Федеральное государственное бюджетное образовательное
учреждение высшего профессионального образования
«Санкт-Петербургский национальный исследовательский университет
информационных технологий, механики и оптики»

Отчёт по лабораторной работе №3

Метод барьеров.

По дисциплине: Методы оптимизации

Выполнил:
Зозуля А.В.
Проверила:
Головкина А.Г.

Москва 2024 г.

Вспомогательная функция $f_t(x, u)$, минимизируемая в методе барьеров:

$f_t(x,u) = tf(x,u) + F(x,u) \Rightarrow|$ Где $F(x,u) =- \sum_{i=1}^{m} ln(-g_i(x,u))$; учитывая $-u \preccurlyeq x \preccurlyeq u \Rightarrow g_1(x,u) = -x_1 - u_1$; $g_2(x,u) = x_2 - u_2 | \Rightarrow$

$\Rightarrow t(\frac{1}{2}||Ax - b||_2^2 + \lambda\langle 1_n, u\rangle) - \sum(ln(u + x) + ln(u - x))$

Система линейных уравнений, задающая ньютоновское направление $d_k = (d_k^x, d_k^u)$:

Так как в методе ньютона решается система $H_{f_t} \cdot d =- \nabla_{f_t}$, где направление спуска представляет из себя направление спуска по переменной x и u т.е., d_x и d_u. Тогда решая полученное уравнение через преобразования Холецкого можно получить направление спуска $d = (d_x, d_u)$:

Гессиан можно найти по следующей формуле:

$H_{f_t} = \begin{bmatrix} \nabla^2_{xx} & \nabla^2_{xu} \\ \nabla^2_{ux} & \nabla^2_{uu} \end{bmatrix}$, где

$\nabla^2_{xx} = tA^TA + diag(\frac{1}{(u+x)^2} + \frac{1}{(u-x)^2})$,

$\nabla^2_{xu} = diag(\frac{1}{(u+x)^2} + \frac{1}{(u-x)^2})$,

$\nabla^2_{uu} = diag(\frac{1}{(u+x)^2} - \frac{1}{(u-x)^2})$.

$(diag(\frac{1}{(u+x)^2} + \frac{1}{(u-x)^2})$ – элементы диагонали где каждый элемент представляет из себя результат выражения $\frac{1}{(u_i+x_i)^2} + \frac{1}{(u_i-x_i)^2})$

И градиент:

$\nabla_{f_t} = (\nabla_x, \nabla_u)^T$, где

$\nabla_x = tA^T(Ax - b) + \frac{1}{u+x} + \frac{1}{u-x}$

$\nabla_u = t\lambda 1_n - \frac{1}{u+x} - \frac{1}{u-x}$

$(\frac{1}{u+x}$ – вектор где каждый элемент представляет из себя $\frac{1}{u_i+x_i})$

Для текущей точки (x, u) и направления d_k определите максимальную допустимую длину шага α.

Должны быть выполнены ограничения: $-u \preccurlyeq x \preccurlyeq u$ и $\alpha \geq 0$. Для ньютоновского направления $d_k = (d_x, d_u)$, по координатно находим минимум. Если выполняется условие $d_{x,i} - d_{u,i} > 0$, то $\alpha \le 0.99 \cdot \frac{x_i - u_i}{d_{u,i} - d_{x,i}}$. А если $-d_{x,i} - d_{u,i} > 0$, то $\alpha \le 0.99 \cdot \frac{-u_i - x_i}{d_{x,i} + d_{u,i}}$.

Какую начальную точку (x0, u0) можно выбрать для метода барьеров?

(x0, u0) – должны быть внутренними допустимыми точками, чтобы их найти запускают метод первой фазы:

$\{$

$g_i(x) \le s$

$Ax = b$

$s \to min$

$\}$

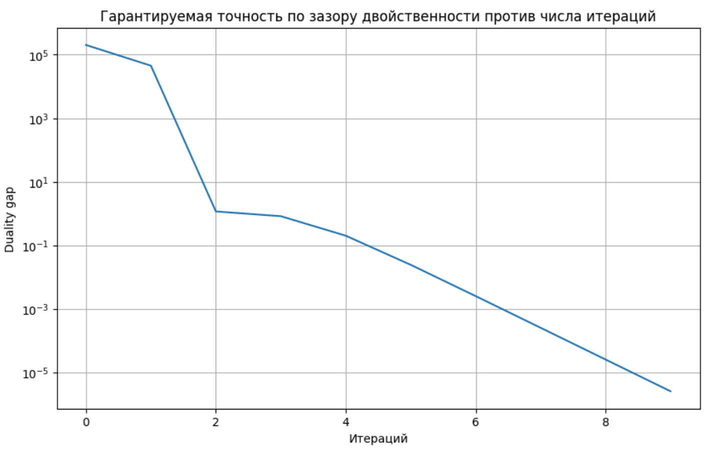
Где x_0: $Ax_0 = b$; $s_0 = max g_i(x) \cdot 1.1$

