

ЧИСЛЕННОЕ СРАВНЕНИЕ ОПТИМИЗАЦИОННЫХ АЛГОРИТМОВ

Е. И. Усков

Работа посвящена численному сравнению эффективности и робастности наиболее успешных и популярных в последние годы оптимизационных пакетов общего назначения: ALGENCAN, реализующего метод модифицированных функций Лагранжа; MINOS, реализующего метод множителей с линеаризованными ограничениями; SNOPT и filterSQP, реализующих метод последовательного квадратичного программирования; KNITRO и IPOPT, реализующих метод внутренней точки.

1. Введение

Будем рассматривать задачу математического программирования

$$f(x) \rightarrow \min, \quad F(x) = 0, \quad G(x) \leq 0, \quad (1)$$

где $f: \mathbb{R}^n \rightarrow \mathbb{R}$ — гладкая функция, а $F: \mathbb{R}^n \rightarrow \mathbb{R}^l$, $G: \mathbb{R}^n \rightarrow \mathbb{R}^m$ — гладкие отображения.

Напомним, что точка $\bar{x} \in \mathbb{R}^n$ называется стационарной точкой задачи (1), если существуют такие $\bar{\lambda} \in \mathbb{R}^l$ и $\bar{\mu} \in \mathbb{R}^m$, что тройка $(\bar{x}, \bar{\lambda}, \bar{\mu})$ удовлетворяет системе Каруша–Куна–Таккера (ККТ) задачи (1):

$$\begin{aligned} \frac{\partial L}{\partial x}(x, \lambda, \mu) &= 0, & F(x) &= 0, \\ \mu &\geq 0, & G(x) &\leq 0, & \langle \mu, G(x) \rangle &= 0. \end{aligned}$$

Здесь $L: \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}$,

$$L(x, \lambda, \mu) = f(x) + \langle \lambda, F(x) \rangle + \langle \mu, G(x) \rangle$$

есть функция Лагранжа задачи (1). При этом $\bar{\lambda}$ и $\bar{\mu}$ называются множителями Лагранжа, отвечающими стационарной точке \bar{x} ; множество пар $(\bar{\lambda}, \bar{\mu})$ таких множителей будем обозначать через $\mathcal{M}(\bar{x})$.

Среди методов решения задачи (1) важнейшее место занимают *методы последовательного квадратичного программирования* (SQP, от английского «Sequential quadratic programming»). Эти методы генерируют траекторию $\{x^k\} \subset \mathbb{R}^n$ следующим образом: по текущему приближению x^k очередное приближение x^{k+1} ищется как стационарная точка задачи квадратичного программирования

$$\begin{aligned} \langle f'(x^k), x - x^k \rangle + \frac{1}{2} \langle H_k(x - x^k), x - x^k \rangle &\rightarrow \min, \\ F(x^k) + F'(x^k)(x - x^k) &= 0, \\ G(x^k) + G'(x^k)(x - x^k) &\leq 0. \end{aligned} \quad (2)$$

Здесь H_k — симметрическая $n \times n$ -матрица, которая в некотором смысле аппроксимирует $\frac{\partial^2 L}{\partial x^2}(\bar{x}, \bar{\lambda}, \bar{\mu})$ при $k \rightarrow \infty$. Например, можно полагать

$$H_k = \frac{\partial^2 L}{\partial x^2}(x^k, \lambda^k, \mu^k),$$

если параллельно с прямой траекторией $\{x^k\}$ генерировать двойственную траекторию $\{(\lambda^k, \mu^k)\}$, скажем, следующим образом: по текущим λ^k и μ^k очередная пара $(\lambda^{k+1}, \mu^{k+1})$ определяется как пара множителей Лагранжа, отвечающих стационарной точке x^{k+1} задачи (2).

