Ерохин В. И., Тамасян Г. Ш.

О НЕКОТОРЫХ ПОДХОДАХ К ПОСТРОЕНИЮ ПРАКТИЧНЫХ АЛГОРИТМОВ КВАДРАТИЧНОГО ПРОГРАММИРОВАНИЯ

Семинар «O&ML»

г. Санкт-Петербург
10 октября 2024 года

ПОСТАНОВКА ЗАДАЧ

Задача 1.

$$\begin{gathered} \frac{1}{2}x^\top Dx + x^\top c \longrightarrow \min, \\ \ell \leqslant x \leqslant w. \end{gathered} \tag{1}$$

Задача 2.

$$\begin{gathered} \frac{1}{2}x^\top Dx + x^\top c \longrightarrow \min, \\ Ax = b, \\ \ell \leqslant x \leqslant w. \end{gathered} \tag{2}$$

Здесь D — симметричная положительно определенная матрица, $\ell < w$.

Далее предполагается, что задачи имеют решения.

КРИТЕРИЙ ОПТИМАЛЬНОСТИ ДЛЯ ЗАДАЧИ (1)

Положим $A = \begin{pmatrix} E \\ -E \end{pmatrix}$, $b = \begin{pmatrix} \ell \\ -w \end{pmatrix}$. Перепишем задачу (1):

$$\frac{1}{2}x^\top Dx + x^\top c \longrightarrow \min,$$
$$Ax \geqslant b.$$

Форма 1: $\begin{cases} b^\top u_* = x_*^\top (Dx_* + c), \\ Dx_* - A^\top u_* = -c, \\ u_* \geqslant \mathbf{0}, \\ \ell \leqslant x_* \leqslant w. \end{cases}$

Форма 2: $\begin{cases} Ax_* - v_* = b, \\ Dx_* - A^\top u_* = -c, \\ v_* \geqslant \mathbf{0},\ u_* \geqslant \mathbf{0}, \\ v_*^\top u_* = 0, \\ \ell \leqslant x_* \leqslant w. \end{cases}$

ПРЕДЛАГАЕМЫЕ РЕШАТЕЛИ ЗАДАЧИ (1)

1. Метод покоординатного спуска (**Spusk**).

2. Метод Дикина (**Dikin**):
   - подход I (на основе преобразования Хилдрета);
   - подход II.

3. Комбинация методов покоординатного спуска и Дикина (**Comb**).

4. Комбинация методов покоординатного спуска и сопряженных градиентов (**MCG**).

МЕТОД ДИКИНА (Dikin)

$$\begin{cases} \frac{1}{2}x^\top Dx + x^\top c \longrightarrow \min\limits_{x \in \Omega}, \\ \Omega = \{x \in \mathbb{R}^n \mid Ax = b,\ x \geqslant \mathbf{0}\}. \end{cases}$$

$$\boxed{\begin{gathered} \Omega_0 = \{x \in \mathbb{R}^n \mid Ax = b,\ x > \mathbf{0}\} \neq \emptyset, \\ A \ -\ \text{матрица полного ранга.} \end{gathered}}$$

Пусть $x_k \in \Omega_0$.

1. Вычислим $g_k = Dx_k + c$ и найдем решение u_k системы

$$A\Lambda_k^2 A^\top u = A\Lambda_k^2 g_k,$$

   где $\Lambda_k = \mathrm{diag}(x_k)$.

2. Находим $y_k = g_k - A^\top u_k$, $\Phi_k = \sum\limits_{k=1}^{n}(x_k[j]y_k[j])^2$. Если $\Phi_k = 0$, то x_k — решение исходной задачи. В противном случае переходим к следующему шагу.

3. Определим направление спуска $s_k = -\Lambda_k^2 y_k$.

4. Вычисляем шаг $\lambda_k = \min\left\{\frac{\alpha}{\sqrt{\Phi_k}}, \frac{\Phi_k}{s_k^\top D s_k}\right\}$, где параметр $\alpha < 1$ и достаточно близок к 1.

5. Строим новое приближение $x_{k+1} = x_k + \lambda_k s_k$.

Подход I (ПРЕОБРАЗОВАНИЕ ХИЛДРЕТА)

$$\begin{cases} \dfrac{1}{2}x^\top Dx + x^\top c \longrightarrow \min, \\ Ax \geqslant b \end{cases} \quad \Rightarrow \quad \begin{cases} \dfrac{1}{2}z^\top Qz + z^\top q \longrightarrow \min, \\ z \geqslant \mathbf{0}, \end{cases}$$

где $A = \begin{pmatrix} E \\ -E \end{pmatrix}$, $b = \begin{pmatrix} \ell \\ -w \end{pmatrix}$,

$$Q = AD^{-1}A^\top = \begin{pmatrix} D^{-1} & -D^{-1} \\ -D^{-1} & D^{-1} \end{pmatrix},$$

$$q = -(b + AD^{-1}c) = \begin{pmatrix} -\ell - D^{-1}c \\ w + D^{-1}c \end{pmatrix}.$$

