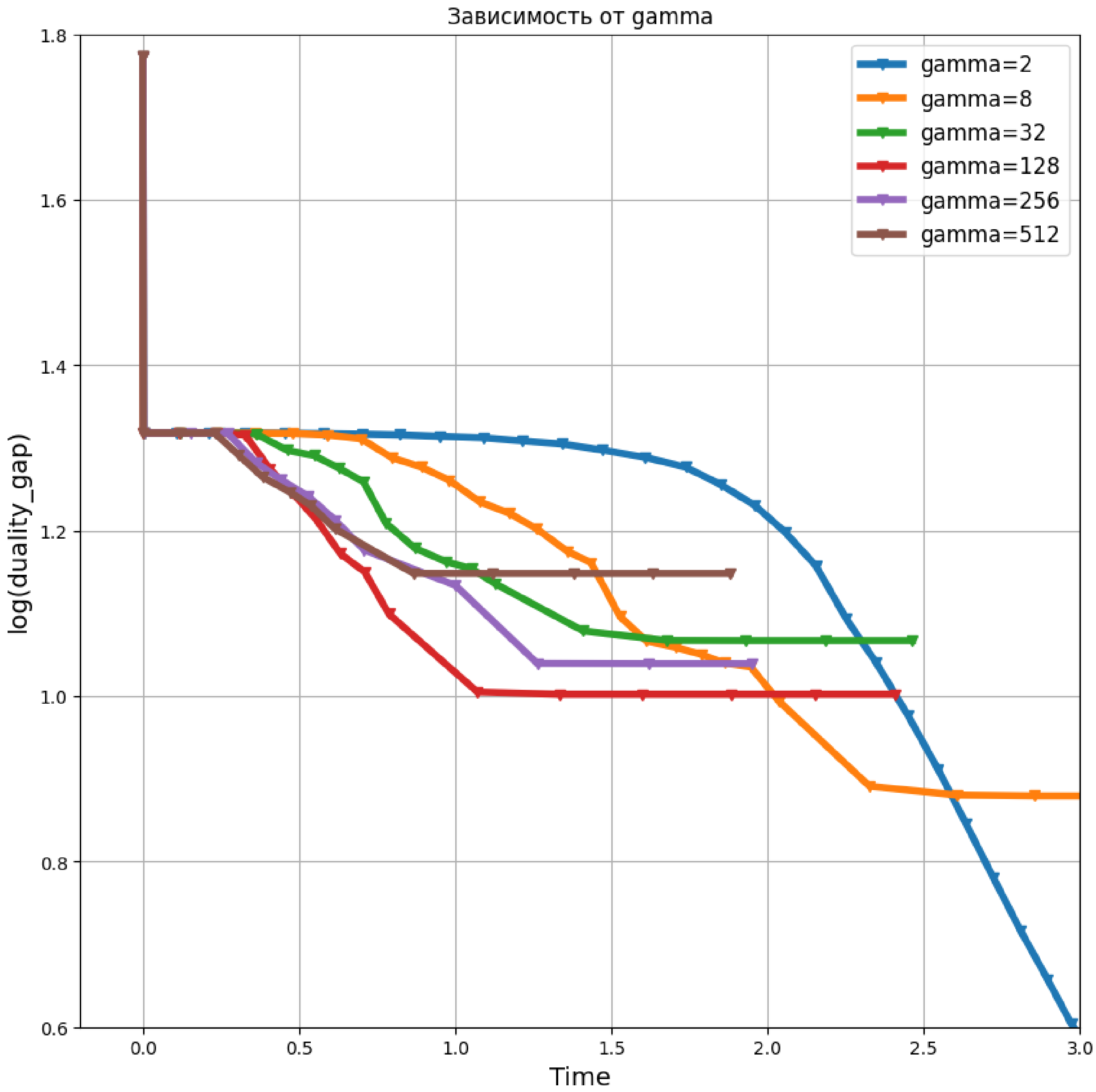

Рис.1 Зависимость зазора двойственности от времени работы

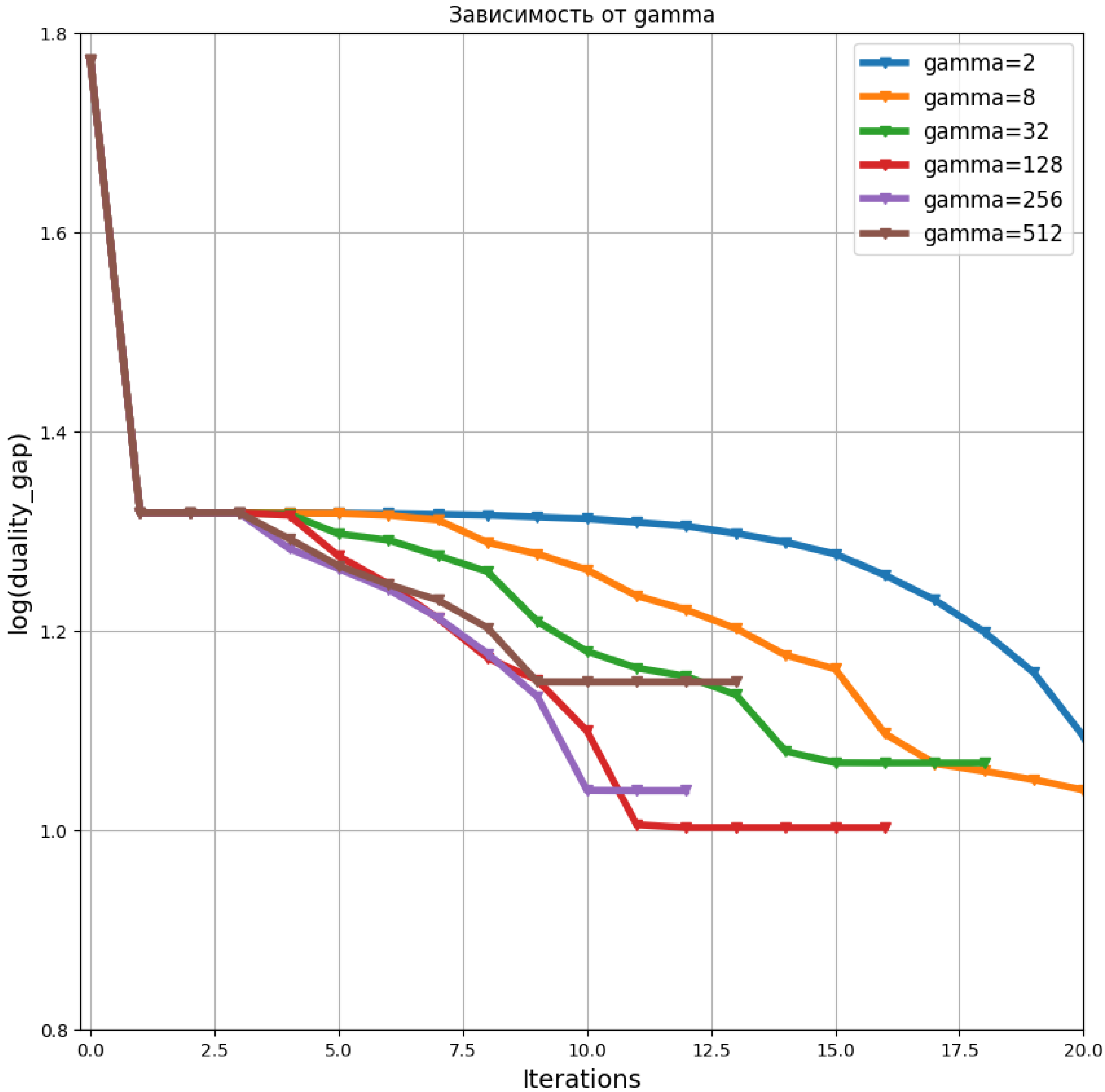

Рис.2 Зависимость зазора двойственности от количества итераций

Вывод: из графиков видно, что при увеличении параметра γ увеличивается скорость сходимости метода, при уменьшении – метод сходится дольше.