Методы последовательного квадратичного программирования лежат в основе ряда весьма успешных солверов, таких как SNOPT [11] и filterSQP [7]. Конкретная реализация метода

в каждом из солверов различается; в частности, используются различные стратегии глобализации сходимости: в SNOPT используется прямо-двойственный одномерный поиск, в то время как в filterSQP глобализация осуществляется с помощью метода доверительной области с применением фильтра.

Другой важный класс образуют *методы модифицированных функций Лагранжа* (или *методы множителей Лагранжа*), в основе которых лежит идея замены исходной задачи (1) последовательностью задач безусловной оптимизации.

Определим семейство модифицированных функций Лагранжа $L_c : \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m \rightarrow \mathbb{R}$ для задачи (1):

$$L_c(x, \lambda, \mu) = f(x) + \frac{1}{2c}(\|\lambda + cF(x)\|_2^2 + \|\max\{0, \mu + cG(x)\}\|_2^2),$$

где $c > 0$ — параметр штрафа, а максимум берется покомпонентно. По текущему двойственному приближению $(\lambda^k, \mu^k) \in \mathbb{R}^l \times \mathbb{R}^m$ и текущему значению параметра штрафа $c_k > 0$ метод множителей генерирует очередное прямо-двойственное приближение $(x^{k+1}, \lambda^{k+1}, \mu^{k+1}) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^m$ следующим образом: x^{k+1} вычисляется как стационарная точка задачи безусловной оптимизации

$$L_{c_k}(x, \lambda^k, \mu^k) \rightarrow \min, \quad x \in \mathbb{R}^n, \quad (3)$$

а λ^{k+1} и μ^{k+1} вычисляются по формулам

$$\lambda^{k+1} = \lambda^k + c_k F(x^{k+1}), \quad \mu^{k+1} = \max\{0, \mu^k + c_k G(x^{k+1})\}.$$

Данный метод является основой ряда известных пакетов, таких как LANCELOT [10] и ALGENCAN [1].

Разновидностью метода множителей является метод множителей с линеаризованными ограничениями, который традиционно вводится для задачи вида (1), в которой $m = n$ и

$G(x) = -x$, $x \in \mathbb{R}^n$, т.е. для задачи с ограничениями равенствами и с условием неотрицательности переменных:

$$f(x) \rightarrow \min, \quad F(x) = 0, \quad x \geq 0. \quad (4)$$

По текущему приближению $(x^k, \lambda^k, \mu^k) \in \mathbb{R}^n \times \mathbb{R}^l \times \mathbb{R}^n$ и значению параметра штрафа $c_k \geq 0$ новое прямое приближение $x^{k+1} \in \mathbb{R}^n$ вычисляется как стационарная точка задачи оптимизации

$$\begin{aligned} f(x) + \langle \lambda^k, F(x) \rangle + \frac{c_k}{2} \|F(x)\|_2^2 &\rightarrow \min, \\ F(x^k) + F'(x^k)(x - x^k) &= 0, \quad x \geq 0, \end{aligned} \quad (5)$$

а новое двойственное приближение $(\lambda^{k+1}, \mu^{k+1}) \in \mathbb{R}^l \times \mathbb{R}^n$ выбирается таким образом, чтобы пара $(\lambda^{k+1} - \lambda^k, \mu^{k+1})$ была отвечающим x^{k+1} множителем Лагранжа. Целевая функция задачи (5) представляет собой модифицированную функцию Лагранжа (за исключением не зависящего от x слагаемого $\|\lambda^k\|^2/(2c_k)$), включающую в себя только ограничения-равенства задачи (4). Именно этот метод лежит в основе хорошо известного пакета MINOS [13].