Возвращение к исходным переменным

$$x_* = D^{-1}(A^\top z_* - c).$$

Подход II

$$\begin{cases} \dfrac{1}{2}x^\top Dx + x^\top c \longrightarrow \min, \\ \ell \leqslant x \leqslant w, \end{cases} \quad \begin{matrix} x = z_1 + \ell \\ \Rightarrow \end{matrix} \quad \begin{cases} \dfrac{1}{2}z_1^\top Dz_1 + z_1^\top (D\ell + c) \longrightarrow \min, \\ \mathbf{0} \leqslant z_1 \leqslant w - \ell, \end{cases}$$

$$\begin{cases} \dfrac{1}{2}z_1^\top Dz_1 + z_1^\top (D\ell + c) \longrightarrow \min \\ z_1 + z_2 = w - \ell, \\ z_1 \geqslant \mathbf{0},\ z_2 \geqslant \mathbf{0}, \end{cases} \quad \Rightarrow \quad \begin{cases} \dfrac{1}{2}z^\top Qz + z^\top q \longrightarrow \min, \\ Az = b, \\ z \geqslant \mathbf{0}, \end{cases}$$

где

$$Q = \begin{pmatrix} D & \mathbf{O} \\ \mathbf{O} & \mathbf{O} \end{pmatrix}, \qquad q = \begin{pmatrix} w - \ell \\ \mathbf{0} \end{pmatrix}, \qquad z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix},$$

$$A = \begin{pmatrix} E & E \end{pmatrix}, \qquad b = w - \ell.$$

Замечание. В качестве начального приближения можно выбрать, например, вектор $z = \dfrac{1}{2}\begin{pmatrix} b \\ b \end{pmatrix} > \mathbf{0}$.

Комбинация методов покоординатного спуска и сопряженных градиентов (MCG)

1. Положить $x = \frac{\ell+w}{2}$. Выполнить процедуру покоординатного спуска: для $j = 1 : n$ последовательно решить совокупность задач вида $f(x+\alpha e_j) \to \min\limits_{\ell_j \leqslant x_j+\alpha \leqslant w_j}$.
2. Построить индексные множества $J = \{j \mid \ell_j < x_j < w_j\}$, $\tilde{J} = \{1 : n\} \setminus J$. Если $J = \emptyset$, то перейти к шагу 4.
3. С помощью соответствующих перестановок представить матрицу Q, векторы q и x в виде
$$\tilde{Q} = \begin{bmatrix} B & C^\top \\ C & D \end{bmatrix}, \ \tilde{q} = \begin{bmatrix} b \\ c \end{bmatrix}, \ \tilde{x} = \begin{bmatrix} y \\ z \end{bmatrix},$$
где $B = \left(Q_{i\in J, j\in J}\right) \in \mathbb{R}^{|J|\times|J|}$, $C = \left(Q_{i\in \tilde{J}, j\in J}\right) \in \mathbb{R}^{\left|\tilde{J}\right|\times|J|}$, $b = (q_{i\in J}) \in \mathbb{R}^{|J|}$,
$c = \left(q_{i\in\tilde{J}}\right) \in \mathbb{R}^{\left|\tilde{J}\right|}$, $y = (x_{i\in J}) \in \mathbb{R}^{|J|}$, $z = \left(x_{i\in\tilde{J}}\right) \in \mathbb{R}^{\left|\tilde{J}\right|}$.

   С помощью метода сопряженных градиентов решить вспомогательную задачу вида $\tilde{f}(y) = \frac{1}{2} y^\top B y + \left(C^\top z + c\right)^\top y \to \min$.

   По решению задачи вектору y^* и фиксированному вектору z восстановить новое значение вектора x.
4. Вычислить $g = \nabla f(x) = Qx+q = (g_j) \in \mathbb{R}^n$ и проверить условие оптимальности:
$$\forall j = 1,\ldots,n \Rightarrow (x_j = l_j \wedge g_j \geqslant 0) \vee (x_j = u_j \wedge g_j \leqslant 0) \vee (l_j < x_j < u_j \wedge g_j = 0)\,.$$
   Если условие выхода выполнено, то положить $x^* = x$, работу алгоритма закончить. В противном случае перейти к шагу 1.

Замечание. Очевидно, что $f(x+\alpha e_j) = \frac{1}{2} Q_{j,j} \cdot \alpha^2 + \left(x^\top Q^{<j>} + q_j\right) \cdot \alpha + f(x) = \varphi(\alpha)$, где $Q^{<j>}$ — j-й столбец матрицы Q. Таким образом, при $Q_{j,j} \neq 0$ задача представляет собой задачу минимизации скалярной квадратичной функции на отрезке $[\ell_j - x_j, w_j - x_j]$.

ГЕНЕРАЦИЯ ЗАДАЧ

Вариант 1. Вектор x^* случайным образом равняется одной из границ «коробки», т.е. $x^* = \ell$ или $x^* = w$.