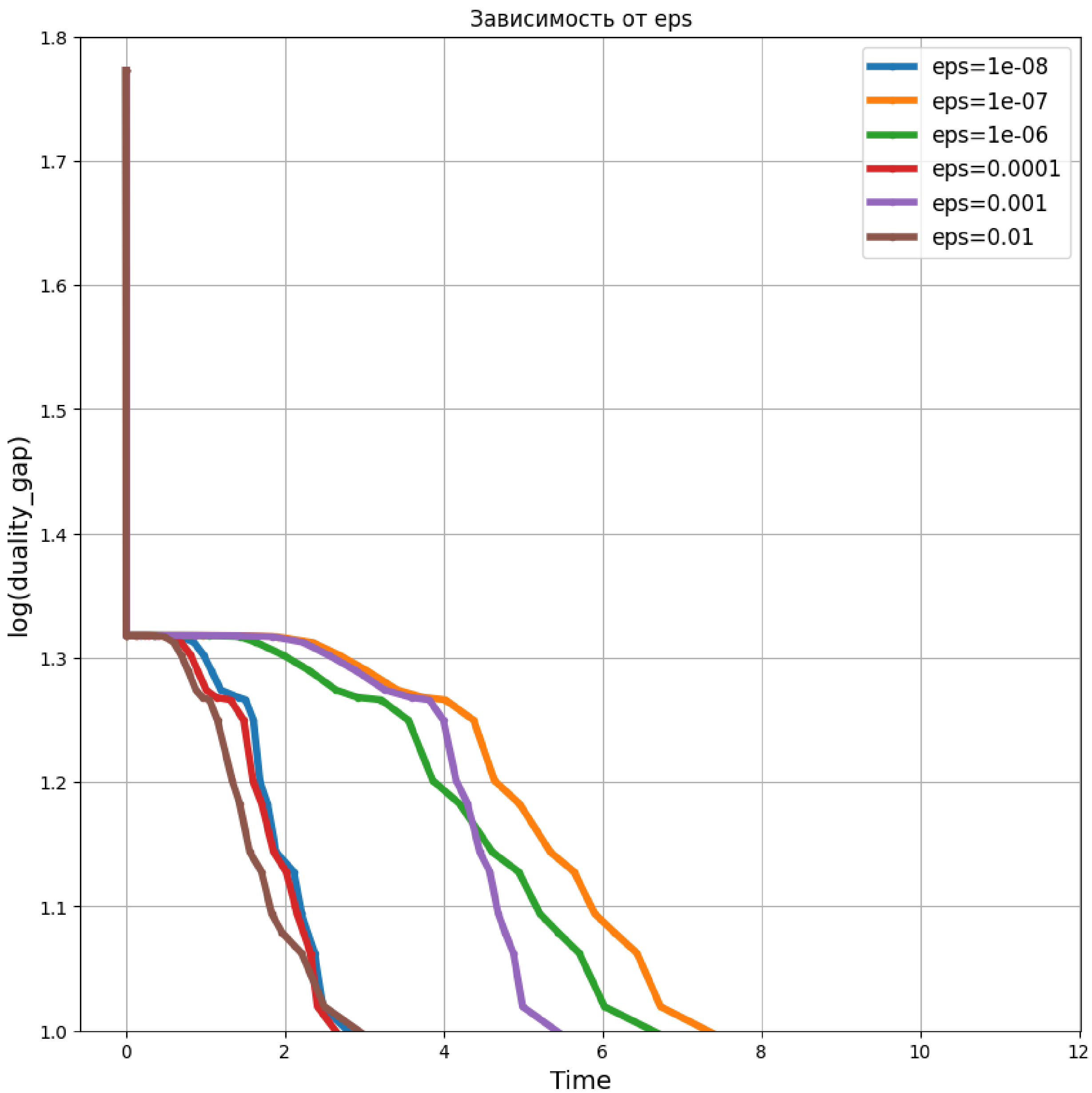

Рис.3 Зависимость зазора двойственности от времени работы

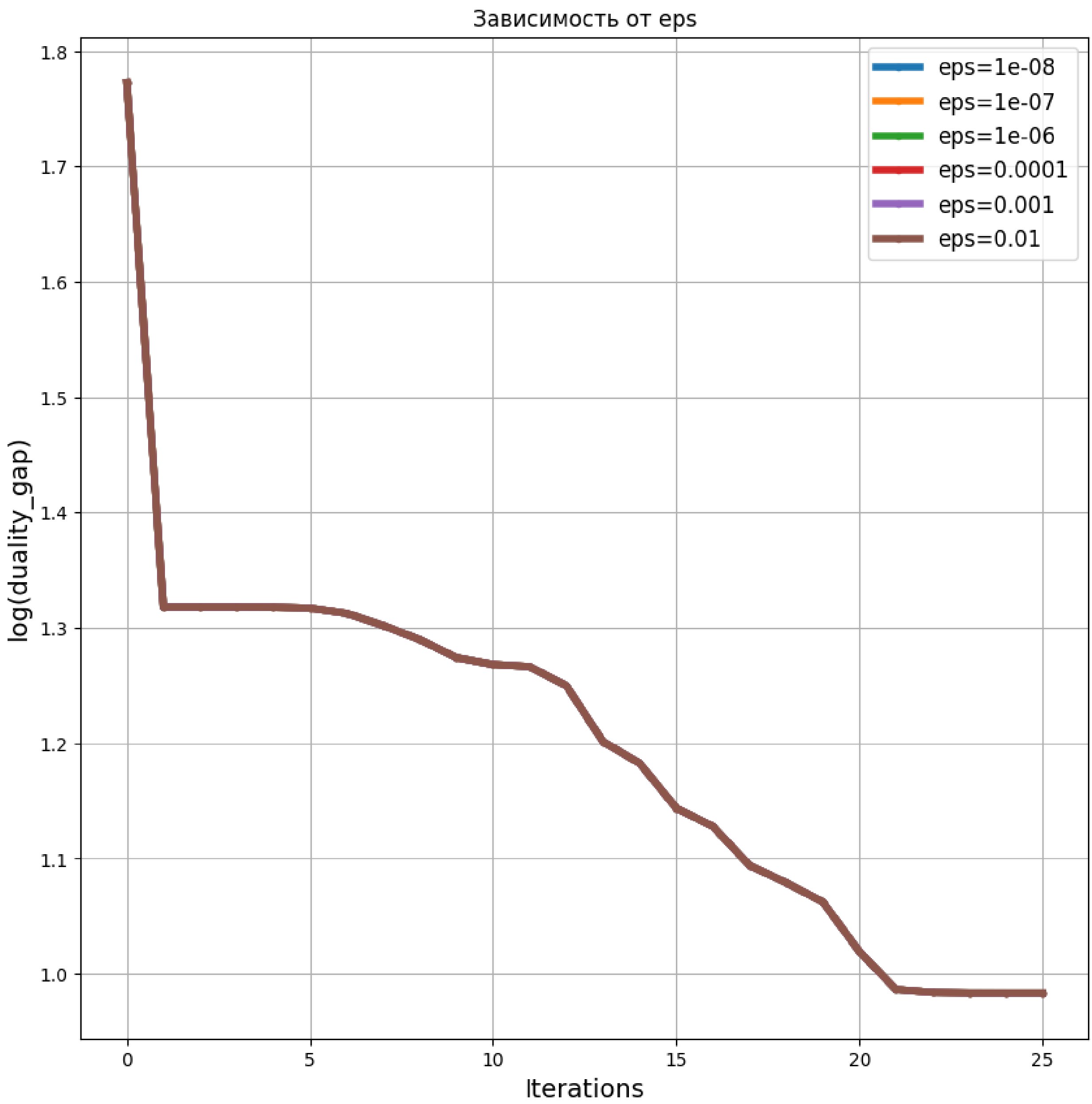

Рис.4 Зависимость зазора двойственности от количества итераций

Вывод: из графиков видим, что ϵ_{inner} влияет на время сходимости, чем меньше коэффициент, тем медленнее сходимость, так как методу Ньютона требуется больше итераций для достижения нужной точности. Но при этом, ϵ_{inner} не влияет на количество итераций.

