Міністерство освіти і науки України
Київський національний університет імені Тараса Шевченка

Кваліфікаційна наукова
праця на правах рукопису

ХАРЬКОВ Олег Сергійович

УДК 517.988 : 519.85

ДИСЕРТАЦІЯ

Модифіковані екстраградієнтні алгоритми для варіаційних нерівностей та їх застосування

113 — «Прикладна математика»

Галузь знань 11 — «Математика та статистика»

Подається на здобуття наукового ступеня доктора філософії

Дисертація містить результати власних досліджень. Використання ідей, результатів і текстів інших авторів мають посилання на відповідне джерело

________________________ О. С. Харьков

Науковий керівник:
Семенов Володимир Вікторович
доктор фізико-математичних наук, професор

Київ — 2024

АНОТАЦІЯ

Харьков О. С. Модифіковані екстраградієнтні алгоритми для варіаційних нерівностей та їх застосування. — Кваліфікаційна наукова праця на правах рукопису.

Дисертація на здобуття ступеня доктора філософії за спеціальністю 113 «Прикладна математика» (11 — «Математика та статистика»). — Київський національний університет імені Тараса Шевченка Міністерства освіти і науки України, Київ, 2023.

Із розвитком комп'ютерних технологій виник новий напрям наукових досліджень — комп'ютерне математичне моделювання, що передбачає побудову моделей досліджуваного об'єкта та організацію серій обчислювальних експериментів. Зазвичай складно отримати точні розв'язки для практичних проблем, тому актуальною є розробка алгоритмів для наближених розрахунків. Отже, наближені алгоритми є вкрай важливою складовою для багатьох програмних продуктів, що впливають на наше повсякденне життя.

Дисертація пов'язана з моделями вигляду варіаційних нерівностей. Вони дають простий та уніфікований засіб формулювання багатьох актуальних задач оптимального керування, математичної фізики та дослідження операцій (пошук сідлових точок та рівноваги за Нешем). Варіаційні нерівності з монотонними операторами є загальним класом задач з опуклою структурою. Окремі задачі опуклої недиференційовної оптимізації можуть ефективно розв'язуватися, якщо їх переформулювати у вигляді сідлових (мінімаксних) задач і застосувати алгоритми розв'язання варіаційних нерівностей. З появою генеруючих змагальних нейронних мереж та інших моделей змагального навчання стійкий інтерес до алгоритмів розв'язання варіаційних нерівностей виник і в середовищі спеціалістів в галузі машинного навчання.

Дослідження алгоритмів розв'язання варіаційних нерівностей та близьких задач триває. Г. М. Корпелевич в 1970-х роках запропонувала екстраградієнтний метод. В 1980 році Л. Д. Попов запропонував для пошуку сідлових точок опукло-угнутих функцій цікаву модифікацію методу Ерроу–Гурвіца, який став джерелом багатьох сучасних алгоритмів. Ефективним сучасним варіантом екстраградієнтного методу є проксимальний дзеркальний метод (MirrorProx) А. С. Немировського. Останнім часом пожвавилася дослідницька активність, пов'язана з підвищенням ефективності та універсальності екстраградієнтних методів. Зокрема, запропоновані адаптивні модифікації проксимального дзеркального методу, які не потребують знання констант Ліпшиця операторів для визначення величини кроку. Велике значення мають методи апроксимації нерухомих точок нерозтягуючих операторів, які дозволяють регуляризовувати існуючі методи або будувати нові.

Але питання про ефективні методи розв'язання варіаційних нерівностей великої розмірності досі актуальне. Наприклад, останнім часом збільшився інтерес до розробки адаптивних алгоритмів, що не використовують значень ліпшицевих констант операторів та на відміну від схем з лінійним пошуком не потребують обчислень значень біфункції в додаткових точках. Подальші спроби зменшити складність виконання ітерації призвели до появи алгоритму операторної екстраполяції. За об'ємом необхідних для здійснення ітераційного кроку обчислень алгоритми мають перевагу над екстаградієнтним методом. Також останнім часом пропонуються алгоритми розв'язання варіаційних задач в банахових просторах та алгоритми, що використовують дивергенцію Брегмана.

Дисертаційна робота складається зі вступу та п'яти розділів, перший з яких містить огляд літератури за темою дисертації.

У другому розділі дисертації розглядались варіаційні нерівності з операторами, що діють в гільбертовому просторі. Для цих задач запропоновано та обгрунтовано нові алгоритми екстраградієнтного типу, які є модифікаціями методу Л. Д. Попова пошуку сідлових точок опукло-угнутих функцій. Також отримано нові результати для відомих варіантів методу.

А саме, отримано сублінійну оцінку ефективності для функції зазору. Доведено сильну збіжність алгоритму екстраполяції з минулого для варіаційних нерівностей з рівномірно монотонними операторами. Доведено лінійну швидкість збіжності алгоритму екстраполяції з минулого для варіаційних нерівностей з операторами, що задовольняють умову типу узагальненої сильної монотонності. Запропоновано та обгрунтовано адаптивний варіант алгоритму. Правило оновлення параметрів не використовує значень ліпшицевих констант операторів та на відміну від правил типу лінійного пошуку не потребує обчислень значень операторів в додаткових точках. Запропоновано регуляризовані варіанти алгоритму та доведено теореми про їх сильну збіжність.

У третьому розділі досліджено нові ітераційні алгоритми для розв'язання варіаційних нерівностей в гільбертових просторах. Дані алгоритми є модификаціями відомого «forward-reflected-backward algorithm».

А саме, доведено слабку збіжність алгоритму операторної екстраполяції та отримано сублінійну оцінку ефективності для функції зазору. Доведено сильну збіжність алгоритму операторної екстраполяції для варіаційних нерівностей з рівномірно монотонними операторами. Доведено лінійну швидкість збіжності алгоритму операторної екстраполяції для варіаційних нерівностей з операторами, що задовольняють умову типу узагальненої сильної монотонності. Запропоновано та обгрунтовано адаптивний варіант алгоритму. Запропоновано регуляризований варіант алгоритму та доведено теорему про його сильну збіжність. Для регуляризації використано схему Гальперна.

У четвертому розділі досліджено варіанти алгоритмів екстраполяції з минулого та операторної екстраполяції з дивергенцією Брегмана для розв'-

язання варіаційних нерівностей з монотонними та ліпшицевими операторами, що діють в скінченновимірному дійсному лінійному просторі. Основні результати: оцінки ефективності в термінах функції зазору.

Також досліджено два нові алгоритми для розв'язання варіаційних нерівностей з монотонними, ліпшицевими операторами, що діють в 2-рівномірно опуклих та рівномірно гладких банахових просторах. Перший алгоритм — варіант методу операторої екстраполяції, що використовує узагальнену проєкцію Альбера замість метричної. Другий алгоритм є адаптивним варіантом першого, де використовується правило поновлення величини кроку, що не вимагає знання ліпшицевих констант та лінійного пошуку. Доведено сублінійну оцінку ефективності в термінах функції зазору та теорему про слабку збіжність першого методу. Доведено слабку збіжність адаптивного варіанту методу, де використовується правило поновлення величини кроку, що не вимагає знання ліпшицевих констант та лінійного пошуку.

У п'ятому розділі дисертації описано експерименти, проведені для перевірки теорем з другого та третього розділів, що містять оцінки швидкості збіжності розглянутих алгоритмів. Розглядались білінійні та квадратичні сідлові задачі. Результати проведених експериментів підтверджують теоретичні оцінки.

Також проведено експерименти з навчанням простих GANs за допомогою двох рандомізованих варіантів методу екстраполяції з минулого з експоненційним усередненням. Результати мають попередній характер. Зауважимо, що задача про теоретичні гарантії збіжності цих варіантів є відкритою.

Проведені дослідження лежать в руслі сучасних досліджень варіаційних нерівностей, інтерес до яких достатньо високий. Результати дисертації призначені для використання при створенні спеціалізованого програмного забезпечення для комп'ютерного моделювання в дослідженні операцій, математичній економіці та математичній фізиці. Також розроблені алгоритми можуть бути використані для розв'язання задач машинного навчання, обробки зображень тощо. Окремі результати, одержані в роботі, було впроваджено у навчальний процес кафедри обчислювальної математики факультету комп'ютерних наук та кібернетики Київського національного університету імені Тараса Шевченка при викладанні дисциплін «Сучасні проблеми аналізу» та «Некласичні задачі оптимального керування». Декілька алгоритмів було розроблено при виконанні наукових проєктів кафедри обчислювальної математики факультету комп'ютерних наук та кібернетики Київського національного університету імені Тараса Шевченка, що виконувались протягом 2020–2023 років.

Ключові слова: варіаційна нерівність, сідлова задача, рівновага, оптимізація, екстраградієнтний метод, метод екстраполяції з минулого, метод операторної екстраполяції, гільбертовий простір, банаховий простір, дивергенція Брегмана, апроксимація, збіжність.

ABSTRACT

Kharkov O. S. Modified Extragradient Algorithms for Variational Inequalities and Their Applications. — Qualifying scientific work as a manuscript.

Doctor of Philosophy Degree Thesis by Speciality 113 «Applied Mathematics» (11 — «Mathematics and Statistics»). — Taras Shevchenko National University of Kyiv of Ministry of Education and Science of Ukraine, Kyiv, 2023.

With the progress of computer technologies, a new direction of scientific research has emerged — computer mathematical modeling, which involves the construction of models of the researched object and the organization of a series of computational experiments. Usually it is difficult to obtain exact solutions for practical problems, so the development of algorithms for approximate calculations is relevant. Therefore, approximate algorithms are an extremely important

component for many software products that affect our daily lives.

The thesis is related to the models of the form of variational inequalities. They provide a simple and unified means of formulating many topical problems of optimal control, mathematical physics and operations research (search for saddle points and Nash equilibrium). Variational inequalities with monotone operators are a general class of problems with a convex structure. Individual problems of convex non-differentiable optimization can be effectively solved if they are reformulated in the form of saddle (min-max) problems and then algorithms for solving variational inequalities are applied. With the invention of generative adversarial neural networks and other models of adversarial learning, sustained interest in algorithms for solving variational inequalities arose also among specialists in the field of machine learning.

Research on algorithms for solving variational inequalities and related problems continues. In the 1970s, G. Korpelevich proposed the extragradient method. In 1980, L. Popov proposed an interesting modification of the Arrow–Hurwitz method for finding saddle points of convex-concave functions, which became the source of many modern algorithms. An effective modern variant of the extragradient method is the proximal mirror method (MirrorProx) by A. Nemirovski. Recently, there has been an increase of research activity related to increasing the efficiency and versatility of extragradient methods. In particular, adaptive modifications of the proximal mirror method are proposed, which do not require knowledge of the Lipschitz constants of operators to determine the step size. The methods of approximation of fixed points of nonexpansive operators are of great importance, which allow to regularize existing methods or to construct new ones.

However the question of effective methods for solving large-dimensional variational inequalities is still relevant. For example, recently there has been increased interest in the development of adaptive algorithms that do not use the values of the Lipschitz constants of operators and, unlike linear search schemes, do not require calculations of bifunction values at additional points. Further attempts to reduce the complexity of iteration led to the appearance of the operator extrapolation algorithm. These algorithms have an advantage over the extragradient method in terms of the amount of calculations required for the iterative step. Recently algorithms for solving variational problems in Banach spaces and algorithms using Bregman divergence have also been proposed.

The thesis consists of an introduction and five chapters, the first of which contains a review of the literature on the thesis topic.

In the second chapter of the thesis, variational inequalities with operators acting in a Hilbert space were considered. For these problems, new algorithms of the extragradient type, which are modifications of L. Popov's method of finding saddle points of convex-concave functions, are proposed and substantiated. New results were also obtained for known variants of the method.

To be more precise, a sublinear efficiency estimate for the gap function is obtained. The strong convergence of the extrapolation from the past algorithm for variational inequalities with uniformly monotone operators is proved. The linear convergence speed of the extrapolation from the past algorithm for variational inequalities with operators satisfying the generalized strong monotonicity type condition is proved. An adaptive version of the algorithm is proposed and substantiated. The parameter update rule does not use the values of the Lipschitz constants of the operators and, unlike the rules of the linear search type, does not require the calculation of the values of the operators at additional points. Regularized variants of the algorithm are proposed and theorems about their strong convergence are proved.

The third chapter explores new iterative algorithms for solving variational inequalities in Hilbert spaces. These algorithms are modifications of the well-known «forward-reflected-backward algorithm».

Namely, the weak convergence of the operator extrapolation algorithm was proved and a sublinear efficiency estimate for the gap function was obtained.

The strong convergence of the operator extrapolation algorithm for variational inequalities with uniformly monotone operators is proved. The linear speed of convergence of the operator extrapolation algorithm for variational inequalities with operators satisfying the generalized strong monotonicity type condition is proved. An adaptive version of the algorithm is proposed and substantiated. A regularized version of the algorithm is proposed and a theorem on its strong convergence is proved. The Halpern scheme was used for regularization.

In the fourth chapter, variants of extrapolation from the past algorithm and operator extrapolation algorithm with Bregman divergence for solving variational inequalities with monotone and Lipschitz-continuous operators, which act in a finite-dimensional real linear space, are investigated. Main results: efficiency estimate for the gap function.

Two new algorithms for solving variational inequalities with monotone Lipschitz operators operating in 2-uniformly convex and uniformly smooth Banach spaces are also investigated. The first algorithm is a variant of the operator extrapolation method that uses a generalized Alber projection instead of a metric one. The second algorithm is an adaptive variant of the first one, which uses the step size update rule, which does not require knowledge of Lipschitz constants and linear search. The sublinear efficiency estimation in terms of the gap function and the weak convergence theorem of the first method are proved. The weak convergence of the adaptive version of the method, which uses the step size renewal rule that does not require knowledge of Lipschitz constants and linear search, is proved.

The fifth chapter of the thesis describes the experiments conducted to verify the theorems from the second and third chapters, containing estimates of the speed of convergence of the considered algorithms. Bilinear and quadratic saddle problems were considered. The results of the conducted experiments confirm the theoretical estimates.

Experiments were also conducted with the training of simple GANs by using two randomized variants of the extrapolation from the past method with exponential averaging. The results are preliminary. Note that the problem of theoretical guarantees of the convergence of these options is open.

The conducted research is in line with modern studies of variational inequalities, the interest of which is quite high. The results of the thesis are intended for use in the creation of specialized software for computer modeling in operations research, mathematical economics and mathematical physics. The developed algorithms can also be used to solve problems of machine learning, image processing, etc. Individual results obtained in the work were implemented in the educational process of the Department of Computational Mathematics of the Faculty of Computer Sciences and Cybernetics of Taras Shevchenko National University of Kyiv when teaching the disciplines «Modern problems of analysis» and «Non-classical problems of optimal control». Several algorithms were developed during the implementation of scientific projects of the Department of Computational Mathematics, Faculty of Computer Sciences and Cybernetics of Taras Shevchenko National University of Kyiv, which were carried out during 2020–2023.

Key words: variational inequality, saddle point problem, equilibrium, optimization, Extragradient method, Extrapolation from the Past method, Operator Extrapolation method, Hilbert space, Banach space, Bregman divergence, approximation, convergence.

Список публікацій здобувача

Статті у наукових фахових виданнях:

1. Semenov V. V., Denisov S. V., Sandrakov G. V., Kharkov O. S. Convergence of the Operator Extrapolation Method for Variational Inequalities in Banach Spaces. Cybernetics and Systems Analysis. 2022. Vol. 58. Issue 5. P. 740–753.