Еще один важный класс методов решения задачи (1) составляют *методы внутренней точки*. Идея методов данного класса заключается в том, чтобы заменить исходную задачу (1) последовательностью задач следующего вида:

$$\begin{aligned} f(x) + \sigma_k \psi(s) &\rightarrow \min, \\ F(x) = 0, \quad G(x) + s &= 0. \end{aligned} \quad (6)$$

Здесь $s \in \mathbb{R}^m$, $\{\sigma_k\} \subset \mathbb{R}_+ \setminus \{0\}$ — монотонно убывающая последовательность барьерных параметров, стремящаяся к нулю, а $\psi: \mathbb{R}_+^m \rightarrow \mathbb{R}$ — функция барьера, которая неограниченно возрастает при приближении к границе неотрицательного

органта \mathbb{R}_+^m . Наиболее часто используется логарифмический барьер

$$\psi(s) = - \sum_{i=1}^m \log s_i.$$

Методы внутренней точки являются привлекательной альтернативой упомянутым выше подходам в случаях большого числа ограничений-неравенств. Большой интерес к методам данного класса объясняется их сравнительно высокой эффективностью, особенно, на задачах большой размерности (более 10^4 переменных или ограничений). В частности, эти методы лежат в основе таких хорошо известных солверов, как IPOPT [12], [17] и KNITRO [9], [5]. Оба солвера реализуют описанную выше идею, однако детали реализации, такие как алгоритм решения задачи (6) при фиксированном σ_k , несколько различаются.

На сегодняшний день перечисленные методы являются наиболее часто используемыми стратегиями численного решения задачи (1), в связи с чем возникает вопрос об их относительной эффективности и робастности. Вместе с тем, существует лишь несколько работ, сравнивающих различные алгоритмы по данным критериям на больших наборах тестовых задач. Например, в [17] авторы сравнивают между собой солверы IPOPT, KNITRO и LOQO, реализующие методы внутренней точки. В [2] KNITRO и LOQO сравниваются с солвером SNOPT, а в [14] к ним добавляется filterSQP. В работе [3] сравниваются реализации метода множителей: LANCELOT и ALGENCAN. Наконец, в работе [8] сравниваются все приведенные в данном разделе алгоритмы, однако эксперименты проводятся только на специальных классах задач, а именно, на задачах с нерегулярными ограничениями.

Целью настоящей работы является численное сравнение перечисленных выше алгоритмов по эффективности и робаст-

ности.

2. Численные результаты

В вычислительном эксперименте, результаты которого представлены в данном разделе, использовались следующие версии солверов: SNOPT 7.2-8, MINOS 5.51, IPOPT 3.8.0, KNITRO 7.0.0, filterSQP 20020316 и ALGENCAN 2.3.7, скомпилированный с использованием библиотеки MA57 [15]. Данная библиотека предназначена для эффективного решения разреженных систем линейных уравнений, что позволяет существенно повысить общую эффективность солверов на задачах большой размерности.

Эксперимент проводился на AMPL-версии [16] тестовой коллекции CUTE [4]. Данная коллекция включает в себя 734 задачи, среди которых есть как очень маленькие, так и очень большие (вплоть до порядка $2 \cdot 10^4$ переменных и ограничений).

Сравнение солверов проводилось по следующим четырем показателям: затраченное процессорное время, количество внешних итераций, а также количество вычислений целевой функции и ограничений. Количество внешних итераций отражает скорость сходимости алгоритма, а остальные характеристики могут рассматриваться как различные показатели его эффективности. Отметим, что MINOS и SNOPT сообщают только количество вычислений для нелинейных функций. Поэтому для задач, в которых целевая функция или все ограничения являются линейными, соответствующие показатели устанавливались равными 1 для всех солверов.

Результаты эксперимента представлены в форме так называемых «performance profiles» (см. [6]). Для каждого из алгоритмов значение функции, график которой изображен на рисунке, в точке $\tau \in [1, \infty)$ есть доля задач в наборе, на которых результат данного алгоритма (т.е. количество итераций