Вариант 2. Вектор x^* случайная внутренняя точка «коробки», т.е. $\ell < x^* < w$.

Вариант 3. Часть компонент (например, 50%) вектора x^* случайным образом оказывается внутри (соответствующих границ) «коробки», а оставшиеся часть на границе.

ГЕНЕРАЦИЯ ЗАДАЧ

$$\begin{cases} Ax - v = b, \\ Dx - A^{\mathsf{T}}u = -c, \\ v \geqslant \mathbf{0},\ u \geqslant \mathbf{0},\ v^{\mathsf{T}}u = 0, \\ \ell \leqslant x \leqslant w, \end{cases} \qquad \text{где} \quad A = \begin{pmatrix} E \\ -E \end{pmatrix}, \quad b = \begin{pmatrix} \ell \\ -w \end{pmatrix}.$$

1) Фиксируем размерность задачи $n \in \mathbb{N}$. Генерируем симметричную положительно определенную матрицу D, например, диагонального преобладания, т. е. $D = \sum_{k=1}^{n} \lambda_k e_k^{\mathsf{T}} e_k + \widehat{D}$, $\lambda_k > n$, $k = 1 : n$, $\widehat{D}$ — случайная симметричная матрица с компонентами распределенными по нормальному закону.
2) Генерируем два вектора $\ell \in R^n$ и $w \in R^n$ такие, что $\ell < w$.
3) Выбираем x_* в зависимости от варианта: $\ell \leqslant x_* \leqslant w$.
4) Вычисляем $v = Ax_* - b = \begin{pmatrix} x_* - \ell \\ w - x_* \end{pmatrix} \geqslant \mathbf{0}$.
5) Генерируем неотрицательный вектор $u \in R^{2n}$. Корректируем компоненты вектора u по правилу: $\begin{cases} u[j] = 0, & \text{если } v[j] > 0, \\ u[j], & \text{если } v[j] = 0. \end{cases}$
6) Вычисляем вектор $c = A^{\mathsf{T}}u - Dx_* = u_1 - u_2 - Dx_*$, где $u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$.

ВЫЧИСЛИТЕЛЬНЫЕ ЭКСПЕРИМЕНТЫ

ЛИТЕРАТУРА

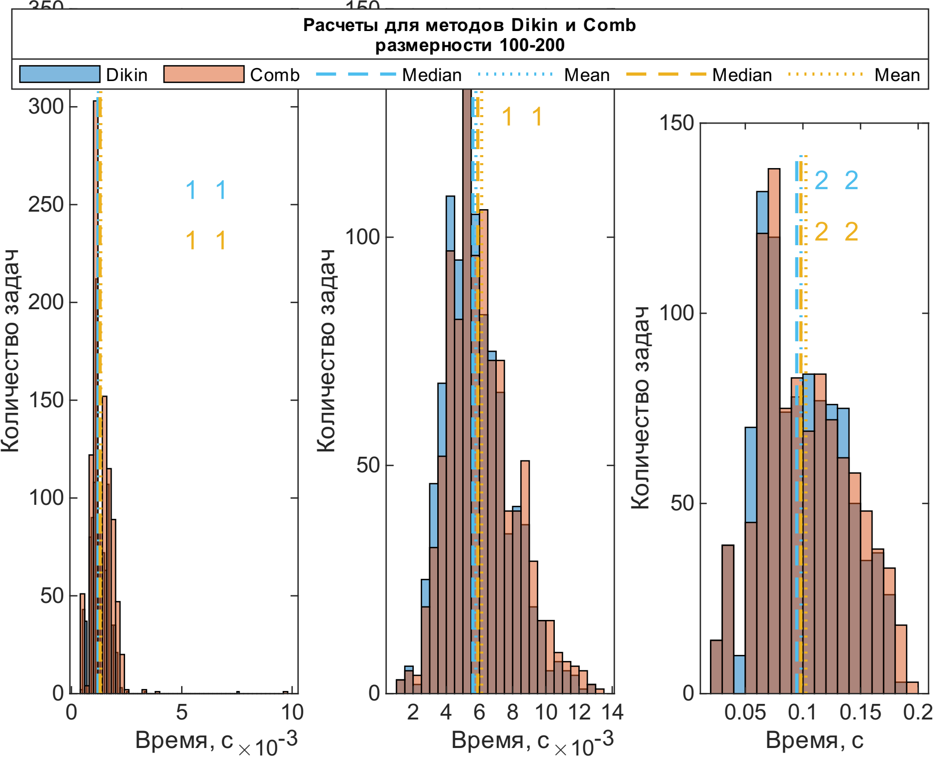
1. Васильев Ф.П. Методы оптимизации: В 2-х кн. Кн. 1. М.: МЦНМО, 2011. 620 с.
2. Гавурин М.К., Малозёмов В.Н. Экстремальные задачи с линейными ограничениями: Учебное пособие. Л.: Изд-во Ленингр. ун-та, 1984. 176 с.
3. Даугавет В.А. Численные методы квадратичного программирования. СПб.: Изд-во С.-Петерб. ун-та, 2004. 128 с.
4. Малоземов В.Н., Плоткин А.В. Двойственность в квадратичном программировании. Задача Сильвестра // Семинар «CNSA & NDO». Избранные доклады. 8 декабря 2021 г. (http://cnsa.cmlaboratory.com/reps21.shtml#1208)
5. Малоземов В.Н. О методе сопряженных градиентов // Семинар «DHA & CAGD». Избранные доклады. 28 апреля 2012 г. (http://dha.spb.ru/reps12.shtml#0428)
6. Дикин И.И., Зоркальцев В.И. Итеративное решение задач математического программирования (алгоритмы метода внутренних точек). Новосибирск: Наука, 1980. 143 с.