2. Харьков О. С. Оцінки ефективності для методів з дивергенцією Брегмана. Журнал обчисл. та прикл. матем. 2023. № 2. С. 83–93.
3. Семенов В. В., Харьков О. С. Алгоритм екстраполяції з минулого для варіаційних нерівностей в гільбертовому просторі. Журнал обчисл. та прикл. матем. 2023. № 2. С. 52–82.
4. Семенов В. В., Харьков О. С. Регуляризований алгоритм операторної екстраполяції. Журнал обчисл. та прикл. матем. 2023. № 1. С. 15–27.
5. Семенов В., Харьков О. Метод операторної екстраполяції для варіаційних нерівностей в банахових просторах. Фізико-математичне моделювання та інформаційні технології. 2023. Вип. 37. С. 118–122.
6. Денисов С. В., Семенов В. В., Харьков О. С. Слабка збіжність методу операторної екстраполяції для варіаційних нерівностей в рівномірно опуклих банахових просторах. Журнал обчисл. та прикл. матем. 2022. № 2. С. 42–49.
7. Семенов В. В., Сірик Д. С., Харьков О. С. Збіжність методу операторної екстраполяції. Доповіді НАН України. 2021. № 4. С. 28–35.
8. Семенов В. В., Сирык Д. С., Харьков О. С. Адаптивный метод операторной экстраполяции. Фізико-математичне моделювання та інформаційні технології. 2021. Вип. 33. С. 143–147.
9. Семенов В. В., Денисов С. В., Сирык Д. С., Харьков О. С. Сходимость метода экстраполяции из прошлого и метода операторной экстраполяции. Міжнародний науково-технічний журнал «Проблеми керування та інформатики». 2021. Том 66. № 3. С. 58–72.
10. Харьков О. С., Ведель Я. І., Семенов В. В. Методи для задач векторного узагальненого оптимального керування системами з розподіленими параметрами. Журнал обчисл. та прикл. матем. 2020. № 2 (134). С. 71–98.

Публікації, що засвідчують апробацію матеріалів дисертації:

11. Kharkov O., Semenov V. The regularized operator extrapolation method. Information Technology and Implementation (Satellite): Conference Proceedings, November 21, 2023, Kyiv, Ukraine. P. 269–270.
12. Харьков О., Семенов В. Швидкість збіжності нових алгоритмів для варіаційних нерівностей. Modeling, control and information technologies: Proceedings of VI International scientific and practical conference. P. 159–160. https://doi.org/10.31713/MCIT.2023.048.
13. Семенов В. В., Харьков О. С. Швидкість збіжності методу операторної екстраполяції. Матеріали Всеукраїнської наукової конференції «Сучасні проблеми прикладної математики та комп'ютерних наук», присвяченої 50-річчю кафедри теорії оптимальних процесів, 7-9 листопада 2023, Львів. С. 203–205.
14. Semenov V., Kharkov O. The regularized operator extrapolation algorithm for monotone variational inequalities. Intelligent Solutions–S: Proceedings of the International Symposium, September 28, 2023, Kyiv–Uzhorod, Ukraine. P. 88–90.
15. Семенов В., Харьков О. Лінійна швидкість збіжності алгоритмів екстраполяції з минулого та операторної екстраполяції для варіаційних нерівностей. Математика та інформаційні технології. Матеріали міжнародної наукової конференції, присвяченої 55-річчю факультету математики та інформатики, 28–30 вересня 2023 р. — Чернівці: Чернівецький нац. ун-т, 2023. С. 301–302.
16. Kharkov O. S., Semenov V. V. The regularized operator extrapolation method for variational inequalities and saddle point problems. XXXVIII International Conference «Problems of decision making under uncertainties (PDMU-2023)». Abstracts. September 11-15, 2023, Polyana, Ukraine. P. 55.

17. Semenov V., Denysov S., Kharkov O. About weak convergence of the operator extrapolation method. XXXVII International Conference «Problems of decision making under uncertainties (PDMU-2022)». Abstracts. November 23-25, 2022, Sheki–Lankaran, Republic of Azerbaijan. P. 98.
18. Semenov V., Siryk D., Kharkov O. About convergence of the operator extrapolation method. XXXVI International Conference «Problems of decision making under uncertainties (PDMU-2021)». Abstracts. May 11-14, 2021, Skhidnytsia, Ukraine. P. 91.
19. Denisov S. V., Kharkov O., Semenov V., Vedel Ya. About regularized adaptive extra-proximal algorithm for equilibrium problems in Hadamard spaces. XXXV International Conference «Problems of decision making under uncertainties (PDMU-2020)». Abstracts. May 11-15, 2020, Baku–Sheki, Republic of Azerbaijan. P. 32.

ЗМІСТ

ВСТУП

Актуальність теми. Із розвитком комп'ютерних технологій виник новий напрям наукових досліджень — комп'ютерне математичне моделювання, що передбачає побудову моделей досліджуваного об'єкта та організацію серій обчислювальних експериментів. Практичні проблеми як правило складні для отримання точних розв'язків, тому актуальною є розробка наближених алгоритмів. Отже, наближені алгоритми є головними складовими багатьох програмних продуктів, що впливають на наше повсякденне життя. Значний внесок у розвиток алгоритмів розв'язання операторних рівнянь та оптимізаційних задач зробили Я. І. Альбер, А. С. Антіпін, А. Б. Бакушинський, Л. М. Брегман, Ю. М. Данилін, Ю. М. Єрмольєв, І. І. Єрьомін, Ю. В. Маліцький, А. С. Немировський, Ю. Є. Нестеров, Є. О. Нурмінський, Б. Т. Поляк, Б. М. Пшеничний, В. В. Семенов, П. І. Стецюк, Н. З. Шор, H. H. Bauschke, D. P. Bertsekas, F. Browder, P. L. Combettes, J. L. Lions, P. L. Lions, P.-E. Mainge, Z. Opial, R. T. Rockafellar, W. Takahashi, P. Tseng, H. K. Xu та ін.

Дисертація пов'язана з моделями вигляду варіаційних нерівностей. Вони дають простий та уніфікований засіб формулювання багатьох актуальних задач оптимального керування, математичної фізики та дослідження операцій (пошук сідлових точок та рівноваги за Нешем). Варіаційні нерівності з монотонними операторами є загальним класом задач з опуклою структурою. Окремі задачі опуклої недиференційовної оптимізації можуть ефективно розв'язуватися, якщо їх переформулювати у вигляді сідлових (мінімаксних) задач і застосувати алгоритми розв'язання варіаційних нерівностей. Нещодавно був розвинутий такий варіант побудови швидких алгоритмів для задач опуклого програмування: за допомогою теорії двоїстості переходимо до деякої опукло-угнутої сідлової задачі та застосовуємо екстраградієнтні алгоритми розв'язання варіаційних нерівностей.

З появою генеруючих змагальних нейронних мереж (generative adversarial networks, GANs) та інших моделей змагального навчання стійкий інтерес до алгоритмів розв'язання варіаційних нерівностей виник і в середовищі спеціалістів в галузі машинного навчання.

Створення та дослідження алгоритмів розв'язання варіаційних нерівностей та близьких задач є напрямом прикладної математики, що активно розвивається. Найвідомішим чисельним методом розв'язання варіаційних нерівностей є екстраградієнтний метод, що запропонований Г. М. Корпелевич в 1970-х роках. В 1980 році Л. Д. Попов запропонував для пошуку сідлових точок опукло-угнутих функцій цікаву модифікацію методу Ерроу–Гурвіца, який став джерелом багатьох сучасних алгоритмів. Алгоритми цього типу відомі серед спеціалістів з машинного навчання під назвою «Extrapolation from the Past».

Сучасним варіантом екстраградієнтного методу є проксимальний дзеркальний метод (MirrorProx) А. С. Немировського. Останнім часом пожвавилася дослідницька активність, пов'язана з підвищенням ефективності та універсальності екстраградієнтних методів (Я. І. Ведель, І. В. Коннов, Ю. В. Маліцький, А. С. Немировський, Ю. Є. Нестеров, І. П. Рязанцева, В. В. Семенов, Y. Censor, A. Gibali, A. Iusem, S. Reich, M. V. Solodov, J. J. Strodiot, B. Svaiter та ін.). Пропонуються адаптивні модифікації проксимального дзеркального методу, які не потребують знання констант Ліпшиця операторів для визначення величини кроку. Подальші спроби зменшити складність виконання ітерації призвели до появи «forward–reflected–backward algorithm» (Ю. В. Маліцький) або алгоритму операторної екстраполяції. За об'ємом необхідних для здійснення ітераційного кроку обчислень алгоритми мають перевагу над екстаградієнтним методом та методом екстраполяції з минулого. Також останнім часом пропонуються алгоритми розв'язання варіаційних задач в банахових просторах та алгоритми, що використовують дивергенцію Брегмана.

Актуальною є задача отримання непокращуваних результатів про збіжність та її характер (оцінки ефективності) для нових схем типу алгоритмів екстраполяції з минулого та операторної екстраполяції. Також є потреба в адаптивних та регуляризованих модифікаціях алгоритмів.

Дисертація є продовженням досліджень Я. І. Ведель, Ю. В. Маліцького, В. В. Семенова, Л. М. Чабак та спрямована на розширення арсеналу методів розв'язання варіаційних нерівностей.

Зв'язок роботи з науковими програмами, планами, темами. Робота виконана у Київському національному університеті імені Тараса Шевченка на кафедрі обчислювальної математики факультету комп'ютерних наук та кібернетики. Дослідження проводились у рамках плану наукових досліджень кафедри обчислювальної математики факультету комп'ютерних наук та кібернетики Київського національного університету імені Тараса Шевченка в межах проєктів:

— Обчислювальні алгоритми і оптимізація для штучного інтелекту, медицини та оборони, № ДР 0122U002026 (науковий керівник — професор С. І. Ляшко, 2022–2024);
— Математичне моделювання та оптимізація динамічних систем для оборони, медицини та екології, № ДР 0119U100337 (науковий керівник — професор С. І. Ляшко, 2019–2021);
— Нові методи дослідження коректності та розв'язання задач дискретної оптимізації, варіаційних нерівностей та їх застосування, №ДР 0119U101608 (наукові керівники — І. В. Сергієнко, С. І. Ляшко, 2019–2020)[1].

[1]Спільний науковий проєкт дослідників Інституту кібернетики ім. В. М. Глушкова НАН України та Київського національного університету імені Тараса Шевченка.

Мета і задачі дослідження. Метою роботи є розробка та обгрунтування нових ефективних методів для розв'язання варіаційних нерівностей. Досягнення мети пов'язано з розв'язком таких конкретних задач:

— розробити та теоретично обгрунтувати адаптивні та регуляризовані варіанти алгоритму екстраполяції з минулого та алгоритму операторної екстраполяції для варіаційних нерівностей в гільбертовому просторі;

— отримати теореми слабкої та сильної збіжності алгоритму екстраполяції з минулого та алгоритму операторної екстраполяції;

— отримати сублінійні та лінійні оцінки швидкості збіжності алгоритмів;

— розробити варіанти алгоритмів з дивергенцією Брегмана;

— розробити та теоретично обгрунтувати варіант алгоритму операторної екстраполяції для варіаційних нерівностей в 2-рівномірно опуклому та рівномірно гладкому банаховому просторі;

— провести чисельні експерименти для підтвердження теоретичних результатів.

Об'єкт дослідження. Монотонні варіаційні нерівності та алгоритми їх наближеного розв'язання.

Предмет дослідження. Збіжність алгоритмів для варіаційних нерівностей в гільбертових та банахових просторах.

Методи дослідження. У роботі використовувалися методи функціонального аналізу, опуклого аналізу, теорія варіаційних нерівностей, схеми дослідження збіжності ітераційних методів оптимізації.

Наукова новизна одержаних результатів. Основні результати, що виносяться на захист, стосуються розробки нових алгоритмів розв'язання задач нелінійного аналізу. Зокрема, *вперше*:

— доведено теореми сильної збіжності алгоритму екстраполяції з минулого та алгоритму операторної екстраполяції для варіаційних нерівностей в гільбертовому просторі, що задовольняють умову рівномірної монотонності;

— розроблено та теоретично обгрунтовано адаптивні та регуляризовані варіанти алгоритму екстраполяції з минулого та алгоритму операторної екстраполяції для варіаційних нерівностей в гільбертовому просторі;

— отримано лінійну оцінку швидкості збіжності алгоритму операторної екстраполяції для варіаційних нерівностей в гільбертовому просторі, що задовольняють умову узагальненої сильної монотонності;

— розроблено варіант алгоритму операторної екстраполяції з дивергенцією Брегмана;

— розроблено та теоретично обгрунтовано варіант алгоритму операторної екстраполяції для варіаційних нерівностей в 2-рівномірно опуклому та рівномірно гладкому банаховому просторі;

удосконалено:

— теореми слабкої збіжності та сублінійні оцінки швидкості збіжності алгоритмів;

— лінійну оцінку швидкості збіжності алгоритму екстраполяції з минулого для варіаційних нерівностей в гільбертовому просторі, що задовольняють умову узагальненої сильної монотонності;

набуло подальшого розвитку:

— для пошуку нормального розв'язку варіаційної нерівності запропоновано та обгрунтовано алгоритми, які суміщають у собі ідеї алгоритму операторної екстраполяції та ітеративної регуляризації з адаптивним правилом оновлення параметру, що не використовує значень ліпшицевої константи оператора.

— побудовано варіант алгоритму екстраполяції з минулого з дивергенцією Брегмана.

Практична цінність і впровадження результатів роботи. Проведені дослідження лежать в руслі сучасних досліджень варіаційних нерівностей, інтерес до яких достатньо високий. Результати дисертації призначені для використання при створенні спеціалізованого програмного забезпечення для комп'ютерного моделювання в дослідженні операцій, математичній економіці та математичній фізиці. Також розроблені методи можуть бути використані для розв'язання задач машинного навчання, обробки зображень тощо. Окремі результати було впроваджено у навчальний процес кафедри обчислювальної математики факультету комп'ютерних наук та кібернетики Київського національного університету імені Тараса Шевченка при викладанні дисциплін «Сучасні проблеми аналізу» та «Некласичні задачі оптимального керування». Декілька алгоритмів було використано при виконанні наукових проєктів кафедри обчислювальної математики факультету комп'ютерних наук та кібернетики Київського національного університету імені Тараса Шевченка, що виконувались протягом 2020-2023 років.

Особистий внесок здобувача. Дисертація є самостійною науковою працею, в якій висвітлені власні ідеї та розробки автора, що дозволили розв'язати поставлені завдання. Теоретичні положення та висновки, сформульовані в роботі, одержані автором самостійно та відображені в опублікованих працях.

В опублікованих у співавторстві роботах здобувачу належить: в [1] — доведення збіжності алгоритмів та оцінка ефективності; в [3] — розділи 3–7; в [4] — алгоритм та теорема збіжності; в [5] — алгоритм, теорема збіжності та оцінка ефективності; в [6] — формулювання та доведення теореми про збіжність; в [7] — формулювання та доведення теореми про збіжність; в [8] — адаптивний алгоритм та теорема 1; в [9] — формулювання алгоритмів та теорема 2; в [10] — формулювання умов оптимальності у вигляді варіаційних нерівностей та варіант методу лінеаризації; в [12] — теореми 1, 2 про лінійну швидкість.

Науковому керівнику В. В. Семенову належать постановки задач, загальне керівництво та участь в обговоренні результатів.

В [1] С. В. Денисову належать варіанти алгоритму для операторних рівнянь, а Г. В. Сандракову — коментарі, приклад задачі у банаховому просторі; в статті [6] С. В. Денисову належать коментарі, огляд літератури та варіанти алгоритму для операторних рівнянь; в статті [7] Д. С. Сірику належать коментарі та ідея доведення основної нерівності; в [8] Д. С. Сірику належать основні нерівності; в [9] С. В. Денисову належать коментарі та огляд літератури, а Д. С. Сірику належить доведення основної нерівності для методу екстраполяції з минулого; в [10] Я. І. Ведель належать огляд літератури та теореми про збіжність алгоритмів.