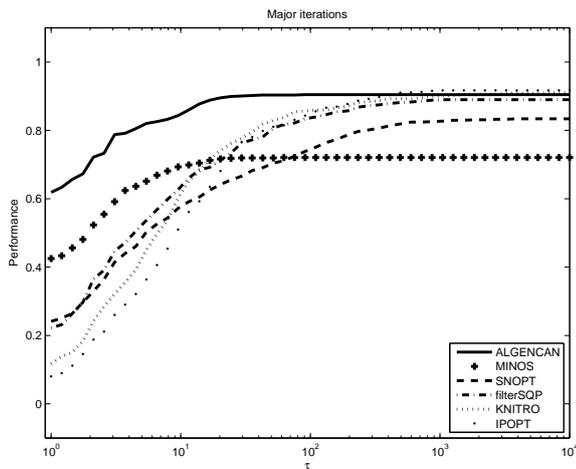


Рис. 1: Внешние итерации

или вычислений функций) был не более чем в τ раз хуже наилучшего результата для данной задачи среди всех сравниваемых алгоритмов. При этом считается, что результат неудачного запуска в бесконечное число раз хуже любого другого результата. В частности, значение этой функции при $\tau = 1$ соответствует доле запусков, на которых результат данного алгоритма был наилучшим. Значение функции при больших τ характеризует робастность алгоритма, т. е. долю его успешных запусков.

На рис. 1–3 представлены результаты для всех задач коллекции. Как видно из рисунков, все солверы, кроме MINOS и SNOPT имеют практически одинаковую робастность, и завершаются успешно примерно на 90% задач. При этом SNOPT незначительно проигрывает остальным солверам, а вот MINOS уступает гораздо более значительно и завершается успешно лишь на 70% задач.

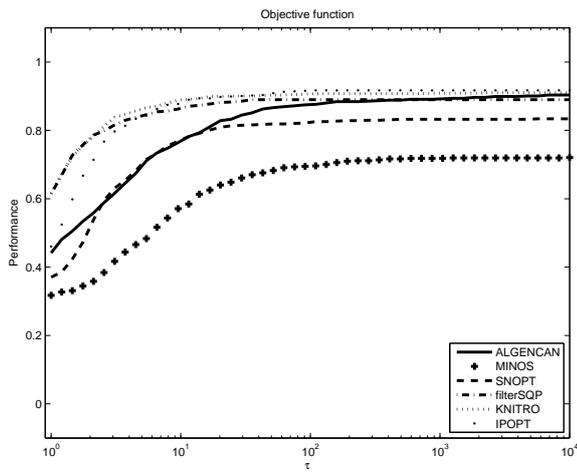


Рис. 2: Вычисления целевой функции

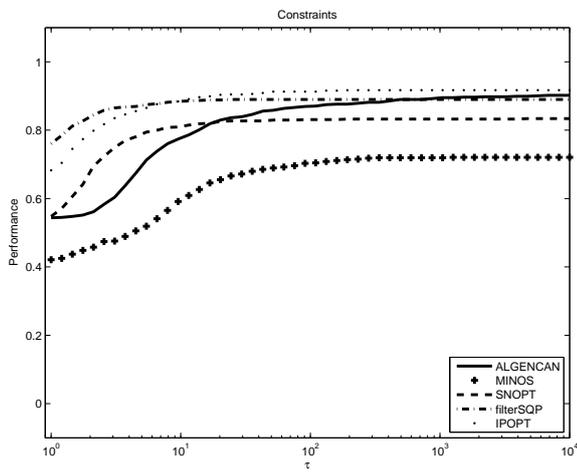


Рис. 3: Вычисления ограничений

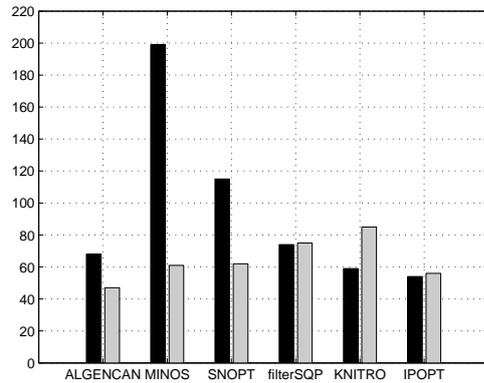


Рис. 4: Неудачные запуски и случаи сходимости к неоптимальным значениям

На рис. 1 показано сравнение солверов по количеству внешних итераций, а на рис. 2 и 3 — по количеству вычислений целевой функции и ограничений. Нетрудно видеть, что ALGENCAN существенно опережает остальные солверы по количеству внешних итераций, что говорит о более высокой скорости сходимости. Тем не менее, из-за сравнительно высокой трудоемкости подзадач вида (3) ALGENCAN уступает по эффективности методам внутренней точки, а также filterSQP.