БЛАГОДАРИМ ЗА ВНИМАНИЕ![*)]

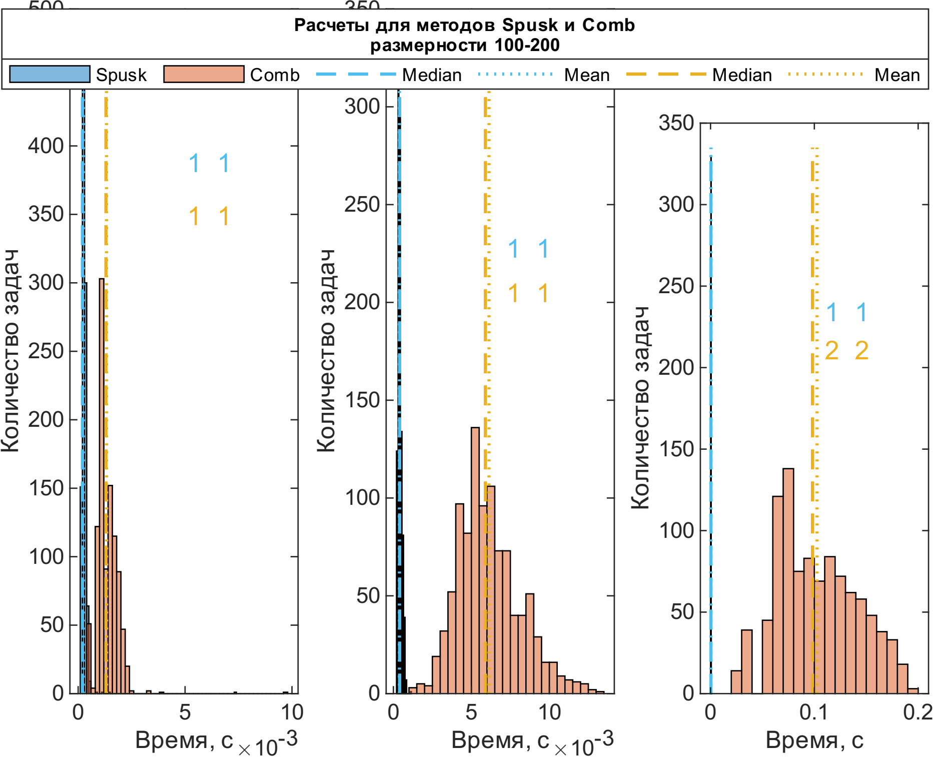
Ерохин Владимир Иванович, д.ф.-м.н., проф.,
с.н.с., ВКА им. Можайского, erohin_v_i@mail.ru

Тамасян Григорий Шаликович, к.ф.-м.н., с.н.с., Институт проблем машиноведения РАН, ВКА им. Можайского, grigoriytamasjan@mail.ru

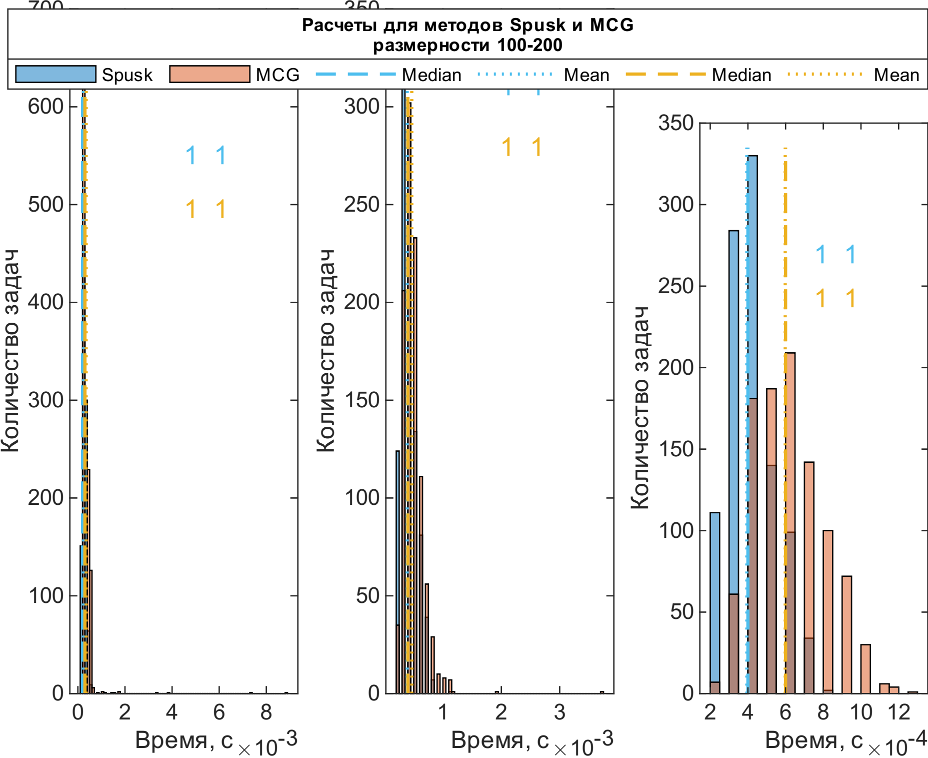
[*)]Работа выполнена в Институте проблем Машиноведения РАН при финансовой поддержке Российского научного фонда (проект № 23-41-00060).