Апробація результатів дисертації. Матеріали дисертаційної роботи доповідались та обговорювались на науковому семінарі кафедри обчислювальної математики факультету комп'ютерних наук та кібернетики Київського національного університету імені Тараса Шевченка (2023), а також на наукових конференціях:

— X International Conference «Information Technology and Implementation – 2023» (Київ, 20-21 листопада 2023);
— 18th International Conference «Mathematical Modeling and Simulation Systems» (MODS2023) (Чернігів, 13-15 листопада 2023);
— VI International Scientific-Practical Conference «Modeling, control and information technology» (Рівне, 9-11 листопада 2023);
— XXVII Всеукраїнська наукова конференція «Сучасні проблеми прикладної математики та комп'ютерних наук», присвячена 50-річчю кафедри теорії оптимальних процесів Львівського національного університету імені Івана Франка (Львів, 7-9 листопада 2023);

— Міжнародна наукова конференція «Математика та інформаційні технології», присвячена 55-річчю факультету математики та інформатики (Чернівці, 28-30 вересня 2023);
— III International Scientific Symposium «Intelligent Solutions» (Satellite) (Київ–Ужгород, 28 вересня 2023);
— Міжнародний науковий симпозіум «Питання оптимізації обчислень (ПОО–XLVIII)», який присвячений 100-річчю від дня народження академіка В. М. Глушкова (Львів, 19-22 вересня 2023);
— XXXVIII International Conferences «Problems of decision making under uncertainties» (Поляна, 11–15 вересня 2023);
— XXXVII International Conferences «Problems of decision making under uncertainties» (Шекі–Ленкорань, Азербайджан, 23-25 листопада 2022);
— IX Міжнародна конференція «Обчислювальна та прикладна математика», присвячена 100-річчю академіка Івана Івановича Ляшка (Київ, 10-11 жовтня 2022);
— Міжнародна наукова конференція «Питання оптимізації обчислень (ПОО–XLVII)», присвячена 30-річчю незалежності України (Львів, 21-24 вересня 2021);
— XXXVI International Conferences «Problems of decision making under uncertainties» (Східниця, 11-14 травня 2021);
— XXXV International Conferences «Problems of decision making under uncertainties» (Баку–Шекі, Азербайджан, 11-15 травня 2020).

Публікації. Основні результати дисертаційної роботи викладено у 10 статтях [1–10], опублікованих у виданнях, що внесені до переліку наукових фахових видань України, та додатково відображено в матеріалах конференцій [11–19]. Стаття [1] опублікована у виданні, що входить до бази даних Scopus.

Структура та обсяг дисертації. Дисертаційна робота складається зі вступу, п'яти розділів, висновків, списку використаних джерел (191 найменування на 20 сторінках), анотацій (на 12 сторінках) та 1 додатку (на 5 сторінках). Дисертація містить 2 таблиці та 31 рисунок. Загальний обсяг дисертації становить 185 стор., обсяг основної частини — 145 стор.

Подяка. Автор висловлює щиру подяку науковому керівнику, доктору фізико-математичних наук, професору Володимиру Вікторовичу Семенову за постійну увагу до роботи, підтримку та допомогу.

РОЗДІЛ 1
ОГЛЯД ЛІТЕРАТУРИ

У цьому розділі наведемо огляд невеликий наукової літератури за темою дисертаційної роботи. Приділемо увагу розвитку алгоритмів для варіаційних нерівностей та близьких задач. В огляді використано матеріали дисертацій [20, 21] та книг [22–24].

1.1. Варіаційні нерівності

Варіаційні нерівності один з центральних об'єктів у прикладному аналізі. Дослідження варіаційних нерівностей розпочалося в 1960-х в роботах Ж.-Л Ліонса, Г. Фікера, Г. Стампак'ї та Ф. Хартмана [25–28]. Ці роботи з'явились при спробах розв'язання проблем варіаційного числення, оптимального керування та теорії крайових задач з односторонніми умовами. Огляд перших застосувань варіаційних нерівностей дають книги Ліонса [29, 30], Кіндерлєрера і Стампак'ї [31] та Байокі і Капело [32]. Відмітимо й грунтовні монографії М. М. Вайнберга [33], В. І. Іваненко та В. С. Мельника [34]. Зв'язок між опуклим аналізом, теорією максимальних монотонних операторів та варіаційними нерівностями встановлений у роботах Ж. Моро [35], Х. Брезіса [36].

Дуже швидко варіаційні нерівності стали ефективним інструментом для розв'язання проблем оптимізації, оптимального керування, економіки, математичної фізики тощо [37–47].

В оптимізації та оптимальному керуванні варіаційні нерівності є перш за все зручним способом запису умов оптимальності [8]. Відома монографія Ліонса [30] є систематичним викладенням теорії оптимального керування системами з розподіленими параметрами з точки зору варіаційних нерівностей.

В статті [41] А. Бенсусан сформулював задачу пошуку рівноваги за Нешем у вигляді варіаційної нерівності. Задача пошуку рівноважного розподілу потоків в транспортній мережі в роботі [42] була подана у вигляді варіаційної нерівності. Починаючи з цих робіт почалося активне дослідження різних задач пошуку рівноваги за допомогою методів теорії варіаційних нерівностей [43, 44]. Актуальним задачам пошуку рівноваги, таким як різні цінові рівноваги, або транспортні рівноваги, що відповідають поведінковим принципам Вардропа, присвячено монографію A. Nagurney [45].

В моделях математичної фізики варіаційні нерівності застосовуються в теорії пружності [47], теорії пружнопластичних матеріалів [47–49], теорії пористих середовищ [29], теорії неньютонівських рідин [47, 50].

Окремі задачі опуклої недиференційовної оптимізації можуть ефективно розв'язуватися, якщо їх переформулювати у вигляді сідлових (мінімаксних) задач і застосувати алгоритми розв'язання варіаційних нерівностей [51]. Нещодавно був розвинутий такий варіант побудови швидких алгоритмів для задач опуклого програмування: за допомогою теорії двоїстості переходимо до деякої опукло-угнутої сідлової задачі (гра Фенхеля) та застосовуємо екстраградієнтні алгоритми розв'язання варіаційних нерівностей [52].

З появою генеруючих змагальних нейронних мереж (generative adversarial networks, GANs) та інших моделей змагального навчання [53, 54] стійкий інтерес до алгоритмів розв'язання варіаційних нерівностей виник і в середовищі спеціалістів в галузі машинного навчання [55–59].

1.2. Алгоритми для варіаційних нерівностей та близьких задач

Перейдемо до огляду основних алгоритмів розв'язання варіаційних нерівностей. Приділимо увагу алгоритмам апроксимації нерухомих точок нерозтягуючих операторів. Це викликано тим, що спираючись на ці абстрактні схеми можна будувати алгоритми для конкретних задач, зокрема, для варіаційних нерівностей. Після цього розглянемо алгоритми для варіаційних нерівностей, які є основними об'єктами дослідження в дисертації. Закінчимо підрозділ обговоренням проксимальних алгоритмів для задач про рівновагу — важливого узагальнення варіаційних нерівностей.

1.2.1. Апроксимація нерухомих точок. Теорія нерухомих точок є ефективним інструментом дослідження нелінійних явищ. Теореми Брауера, Шаудера та Какутані про існування нерухомих точок мають неконструктивний характер. Важливими є класи операторів, для яких можна конструктивно знаходити нерухомі точки. Найвідомішим прикладом такого класу є стискаючі відображення. Дійсно, в класичному доведенні теореми Банаха явно будується послідовність, що збігається до нерухомої точки. Більш широким є клас нерозтягуючих операторів, теорія яких плідно розвивається з 1965 р. Вступ до цієї теорії з описом основних алгоритмів апроксимації нерухомих точок є в книзі Bauschke та Combettes [22] (див. також [24]).

У статті [60] Браудер доводить теорему про існування нерухомої точки нерозтягуючого оператора, що діє в опуклій замкненій та обмеженій підмножині гільбертового простору, а в [61] він доводить подібну теорему для рівномірно опуклих банахових просторів. Ці результати, причому в той самий час, були одержані в роботах Кірка [62] та Гьоде [63].

Основними методами, що забезпечують наближення нерухомих точок нерозтягуючих операторів (в припущенні їх існування) є метод Красносельського–Манна [64–66], Гальперна [67–73], гібридний метод Nakajo–Takahashi [74] та метод (shrinking algorithm) Takahashi–Takeuchi–Kubota [75].

Нехай $C \subseteq H$ — непорожня опукла замкнена підмножина гільбертового простору H, $T : C \to C$ — нерозтягуючий оператор, тобто

$$\|Tx - Ty\| \leqslant \|x - y\| \quad \forall x, y \in C,$$

причому $F(T) = \{x \in C : \, x = Tx\} \neq \varnothing$.

Метод Красносельського–Манна генерує послідовність (x_n) за допомогою схеми

$$x_{n+1} = \alpha x_n + (1 - \alpha) T x_n,$$

де $\alpha \in (0, 1)$. Відомо, що (x_n) слабко збігається до деякої точки з $F(T)$ [66].

У 1967 р. Гальперн [67] дослідив збіжність методу

$$x_{n+1} = \alpha_n y + (1 - \alpha_n) T x_n, \tag{1.1}$$

де $y \in C$, а (α_n) — послідовність чисел з $(0, 1)$. Він показав, що умови

$$\lim_{n\to\infty} \alpha_n = 0, \quad \sum_{n=1}^{\infty} \alpha_n = +\infty \tag{1.2}$$

є необхідними для сильної збіжності послідовності (x_n) до нерухомої точки нерозтягуючого оператора T, та отримав достатні умови сильної збіжності. Зокрема, Гальперн довів збіжність (1.1) при $\alpha_n = \frac{1}{(n+1)^q}$, де $q \in (0, 1)$.

У 1977 р. П.–Л. Ліонс [68] покращив результат Гальперна, довівши сильну збіжність (1.1), якщо послідовність (α_n) задовольняє умови (1.2) та

$$\lim_{n\to\infty} (\alpha_{n+1} - \alpha_n)/\alpha_{n+1}^2 = 0. \tag{1.3}$$

У 1992 р. Віттман [69] довів сильну збіжність методу (1.1) до нерухомої точки оператора T за виконання умов (1.2) та

$$\sum_{n=0}^{\infty} |\alpha_{n+1} - \alpha_n| < +\infty. \tag{1.4}$$

Нарешті, у 2002 р. H. K. Xu [70, 71] довів теорему про сильну збіжність методу (1.1) за виконання умов (1.2) та

$$\lim_{n\to\infty} (\alpha_{n+1} - \alpha_n)/\alpha_{n+1} = 0. \tag{1.5}$$

Досі відкритим залишається питання [72]: чи є (1.2) достатньою умовою сильної збіжності методу (1.1) для довільного нерозтягуючого оператора $T : C \to C$?

Відомо, що у загальному випадку метод Красносельського–Манна лише слабко збігається. Існує кілька сильно збіжних його модифікацій. Однією з найпопулярніших є CQ–метод (відомий під назвою «гібридний метод»), що запропонований у 2003 р. K. Nakajo та W. Takahashi [74]. У 2008 р. W. Takahashi, Y. Takeuchi, та R. Kubota [75] запропонували алгоритм (shrinking algorithm).

1.2.2. Алгоритми екстраградієнтного типу для варіаційних нерівностей. В дисертації розглядаються варіаційні нерівності вигляду:

$$\text{знайти } x \in C : \ (Ax, y - x) \geqslant 0 \quad \forall y \in C, \tag{1.6}$$

де C — непорожня опукла і замкнена підмножина дійсного гільбертового простору H, $A : H \to H$ – деякий оператор. Як правило, вважаємо оператор A є монотонним (або псевдомонотонним[1]) та ліпшицевим зі сталою $L > 0$.

Найпростішим методом розв'язання варіаційних нерівностей є аналог методу градієнтного спуску, що у випадку сідлової задачі відомий як метод градієнтного спуску–підйому [24, 76, 77]. Теоретичним базисом методу є рівносильність варіаційної нерівності (1.6) та задачі пошуку нерухомої точки оператора $P_C (I - \lambda A)$, де $\lambda > 0$ і P_C — оператор метричної проекції на множину C [31]. Дослідження цього методу починається з робіт [78, 79],

[1]Оператор $A : C \to H$ псевдомонотонний, якщо $\forall x,\ y \in C$ з $(Ax, y - x) \geqslant 0$ випливає $(Ay, x - y) \leqslant 0$.

в яких незалежно запропоновано метод проекції градієнта для задачі умовної оптимізації. Для (1.6) аналог цього методу генерує послідовність (x_n) за правилом

$$x_{n+1} = P_C\left(x_n - \lambda A x_n\right), \tag{1.7}$$

де $\lambda > 0$. Для гарантування збіжності (x_n) до розв'язку (1.6) необхідно накласти на A більш сильні умови, ніж просто монотонність. Зокрема, якщо A — сильно монотонний[2] або обернено сильно монотонний (ко-коерцитивний)[3], то послідовність (x_n) збігається до розв'язку (1.6) (у першому випадку — сильно, у другому — слабко). Величина λ у першому випадку обирається з інтервалу $\left(0, 2\mu/L^2\right)$, де $\mu > 0$, $L > 0$ — сталі сильної монотонності та Ліпшиця оператора A, відповідно, у другому випадку — з $(0, 2\alpha)$, де $\alpha > 0$ стала ко-коерцитивності A.

У випадку лише монотонності оператора A для схеми (1.7) можна гарантувати лише ергодичну збіжність [77].

У 1976 р. Г. М. Корпелевич запропонувала для розв'язання (1.6) в евклідовому просторі $\mathbb{R}^m$, так званий, екстраградієнтний алгоритм [80] (див. також роботу А. С. Антіпіна [81]):

$$\begin{cases} y_n = P_C\left(x_n - \lambda A x_n\right), \\ x_{n+1} = P_C\left(x_n - \lambda A y_n\right), \end{cases} \tag{1.8}$$

де $\lambda \in (0, 1/L)$. Вона довела збіжність послідовностей (x_n) і (y_n) до деякого розв'язку нерівності (1.6). Узагальненню і дослідженню цього алгоритму присвячена велика кількість публікацій. Зокрема, у 2006 р. N. Nadezhkina та W. Takahashi [82] розглянули екстраградієнтний алгоритм для варіаційних нерівностей в нескінченновимірному гільбертовому просторі та довели його слабку збіжність. А у 2010 р. в роботі [83] запропоновано такий дво-

[2]Оператор $A : C \to H$ сильно монотонний, якщо $\exists\ \mu > 0$: $(Ax - Ay, x - y) \geqslant \mu \left\|x - y\right\|^2 \quad \forall x, y \in C$.

[3]Оператор $A : C \to H$ називають обернено сильно монотонним (ко-коерцитивним), якщо $\exists\ \alpha > 0$: $(Ax - Ay, x - y) \geqslant \alpha \left\|Ax - Ay\right\|^2 \quad \forall x, y \in C$.

кроковий екстраградієнтний алгоритм

$$\begin{cases} y_n = P_C\left(x_n - \lambda A x_n\right), \\ z_n = P_C\left(y_n - \lambda A y_n\right), \\ x_{n+1} = P_C\left(x_n - \lambda A z_n\right), \end{cases}$$

де $\lambda > 0$. Близьким до (1.8) є метод дуальної екстраполяції [84].

«Обчислювально дешеві» варіанти екстраградієнтного алгоритму з одним проєктуванням на множину C пропонувались у роботах [85–87]. У [85] Y. Censor, A. Gibali та S. Reich запропонували субградієнтний екстраградієнтний алгоритм, що має вигляд:

$$\begin{cases} y_n = P_C\left(x_n - \lambda A x_n\right), \\ T_n = \left\{z \in H :\ \left(x_n - \lambda A x_n - y_n, z - y_n\right) \leqslant 0\right\}, \\ x_{n+1} = P_{T_n}\left(x_n - \lambda A y_n\right), \end{cases} \tag{1.9}$$

де $\lambda \in (0, 1/L)$. Друга проекція в (1.9) має явну форму. Тому цей метод є привабливим, коли проектування на допустиму множину C є складною задачею. Зазначимо, що незалежно алгоритм (1.9) був запропонований Войтовою, Семеновим і Ляшком [86] для задач рівноважного програмування. В роботах [85, 86] доведена слабка збіжність породжених цим алгоритмом послідовностей (x_n) і (y_n) до деякого розв'язку варіаційної нерівності (1.6). Цікавою є така схема, запропонована в 2000 р. P. Tseng [87]

$$\begin{cases} y_n = P_C\left(x_n - \lambda A x_n\right), \\ x_{n+1} = y_n + \lambda\left(A x_n - A y_n\right), \end{cases} \tag{1.10}$$

де $\lambda \in (0, 1/L)$. Послідовності (x_n) і (y_n) слабко збіжні до деякого розв'язку нерівності (1.6). Обидва алгоритми (1.9) і (1.10) мають однакову складність ітераційного кроку: одне проектування на C та обчислення двох значень оператора A.