2.2 Эксперимент: исследование поведения метода для различных значений размерности пространства n, размера выборки m и коэффициента регуляризации λ

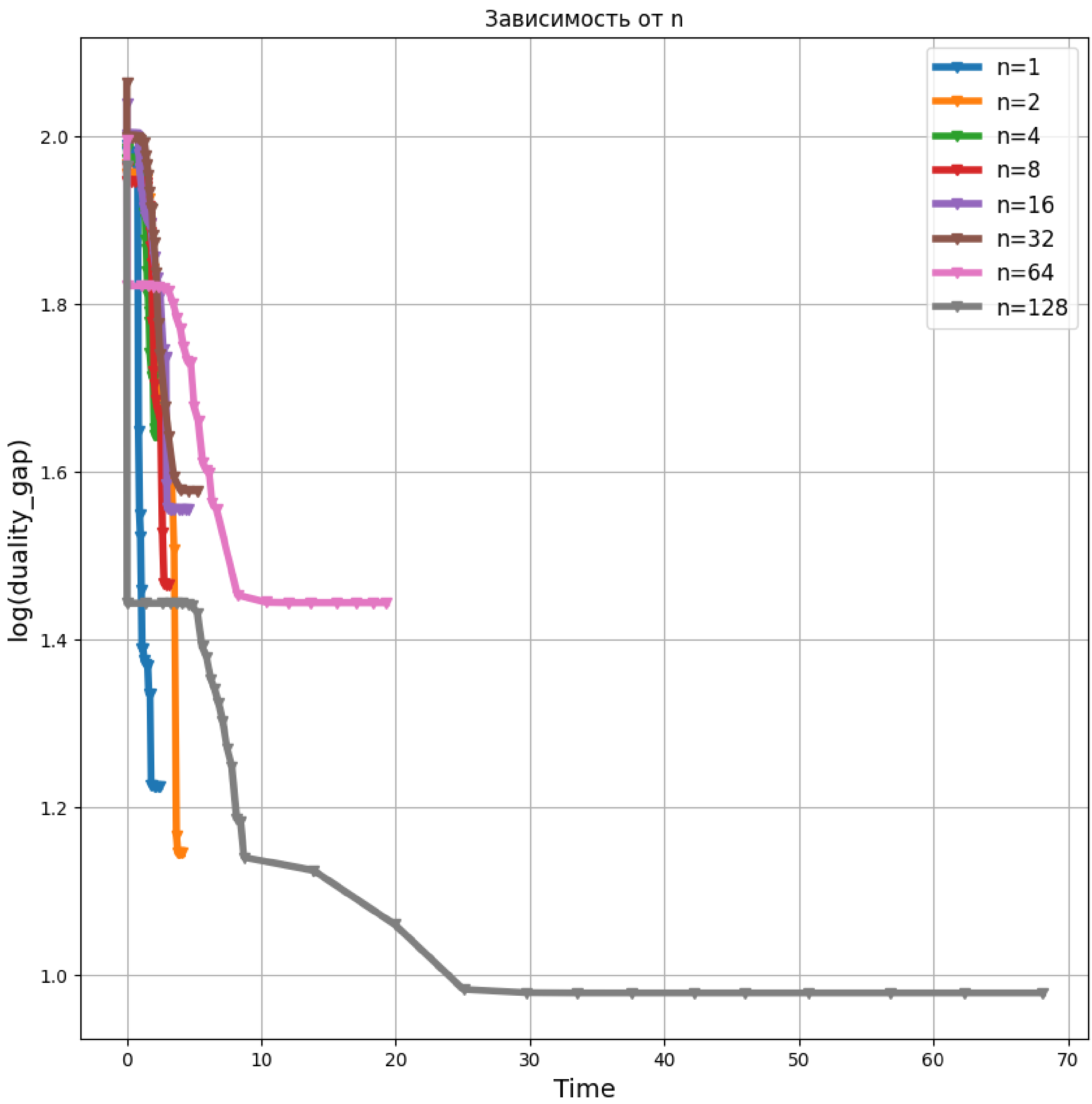

Рис.5 Зависимость зазора двойственности от времени работы при разной размерности задачи