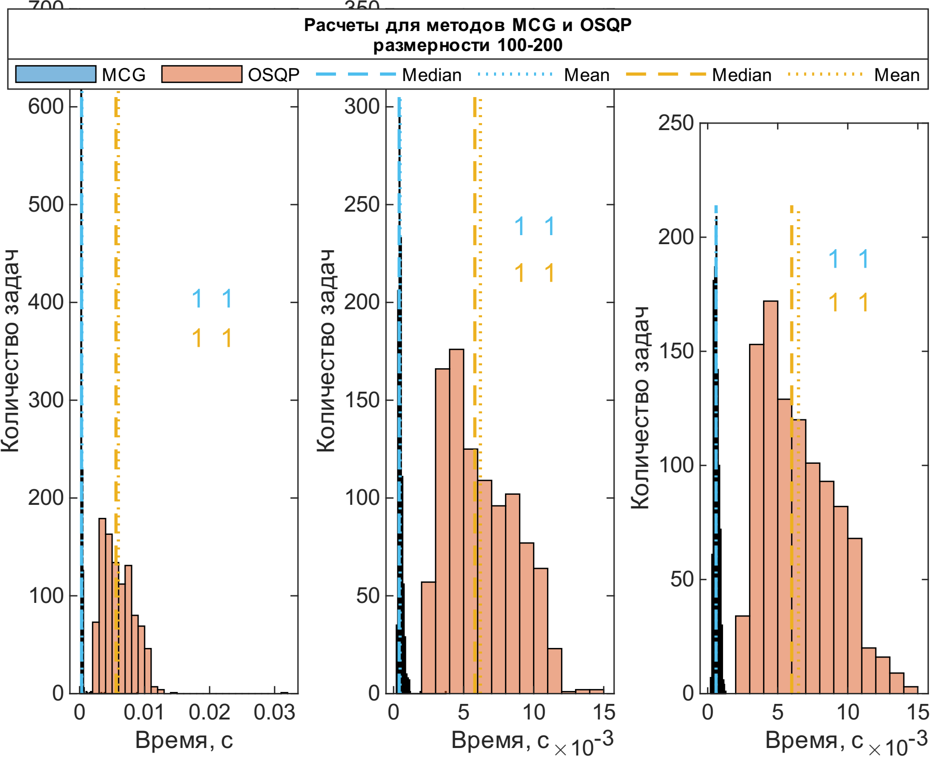
Расчеты для методов Dikin и Comb
размерности 100-200
Dikin
Comb
Median
Mean
Median
Mean
Количество задач
Время, с
$\times 10^{-3}$
1 1
1 1
2 2
2 2