Ефективним сучасним варіантом екстраградієнтного методу є проксимальний дзеркальний метод (MirrorProx) А. С. Немировського [51]. Даний метод можна проінтерпретувати як варіант екстраградієнтного методу з проектуванням відносно дивергенції Брегмана. Дивергенція Брегмана (породжена функцією g) на множині C задається формулою

$$V(a,b) = \varphi(a) - \varphi(b) - (\nabla\varphi(b), a-b) \quad \forall a,b \in C,$$

де φ — неперервно диференційовна та сильно опукла на множині C функція. Алгоритм MirrorProx має вигляд:

$$\begin{cases} y_n = \Pi_C\left((\nabla\varphi)^{-1}(\nabla g(x_n) - \lambda A x_n)\right), \\ x_{n+1} = \Pi_C\left((\nabla\varphi)^{-1}(\nabla g(x_n) - \lambda A y_n)\right), \end{cases}$$

де Π_C — оператор проектування Брегмана на множину C [88], що задається правилом $\Pi_C x = \operatorname{argmin}_{y\in C} V(y,x)$.

Адаптивні варіанти методів Корпелевич та Немировського, які не потребують знання констант Ліпшиця операторів для визначення величини кроку, досліджені в роботах [21, 89–95]. Ідея з [89] модифікації (1.8) полягала в тому, що на кожному кроці в екстраградієнтному методі потрібно перераховувати $y_n = P_C(x_n - \lambda_n A x_n)$ доти, поки

$$\|Ay_n - Ax_n\| \leqslant \frac{\theta}{\lambda_n}\|y_n - x_n\|,$$

де $\theta \in (0,1)$. Після цього може бути обчислений x_{n+1}, як

$$x_{n+1} = P_C(x_n - \lambda_n A y_n).$$

Як показано в [90] такий підхід може бути більш успішним, ніж звичайний екстраградієнтний метод. У роботах [21,91] дана стратегія використана для побудови модифікацій (1.9) для варіаційних нерівностей (1.6) з монотонними операторами, рівномірно неперервними на обмежених множинах. В [94, 95] досліджувався такий адаптивний варіант екстраградієнтного методу

$$\begin{cases} y_n = P_C(x_n - \lambda_n A x_n), \\ x_{n+1} = P_C(x_n - \lambda_n A y_n), \end{cases}$$

де

$$\lambda_{n+1} = \begin{cases} \min\left\{\lambda_n, \tau\frac{\|x_n-y_n\|}{\|Ax_n-Ay_n\|}\right\}, & \text{якщо} \quad Ax_n \neq Ay_n, \\ \lambda_n, & \text{інакше.} \end{cases}$$

В дисертації, відштовхуючись від логіки цієї схеми, будуть запропоновані нові адаптивні методи для варіаційних нерівностей.

В нескінченновимірному гільбертовому просторі екстраградієнтний метод не має сильної збіжності. Важливим є питання побудови алгоритмів, які б гарантовано сильно збігались. Для побудови сильно збіжних алгоритмів до базового методу застосовують техніку ітеративної регуляризації Бакушинського [96] або доповнюють базовий метод однією з сильно збіжних схем апроксимації нерухомих точок [67,74,75]. Наприклад, у 2006 р. N. Nadezhkina та W. Takahashi [97] запропонували метод, що поєднує екстраградієнтний метод Корпелевич (1.8) з гібридним методом [74]. В роботі [97] доведено, що послідовність (x_n), згенерована за допомогою

$$\begin{cases} y_n = P_C(x_n - \lambda Ax_n), \\ z_n = P_C(x_n - \lambda Ay_n), \\ C_n = \left\{z \in H: \ \|z - z_n\| \leqslant \|z - x_n\|\right\}, \\ Q_n = \left\{z \in H: \ (x_n - z, x_1 - x_n) \geqslant 0\right\}, \\ x_{n+1} = P_{C_n \cap Q_n} x_1, \end{cases}$$

де $x_1 \in C$, $\lambda \in (0, 1/L)$, сильно збігається до розв'язку (1.6). Поєднавши ідеї [85] та [97] Y. Censor, A. Gibali та S. Reich у роботі [98] запропонували сильно збіжний метод ($x_1 \in H$, $\lambda \in (0, 1/L)$)

$$\begin{cases} y_n = P_C\left(x_n - \lambda Ax_n\right), \\ T_n = \left\{z \in H: \ (x_n - \lambda Ax_n - y_n, z - y_n) \leqslant 0\right\}, \\ z_n = P_{T_n}\left(x_n - \lambda Ay_n\right), \\ C_n = \left\{z \in H: \ \|z - z_n\| \leqslant \|z - x_n\|\right\}, \\ Q_n = \left\{z \in H: \ (x_n - z, x_1 - x_n) \geqslant 0\right\}, \\ x_{n+1} = P_{C_n \cap Q_n} x_1. \end{cases}$$

В роботах [21,99–101] для регуляризації екстраградієнтних схем використано метод Гальперна (1.1).

У 1980 р. Л. Д. Попов [102] запропонував метод, подібний до екстраградієнтного, але з використанням в ітераційному кроці лише одного значення оператора A

$$\begin{cases} x_{n+1} = P_C\left(x_n - \lambda A y_n\right), \\ y_{n+1} = P_C\left(x_{n+1} - \lambda A y_n\right), \end{cases} \tag{1.11}$$

де $\lambda \in (0, 1/3L)$. У роботі [102] обґрунтовано збіжність (1.11) для скінченновимірного простору. У 2014 р. Маліцький і Семенов [103] довели слабку збіжність (1.11) для варіаційних нерівностей в нескінченновимірному гільбертовому просторі та запропонували модифікацію алгоритму Попова з одним метричним проектуванням на допустиму множину C

$$\begin{cases} T_n = \left\{z \in H:\ (x_n - \lambda A y_{n-1} - y_n, z - y_n) \leqslant 0\right\}, \\ x_{n+1} = P_{T_n}\left(x_n - \lambda A y_n\right), \\ y_{n+1} = P_C\left(x_{n+1} - \lambda A y_n\right), \end{cases}$$

де $\lambda \in (0, 1/3L)$. Модифікаціям алгоритму (1.11), що використовують дивергенцію Брегмана, присвячено роботи [20,21,104–109]. Збіжність за скінченну кількість ітерацій методу Попова за умови гостроти рівноваги доведена в [110–112].

Як зазначено, ітерація даного алгоритму дешевша за ітерацію екстраградієнтного алгоритму за кількістю обчислень значень оператора: одне проти двох. Даний алгоритм для варіаційних нерівностей став відомим серед спеціалістів з машинного навчання під назвою «Extrapolation from the Past» [55].

У 2015 р. Ю. В. Маліцький [113] опублікував, так званий, проективний відбиваючий алгоритм

$$\begin{cases} y_n = 2x_n - x_{n-1}, \\ x_{n+1} = P_C\left(x_n - \lambda A y_n\right), \end{cases}$$

де $\lambda \in \left(0, \frac{\sqrt{2}-1}{L}\right)$, та довів його слабку збіжність. Стохастичний варіант досліджено в [114].

Подальший розвиток даного кола ідей та спроби зменшити складність виконання ітерації з збереженням характеру збіжності призвели до появи нового «forward–reflected–backward algorithm» для розв'язання операторних включень [115]. Точніше, для розв'язання нерівності (1.6) в роботах [115, 116] Ю. В. Маліцьким з колегами запропоновано такі схеми:

$$x_{n+1} = P_C\left(x_n - \lambda A x_n - \lambda\left(A x_n - A x_{n-1}\right)\right), \quad \lambda \in (0, 1/2L),$$

$$x_{n+1} = P_C\left(x_n - \lambda A x_n\right) - \lambda\left(A x_n - A x_{n-1}\right), \quad \lambda \in (0, 1/3L).$$

Перша з схем — «forward–reflected–backward algorithm» — відома під назвою «optimistic gradient descent ascent» [55, 117, 118] та «алгоритм операторної екстраполяції» [1,4–9,119,120]. За об'ємом необхідних для здійснення ітераційного кроку обчислень алгоритми мають перевагу над екстаградієнтним методом Корпелевич та методом екстраполяції з минулого.

Актуальною є задача розробки сильно збіжного варіанту алгоритму операторної екстраполяції. У дисертації за допомогою схеми Гальперна побудовано подібний алгоритм. Для алгоритму екстраполяції з минулого сильно збіжна модифікація для пошуку нормального розв'язку (розв'язку найменшої норми) досліджена в роботі [121].

Останнім часом отримано багато результатів для алгоритмів розв'язання варіаційних задач в банахових просторах [120, 122–131]. Зокрема, побудовані та теоретично обґрунтовані аналоги алгоритмів Корпелевич, Tseng'а та Попова для задач в рівномірно опуклих банахових просторах [127, 129, 130]. У роботах [120, 128, 131] запропоновано адаптивні варіанти «forward–reflected–backward algorithm» для варіаційних нерівностей в 2-рівномірно опуклому та рівномірно гладкому банаховому просторі.

У дисертації для методу операторної екстраполяції з проєкцією Альбера доведено доведено оцінку ефективності в термінах функції зазору.

Також за допомогою ляпуновського аналізу отримано теореми про слабку збіжність методів типу операторної екстраполяції (зокрема, адаптивного).

1.2.3. Алгоритми для задач рівноважного програмування. Розглянемо задачі рівноважного програмування у такій постановці [132]:

$$\text{знайти } \ x \in C : \ F(x, y) \geqslant 0 \ \ \forall\, y \in C, \tag{1.12}$$

де C — підмножина гільбертового простору H, $F : C \times C \to \mathbb{R}$ — біфункція.

Рівноважні постановки та варіаційні нерівності дуже близькі. Ідеї з однієї області плідно використовуються в іншій.

Формулювання задачі про рівновагу, яке вважають класичним, було наведено ще в роботах Х. Нікайдо та К. Ісоди, виконаних в 1950-х роках [133] та пов'язаних з доведенням існування точок рівноваги за Нешем в некооперативних іграх. Увагу дослідників до задач рівноважного програмування у 1990-х привернули роботи W. Oettli [134, 135]. Пріоритетні результати, повязані з ітераційними проксимальними методами рівноважного програмування, належать А. С. Антіпіну [136–139].

Припустимо, що множина C опукла замкнена, а біфункція F задовольняє умови:

1) $F(x, x) = 0$ для всіх $x \in C$;
2) $\forall\, x, y \in C$ з $F(x, y) \geqslant 0$ випливає $F(y, x) \leqslant 0$ (псевдомонотонність);
3) $\forall\, x \in C$ функція $F(x, \cdot)$ напівнеперервна знизу та опукла на C;
4) $\forall\, y \in C$ функція $F(\cdot, y)$ слабко напівнеперервна зверху на C;
5) $\forall\, x,\, y,\, z \in C$ має місце

$$F(x, y) \leqslant F(x, z) + F(z, y) + a \left\| x - z \right\|^2 + b \left\| z - y \right\|^2,$$

де a, b — додатні константи (ліпшицевість).

У 2008 р. Quoc, Muu та Hien, грунтуючись на ідеях [140], запропонували у роботі [141] аналог екстраградієнтного методу

$$\begin{cases} y_n = \operatorname{prox}_{\lambda_n \cdot F(x_n, \cdot)} x_n, \\ x_{n+1} = \operatorname{prox}_{\lambda_n \cdot F(y_n, \cdot)} x_n, \end{cases} \tag{1.13}$$

де $\lambda_n > 0$, а prox_g — проксимальний оператор [22], що відповідає власній опуклій напівнеперервній знизу функції g:

$$H \ni x \mapsto \operatorname{prox}_g x = \operatorname{argmin}_{y \in \operatorname{dom} g} \left(g(y) + \frac{1}{2} \|y - x\|_2^2 \right) \in \operatorname{dom} g.$$

Автори [141] довели при певних умовах збіжність методу (1.13) та його аналогу з відстанню Брегмана замість евклідової. Дана робота отримала продовження в багатьох дослідників [142–150, 152–155]. Наприклад, автори [143] поєднавши ідеї [141] та [74,75] запропонували сильно збіжні методи. А в роботі [148], відштовхуючись від двокрокового екстраградієнтного алгоритму Зикіної та Меленьчука [83], запропоновано та досліджено такий алгоритм

$$\begin{cases} y_n = \operatorname{prox}_{\lambda_n \cdot F(x_n, \cdot)} x_n, \\ z_n = \operatorname{prox}_{\lambda_n \cdot F(y_n, \cdot)} y_n, \\ x_{n+1} = \operatorname{prox}_{\lambda_n \cdot F(z_n, \cdot)} x_n, \end{cases}$$

де $\lambda_n > 0$. Нарешті, у роботі Я. І. Ведель та В. В. Семенова [150] (див. також [151]) запропоновано новий ітераційний двоетапний проксимальний метод розв'язання задачі рівноважного програмування в гільбертовому просторі

$$\begin{cases} x_{n+1} = \operatorname{prox}_{\lambda \cdot F(y_n, \cdot)} x_n, \\ y_{n+1} = \operatorname{prox}_{\lambda \cdot F(y_n, \cdot)} x_{n+1}, \end{cases} \tag{1.14}$$

де $\lambda > 0$. Цей метод є розвитком методу Попова (1.11). Показано, що збіжність гарантує умова

$$\lambda \in \left(0, \frac{1}{2(2a + b)} \right),$$

тобто явно використовувалася інформація про константи умови типу ліпшицевості біфункції F. Адаптивні модифікації методу (1.14) запропоновано та досліджено в роботах [152–154].

Роботи [21, 106, 107, 155, 156] присвячено дослідженню модифікації методу (1.14) з використанням відстані Брегмана для більш ефективного врахування геометрії допустимої множини задачі.

Алгоритми для варіаційних нерівностей на множині розв'язків задачі про рівновагу та дворівневих варіаційних нерівностей розглядались у роботах [99, 157–162].

Останнім часом зростає активність в дослідженні задач про рівновагу в метричних просторах Адамара [163–165]. В роботах [101, 152–154, 166–168] для псевдомонотонних задач про рівновагу в просторах Адамара запропоновано та досліджено аналоги алгоритмів (1.13), (1.14).

1.3. Застосування варіаційних нерівностей у машинному навчанні

Розглянемо декілька задач, що зв'язані з машинним навчанням та ведуть до варіаційних нерівностей.

1.3.1. Сідлові формулювання. Задача мінімізації функції максимуму

$$\max_{i=1,\dots,m} f_i(x) \to \min_{x\in C},$$

де $C \subseteq \mathbb{R}^n$, може бути розглянута у вигляді сідлової

$$\min_{x\in C} \max_{y\in\Delta^m} \left(\sum_{i=1}^m y_i f_i(x) \right), \tag{1.15}$$

де $\Delta^m = \{y \in \mathbb{R}^m : \ \ y_i \geqslant 0, \ \ \sum_{i=1}^m y_i = 1\}$ [169]. До (1.15) задачі можна застосовувати алгоритми екстраградієнтного типу.

Деякі задачі аналізу даних та машинного навчання можна сформулювати у вигляді

$$F(Kx) + G(x) \to \min_{x\in\mathbb{R}^n}, \tag{1.16}$$

де $F : \mathbb{R}^m \to \mathbb{R} \cup \{+\infty\}$, $G : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ — власні замкнені та опуклі функції, $K : \mathbb{R}^n \to \mathbb{R}^m$ — лінійний оператор [170].