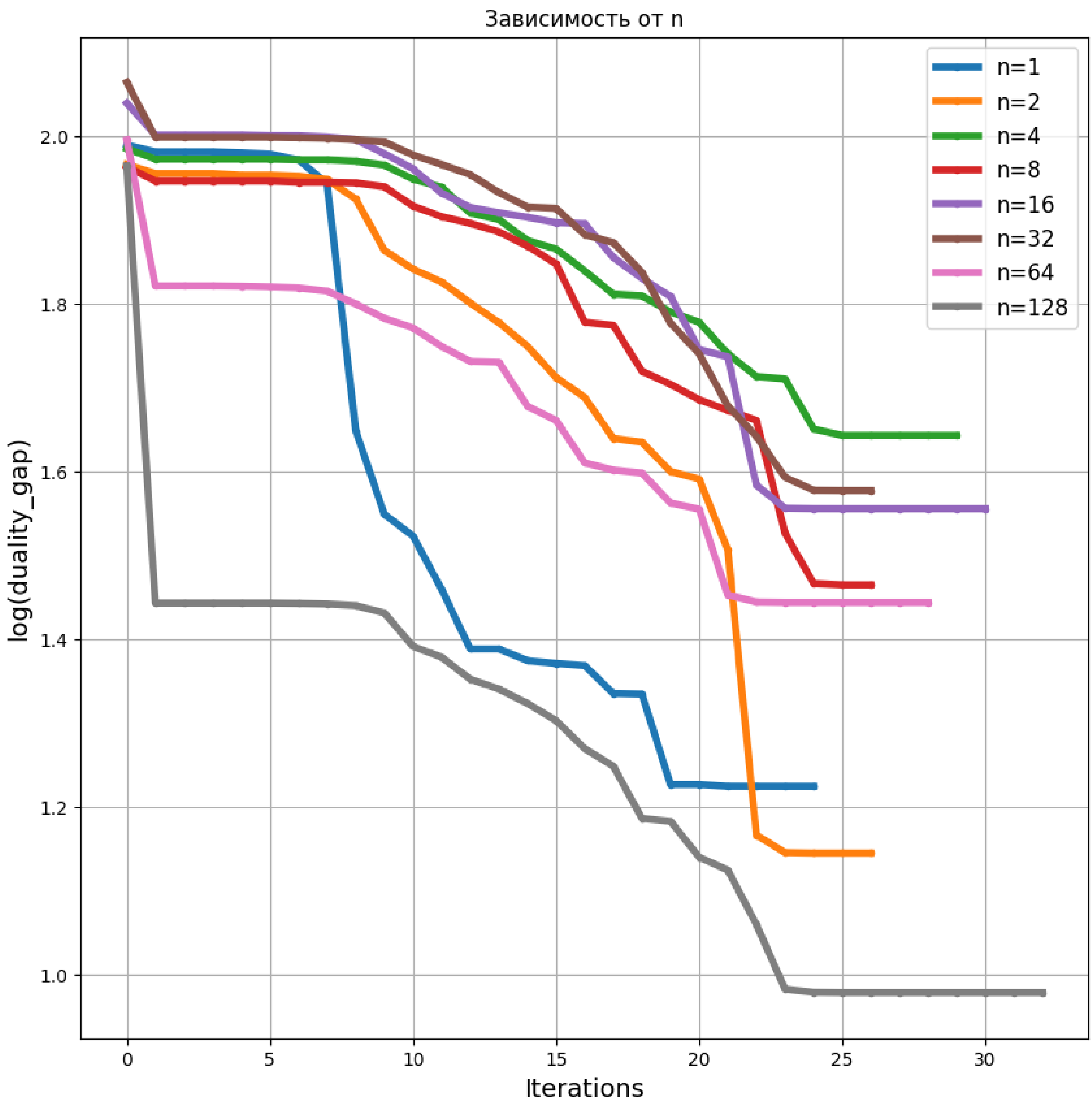

Рис.6 Зависимость зазора двойственности от итераций при разной размерности задачи

Вывод: из графиков видим, чем больше размерность задачи, тем больше время работы метода и количество итераций