Диаграмма на рис. 4 иллюстрирует результаты сравнения солверов по количеству неудачных запусков (отмечены черным цветом) и случаев успешной сходимости к неоптимальным значениям (отмечены серым цветом). Значение целевой функции, найденное солвером, считалось неоптимальным, если оно превосходило наилучшее найденное солверами значение более чем на 10^{-2} . Из диаграммы видно, что количество случаев сходимости к оптимальному значению примерно одинаковое для всех солверов, кроме MINOS и SNOPT, которые

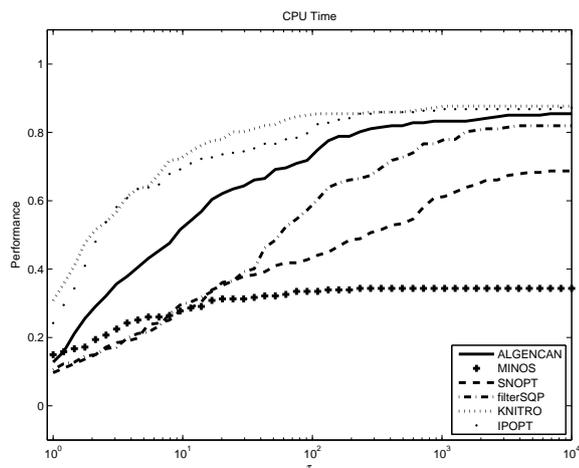


Рис. 5: Затраченное процессорное время (большие задачи)

несколько уступают остальным солверам.

В дальнейшем эксперименте тестовая коллекция была разбита на две части, в первую из которых были включены большие задачи (в которых число переменных или ограничений не менее 500), во вторую — все остальные.

На рис. 5 и 6 представлены результаты сравнения по времени и количеству вычислений целевой функции на больших задачах. Как видно из рисунков, наибольшей эффективностью на больших задачах обладают методы внутренней точки. При этом ALGENCAN обгоняет filterSQP по времени, однако проигрывает по количеству вычислений целевой функции.

Рис. 7 иллюстрирует результаты сравнения по количеству вычислений целевой функции на небольших задачах. В этом случае наиболее эффективным является filterSQP, однако следующими по эффективности по-прежнему являются методы внутренней точки. Заметим также, что, как и следовало ожи-

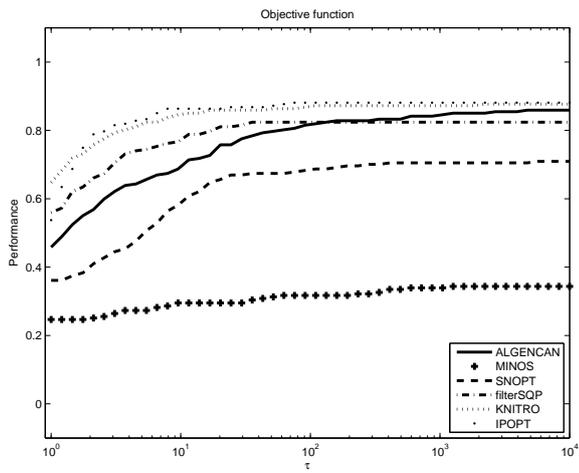


Рис. 6: Вычисления целевой функции (большие задачи)