Тогда (x_0, s_0) – гарантированно внутренняя точка. Достаточно чтобы выполнялось $s < 0$.

Поэкспериментируйте с реализованным методом барьеров для задачи LASSO:

(a) Исследуйте, насколько сильно метод чувствителен к выбору параметров γ (скорость увеличения t_k) и ϵ_{inner} (точность для метода Ньютона).

Для γ = 10, ϵ_{inner} = 1e-05 :



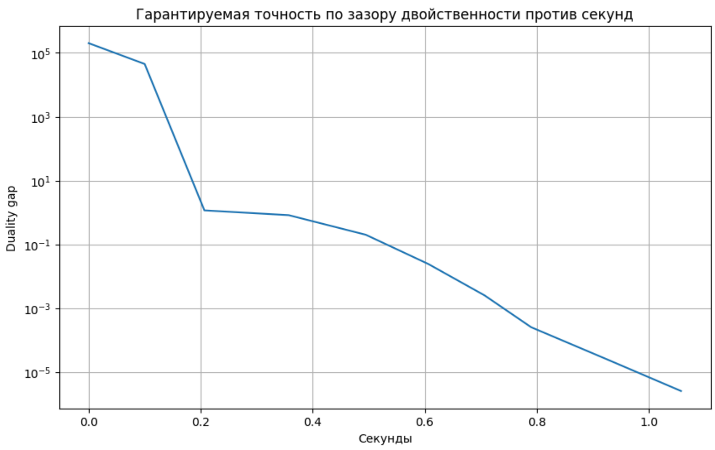


Рис 1. Точность по зазору двойственности против числа итераций и гарантированная точность по зазору двойственности против реального времени работы

Для γ = 10, ϵ_{inner} = 1e-09 :

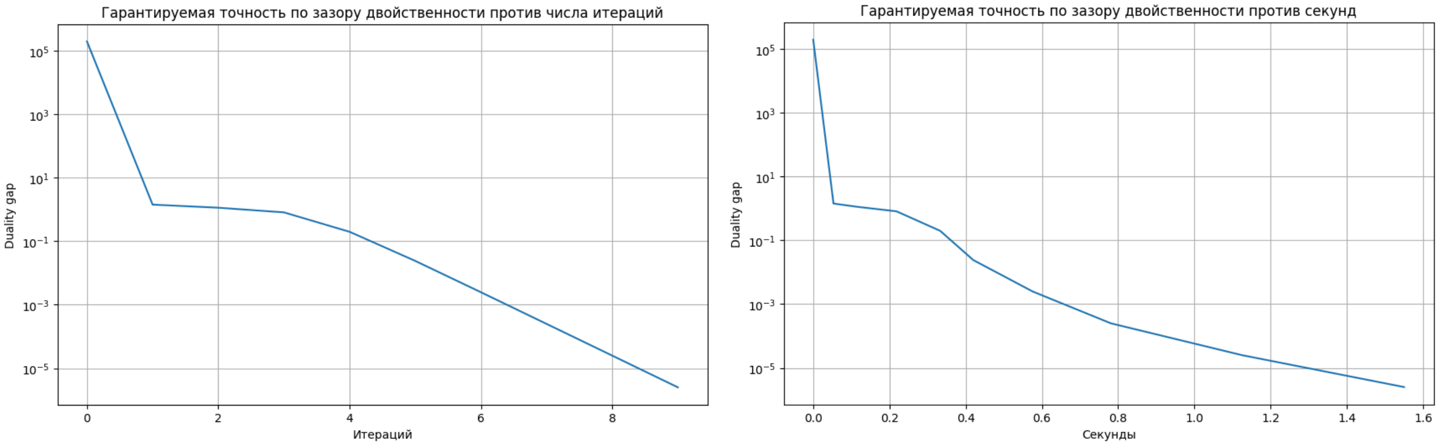


Рис 2. Точность по зазору двойственности против числа итераций и гарантированная точность по зазору двойственности против реального времени работы