Задачу (1.16) та її двоїсту

$$-G^*(-K^*y) - F^*(y) \to \max_{y \in \mathbb{R}^m}$$

можна сформулювати у вигляді сідлової задачі (прямо-двоїсте формулювання)

$$\min_{x \in \mathbb{R}^n} \max_{y \in \mathbb{R}^m} \left(\langle Kx, y \rangle + G(x) - F^*(y) \right). \tag{1.17}$$

Тут $\varphi^*(y) = \sup_{x \in \mathbb{R}^n} \left(\langle y, x \rangle - \varphi(x) \right)$ — спряжена функція до $\varphi : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$. Деталі можна знайти в [39, 170].

Припустимо, що задача (1.18) має розв'язок $(\bar{x}, \bar{y}) \in \mathbb{R}^n \times \mathbb{R}^m$. Тоді він задовольняє операторне включення (що рівносильне варіаційній нерівності)

$$\begin{cases} 0 \in K^*\bar{y} + \partial G(\bar{x}), \\ 0 \in -K\bar{x} + \partial F^*(\bar{y}), \end{cases} \tag{1.18}$$

де ∂G та ∂F^* — субдиференціали функцій G та F^* [39, 169, 170]. А до (1.18) підходимо з алгоритмами для варіаційних нерівностей.

У вигляді (1.16) формулюються ROF та TV-L1 моделі в задачах відновлення зашумлених зображень [171, 172].

Наведемо дві задачі вигляду (1.16).

Приклад 1.1 (ℓ_2-регуляризована ℓ_2-регресія). Розглянемо задачу

$$\frac{1}{2}\|Kx - b\|_2^2 + \frac{\lambda}{2}\|x\|_2^2 \to \min_{x \in \mathbb{R}^n},$$

де $b \in \mathbb{R}^m$, $K : \mathbb{R}^n \to \mathbb{R}^m$ — лінійний оператор, $\lambda > 0$. Сідлове формулювання має вигляд

$$\min_{x \in \mathbb{R}^n} \max_{y \in \mathbb{R}^m} \left(\langle Kx, y \rangle + \frac{\lambda}{2}\|x\|_2^2 - \frac{1}{2}\|y\|_2^2 - \langle b, y \rangle \right).$$

Отримуємо таку умову оптимальності

$$\begin{cases} 0 = K^*\bar{y} + \lambda\bar{x}, \\ 0 = -K\bar{x} + \bar{y} + b. \end{cases}$$

Приклад 1.2 (ℓ_1-регресія з умовою невід'ємності). Розглянемо задачу

$$\|Kx - b\|_1 \to \min_{x \in \mathbb{R}^n}, \quad x \geqslant 0,$$

де $b \in \mathbb{R}^m$, $K : \mathbb{R}^n \to \mathbb{R}^m$ — лінійний оператор. Сідлове формулювання має вигляд

$$\min_{x \in \mathbb{R}^n} \max_{y \in \mathbb{R}^m} \left(\langle Kx, y \rangle + \delta_{\mathbb{R}^n_+}(x) - \delta_{B_\infty(\mathbb{R}^m)}(y) - \langle b, y \rangle \right),$$

де $\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : \ x_i \geqslant 0\}$, $B_\infty(\mathbb{R}^m) = \{y \in \mathbb{R}^m : \ \|y\|_\infty = \max_i |y_i| \leqslant 1\}$, δ_M — індикаторна функція множини M. Отримуємо таку умову оптимальності

$$\begin{cases} 0 \in K^*\bar{y} + N_{\mathbb{R}^n_+}\bar{x}, \\ 0 \in -K\bar{x} + b + N_{B_\infty(\mathbb{R}^m)}\bar{y}, \end{cases}$$

де $N_C z$ — нормальний конус множини C в точці z [169].

1.3.2. Робастне навчання. Мінімаксні задачі мають довгу історію успішного застосування у робастній оптимізації [173]. Вони також знайшли застосування в навчанні глибоких нейронних мереж, які виявились вразливими до суперечливих прикладів — вхідних даних, які майже неможливо відрізнити від навчальних даних і все ж неправильно класифікованих мережею.

Робастне навчання [174] має на меті подолання цих проблем шляхом розв'язання мінімаксної задачі

$$\min_x \mathbb{E}_{(u,v)} \left(\max_{y \in S} L\left(u + y, v, x\right) \right),$$

де u — вектор ознак об'єкту, v — мітка класу цього об'єкту, x — вектор параметрів моделі, що тренується, y — спотворення (шум) для змагальної модифікації об'єкту, S — множина допустимих спотворень.

1.3.3. Породжуючі змагальні мережі. Розглянемо ігрову задачу для побудови породжуючої змагальної мережі (Generative Adversarial Network (GAN)). Така нейронна мережа є комбінацією двох нейронних мереж, одна з яких — генератор G — генерує зразки, а друга — дискримінатор D — відрізняє справжні зразки з деякої бази даних від штучно згенерованих генератором G.

Оскільки нейронні мережі G та D мають протилежні задачі — генерувати зразки, що не відрізняються від справжніх, та відсіювати штучні зразки, то між ними виникає антагоністична гра. Породжуючі змагальні мережі були вперше запропоновані в роботі [53].

Наведемо математичний опис цієї гри. Нехай задано ймовірносний розподіл p_{data} на множині всіх зразків. Зразки зображаються векторами простору $\mathbb{R}^d$.

Нехай X та Y — опуклі підмножини просторів $\mathbb{R}^k$ та $\mathbb{R}^m$, а відображення G_x — генератор, де $x \in X$ — вектор його параметрів. Розподіл p_x, що моделюється на множині усіх зразків, породжується генератором з рівномірного (або нормального) розподілу μ:

$$p_x(A) = \mu\left(G_x^{-1}(A)\right).$$

Нехай функція D_y — дискримінатор, вона визначена на множині зразків та приймає значення в сегменті $[0, 1]$, $y \in Y$ — вектор її параметрів.

Генератор та дискримінатор реалізуються у вигляді нейронних мереж певної архітектури.

Значення $D_y(u)$ розуміється як ймовірність того, що зразок згенерований розподілом p_{data}.

Навчаємо нейронну мережу D_y так, щоб максимізувати ймовірність привильної класифікації зразків. Одночасно навчаємо нейронну мережу G_x так, щоб мінімізувати середнє значення величини

$$\log\left(1 - D_y\left(G_x(z)\right)\right)$$

при $z \sim \mu$.

Між генератором та дискримінатором виникає антагоністична гра з деякою платіжною функцією $L : \mathbb{R}^k \times \mathbb{R}^m \to \mathbb{R}$.

В роботі [53] запропоновано таку платіжну функцію

$$L(x, y) = \mathbb{E}_{u \sim p_{data}} (\log D_y(u)) + \mathbb{E}_{z \sim \mu} (\log (1 - D_y (G_x(z)))) . \tag{1.19}$$

Зауваження 1.1. Існують інші способи обрати цільову функцію для навчання [54].

Для побудови оптимальних мереж розв'язується задача пошуку сідлової точки функції (1.19) на $X \times Y$:

$$\min_{x \in X} \max_{y \in Y} L(x, y) =$$

$$= \min_{x \in X} \max_{y \in Y} \left(\mathbb{E}_{u \sim p_{data}} (\log D_y(u)) + \mathbb{E}_{z \sim \mu} (\log (1 - D_y (G_x(z)))) \right) . \tag{1.20}$$

Зауваження 1.2. Функція $y \mapsto L(x, y)$ є угнутою, якщо дискримінатор D_y реалізовано у вигляді нейронної мережі з одним внутрішнім шаром та пороговою функцією $D_y(u) = \frac{1}{1+e^{-(y,u)}}$.

Локальні необхідні умови, що характеризують розв'язок $(x^*, y^*) \in X \times Y$ задачі (1.20) мають вигляд:

$$(\nabla_x L(x^*, y^*), x - x^*) + (-\nabla_y L(x^*, y^*), y - y^*) \geqslant 0 \quad \forall (x, y) \in X \times Y. \tag{1.21}$$

За відсутності обмежень ($X = \mathbb{R}^k$, $Y = R^m$) варіаційна нерівність (1.21) набуває вигляду:

$$\nabla_x L(x^*, y^*) = 0 \quad \text{та} \quad - \nabla_y L(x^*, y^*) = 0.$$

В роботах [55–59] алгоритми для варіаційних нерівностей застосовувались для навчання генеруючих змагальних нейронних мереж. Зрозуміло, що створення більш швидких та стійких алгоритмів для варіаційних нерівностей (особливо для стохастичних постановок) веде до прогресу в навчанні породжуючих змагальних мереж.

1.4. Висновки до розділу 1

В розділі був проведений огляд літератури за темою дисертаційної роботи. Приділено увагу розвитку алгоритмів для варіаційних нерівностей та апроксимації нерухомих точок. Розглянуто застосування варіаційних нерівностей у машинному навчанні. Також розглянуто деякі моменти, які виникають в практичній реалізації алгоритмів. Акцентовано увагу на проблемних питаннях, вирішенню яких присвячено дисертацію.

РОЗДІЛ 2
АЛГОРИТМ ЕКСТРАПОЛЯЦІЇ З МИНУЛОГО ДЛЯ ВАРІАЦІЙНИХ НЕРІВНОСТЕЙ В ГІЛЬБЕРТОВОМУ ПРОСТОРІ

У цьому розділі розглядаються варіаційні нерівності з операторами, що діють в гільбертовому просторі. Для цих задач запропоновано та обгрунтовано нові алгоритми екстраградієнтного типу, які є модифікаціями методу Л. Д. Попова пошуку сідлових точок опукло-угнутих функцій. Також отримано нові результати для відомих варіантів методу.

Розділ побудовано таким чином. В підрозділі 2.1 розглянуто варіаційні нерівності в гільбертовому просторі, наведено основні припущення та необхідний мінімум відомостей, що відіграють важливу роль у доведеннях основних результатів дисертації. В підрозділі 2.2 розглянуто алгоритм екстраполяції з минулого та прокоментовано літературу, що присвячено дослідженню його збіжності. В підрозділі 2.3 доведено слабку збіжність алгоритму екстраполяції з минулого та отримано сублінійну оцінку ефективності для функції зазору. Сильну збіжність алгоритму екстраполяції з минулого для варіаційних нерівностей з рівномірно монотонними операторами доведено в підрозділі 2.4. Лінійну швидкість збіжності алгоритму екстраполяції з минулого для варіаційних нерівностей з операторами, що задовольняють умову типу узагальненої сильної монотонності, доведено в підрозділі 2.5. В підрозділі 2.6 запропоновано та обгрунтовано адаптивний варіант алгоритму. Нарешті, в підрозділі 2.7 запропоновано регуляризовані варіанти алгоритму та доведено теореми про їх сильну збіжність.

2.1. Постановка задачі та допоміжні відомості

Нехай H — дійсний гільбертовий простір з скалярним добутком $(\cdot,\cdot)$ та породженою нормою $\|\cdot\|$. Сильну та слабку збіжність в H послідовності (x_n) до x позначимо $x_n \to x$ та $x_n \rightharpoonup x$, відповідно.

Нехай C — непорожня опукла і замкнена підмножина простору H та $A : H \to H$ – деякий оператор.

Означення 2.1. Варіаційною нерівністю називаємо таку задачу:

$$\text{знайти } x \in C: \ (Ax, y - x) \geqslant 0 \quad \forall y \in C. \tag{2.1}$$

Множину розв'язків варіаційної неравності (2.1) позначимо через S.

Задача (2.1) — зручна загальна форма запису різних задач, що виникають в математичній фізиці, дослідженні операцій [24,30–32,45,47]. Зокрема, у вигляді варіаційної нерівності можуть бути сформульовані задачі розв'язання рівнянь, знаходження екстремуму функціоналів, знаходження точок рівноваги тощо. Наведемо ряд типових постановок.

(1) Задача розв'язання операторного рівняння

$$\text{знайти } x \in H: \ Ax = 0$$

рівносильна варіаційній нерівності

$$\text{знайти } x \in H: \ (Ax, y - x) \geqslant 0 \quad \forall y \in H.$$

(2) Нехай f — диференційовна опукла функція, C — опукла замкнена множина. Критерій оптимальності першого порядку для задачі

$$f \to \min_C$$

має вигляд

$$x \in C \quad \text{та} \quad (\nabla f(x), y - x) \geqslant 0 \quad \forall y \in C.$$

(3) Нехай $F : X \times Y \to \mathbb{R}$ — диференційовна опукло-угнута функція, X, Y — опуклі замкнені множини. Точка $(x^*, y^*) \in X \times Y$ називається сідловою точкою функції F, якщо

$$F(x^*, y) \leqslant F(x^*, y^*) \leqslant F(x, y^*) \quad \forall x \in X \ \forall y \in Y. \tag{2.2}$$

Задача пошуку сідлової точки (2.2) рівносильна варіаційній нерівності:

$$\left(\begin{pmatrix} \nabla_x F(x^*, y^*) \\ -\nabla_y F(x^*, y^*) \end{pmatrix}, \begin{pmatrix} x - x^* \\ y - y^* \end{pmatrix} \right) \geqslant 0 \quad \forall (x, y) \in X \times Y.$$

(4) Нехай X, Y — опуклі замкнені множини, $f : X \times Y \to \mathbb{R}$ ($g : X \times Y \to \mathbb{R}$) — диференційовна за x (по y) при фіксованому y (при фіксованому x) функція. Точка $(x^*, y^*) \in X \times Y$ називається рівновагою Неша, якщо

$$f(x^*, y^*) \leqslant f(x, y^*) \quad \forall x \in X, \quad g(x^*, y^*) \leqslant g(x^*, y) \quad \forall y \in Y. \tag{2.3}$$

Якщо функція $f(\cdot, y)$ опукла на X для всіх $y \in Y$, а функція $g(x, \cdot)$ опукла на Y для всіх $x \in X$, то задача пошуку рівноваги Неша (2.3) рівносильна варіаційній нерівності:

$$\left(\begin{pmatrix} \nabla_x f(x^*, y^*) \\ \nabla_y g(x^*, y^*) \end{pmatrix}, \begin{pmatrix} x - x^* \\ y - y^* \end{pmatrix} \right) \geqslant 0 \quad \forall (x, y) \in X \times Y.$$

Нагадаємо основні означення [22, 24, 76].

Означення 2.2. Оператор $A : H \to H$ називаємо псевдомонотонним на множині $C \subseteq H$, якщо для x, $y \in C$

$$(Ax, y - x) \geqslant 0 \quad \Rightarrow \quad (Ay, x - y) \leqslant 0.$$

Означення 2.3. Оператор $A : H \to H$ називаємо монотонним, якщо

$$(Ax - Ay, x - y) \geqslant 0 \quad \forall x, \ y \in H.$$

Означення 2.4. Оператор $A : H \to H$ називаємо обернено сильно монотонним (ко-коерцитивним) на множині $C \subseteq H$, якщо існує така стала $\alpha > 0$, що

$$(Ax - Ay, x - y) \geqslant \alpha \|Ax - Ay\|^2 \quad \forall x,\ y \in C.$$

У цьому випадку кажуть, що оператор A — α-обернено сильно монотонним (α-ко-коерцитивним).

Означення 2.5. Оператор $A : H \to H$ називаємо рівномірно монотонним на множині $C \subseteq H$, якщо існує така зростаюча функція $\phi : [0, +\infty) \to [0, +\infty)$, $\phi(0) = 0$ та $\phi(t) > 0$ для $t > 0$, що

$$(Ax - Ay, x - y) \geqslant \phi\left(\|x - y\|\right) \quad \forall x,\ y \in C.$$

Означення 2.6. Оператор $A : H \to H$ називаємо сильно монотонним на множині $C \subseteq H$, якщо існує така стала $\mu > 0$, що

$$(Ax - Ay, x - y) \geqslant \mu \|x - y\|^2 \quad \forall x,\ y \in C.$$

У цьому випадку кажуть, що оператор A — μ-сильно монотонний.