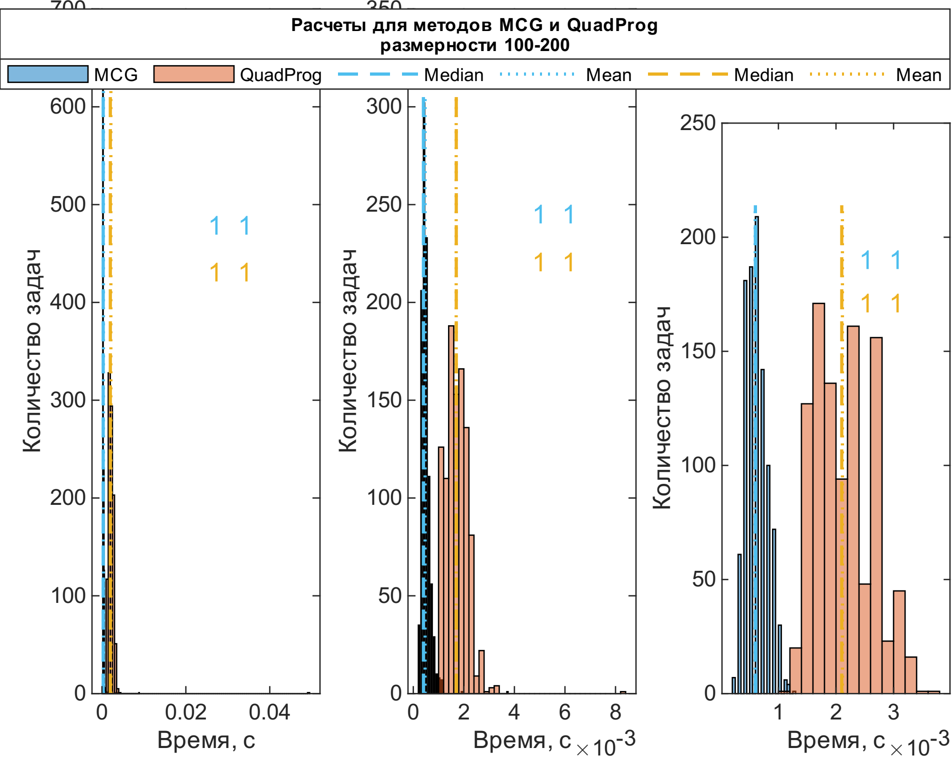
Расчеты для методов Spusk и Comb
размерности 100-200
Spusk
Comb
Median
Mean
Median
Mean
1 1
1 1
1 1
1 1
1 1
2 2
Количество задач
Время, с
$\times 10^{-3}$



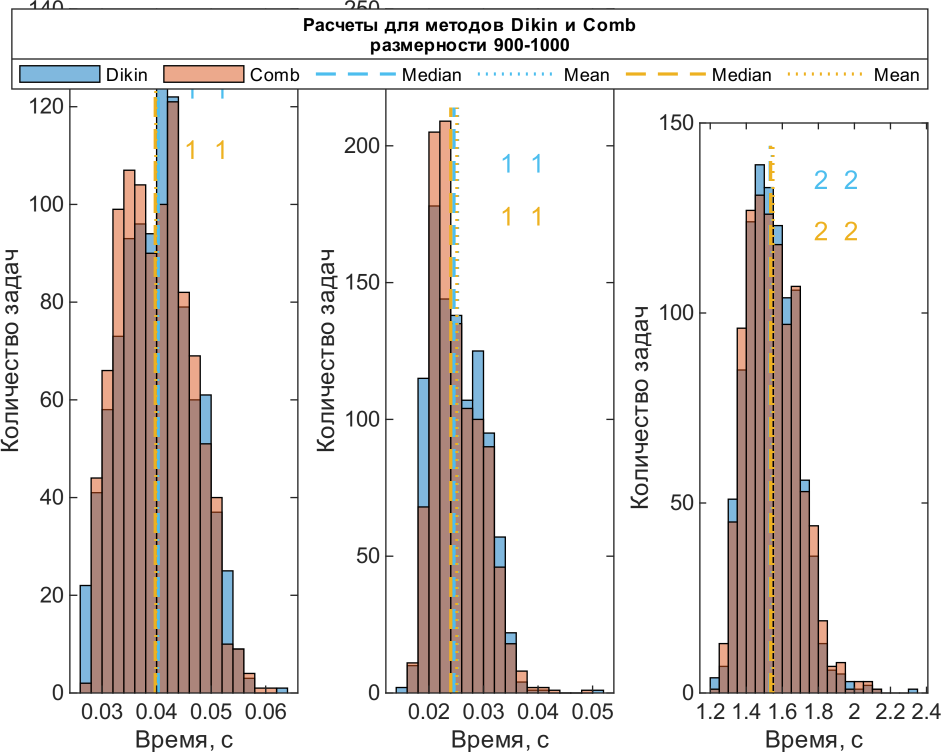
Расчеты для методов Spusk и MCG
размерности 100-200
Spusk
MCG
Median
Mean
Median
Mean
Количество задач
Время, с
$\times 10^{-3}$
$\times 10^{-3}$
$\times 10^{-4}$