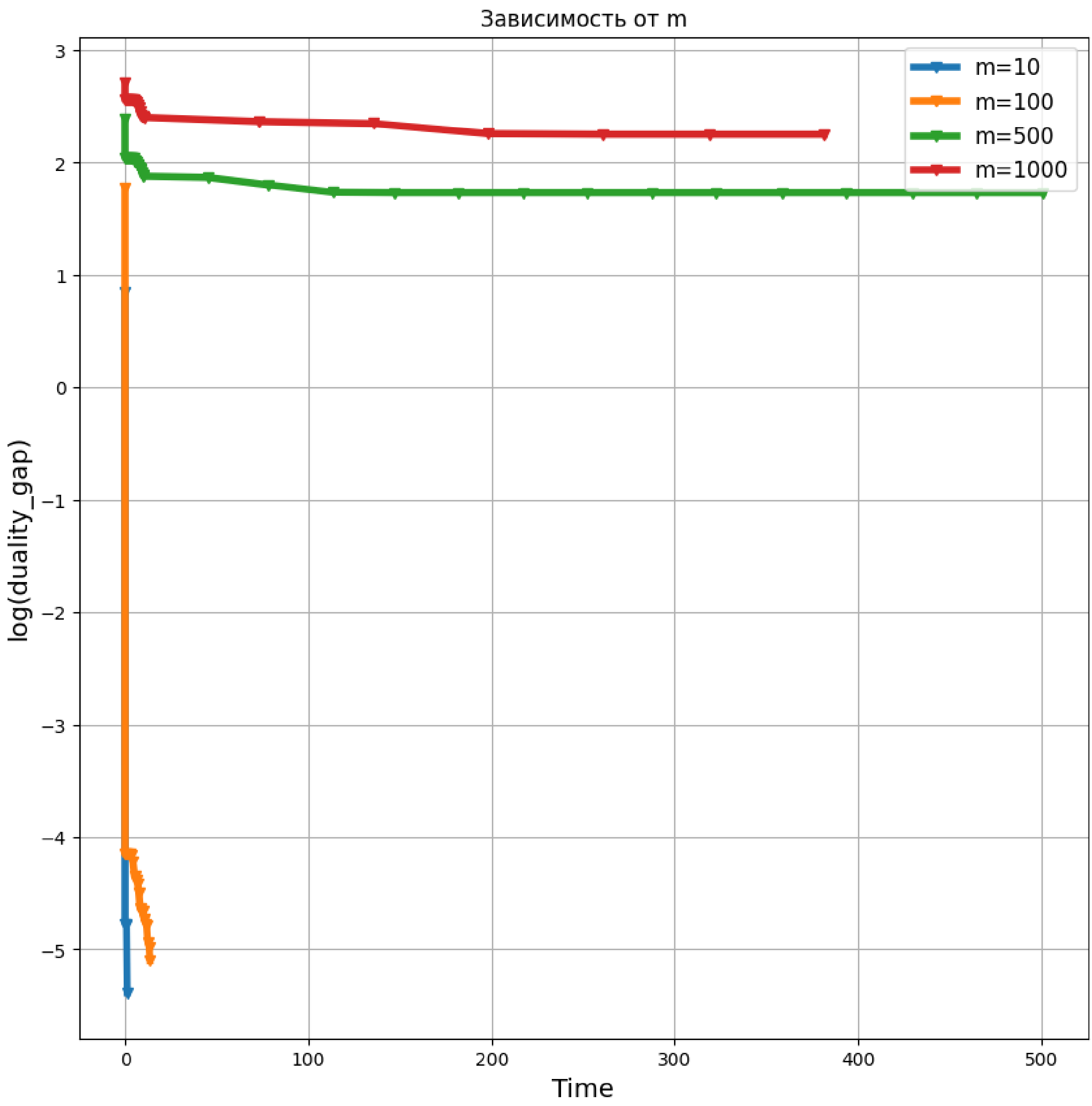

Рис.7 Зависимость зазора двойственности от времени работы при разных размерностях выборки