Для $\gamma = 100$, $\epsilon_{\text{inner}} = 1\text{e-}05$:

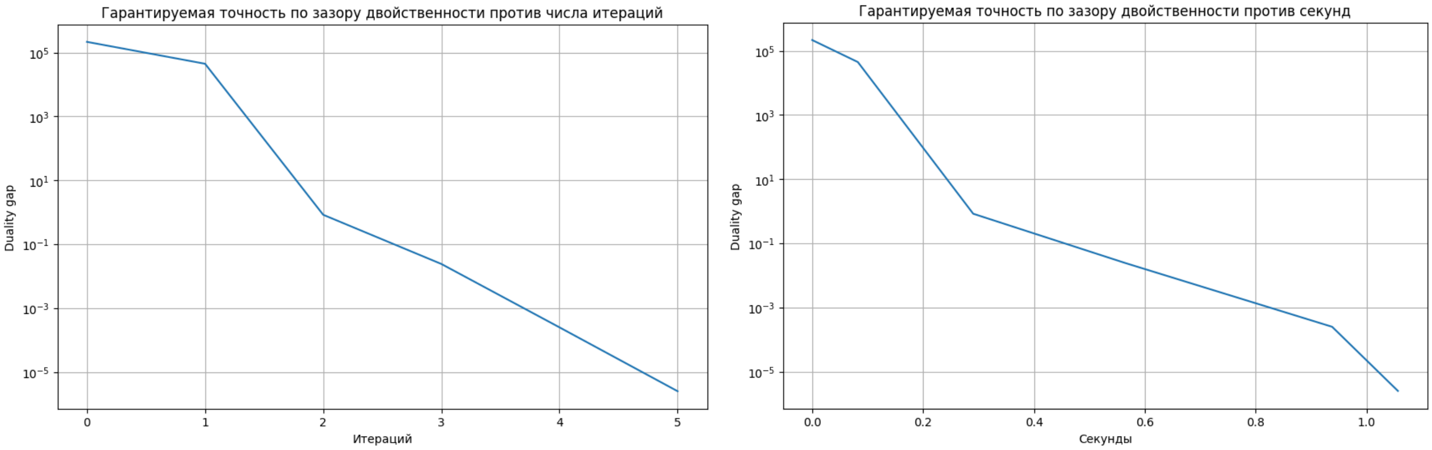


Рис 3. Точность по зазору двойственности против числа итераций и гарантированная точность по зазору двойственности против реального времени работы

Для $\gamma = 100$, $\epsilon_{\text{inner}} = 1\text{e-}09$:

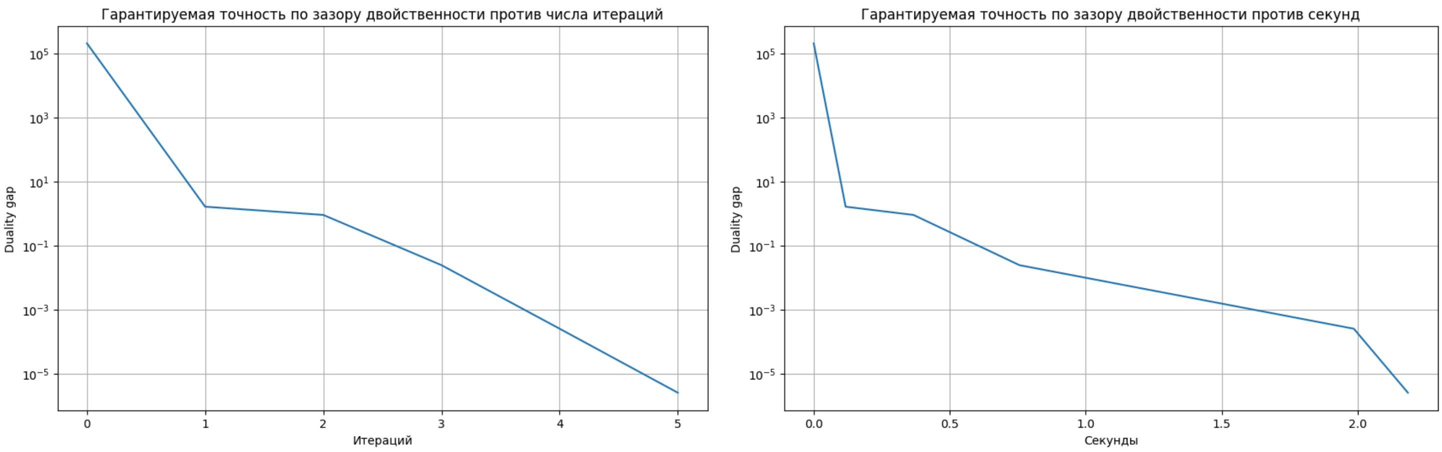


Рис 4. Точность по зазору двойственности против числа итераций и гарантированная точность по зазору двойственности против реального времени работы

