Федеральное государственное автономное образовательное учреждение
высшего образования
«Московский физико-технический институт
(национальный исследовательский университет)»
Физтех-школа Прикладной Математики и Информатики
Кафедра интеллектуальных систем

Направление подготовки / специальность: 03.04.01 Прикладные математика и физика
Направленность (профиль) подготовки: Математическая физика, компьютерные технологии и математическое моделирование в экономике

РЕШЕНИЕ СЕДЛОВЫХ ЗАДАЧ С НЕБОЛЬШОЙ РАЗМЕРНОСТЬЮ ОДНОЙ ИЗ ГРУПП ПЕРЕМЕННЫХ

(магистерская диссертация)

Студент:
Гладин Егор Леонидович

____
(подпись студента)

Научный руководитель:
Гасников Александр Владимирович,
д-р физ.-мат. наук, доц.

____
(подпись научного руководителя)

Консультант *(при наличии)***:**
____
____
(подпись консультанта)

Москва 2021

Решение седловых задач с небольшой размерностью одной из групп переменных

Егор Гладин

Июнь 8, 2021

Реферат

Настоящая диссертация посвящена выпукло-вогнутым седловым задачам с небольшой размерностью одной из групп переменных. Седловые задачи возникают в машинном обучении, компьютерной графике, теории оптимального транспорта и других областях. Предлагаемый подход основан на сведении рассматриваемой проблемы к внешней задаче минимизации с неточным оракулом, который вычисляется на основе приближённого решения внутренней задачи максимизации. Для максимизации используется быстрый градиентный метод, тогда как внешняя задача решается с помощью метода Вайды.

Вспомогательным, но важным результатом диссертации является теорема о сходимости метода Вайды с неточным оракулом. С одной стороны, она обосновывает использование этого метода в предлагаемом подходе, а с другой, может быть применена ко многим другим проблемам. Предлагаемый подход накладывает меньше предположений о задаче, чем многие существующие методы. Несмотря на это, его скорость сходимости не уступает современным алгоритмам.

Дополнительно в работе рассматриваются седловые задачи со смешанным оракулом, под которым понимается возможность вычислять градиент только по отношению ко внешнему блоку переменных, в то время как для внутренней задачи доступен лишь оракул нулевого порядка. В этом случае внутренняя задача решается с помощью ускоренного безградиентного метода. Смешанные оракулы относительно мало исследованы, поэтому их использование является ценным вкладом в область оптимизации.

Научный руководитель:
Имя: Александр В. Гасников
Ученое звание, степень: Доктор физико-математических наук
Должность: Профессор

Contents

Bibliography

[1] *Cox B., Juditsky A., Nemirovski A.* Decomposition Techniques for Bilinear Saddle Point Problems and Variational Inequalities with Affine Monotone Operators // J. Optim. Theory Appl. 2017. V. 172. P. 402–435.

[2] *Hien L.T.K., Zhao R., Haskell W.B.* An inexact primal-dual framework for large-scale non bilinear saddle point problem // arxiv e-print 2019. `https://arxiv.org/pdf/1711.03669.pdf`.

[3] *Lan G.* First-order and Stochastic Optimization Methods for Machine Learning. Switzerland: Springer Series in the Data Sciences, 2020.

[4] *Lin T., Jin C., Jordan M.I.* Near-Optimal Algorithms for Minimax Optimization // arxiv e-print 2020. `https://arxiv.org/pdf/2002.02417.pdf`.

[5] *Nemirovski A.* Information-based complexity of convex programming. Lecture Notes, 1995.

[6] *Nemirovski A., Onn Shmuel, Rothblum U. G.* Accuracy certificates for computational problems with convex structure // Math. Oper. Res. 2010. V. 35. № 1. P. 52–78.

[7] *Nesterov Yu.* Excessive Gap Technique in Nonsmooth Convex Minimization // SIAM J. Optim. 2005. V. 16. № 1. P. 235–249.

[8] *Yuanhao W., Jian Li.* Improved Algorithms for Convex-Concave Minimax Optimization // arxiv e-print 2020. `https://arxiv.org/pdf/2006.06359.pdf`.

[9] *Zhang J., Hong M., Zhang S.* On Lower Iteration Complexity Bounds for the Saddle Point Problems // arxiv e-print 2019. `https://arxiv.org/pdf/1912.07481.pdf`.

[10] *Алкуса М., Двинских Д., Стонякин Ф., Гасников А., Ковалев Д.* Ускоренные методы для седловых задач // ЖВМ и МФ., 2020. Т. 60. № 11. С. 1843–1866 (в печати). `https://arxiv.org/pdf/1906.03620.pdf`.

[11] *Гасников А.В., Двинских Д.М., Двуреченский П.Е. Камзолов Д.И., Матюхин В.В., Пасечнюк Д.А., Тупица Н.К., Чернов А. В.* Ускоренный метаалгоритм для задач выпуклой оптимизации // ЖВМ и МФ., 2020. Т. 60. № 12. `https://arxiv.org/pdf/2004.08691.pdf`.

[12] *Гасников А.В., Двуреченский П.Е., Нестеров Ю.Е.* Стохастические градиентные методы с неточным оракулом // Тр. МФТИ. М., 2016. Т. 8. № 1. С. 41–91.

[13] Aleksandr Beznosikov, Abdurakhmon Sadiev, and Alexander Gasnikov. Gradient-free methods for saddle-point problem. *arXiv preprint arXiv:2005.05913*, 2020.