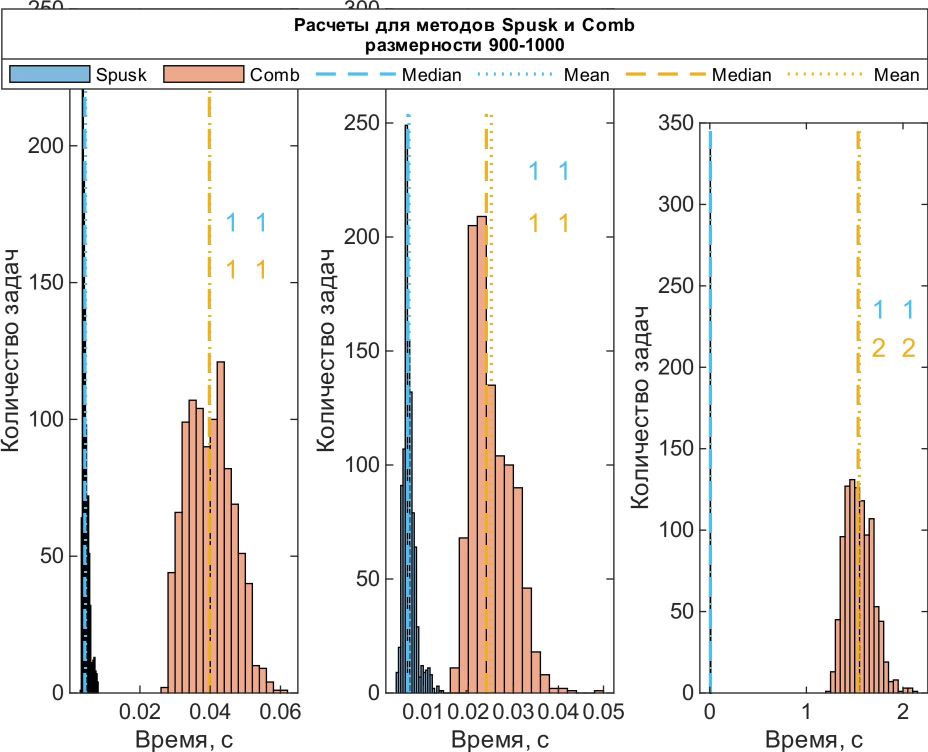
Расчеты для методов MCG и OSQP
размерности 100-200
MCG
OSQP
Median
Mean
Median
Mean
Количество задач
Время, с
Количество задач
Время, с
$\times 10^{-3}$
Количество задач
Время, с
$\times 10^{-3}$



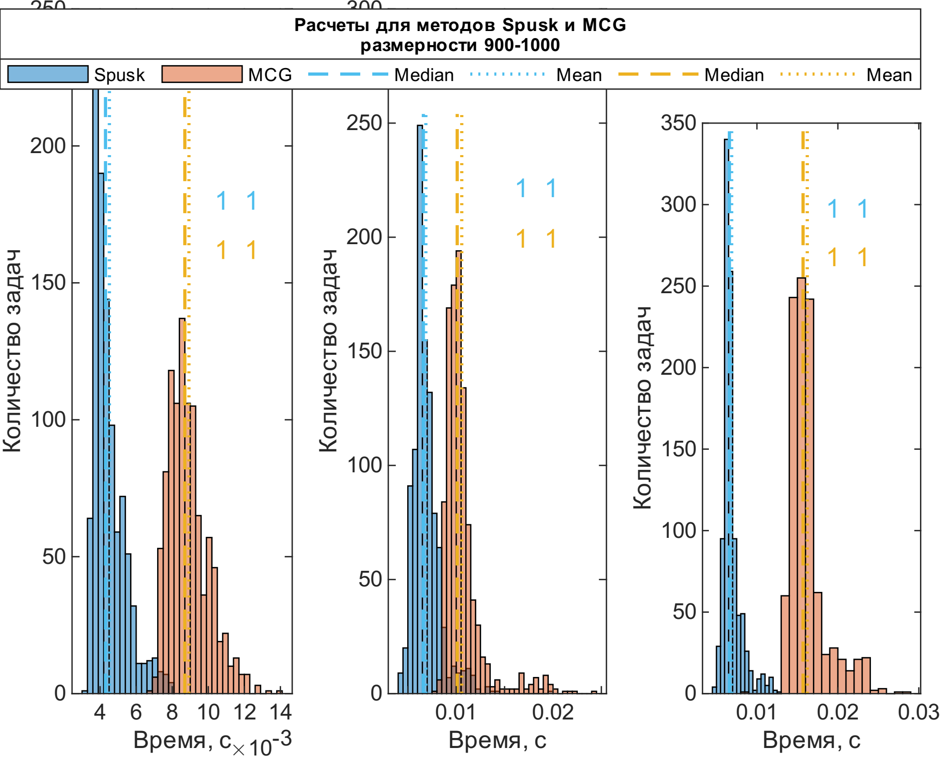
Расчеты для методов MCG и QuadProg
размерности 100-200
MCG
QuadProg
Median
Mean
Median
Mean
Количество задач
Время, с
Количество задач
Время, с
$\times 10^{-3}$
Количество задач
Время, с
$\times 10^{-3}$