Скрізь у цьому розділі будемо вважати, що оператор $A : H \to H$ — ліпшицевий на множині C (з константою $L > 0$), тобто

$$\|Ax - Ay\| \leqslant L \|x - y\| \quad \forall x, y \in C,$$

та $S \neq \varnothing$.

Для замкненої опуклої множини C та ліпшицевого сильно монотонного оператора A множина S не порожня та складається з одного елемента [31].

Нагадаємо декілька відомих та потрібних нам фактів.

Нехай P_C — оператор метричного проектування на замкнену опуклу підмножину $C \subseteq H$, тобто $P_C x$ — єдиний елемент C, що володіє властивістю

$$\|P_C x - x\| = \min_{z \in C} \|z - x\|.$$

Елемент $P_C x$ можна охарактеризувати таким чином [24, 31]:

$$y = P_C x \quad \Leftrightarrow \quad y \in C \text{ та } (y - x, z - y) \geqslant 0 \ \forall z \in C, \tag{2.4}$$

$$y = P_C x \quad \Leftrightarrow \quad y \in C \text{ та } \|y - z\|^2 \leqslant \|x - z\|^2 - \|y - x\|^2 \ \forall z \in C. \tag{2.5}$$

Оператор метричного проектування P_C є нерозтягуючий, тобто

$$\|P_C x - P_C y\| \leqslant \|x - y\| \quad \forall x, y \in H,$$

точніше, обернено сильно монотонним (1-ко-коерцитивним)

$$\|P_C x - P_C y\| \leqslant (P_C x - P_C y, x - y) \quad \forall x, y \in H.$$

Варіаційну нерівність (2.1) можна сформулювати як задачу пошуку нерухомої точки [24, 31]:

$$x = P_C \left(x - \lambda A x \right), \tag{2.6}$$

де $\lambda > 0$.

Формулювання (2.6) корисне, оскільки веде до ітераційної схеми

$$x_{n+1} = P_C \left(x_n - \lambda A x_n \right), \tag{2.7}$$

яка сильно збіжна для ліпшицевих сильно монотонних операторів та слабко збіжна для обернено сильно монотонних (ко-коерцитивних) операторів [22, 24]. Але для ліпшицевих монотонних операторів схема (2.7) в загальному випадку не збігається.

Зауваження 2.1. Варіаційна нерівність (2.1) записується у формі операторного включення [22]:

$$\text{знайти } x \in H: \ \ 0 \in Ax + N_C x,$$

де $N_C x$ — нормальний конус множини C в точці x:

$$N_C x = \begin{cases} \{z \in H: \ (z, y - x) \leqslant 0 \ \forall y \in C\}, & \text{якщо } x \in C, \\ \varnothing, & \text{інакше.} \end{cases}$$

Розглянемо дуальну варіаційну нерівність:

$$\text{знайти } x \in C: \quad (Ay, y - x) \geqslant 0 \quad \forall y \in C. \tag{2.8}$$

Множину розв'язків варіаційної нерівності (2.8) позначимо S^d. Відомо, що множина S^d опукла та замкнена [31]. Нерівність (2.8) називають слабким або дуальним формулюванням варіаційної нерівності (2.1) (або нерівністю типу Мінті), а розв'язки нерівності (2.8) — слабкими розв'язками варіаційної нерівності (2.1). Для монотонних (псевдомонотонних) операторів A завжди маємо

$$S \subseteq S^d.$$

А коли оператор A монотонний та неперервний маємо (лема Мінті, [31])

$$S^d = S.$$

Однією з основних теоретичних задач є оцінка числа ітерацій алгоритму, що необхідне для отримання наближеного розв'язку заданої якості. Якість наближеного розв'язку $x \in C$ варіаційної нерівності (2.1) у монотонному випадку будемо оцінювати за допомогою невід'ємної функції зазору [51, 84]

$$\operatorname{gap}(x) = \sup_{y \in C} (Ay, x - y). \tag{2.9}$$

Очевидно, що для коректності означення функції зазору (2.9) необхідна обмеженість допустимої множини C.

Лема 2.1 (Nesterov Y., [84]). *Нехай оператор A — монотонний. Якщо $x \in C$ — розв'язок* (2.1), *то*

$$\operatorname{gap}(x) = 0.$$

Навпаки, якщо для $x \in C$ маємо

$$\operatorname{gap}(x) = 0,$$

то x — розв'язок (2.1).

Нагадаємо відомі леми про числові нерівності.

Лема 2.2. *Нехай невід'ємні послідовності* (α_n), (β_n), *такі, що*

$$\alpha_{n+1} \leqslant \alpha_n - \beta_n, \quad n \geqslant 1.$$

Тоді існує границя $\lim_{n\to\infty} \alpha_n \in \mathbb{R}$ *та* $\sum_{n=1}^{\infty} \beta_n < +\infty$.

Лема 2.3. *Нехай послідовність невід'ємних чисел* (ξ_n) *задовольняє рекурентній нерівності*

$$\xi_{n+1} \leqslant (1 - \alpha_n)\, \xi_n + \alpha_n \beta_n, \quad n \geqslant 1,$$

де послідовності (α_n) *та* (β_n) *мають властивості:* $\alpha_n \in (0,1)$ *та* $\beta_n \leqslant \beta$, *де* $\beta \geqslant 0$. *Тоді*

$$\xi_n \leqslant e^{-\sum_{k=1}^{n-1} \alpha_k} \xi_1 + \beta.$$

Лема 2.4. *Нехай послідовність невід'ємних чисел* (ξ_n) *задовольняє рекурентній нерівності*

$$\xi_{n+1} \leqslant (1 - \alpha_n)\, \xi_n + \alpha_n \beta_n + \gamma_n, \quad n \geqslant 1,$$

де послідовності (α_n) *та* (β_n) *мають властивості:*

1) $\alpha_n \in (0,1)$, $\sum_{n=1}^{\infty} \alpha_n = +\infty$;

2) $\limsup_{n\to\infty} \beta_n \leqslant 0$;

3) $\gamma_n \in [0, +\infty)$, $\sum_{n=1}^{\infty} \gamma_n < +\infty$.

Тоді $\lim\limits_{n\to\infty} \xi_n = 0$.

Лема 2.5 (Mainge P.-E., [175]). *Нехай числова послідовність* (α_n) *має підпослідовність* (α_{n_k}) *з властивістю* $\alpha_{n_k} < a_{n_k+1}$ *для всіх* $k \geqslant 1$. *Тоді існує така неспадна послідовність* (m_k) *натуральних чисел, що*

$$m_k \to +\infty$$

та

$$\alpha_{m_k} \leqslant \alpha_{m_k+1}, \quad \alpha_k \leqslant \alpha_{m_k+1}$$

для всіх $k \geqslant n_1$.

При доведенні слабкої збіжності послідовностей елементів гільбертового простору будемо використовувати відому лему Опяла.

Лема 2.6 (Z. Opial, [66]). *Нехай послідовність* (x_n) *елементів гільбертового простору* H *слабко збігається до точки* $x \in H$. *Тоді для всіх* $y \in H \setminus \{x\}$ *маємо*

$$\liminf_{n\to\infty} \|x_n - x\| < \liminf_{n\to\infty} \|x_n - y\|.$$

2.2. Алгоритм екстраполяції з минулого

Для розв'язання задачі (2.1) розглянемо такий ітераційний алгоритм.

Алгоритм 2.1. Для $x_1 = y_0 \in C$ генерируємо послідовність елементів x_n, $y_n \in C$ за допомогою ітераційної схеми

$$\begin{cases} y_n = P_C\left(x_n - \lambda_n A y_{n-1}\right), \\ x_{n+1} = P_C\left(x_n - \lambda_n A y_n\right), \end{cases}$$

де $\lambda_n > 0$.

Зауваження 2.2. Частнинний випадок алгоритму 2.1 запропонований Л. Д. Поповим [102] для пошуку сідлових точок опукло-угнутих функцій, що задані в скінченомірному евклідовому просторі. Останнім часом цей метод став відомим у середовищі спеціалістів з машинного навчання під назвою «Extrapolation from the Past» (екстраполяція з минулого) [55]. Дослідженню та розробці модифікацій цього методу присвячено роботи [20, 21, 103–108, 110–112, 117, 118, 121, 150–156, 160–162, 166, 167]. Модифікація алгоритму Попова з одним метричним проектуванням на допустиму множину запропонована в [103], варіантам методу для задач про рівновагу присвячено роботи [111, 112, 150, 151], модифікації з бегманівською відстанню досліджено в [104–108, 155, 156], методи для задач в просторах Адамара запропоновані в [152–154, 166, 167]. Регуляризовані методи та модифікації для дворівневих задач досліджено в [160–162]. У дисертації уточнюються та узагальнюються результати вказаних робіт.

При виконанні для деякого $n \in \mathbb{N}$ в алгоритмі 2.1 рівностей

$$y_n = y_{n-1} = x_n \quad \text{або} \quad x_{n+1} = x_n = y_n \tag{2.10}$$

має місце включення $y_n \in S$. Дійсно, рівність

$$x_{n+1} = P_C\left(x_n - \lambda_n A y_n\right)$$

рівносильна нерівності

$$(Ay_n, y - x_{n+1}) + \frac{(x_{n+1} - x_n, y - x_{n+1})}{\lambda_n} \geqslant 0 \quad \forall y \in C.$$

З другої рівності (2.10) випливає

$$(Ay_n, y - y_n) \geqslant 0 \;\; \forall y \in C,$$

тобто, $y_n \in S$.

Аналогічно, з

$$(Ay_{n-1}, y - y_n) + \frac{(y_n - x_n, y - y_n)}{\lambda_n} \geqslant 0 \quad \forall y \in C$$

при першій рівності в (2.10) отримуємо $y_n \in S$.

Далі припустимо, що для всіх номерів $n \in \mathbb{N}$ умова (2.10) не має місця.

2.3. Слабка збіжність та сублінійна оцінка

Припустимо додатково, що оператор A є псевдомонотонним.

Лема 2.7. *Для породжених алгоритмом 2.1 послідовностей* (x_n), (y_n) *та точки* $z \in S$ *виконується нерівність*

$$\begin{aligned} \|x_{n+1} - z\|^2 \leqslant \|x_n - z\|^2 - (1 - \lambda_n L)\|x_{n+1} - y_n\|^2 - \\ - (1 - 2\lambda_n L)\|y_n - x_n\|^2 + 2\lambda_n L\|x_n - y_{n-1}\|^2. \end{aligned} \tag{2.11}$$

Доведення. Маємо

$$\begin{aligned}\|x_{n+1}-z\|^2 = \|x_n-z\|^2 - \|x_n-x_{n+1}\|^2 + 2\left(x_{n+1}-x_n, x_{n+1}-z\right) = \\ = \|x_n-z\|^2 - \|x_n-y_n\|^2 - \|y_n-x_{n+1}\|^2 - \\ - 2\left(x_n-y_n, y_n-x_{n+1}\right) + 2\left(x_{n+1}-x_n, x_{n+1}-z\right). \quad (2.12)\end{aligned}$$

З правила породження точок x_{n+1} та y_n випливає

$$\lambda_n(Ay_n, z-x_{n+1}) \geqslant (x_{n+1}-x_n, x_{n+1}-z), \quad (2.13)$$

$$\lambda_n(Ay_{n-1}, x_{n+1}-y_n) \geqslant -(x_n-y_n, y_n-x_{n+1}). \quad (2.14)$$

Використавши нерівності (2.13), (2.14) для оцінки скалярних добутків в (2.12), отримуємо

$$\begin{aligned}\|x_{n+1}-z\|^2 \leqslant \|x_n-z\|^2 - \|x_n-y_n\|^2 - \|y_n-x_{n+1}\|^2 + \\ + 2\lambda_n\left\{(Ay_{n-1}, x_{n+1}-y_n) + (Ay_n, z-x_{n+1})\right\} = \\ = \|x_n-z\|^2 - \|x_n-y_n\|^2 - \|y_n-x_{n+1}\|^2 + \\ + 2\lambda_n\left\{(Ay_n, z-y_n) + (Ay_{n-1}, x_{n+1}-y_n) + (Ay_n, y_n-x_{n+1})\right\}. \quad (2.15)\end{aligned}$$

З псевдомонотонності оператора A та включення $z \in S$ випливає

$$(Ay_n, z-y_n) \leqslant 0,$$

а ліпшицевість F гарантує виконання нерівності

$$\begin{aligned}(Ay_{n-1}-Ay_n, x_{n+1}-y_n) \leqslant L\|y_{n-1}-y_n\|\,\|x_{n+1}-y_n\| \leqslant \\ \leqslant \frac{L}{2}\|y_{n-1}-y_n\|^2 + \frac{L}{2}\|y_n-x_{n+1}\|^2.\end{aligned}$$

Використавши вищенаведені оцінки в (2.15), отримуємо

$$\begin{aligned}\|x_{n+1}-z\|^2 \leqslant \|x_n-z\|^2 - \|x_n-y_n\|^2 - \|y_n-x_{n+1}\|^2 + \\ + \lambda_n L\|y_{n-1}-y_n\|^2 + \lambda_n L\|y_n-x_{n+1}\|^2. \quad (2.16)\end{aligned}$$

Член $\|y_{n-1}-y_n\|^2$ оцінимо таким чином

$$\|y_{n-1}-y_n\|^2 \leqslant 2\|y_{n-1}-x_n\|^2 + 2\|y_n-x_n\|^2.$$

Ураховуючи цю оцінку в (2.16), приходимо до нерівності

$$\|x_{n+1}-z\|^2 \leqslant \|x_n-z\|^2 - \|x_n-y_n\|^2 - \|y_n-x_{n+1}\|^2 + \\ + 2\lambda_n L\|y_{n-1}-x_n\|^2 + 2\lambda_n L\|y_n-x_n\|^2 + \lambda_n L\|y_n-x_{n+1}\|^2,$$

тобто, до нерівності (2.11). □

Зауваження 2.3. Для монотонних операторів нерівність леми 2.7 отримана в [20, 103].

З теореми 1 роботи [111] випливає

Теорема 2.1. *Нехай H — гільбертовий простір, $C \subseteq H$ — непорожня опукла замкнена множина, $A : H \to H$ — монотонний та L-ліпшицевий на множині C оператор та $S \neq \varnothing$. Припустимо, що*

$$\lambda_n \in \left[\underline{\lambda}, \overline{\lambda}\right] \subseteq \left(0, \tfrac{1}{3L}\right).$$

Тоді породжені алгоритмом 2.1 послідовності (x_n), (y_n) слабко збігаються до розв'язку $\bar{z} \in S$ варіаційної нерівності (2.1), причому

$$\lim_{n\to\infty} \|x_n - y_n\| = 0.$$

У випадку обмеженості допустимої множини C доведемо, що алгоритму 2.1 необхідно зробити $O\left(\frac{LD^2}{\varepsilon}\right)$ ітерацій для отримання точки $x \in C$ з

$$\operatorname{gap}(x) \leqslant \varepsilon,$$

де $\varepsilon > 0$, $D = \sup_{a,b\in C}\|a-b\| < +\infty$.