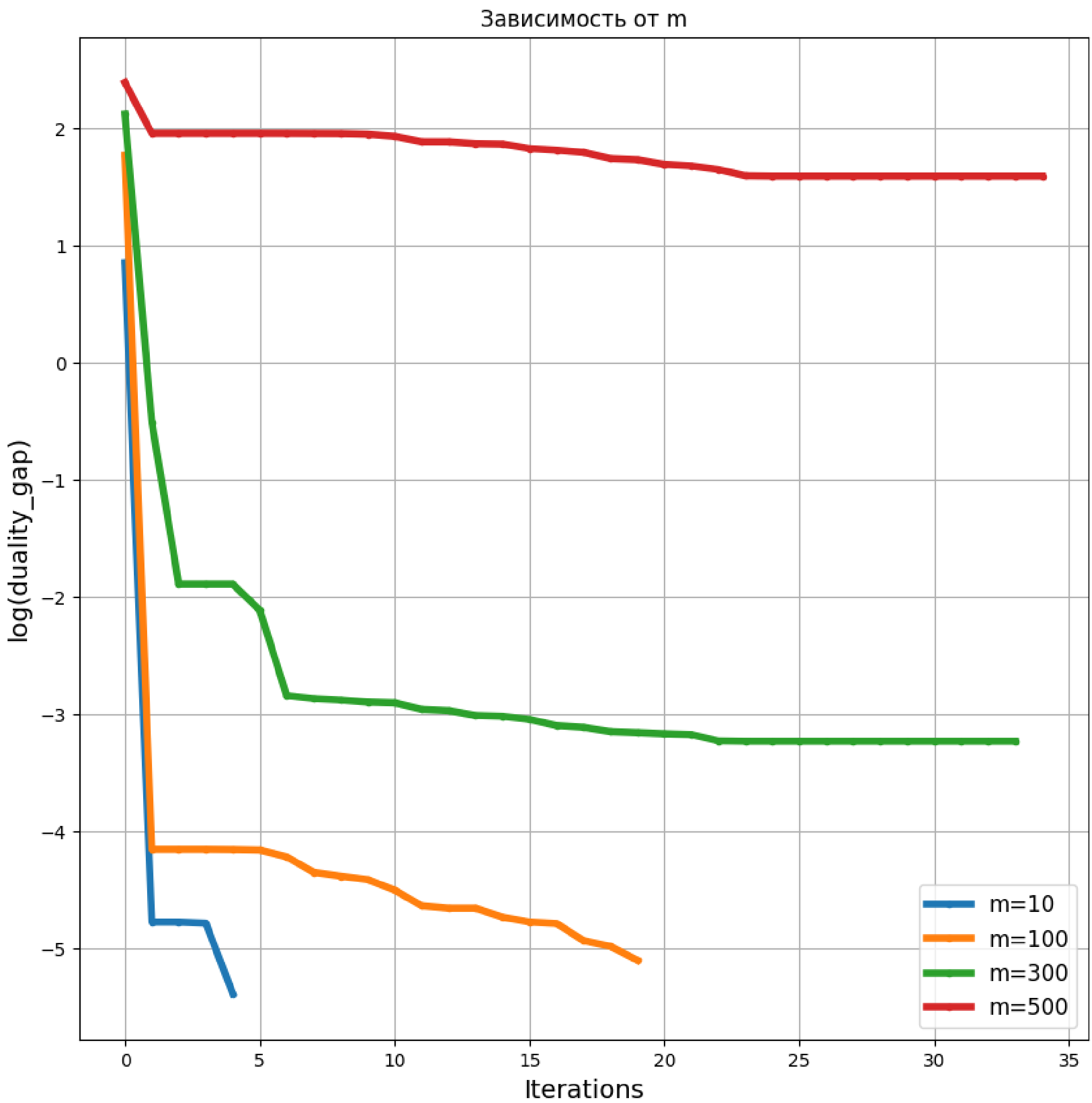

Рис.8 Зависимость зазора двойственности от итераций при разных размерностях выборки

Вывод: из графиков видим, чем больше размер выборки, тем больше время работы метода и количество итераций.

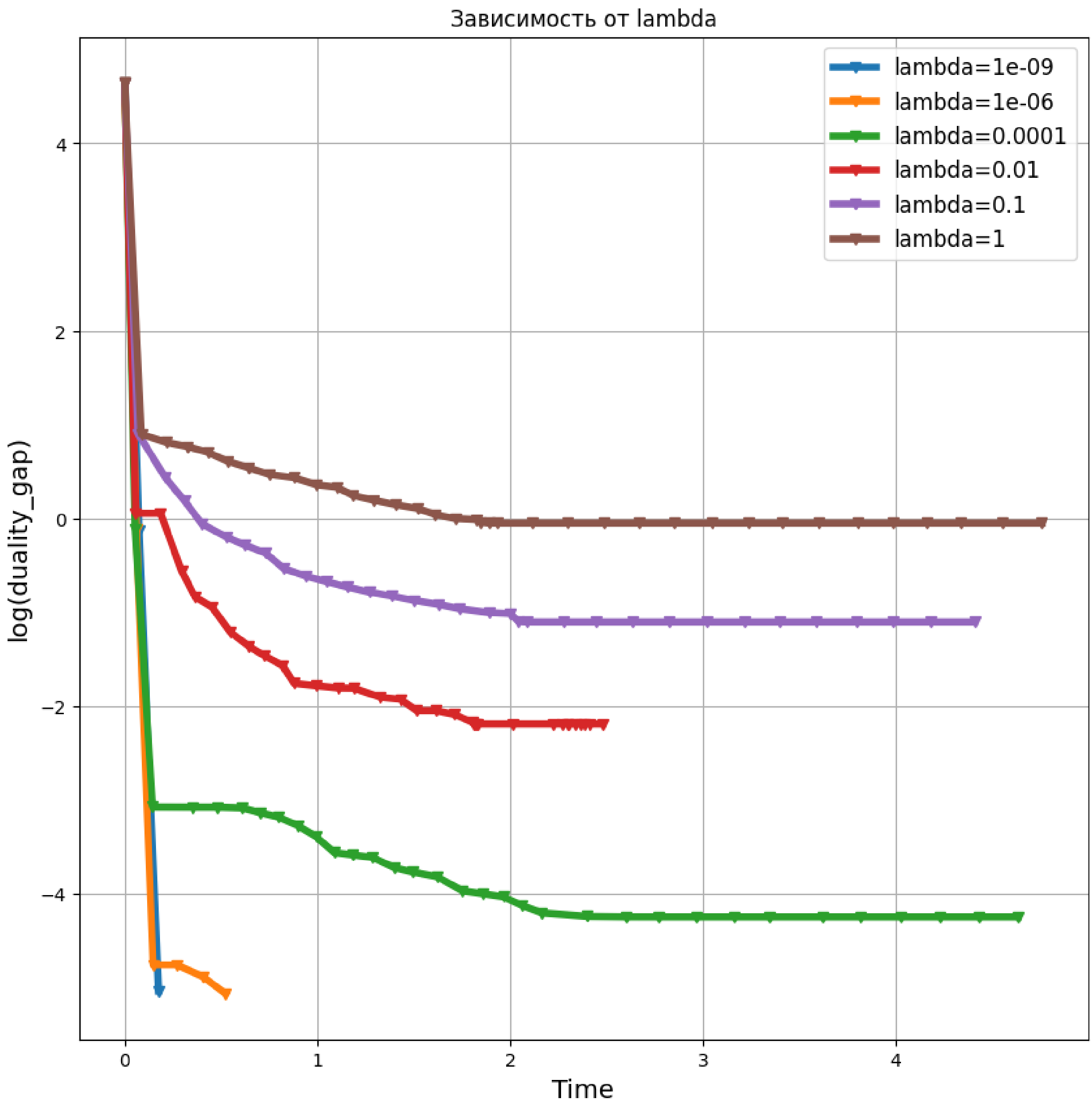

Рис.9 Зависимость зазора двойственности от времени работы при разных коэффициентах регуляризации lambda