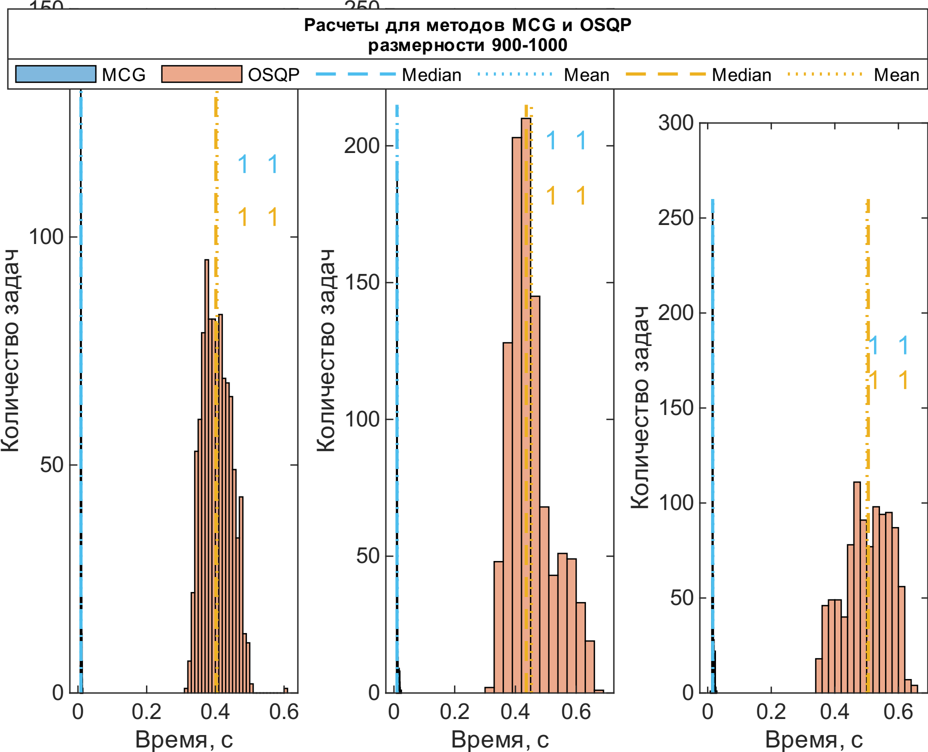
Расчеты для методов Dikin и Comb
размерности 900-1000
Dikin
Comb
Median
Mean
Median
Mean
Количество задач
Время, с
1 1
1 1
1 1
2 2
2 2

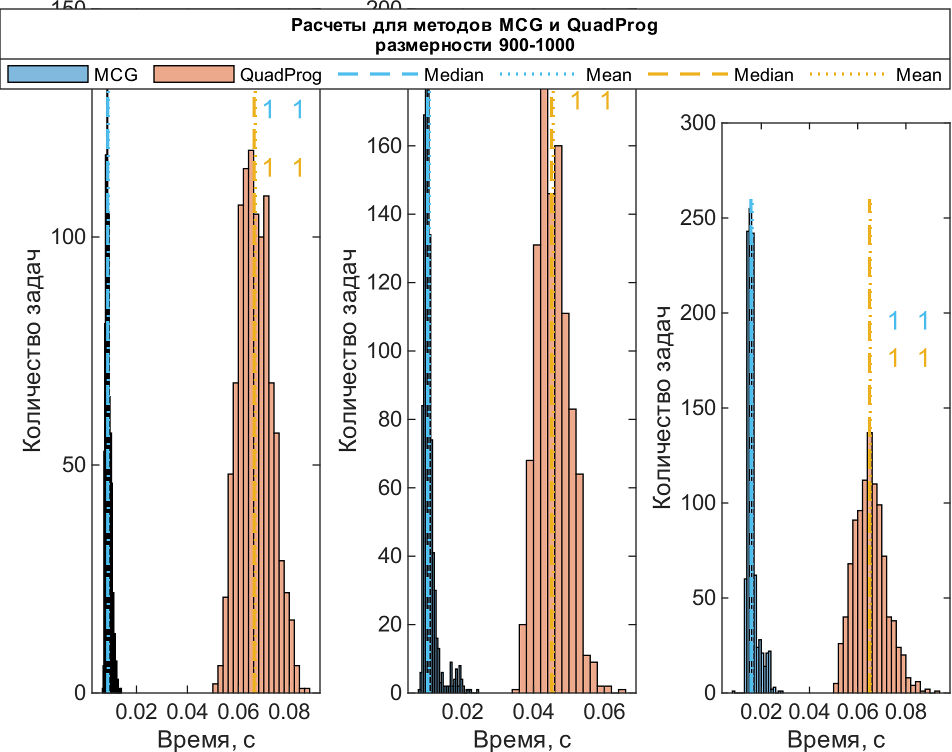


Расчеты для методов Spusk и Comb
размерности 900-1000
Spusk
Comb
Median
Mean
Median
Mean
Количество задач
Время, с
1 1
1 1
2 2



Расчеты для методов Spusk и MCG
размерности 900-1000
Spusk
MCG
Median
Mean
Median
Mean
Количество задач
Время, с
$\times 10^{-3}$
Количество задач
Время, с
Количество задач
Время, с


Расчеты для методов MCG и OSQP
размерности 900-1000
MCG
OSQP
Median
Mean
Median
Mean
Количество задач
Время, с
0
50
100
150
200
250
300
0.2
0.4
0.6
1



Расчеты для методов MCG и QuadProg
размерности 900-1000
MCG
QuadProg
Median
Mean
Median
Mean
Количество задач
Время, с
0.02
0.04
0.06
0.08
0
50
100
20
40
60
80
120
140
160
150
200
250
300