Вывод: при уменьшении параметра ϵ_{inner} метод дает более хорошее решение за одну внешнюю итерацию и быстрее сходится к значению “перегиба” функции, затем точность и скорость не улучшается по сравнению с большим ϵ_{inner}, но увеличивается реальное время работы метода. Это можно объяснить тем что с меньшим параметром ϵ_{inner} метод Ньютона возвращает достаточно хороший результат еще на первой итерации метода барьеров (близкий к истинному решению вспомогательной задачи), при этом условие останова метода барьеров является менее требовательным по точности чем условие останова у метода Ньютона: (Зазор двойственности) $< \varepsilon = 10^{-5}$, когда как Ньютон: $||\nabla f_{t_k}(y_l)||_2^2 \leq \epsilon_{inner}||\nabla f_{t_k}(x_k)||_2^2$ где $\epsilon_{inner} = 10^{-9}$. Параметр γ – уменьшает количество итераций до нахождения оптимума, по реальному времени работы одинаково. Уменьшение количества итераций объясняется тем что чем больше γ, т.е., множитель для коэффициента τ тем меньше требуется внешних итераций, но потребуется больше внутренних итераций (метода Ньютона) для достижения оптимального решения задачи.

(b) Как меняется поведение метода для различных значений размерности пространства n, размера выборки m и коэффициента регуляризации λ?

Для $n = 10$, $m = 100$, $\lambda = 0.1$:

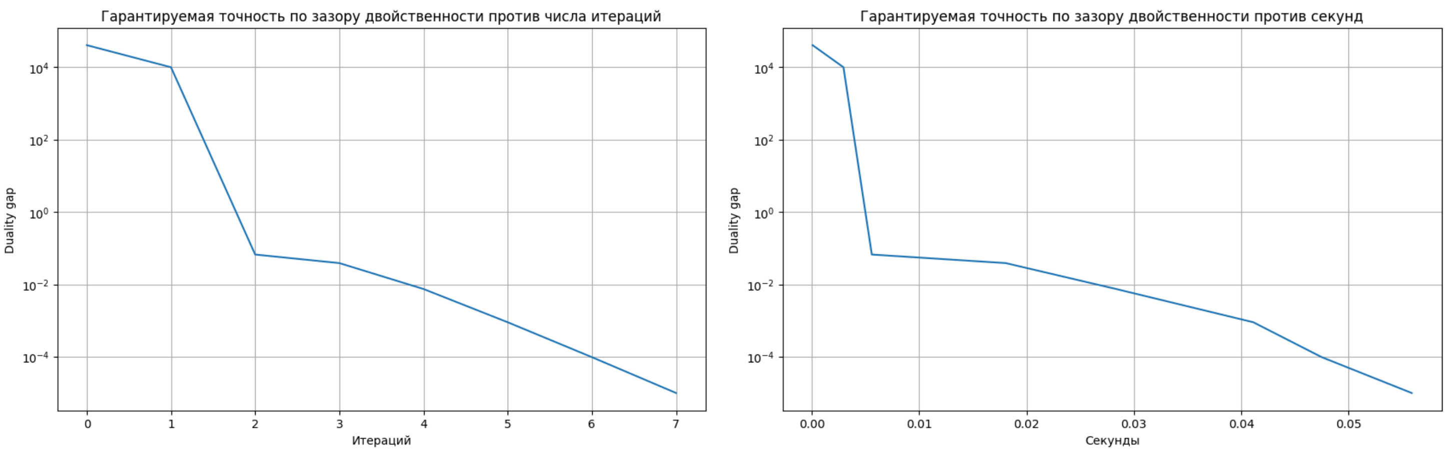


Рис 5. Точность по зазору двойственности против числа итераций и гарантированная точность по зазору двойственности против реального времени работы

Для $n = 10$, $m = 10000$, $\lambda = 0.1$:

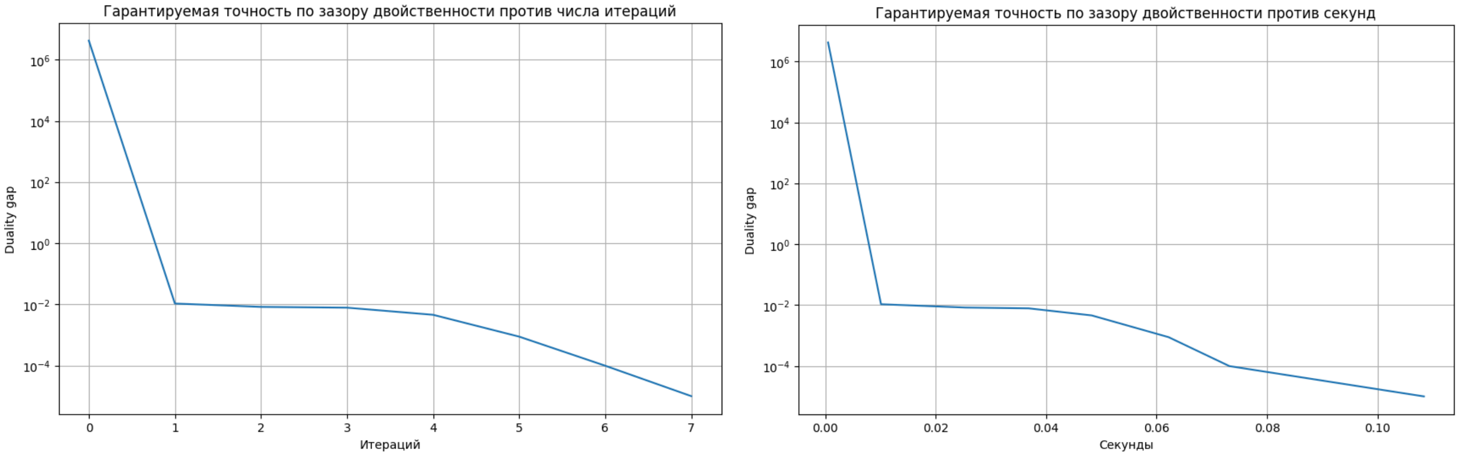


Рис 6. Точность по зазору двойственности против числа итераций и гарантированная точность по зазору двойственности против реального времени работы

Для n = 1000, m = 10000, λ = 0.1:

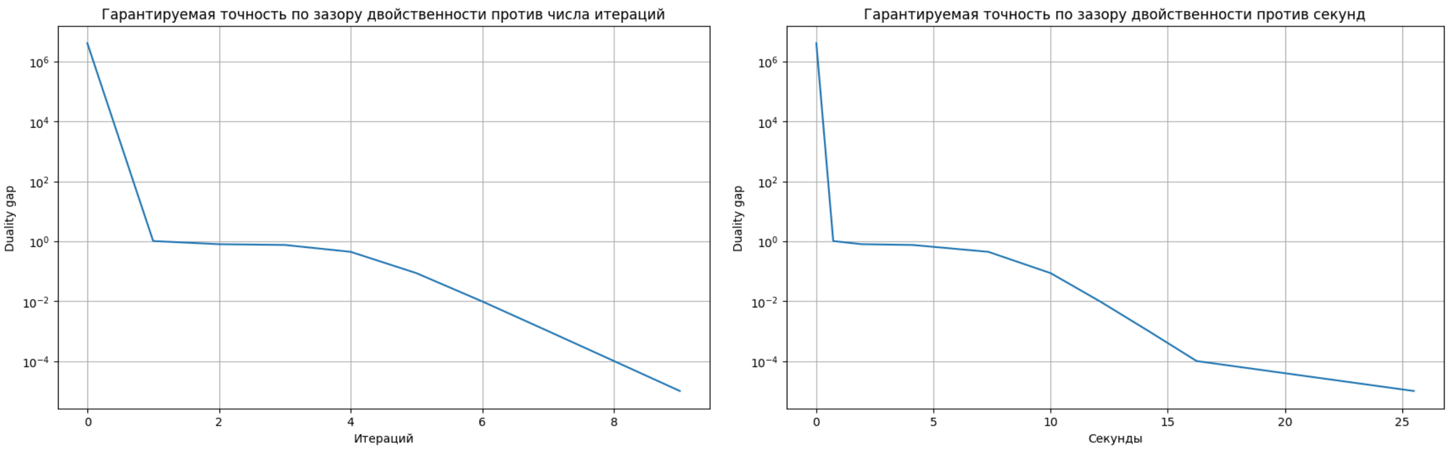


Рис 7. Точность по зазору двойственности против числа итераций и гарантированная точность по зазору двойственности против реального времени работы