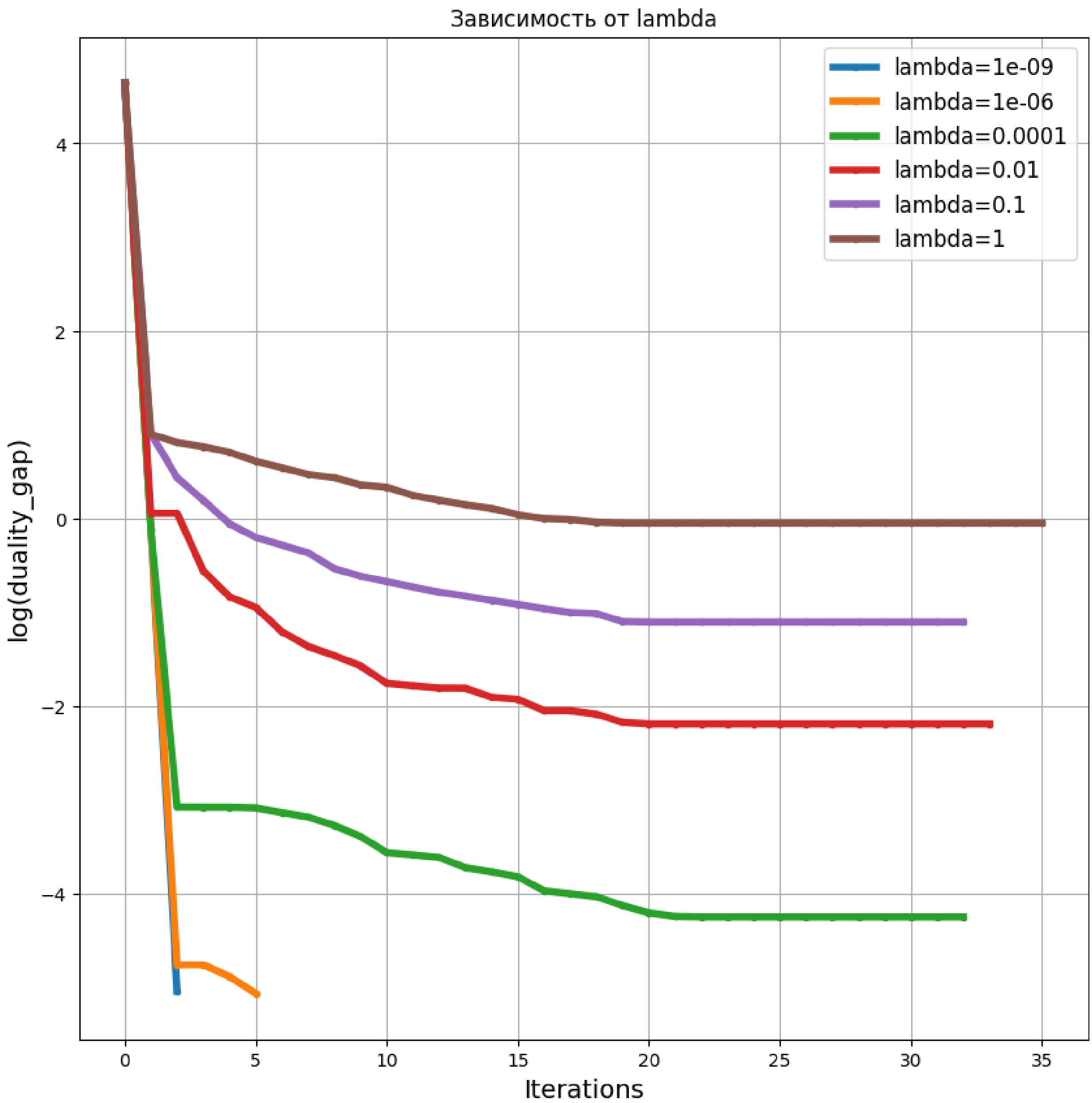

Рис.10 Зависимость зазора двойственности от итераций при разных коэффициентах регуляризации lambda

Вывод: из графиков видим, при маленьких коэффициентах регуляризации, наша задача быстрее сходится, так как при увеличении коэффициента, мы отдаем больше предпочтения на сумму весов. Следовательно, при больших коэффициентах, методу потребуется больше времени и итераций.

Вывод

Метод барьеров показывает хорошие результаты и скорость сходимости. За счет того, что используется метод Ньютона, метод работает эффективно на задачах малой размерности. Основываясь на результатах исследований, можно сказать, что параметр γ влияет на скорость работы, чем выше его значение, тем быстрее он отрабатывает. Параметр ϵ_{inner} влияет на время на работы, но не на количество итераций, параметр больше, сходимость быстрее. Время работы и количество итераций увеличивается с увеличением параметров n, m. При низких значениях коэффициента регуляризации λ, метод может сойтись за несколько итераций.