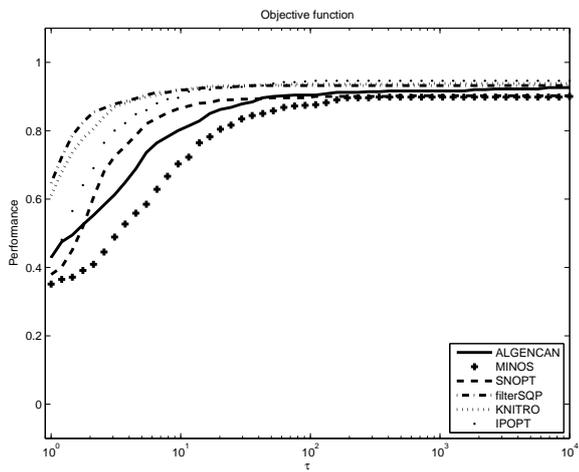


Рис. 7: Вычисления целевой функции (небольшие задачи)

дать, на небольших задачах робастность всех солверов (в особенности SNOPT и MINOS) существенно выше, чем на больших.

Таким образом, методы внутренней точки оказываются наиболее эффективными на задачах большой размерности, уступая на небольших задачах методам последовательного квадратичного программирования. Методы множителей, несмотря на более высокую скорость сходимости, несколько уступают по эффективности обоим перечисленным методам.

Работа выполнена при финансовой поддержке РФФИ (проект 10-01-00251).

Литература

1. ALGENCAN. <http://www.ime.usp.br/~egbirgin/tango/>.
2. *Benson H.Y., Shanno D.F., Vanderbei R.J.* A comparative study of large scale nonlinear optimization algorithms. Technical Report ORFE 01-04. Princeton University, New Jersey, 2001.
3. *Birgin E.G., Martínez J.M.* Augmented Lagrangian method with nonmonotone penalty parameters for constrained optimization // *Comput. Optim. Appl.* 2012. V. 51. № 3. P. 941–965.
4. *Bongartz I., Conn A.R., Gould N.I.M., Toint Ph.L.* CUTE: Constrained and Unconstrained Testing Environment // *ACM Transactions on Mathematical Software.* 1995. V. 21. № 1. P. 123–160.
5. *Byrd R.H., Nocedal J., Waltz R.A.* KNITRO: an integrated package for nonlinear optimization // *Large-Scale Nonlinear Optimization.* New York, Springer-Verlag, 2006. P. 35–59.

6. *Dolan E., Moré J.* Benchmarking optimization software with performance profiles // *Math. Program.* 2002. V. 91. № 2. P. 201–213.
7. *Fletcher R., Leyffer S.* User manual for filterSQP. University of Dundee Numerical Analysis Report NA 181, 1998.
8. *Izmailov A.F., Solodov M.V., Uskov E.I.* Augmented Lagrangian methods applied to optimization problems with degenerate constraints, including problems with complementarity constraints. IMPA preprint A698, 2011.
9. KNITRO. <http://www.ziena.com/knitro.htm>.
10. LANCELOT.
<http://www.cse.scitech.ac.uk/nag/lancelot/lancelot.shtml>.
11. *Gill P.E., Murray W., Saunders M.A.* SNOPT: An SQP algorithm for large-scale constrained optimization // *SIAM J. Optim.* 2002. V. 12. P. 979–1006.
12. IPOPT. <https://projects.coin-or.org/Ipopt>.
13. *Murtagh B.A., Saunders M.A.* MINOS 5.5 user's guide. Technical Report SOL 83.20. Stanford University, 1983.
14. *Nocedal J., Morales J.L., Waltz R., Liu G., Goux J.P.* Assessing the potential of interiorpoint methods for nonlinear optimization. Technical Report OTC-2001-6. Optimization Technology Center, 2001.
15. The HSL Mathematical Software Library.
<http://www.hsl.rl.ac.uk>.
16. *Vanderbei R.J.* AMPL models.
<http://www.sor.princeton.edu/rvdb/ampl/nlmodels>.

17. *Wächter A., Biegler L.T.* On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming // *Math. Program.* 2006. V. 106. P 25–57.