[14] Antonin Chambolle and Thomas Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *Journal of Mathematical Imaging and Vision*, 40(1):120–145, 2011.

[15] Chao-Ping Chen and Long Lin. Inequalities for the volume of the unit ball in $\mathbb{R}^n$ r n. *Mediterranean Journal of Mathematics*, 11(2):299–314, 2014.

[16] Andrew R. Conn, Katya Scheinberg, and Luis N. Vicente. *Introduction to Derivative-Free Optimization*. Society for Industrial and Applied Mathematics, 2009.

[17] Pavel Dvurechensky, Eduard Gorbunov, and Alexander Gasnikov. An accelerated directional derivative method for smooth stochastic convex optimization. *European Journal of Operational Research*, 290(2):601–621, Apr 2021.

[18] Egor Gladin, Ilya Kuruzov, Fedor Stonyakin, Dmitry Pasechnyuk, Mohammad Alkousa, and Alexander Gasnikov. Solving strongly convex-concave composite saddle point problems with a small dimension of one of the variables. *arXiv preprint arXiv:2010.02280*, 2020.

[19] Egor Gladin, Abdurakhmon Sadiev, Alexander Gasnikov, Pavel Dvurechensky, Aleksandr Beznosikov, and Mohammad Alkousa. Solving smooth min-min and min-max problems by mixed oracle algorithms. *arXiv preprint arXiv:2103.00434*, 2021.

[20] Egor Gladin and Karina Zaynullina. Ellipsoid method for convex stochastic optimization in small dimension. *arXiv preprint arXiv:2011.04462*, 2020.

[21] Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial networks, 2014.

[22] Ian J. Goodfellow, Jonathon Shlens, and Christian Szegedy. Explaining and harnessing adversarial examples, 2014.

[23] G. M. Korpelevich. The extragradient method for finding saddle points and other problems. 1976.

[24] Guanghui Lan, Zhize Li, and Yi Zhou. A unified variance-reduced accelerated gradient method for convex optimization. In H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, and R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc., 2019.

[25] Sijia Liu, Songtao Lu, Xiangyi Chen, Yao Feng, Kaidi Xu, Abdullah Al-Dujaili, Minyi Hong, and Una-May O'Reilly. Min-max optimization without gradients: Convergence and applications to adversarial ml, 2019.

[26] Aleksander Madry, Aleksandar Makelov, Ludwig Schmidt, Dimitris Tsipras, and Adrian Vladu. Towards deep learning models resistant to adversarial attacks. In *6th International Conference on Learning Representations, ICLR 2018, Vancouver, BC, Canada, April 30 - May 3, 2018, Conference Track Proceedings*, 2018.

[27] Nina Narodytska and Shiva Prasad Kasiviswanathan. Simple black-box adversarial attacks on deep neural networks. In *CVPR Workshops*, pages 1310–1318. IEEE Computer Society, 2017.

[28] A. Nedić and A. Ozdaglar. Subgradient methods for saddle-point problems. *Journal of Optimization Theory and Applications*, 142(1):205–228, Jul 2009.

[29] Arkadi Nemirovski. Prox-method with rate of convergence o (1/ t) for variational inequalities with lipschitz continuous monotone operators and smooth convex-concave saddle point problems. *SIAM Journal on Optimization*, 15:229–251, 01 2004.

[30] Yurii Nesterov. *Lectures on convex optimization*, volume 137. Springer International Publishing, 2018.

[31] Yurii E Nesterov. A method for solving the convex programming problem with convergence rate o (1/k^ 2). In *Dokl. akad. nauk Sssr*, volume 269, pages 543–547, 1983.

[32] Lerrel Pinto, James Davidson, Rahul Sukthankar, and Abhinav Gupta. Robust adversarial reinforcement learning. volume 70 of *Proceedings of Machine Learning Research*, pages 2817–2826, International Convention Centre, Sydney, Australia, 06–11 Aug 2017. PMLR.

[33] Boris T Polyak. Introduction to optimization. *Inc., Publications Division, New York*, 1987.

[34] James Renegar. A polynomial-time algorithm, based on newton's method, for linear programming. *Mathematical programming*, 40(1):59–93, 1988.

[35] Abdurakhmon Sadiev, Aleksandr Beznosikov, Pavel Dvurechensky, and Alexander Gasnikov. Zeroth-order algorithms for smooth saddle-point problems. *arXiv:2009.09908*, 2020.

[36] Maurice Sion et al. On general minimax theorems. *Pacific Journal of mathematics*, 8(1):171–176, 1958.

[37] Florian Tramèr, Alexey Kurakin, Nicolas Papernot, Ian Goodfellow, Dan Boneh, and Patrick McDaniel. Ensemble adversarial training: Attacks and defenses, 2017.

[38] Pravin M Vaidya. A new algorithm for minimizing convex functions over convex sets. In *30th Annual Symposium on Foundations of Computer Science*, pages 338–343. IEEE Computer Society, 1989.

[39] Pravin M Vaidya. A new algorithm for minimizing convex functions over convex sets. *Mathematical programming*, 73(3):291–341, 1996.

[40] Zhongruo Wang, Krishnakumar Balasubramanian, Shiqian Ma, and Meisam Razaviyayn. Zeroth-order algorithms for nonconvex minimax problems with improved complexities, 2020.