Для n = 10, m = 100, λ = 5:

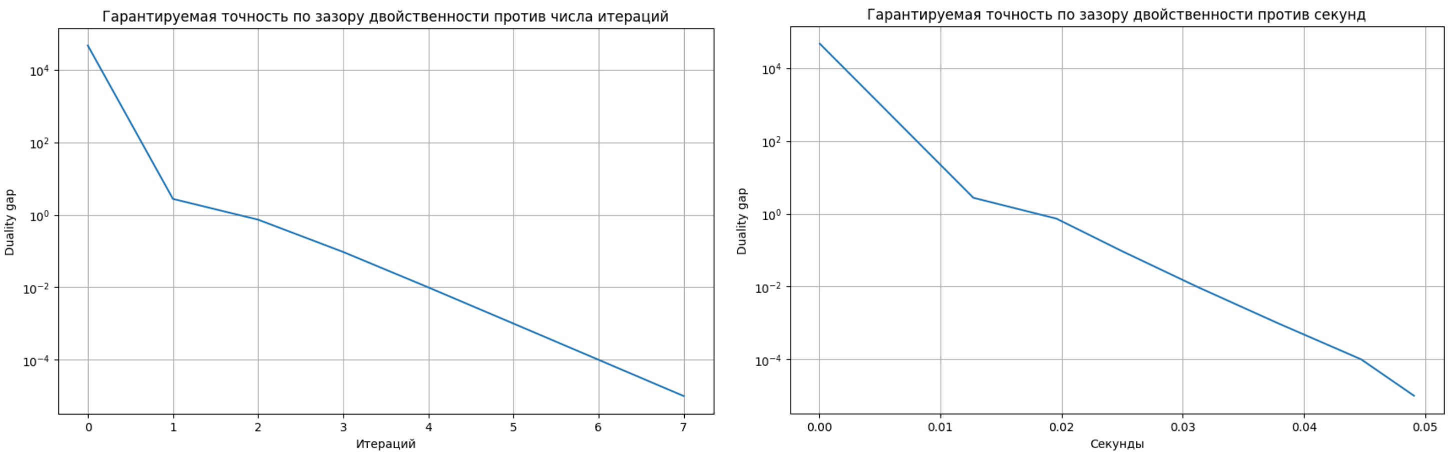


Рис 8. Точность по зазору двойственности против числа итераций и гарантированная точность по зазору двойственности против реального времени работы

Для n = 10, m = 10000, λ = 5:

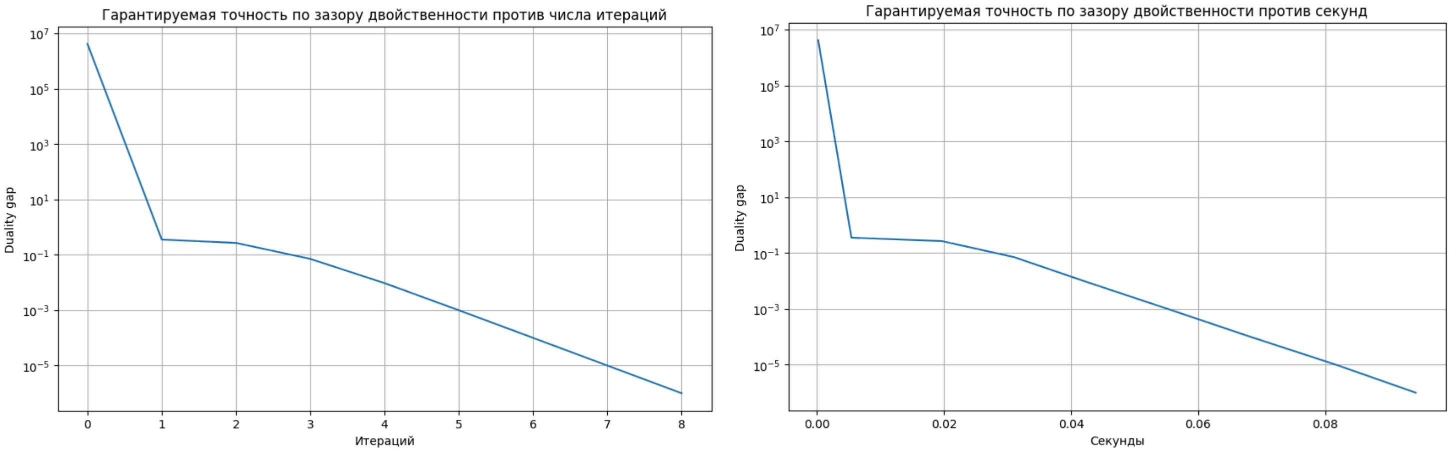


Рис 9. Точность по зазору двойственности против числа итераций и гарантированная точность по зазору двойственности против реального времени работы

Для n = 1000, m = 10000, λ = 5:

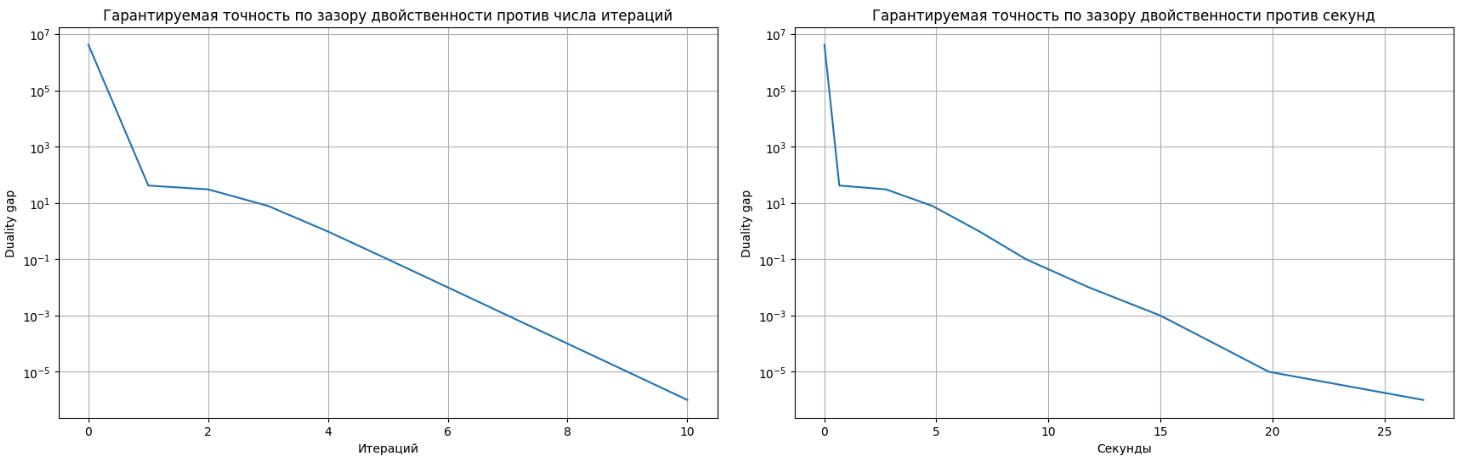


Рис 10. Точность по зазору двойственности против числа итераций и гарантированная точность по зазору двойственности против реального времени работы

Вывод: при увеличении размера выборки m за одну итерацию метод достигает более хорошего решения вспомогательной задачи, количество итераций до “перегиба” уменьшается, общее количество требуемых для нахождения оптимального решения итераций не изменяется, увеличивается реальное время работы алгоритма. С увеличением размерности пространства n увеличивается общее количество требуемых для оптимального решения итераций, время работы реального алгоритма значительно увеличивается, следовательно размерность пространства намного сильнее влияет на реальное время работы алгоритма, по сравнению с размерами выборки m. Это можно объяснить тем, что “узкое горлышко” работы алгоритма заключается в подборе оптимальных значений x и u, при увеличении параметра m, размерность оптимальных значений не изменяется, когда как при увеличении n их размерность увеличивается 2x раз. Увеличение параметра λ почти не меняет количество требуемых для достижения оптимума итераций и реальной продолжительности работы программы, за первую итерацию решение получается менее точным. Учитывая что коэффициент регуляризации является весовым коэффициентом для для переменной u (который задает ограничения $-u \preccurlyeq x \preccurlyeq u$) то изменяя параметр регуляризации изменяется и точность решения вспомогательной задачи.