Нехай $y \in C$. Маємо

$$\|x_{n+1}-y\|^2 \leqslant \|x_n-y\|^2 - (1-\lambda_n L)\|x_{n+1}-y_n\|^2 -$$

$$- (1 - 2\lambda_n L) \|y_n - x_n\|^2 + 2\lambda_n L \|x_n - y_{n-1}\|^2 + 2\lambda_n (Ay_n, y - y_n). \quad (2.17)$$

З монотонності оператора A та (2.17) випливає

$$\|x_{n+1} - y\|^2 \leqslant \|x_n - y\|^2 - (1 - \lambda_n L) \|x_{n+1} - y_n\|^2 - \\ - (1 - 2\lambda_n L) \|y_n - x_n\|^2 + 2\lambda_n L \|x_n - y_{n-1}\|^2 + 2\lambda_n (Ay, y - y_n). \quad (2.18)$$

Перепишемо (2.18) у вигляді

$$2\lambda_n (Ay, y_n - y) \leqslant \left(\|x_n - y\|^2 + 2\lambda_n L \|x_n - y_{n-1}\|^2 \right) - \\ - \left(\|x_{n+1} - y\|^2 + 2\lambda_{n+1} L \|x_{n+1} - y_n\|^2 \right) - \\ - (1 - 2\lambda_{n+1} L - \lambda_n L) \|x_{n+1} - y_n\|^2 - \\ - (1 - 2\lambda_n L) \|y_n - x_n\|^2 . \quad (2.19)$$

Припустимо, що $\lambda_n \in \left(0, \frac{1}{3L}\right]$. Тоді з (2.19) випливає

$$2\lambda_n (Ay, y_n - y) \leqslant \left(\|x_n - y\|^2 + 2\lambda_n L \|x_n - y_{n-1}\|^2 \right) - \\ - \left(\|x_{n+1} - y\|^2 + 2\lambda_{n+1} L \|x_{n+1} - y_n\|^2 \right) . \quad (2.20)$$

Просумувавши (2.20) по n від 1 до N отримаємо

$$2 \sum_{n=1}^{N} \lambda_n (Ay, y_n - y) \leqslant \|x_1 - y\|^2 + 2\lambda_1 L \|x_1 - y_0\|^2 ,$$

та

$$(Ay, z_N - y) \leqslant \frac{\|x_1 - y\|^2}{2 \sum_{n=1}^{N} \lambda_n}, \quad (2.21)$$

де $z_N = \frac{\sum_{n=1}^{N} \lambda_n y_n}{\sum_{n=1}^{N} \lambda_n}$. Переходимо до супремуму за $y \in C$ в (2.20)

$$\operatorname{gap}(z_N) \leqslant \frac{\sup_{y \in C} \|x_1 - y\|^2}{2 \sum_{n=1}^{N} \lambda_n}.$$

Таким чином, має місце

Теорема 2.2. *Нехай H — гільбертовий простір, $C \subseteq H$ — непорожня опукла замкнена обмежена множина, $A : H \to H$ — монотонний та L-ліпшицевий на множині C оператор. Нехай (y_n) — послідовність, що породжена алгоритмом 2.1 з $\lambda_n = \frac{1}{3L}$, тобто,*

$$\begin{cases} x_1 = y_0 \in C, \\ y_n = P_C\left(x_n - \frac{1}{3L} A y_{n-1}\right), \\ x_{n+1} = P_C\left(x_n - \frac{1}{3L} A y_n\right). \end{cases}$$

Тоді для послідовності середніх $z_N = \frac{1}{N}\sum_{n=1}^{N} y_n$ має місце оцінка

$$\operatorname{gap}(z_N) \leqslant \frac{3L \sup_{y\in C} \|x_1 - y\|^2}{2N}.$$

2.4. Сильна збіжність

Припустимо, що оператор $A : H \to H$ є рівномірно монотонним на обмежених підмножинах множини $C \subseteq H$. Тоді варіаційна нерівність (2.1) має єдиний розв'язок $\bar{z} \in C$.

Покажемо, що породжені алгоритмом 2.1 послідовності (x_n), (y_n) сильно збігаються до $\bar{z}$.

Осклільки множина $\{y_n\} \cup \{\bar{z}\} \subseteq C$ обмежена, то якщо існує така зростаюча функція $\phi : [0, +\infty) \to [0, +\infty)$, $\phi(0) = 0$ та $\phi(t) > 0$ для $t > 0$, що

$$(Ay_n, \bar{z} - y_n) \leqslant -\phi\left(\|\bar{z} - y_n\|\right).$$

Замість нерівності леми 2.7 запишемо її уточнену версію

$$\|x_{n+1} - \bar{z}\|^2 + 2\lambda_n \phi\left(\|\bar{z} - y_n\|\right) \leqslant \|x_n - \bar{z}\|^2 - (1 - \lambda_n L)\|x_{n+1} - y_n\|^2 - \\ - (1 - 2\lambda_n L)\|y_n - x_n\|^2 + 2\lambda_n L \|x_n - y_{n-1}\|^2. \quad (2.22)$$

Перепишемо (2.22) у вигляді

$$
\begin{aligned}
2\lambda_n \phi\left(\|\bar{z}-y_n\|\right) \leqslant \left(\|x_n-\bar{z}\|^2 + 2\lambda_n L\|x_n-y_{n-1}\|^2\right) - \\
-\left(\|x_{n+1}-\bar{z}\|^2 + 2\lambda_{n+1} L\|x_{n+1}-y_n\|^2\right) - \\
-\left(1-2\lambda_{n+1}L-\lambda_n L\right)\|x_{n+1}-y_n\|^2 - \\
-\left(1-2\lambda_n L\right)\|y_n-x_n\|^2. \quad (2.23)
\end{aligned}
$$

Нерівність (2.23) та припущення $\lambda_n \in \left[\underline{\lambda}, \overline{\lambda}\right] \subseteq \left(0, \frac{1}{3L}\right)$ дають

$$
\sum_{n=1}^{\infty} \lambda_n \phi\left(\|\bar{z}-y_n\|\right) < +\infty \quad \text{та} \quad \lim_{n\to\infty} \phi\left(\|\bar{z}-y_n\|\right) = 0.
$$

Отже, $\|\bar{z}-y_n\| \to 0$ та $\|\bar{z}-x_n\| \to 0$.

Таким чином, має місце

Теорема 2.3. *Нехай H — гільбертовий простір, $C \subseteq H$ — непорожня опукла замкнена множина, $A : H \to H$ — рівномірно монотонний на обмежених підмножинах множини $C \subseteq H$ та L-ліпшицевий на множині C оператор, $\bar{z} \in C$ — єдиний розв'язок варіаційної нерівності (2.1). Припустимо, що $\lambda_n \in \left[\underline{\lambda}, \overline{\lambda}\right] \subseteq \left(0, \frac{1}{3L}\right)$. Тоді породжені алгоритмом 2.1 послідовності (x_n), (y_n) сильно збігаються до $\bar{z}$.*

2.5. Лінійна швидкість збіжності

Припустимо, що існує єдиний розв'язок $z \in C$ варіаційної нерівності (2.1), а оператор A задовольняє умову

$$
(Ax, x-z) \geqslant \mu\|x-z\|^2 \quad \forall x \in C \tag{2.24}
$$

для деякого $\mu > 0$.

Зауваження 2.4. Умова (2.24) випливає з сильної монотонності A. Але існують немонотонні оператори з (2.24). Наприклад [176],

$$
Ax = (2-\|x\|)x, \quad x \in C = \left\{x \in \ell_2 : \|x\| \leqslant \frac{3}{2}\right\}.
$$

Розглянемо варіант алгоритму 2.1 з лінійною швидкістю збіжності для варіаціної нерівності (2.1), де ліпшицевий оператор A задовольняє (2.24).

Алгоритм 2.2. Для $x_1 = y_0 \in C$ генерируємо послідовність елементів x_n, $y_n \in C$ за допомогою ітераційної схеми

$$\begin{cases} y_n = P_C\left(x_n - \frac{1}{4L}Ay_{n-1}\right), \\ x_{n+1} = P_C\left(x_n - \frac{1}{4L}Ay_n\right). \end{cases}$$

Покажемо, що

$$\|x_{n+1} - z\|^2 = O\left(\left(1 - \frac{\mu}{4L}\right)^n\right), \quad \|y_n - x_{n+1}\|^2 = \left(\left(1 - \frac{\mu}{4L}\right)^n\right).$$

Запишемо нерівність (2.15):

$$\|x_{n+1} - z\|^2 \leqslant \|x_n - z\|^2 - \|x_n - y_n\|^2 - \|y_n - x_{n+1}\|^2 + \\ + 2\lambda\left\{(Ay_n, z - y_n) + (Ay_{n-1} - Ay_n, x_{n+1} - y_n)\right\}.$$

З умови (2.24) та нерівності $\|a + b\|^2 \leqslant 2\|a\|^2 + 2\|b\|^2$ випливає

$$(Ay_n, y_n - z) \geqslant \mu\|y_n - z\|^2 \geqslant \mu\left(\frac{1}{2}\|x_n - z\|^2 - \|y_n - x_n\|^2\right),$$

а ліпшицевість A гарантує виконання нерівності

$$(Ay_{n-1} - Ay_n, x_{n+1} - y_n) \leqslant L\|y_{n-1} - y_n\|\,\|x_{n+1} - y_n\| \leqslant \\ \leqslant \frac{L}{2}\|y_{n-1} - y_n\|^2 + \frac{L}{2}\|y_n - x_{n+1}\|^2.$$

Використавши вищенаведені оцінки, отримуємо

$$\|x_{n+1} - z\|^2 \leqslant (1 - \lambda\mu)\|x_n - z\|^2 - (1 - 2\lambda\mu)\|x_n - y_n\|^2 - \\ - \|y_n - x_{n+1}\|^2 + \lambda L\|y_{n-1} - y_n\|^2 + \lambda L\|y_n - x_{n+1}\|^2. \quad (2.25)$$

Член $\|y_{n-1} - y_n\|^2$ оцінимо таким чином

$$\|y_{n-1} - y_n\|^2 \leqslant 2\|y_{n-1} - x_n\|^2 + 2\|y_n - x_n\|^2.$$

Ураховуючи цю оцінку в (2.25), приходимо до нерівності

$$\begin{aligned}\|x_{n+1}-z\|^2 \leqslant (1-\lambda\mu)\|x_n-z\|^2-(1-2\lambda L-2\lambda\mu)\|x_n-y_n\|^2-\\ -(1-2\lambda L)\|y_n-x_{n+1}\|^2+2\lambda L\|y_{n-1}-x_n\|^2.\end{aligned}$$

Перепишемо останню нерівність у вигляді

$$\begin{aligned}\|x_{n+1}-z\|^2+\frac{2\lambda L}{1-\lambda\mu}\|y_n-x_{n+1}\|^2 \leqslant\\ \leqslant (1-\lambda\mu)\left(\|x_n-z\|^2+\frac{2\lambda L}{1-\lambda\mu}\|y_{n-1}-x_n\|^2\right)-\\ -(1-2\lambda L-2\lambda\mu)\|x_n-y_n\|^2-\\ -\left(1-2\lambda L-\frac{2\lambda L}{1-\lambda\mu}\right)\|y_n-x_{n+1}\|^2.\end{aligned}$$

Для $\lambda=\frac{1}{4L}$ отримуємо

$$\begin{aligned}\|x_{n+1}-z\|^2+\frac{2L}{4L-\mu}\|y_n-x_{n+1}\|^2 \leqslant\\ \leqslant\left(1-\frac{\mu}{4L}\right)\left(\|x_n-z\|^2+\frac{2L}{4L-\mu}\|y_{n-1}-x_n\|^2\right).\end{aligned}$$

Таким чином, має місце

Теорема 2.4. *Нехай H — гільбертовий простір, $C\subseteq H$ — непорожня опукла замкнена множина, $A: H\to H$ — L-ліпшицевий на множині C оператор, існує єдиний розв'язок $z\in C$ варіаціної нерівності (2.1) та виконується умова (2.24). Тоді для породжених алгоритмом 2.2 послідовностей (x_n), (y_n) виконується оцінка*

$$\|x_{n+1}-z\|^2+\frac{1}{2}\|y_n-x_{n+1}\|^2 \leqslant\left(1-\frac{\mu}{4L}\right)^n\|x_1-z\|^2,\quad n\geqslant 1.$$

Зауваження 2.5. З теореми 2.4 випливають оцінки

$$\|x_{n+1}-z\|^2\leqslant e^{-\frac{\mu}{4L}n}\|x_1-z\|^2,\quad \|y_n-x_{n+1}\|^2\leqslant e^{-\frac{\mu}{4L}n}2\|x_1-z\|^2.$$

Зауваження 2.6. Для сильно монотонних та ліпшицевих операторів та алгоритму 2.2 оцінка $\|x_{n+1}-z\|^2=O\left(e^{-\frac{\mu}{4L}n}\right)$ отримана в [55].

2.6. Адаптивний алгоритм екстраполяції з минулого

Для наближеного розв'язання задачі (2.1) розглянемо алгоритм екстраполяції з минулого з адаптивним вибором λ_n.

Алгоритм 2.3. Обираємо елементи $x_1 = y_0 \in C$, $\tau \in \left(0, \frac{1}{3}\right)$, $\lambda_1 \in (0, +\infty)$. Покладаємо $n = 1$.

1. Обчислити

$$y_n = P_C(x_n - \lambda_n A y_{n-1}).$$

2. Обчислити

$$x_{n+1} = P_C(x_n - \lambda_n A y_n).$$

Якщо $x_{n+1} = x_n = y_n$, то зупинити та x_n — розв'язок. Інакше перейти на крок **3**.

3. Обчислити

$$\lambda_{n+1} = \begin{cases} \lambda_n, \ \text{якщо} \ (Ay_{n-1} - Ay_n, x_{n+1} - y_n) \leqslant 0, \\ \min\left\{\lambda_n, \frac{\tau}{2}\frac{\|y_{n-1}-y_n\|^2 + \|x_{n+1}-y_n\|^2)}{(Ay_{n-1}-Ay_n, x_{n+1}-y_n)}\right\}, \quad \text{інакше.} \end{cases}$$

Покласти $n := n + 1$ та перейти на **1**.

Зауваження 2.7. Схема, що подібна алгоритму 2.3, запропонована для задач про рівновагу в просторах Адамара в [152, 153].

Зауваження 2.8. Результати, аналогічні наведеним нижче, мають місце для модифікації алгоритму 2.3 з заміною інструкції перерахунку λ_n на таку

$$\lambda_{n+1} = \begin{cases} \min\left\{\lambda_n, \tau\frac{\|y_{n-1}-y_n\|}{\|Ay_{n-1}-Ay_n\|}\right\}, \quad \text{якщо} \quad Ay_{n-1} \neq Ay_n, \\ \lambda_n, \qquad \text{інакше.} \end{cases}$$

де $\tau \in \left(0, \frac{1}{3}\right)$ [109].

Параметр λ_{n+1} залежить від розташування точок y_{n-1}, y_n, x_{n+1}, значень Ay_{n-1}, Ay_n. Інформація про константу L не використовується. Очевидно, що послідовність (λ_n) незростаюча. Також вона обмежена знизу числом

$$\min\left\{\lambda_1, \frac{\tau}{L}\right\}.$$

Дійсно, маємо

$$(Ay_{n-1} - Ay_n, x_{n+1} - y_n) \leqslant \frac{L}{2}\left(\|y_{n-1} - y_n\|^2 + \|y_n - x_{n+1}\|^2\right).$$

Для послідовностей (x_n), (y_n), породжених алгоритмом 2.3, мають місце нерівності

$$(Ay_{n-1}, y_n - y) \leqslant \\ \leqslant \frac{1}{2\lambda_n}\left(\|y - x_n\|^2 - \|x_n - y_n\|^2 - \|y_n - y\|^2\right) \quad \forall y \in C, \quad (2.26)$$

$$(Ay_n, x_{n+1} - y) \leqslant \\ \leqslant \frac{1}{2\lambda_n}\left(\|y - x_n\|^2 - \|x_n - x_{n+1}\|^2 - \|x_{n+1} - y\|^2\right) \quad \forall y \in C. \quad (2.27)$$

Зауваження 2.9. Зокрема, нерівність (2.27) дає обґрунтування правила зупинки алгоритму 2.3. Дійсно, при $x_{n+1} = x_n = y_n$ із (2.27) випливає

$$(Ay_n, y_n - y) \leqslant 0 \qquad \forall y \in C,$$

тобто $x_n = y_n \in S$.

Зауваження 2.10. Можна використовувати для зупинки алгоритму 2.3 правило $x_n = y_n = y_{n-1}$, яке гарантує $x_n \in S$.

Доведемо основну оцінку, яка пов'язує відстані між породженими алгоритмом 2.3 точками і довільним елементом множини розв'язків S.

Припустимо додатково, що оператор A є псевдомонотонним.

Лема 2.8. *Для послідовностей (x_n), (y_n), породжених алгоритмом 2.3, має місце нерівність*

$$\|x_{n+1} - z\|^2 \leqslant \|x_n - z\|^2 - \left(1 - \tau\frac{\lambda_n}{\lambda_{n+1}}\right)\|x_{n+1} - y_n\|^2 - \\ - \left(1 - 2\tau\frac{\lambda_n}{\lambda_{n+1}}\right)\|x_n - y_n\|^2 + 2\tau\frac{\lambda_n}{\lambda_{n+1}}\|x_n - y_{n-1}\|^2, \quad (2.28)$$

де $z \in S$.

Доведення. Нехай $z \in S$. Із псевдомонотонності оператора A випливає

$$(Ay_n, z - y_n) \leqslant 0. \tag{2.29}$$

Із (2.29) і (2.27) випливає

$$2\lambda_n(Ay_n, x_{n+1} - y_n) \leqslant \|z - x_n\|^2 - \|x_n - x_{n+1}\|^2 - \|x_{n+1} - z\|^2. \tag{2.30}$$

Склавши нерівності (2.30) та нерівність, що випливає з (2.26)

$$2\lambda_n(Ay_{n-1}, y_n - x_{n+1}) \leqslant \|x_{n+1} - x_n\|^2 - \|x_n - y_n\|^2 - \|y_n - x_{n+1}\|^2$$

маємо

$$\begin{aligned} 2\lambda_n(Ay_{n-1} - Ay_n, y_n - x_{n+1}) \leqslant \|z - x_n\|^2 - \\ - \|z - x_{n+1}\|^2 - \|x_n - y_n\|^2 - \|x_{n+1} - y_n\|^2. \end{aligned} \tag{2.31}$$

З правила обчислення λ_{n+1} випливає нерівність

$$(Ay_{n-1} - Ay_n, x_{n+1} - y_n) \leqslant \frac{\tau}{2\lambda_{n+1}} \left(\|y_{n-1} - y_n\|^2 + \|y_n - x_{n+1}\|^2\right). \tag{2.32}$$

Для оцінки виразу $(Ay_{n-1} - Ay_n, y_n - x_{n+1})$ в (2.31) скористаємося (2.32). Отримаємо

$$\begin{aligned} \|z - x_{n+1}\|^2 \leqslant \|z - x_n\|^2 - \|x_n - y_n\|^2 - \|x_{n+1} - y_n\|^2 + \\ + \tau \frac{\lambda_n}{\lambda_{n+1}} \big(\|y_{n-1} - y_n\|^2 + \|y_n - x_{n+1}\|^2\big). \end{aligned}$$

Оскільки

$$\|y_{n-1} - y_n\|^2 \leqslant 2\left\|y_{n-1} - x_n\right\|^2 + 2\left\|y_n - x_n\right\|^2,$$

то

$$\begin{aligned} \|z - x_{n+1}\|^2 \leqslant \|z - x_n\|^2 - \|x_n - y_n\|^2 - \|x_{n+1} - y_n\|^2 + \\ + 2\tau \frac{\lambda_n}{\lambda_{n+1}} \left\|y_{n-1} - x_n\right\|^2 + 2\tau \frac{\lambda_n}{\lambda_{n+1}} \left\|y_n - x_n\right\|^2 + \tau \frac{\lambda_n}{\lambda_{n+1}} \|y_n - x_{n+1}\|^2, \end{aligned}$$

що і треба було довести. $\square$

Сформулюємо основний результат.

Теорема 2.5. *Нехай H — гільбертовий простір, $C \subseteq H$ — непорожня опукла замкнена множина, $A : H \to H$ — монотонний та L-ліпшицевий на множині C оператор та $S \neq \varnothing$. Тоді породжені алгоритмом 2.3 послідовності (x_n), (y_n) слабко збігаються до розв'язку $\bar{z} \in S$ варіаційної нерівності (2.1), причому*

$$\lim_{n\to\infty} \|x_n - y_n\| = \lim_{n\to\infty} \|x_{n+1} - y_n\| = 0.$$

Доведення. Нехай $z' \in S$. Покладемо

$$\alpha_n = \|x_n - z'\|^2 + 2\tau \frac{\lambda_n}{\lambda_{n+1}} \|x_n - y_{n-1}\|^2,$$

$$\beta_n = \left(1 - 2\tau \frac{\lambda_n}{\lambda_{n+1}}\right) \|y_n - x_n\|^2 + \left(1 - 2\tau \frac{\lambda_{n+1}}{\lambda_{n+2}} - \tau \frac{\lambda_n}{\lambda_{n+1}}\right) \|x_{n+1} - y_n\|^2.$$

Нерівність (2.28) набуває вигляду

$$\alpha_{n+1} \leqslant \alpha_n - \beta_n.$$

Оскільки існує $\lim\limits_{n\to\infty} \lambda_n > 0$, то

$$1 - 2\tau \frac{\lambda_n}{\lambda_{n+1}} \to 1 - 2\tau \in (0,1) \quad \text{і} \quad 1 - 2\tau \frac{\lambda_{n+1}}{\lambda_{n+2}} - \tau \frac{\lambda_n}{\lambda_{n+1}} \to 1 - 3\tau \in (0,1)$$

при $n \to \infty$. З леми 2.2 можемо зробити висновок, що існує границя

$$\lim_{n\to\infty} \left(\|x_n - z'\|^2 + 2\tau \frac{\lambda_n}{\lambda_{n+1}} \|x_n - y_{n-1}\|^2 \right)$$

та збігається числовий ряд

$$\sum_n \left(\left(1 - 2\tau \frac{\lambda_n}{\lambda_{n+1}}\right) \|y_n - x_n\|^2 + \left(1 - 2\tau \frac{\lambda_{n+1}}{\lambda_{n+2}} - \tau \frac{\lambda_n}{\lambda_{n+1}}\right) \|x_{n+1} - y_n\|^2 \right).$$

Звідки отримуємо

$$\lim_{n\to\infty} \|y_n - x_n\| = \lim_{n\to\infty} \|x_{n+1} - y_n\| = \lim_{n\to\infty} \|x_{n+1} - x_n\| = 0 \tag{2.33}$$

і збіжність числових послідовностей $\big(\|x_n - z'\|\big)$, $\big(\|y_n - z'\|\big)$ для всіх $z' \in S$. Зокрема, послідовності (x_n), (y_n) обмежені.

Розглянемо підпослідовність (x_{n_k}), слабко збіжну до деякої точки $\bar{z} \in C$. Тоді з (2.33) випливає, що й $y_{n_k} \rightharpoonup \bar{z}$. Покажемо, що $\bar{z} \in S$. Маємо

$$(Ay_n, y - y_n) \geqslant (Ay_n, x_{n+1} - y_n) + \frac{(x_{n+1} - x_n, x_{n+1} - y)}{\lambda_n} \quad \forall y \in C. \tag{2.34}$$

З (2.33) випливає

$$\lim_{n\to\infty}(Ay_n, x_{n+1} - y_n) = 0. \tag{2.35}$$

Здійснивши граничний перехід в (2.34) з урахуванням (2.33), (2.35) та умови монотонності оператора A, отримаємо

$$\begin{aligned}(Ay, y - \bar{z}) = \lim_{k\to\infty}(Ay, y - y_{n_k}) \geqslant \limsup_{k\to\infty}(Ay_{n_k}, y - y_{n_k}) \geqslant \\ \geqslant \lim_{k\to\infty}\left\{(Ay_{n_k}, x_{n_k+1} - y_{n_k}) + \frac{(x_{n_k+1} - x_{n_k}, x_{n_k+1} - y)}{\lambda_{n_k}}\right\} = 0 \quad \forall y \in C,\end{aligned}$$

тобто, $\bar{z} \in S$.

Покажемо тепер, що $x_n \rightharpoonup \bar{z} \in S$. Тоді з $x_n - y_n \to 0$ випливає, що й послідовність (y_n) слабко збігається $\bar{z}$. Міркуємо від супротивного. Нехай існує така підпослідовність (x_{m_k}), що $x_{m_k} \rightharpoonup \tilde{z}$ та $\tilde{z} \neq \bar{z}$. Ясно, що $\tilde{z} \in S$. Застосуємо двічі лему 2.6. Маємо

$$\begin{aligned}\lim_{n\to\infty}\|x_n - \bar{z}\| = \lim_{k\to\infty}\|x_{n_k} - \bar{z}\| < \lim_{k\to\infty}\|x_{n_k} - \tilde{z}\| = \lim_{n\to\infty}\|x_n - \tilde{z}\| = \\ = \lim_{k\to\infty}\|x_{m_k} - \tilde{z}\| < \lim_{k\to\infty}\|x_{m_k} - \bar{z}\| = \lim_{n\to\infty}\|x_n - \bar{z}\|,\end{aligned}$$

що неможливо. Отже, $x_n \rightharpoonup \bar{z}$. □

2.7. Регуляризований алгоритм

Для непорожньої опуклої замкненої множини $C \subseteq H$, монотонного ліпшицевого оператора $A : H \to H$ та $z \in H$ розглянемо задачу:

$$\text{знайти } \ x^* = P_S z, \tag{2.36}$$

де $S = \{x \in C : \ (Ax, y - x) \geqslant 0 \ \ \forall\, y \in C\}$.

Будемо звичайно припускати, що $S \neq \varnothing$. Задача (2.36) має єдиний розв'язок [31].

Зауваження 2.11. Відомим окремим випадком (2.36) є задача пошуку нормального розв'язку варіаційної нерівності (при $z = 0$).

Апроксимуємо задачу (2.36) однорівневою та більш регулярною варіаційною нерівністю.

Розглянемо допоміжну задачу:

$$\text{знайти } \ x \in C : \quad (Ax, y - x) + \varepsilon(x - z, y - x) \geqslant 0 \quad \forall y \in C, \tag{2.37}$$

де $\varepsilon > 0$.

Зауваження 2.12. Варіаційну нерівність (2.37) називають апроксимацією Тихонова–Браудера задачі (2.36) [96]. Для розв'язання екстремальних задач подібна апроксимація була запропонована А. М. Тихоновим для побудови регуляризуючих алгоритмів, а пізніше F. Browder [177,178] застосував подібну схему для стійкої апроксимації нормального розв'язку варіаційної нерівності або проекції заданої точки на множину нерухомих точок нерозтягуючих операторів.

З результатів [179] випливає існування та єдиність розв'язку $x_\varepsilon \in C$ задачі (2.37) для довільного $\varepsilon > 0$.

Елементи $x_\varepsilon \in C$ мають декілька важливих властивостей [162].

Лема 2.9. *Справедливі такі нерівності:*

(i) $\|x_\varepsilon\| \leqslant \|x^* - z\| + \|x^*\|$ *для всіх* $\varepsilon > 0$*;*

(ii) $\|x_\varepsilon - x_\delta\| \leqslant \frac{|\varepsilon - \delta|}{\varepsilon} 2\|x^* - z\|$ *для всіх* ε, $\delta > 0$.

При прямуванні малого додатнього параметру ε до нуля елементи x_ε сильно збігаються до розв'язку задачі (2.36).

Лема 2.10. *Має місце*

$$\lim_{\varepsilon \to 0} \|x_\varepsilon - x^*\| = 0.$$

Перейдемо до опису алгоритму розв'язання задачі (2.36).

Відштовхуючись від алгоритму 2.1, для розв'язання задачі (2.36) пропонуємо такий

Алгоритм 2.4. Для $z \in H$, $x_1 = y_0 \in C$ генерируємо послідовність елементів x_n, $y_n \in C$ за допомогою ітераційної схеми

$$\begin{cases} y_n = P_C\left(\alpha_n\lambda_n z + (1-\alpha_n\lambda_n)x_n - \lambda_n A y_{n-1}\right), \\ x_{n+1} = P_C\left(\alpha_n\lambda_n z + (1-\alpha_n\lambda_n)x_n - \lambda_n A y_n\right), \end{cases}$$

де $\lambda_n > 0$, $\alpha_n > 0$.

Відносно параметрів алгоритму 2.4 будемо припускати, що виконані такі умови:

(A1) $\lambda_n \in \left[\underline{\lambda}, \overline{\lambda}\right] \subseteq \left(0, \frac{1}{3L}\right)$, де $L > 0$ — стала Ліпшиця оператора A;

(A2) $\lim_{n\to\infty} \alpha_n = 0$;

(A3) $\sum_{n=1}^{\infty} \alpha_n = +\infty$;

(A4) $\lim_{n\to\infty} \frac{\alpha_{n+1}-\alpha_n}{\alpha_n^2} = 0$.

Зауваження 2.13. В якості допустимої послідовності (α_n) можна обрати таку:

$$\alpha_n = \frac{1}{(n+1)^p}, \;\; p \in (0,1).$$

Зауваження 2.14. В [99] для був запропонований та обгрунтований близький алгоритм:

$$\begin{cases} x_1 \in C, \\ y_n = P_C(x_n - \lambda_n A x_n), \\ z_n = P_C(x_n - \lambda_n A y_n), \\ x_{n+1} = \alpha_n z + (1-\alpha_n) z_n. \end{cases}$$

Алгоритм 2.4 поєднує у собі ідеї методу екстраполяції з минулого (алгоритм 2.1) та ітеративної регуляризації [96]. Доведення його сильної збіжності проведемо за такою схемою. Нехай x_{α_n} — розв'язок задачі (2.37) при $\varepsilon = \alpha_n$. Оскільки

$$\|x_n - x^*\| \leqslant \|x_n - x_{\alpha_n}\| + \|x_{\alpha_n} - x^*\|, \quad \lim_{n\to\infty} \|x_{\alpha_n} - x^*\| = 0,$$

то достатньо показати, що породжена алгоритмом 2.4 послідовність (x_n) має властивість

$$\lim_{n\to\infty} \|x_n - x_{\alpha_n}\| = 0.$$

Доведення збіжності алгоритму 2.4 почнемо з доведення важливої нерівності для згенерованих ним послідовностей (x_n), (y_n) та елементів x_{α_n}.

Лема 2.11. *Для породжених алгоритмом 2.4 послідовностей* (x_n), (y_n) *та елементів* x_{α_n} *виконується нерівність*

$$\begin{aligned}\|x_{n+1} - x_{\alpha_n}\|^2 \leqslant (1 - \alpha_n\lambda_n) \|x_n - x_{\alpha_n}\|^2 - \\ - (1 - 2\lambda_n L - \alpha_n\lambda_n) \|x_n - y_n\|^2 - (1 - \lambda_n L) \|y_n - x_{n+1}\|^2 + \\ + 2\lambda_n L \|y_{n-1} - x_n\|^2 . \quad (2.38)\end{aligned}$$

Доведення. Маємо

$$\begin{aligned}\|x_{n+1} - x_{\alpha_n}\|^2 = \|x_n - x_{\alpha_n}\|^2 - \|x_n - x_{n+1}\|^2 + 2(x_{n+1} - x_n, x_{n+1} - x_{\alpha_n}) = \\ = \|x_n - x_{\alpha_n}\|^2 - \|x_n - y_n\|^2 - \|y_n - x_{n+1}\|^2 - \\ - 2(x_n - y_n, y_n - x_{n+1}) + 2(x_{n+1} - x_n, x_{n+1} - x_{\alpha_n}) . \quad (2.39)\end{aligned}$$

З визначення точок x_{n+1} та y_n випливає

$$\lambda_n(Ay_n, x_{\alpha_n} - x_{n+1}) \geqslant (x_{n+1} - x_n + \alpha_n\lambda_n(x_n - z), x_{n+1} - x_{\alpha_n}), \qquad (2.40)$$

$$\lambda_n(Ay_{n-1}, x_{n-1} - y_n) \geqslant -(x_n - \alpha_n\lambda_n(x_n - z) - y_n, y_n - x_{n+1}). \qquad (2.41)$$

Використовуючи нерівності (2.40), (2.41) для оцінки скалярних добутків в (2.39), отримуємо

$$\|x_{n+1} - x_{\alpha_n}\|^2 \leqslant \|x_n - x_{\alpha_n}\|^2 - \|x_n - y_n\|^2 - \|y_n - x_{n+1}\|^2 +$$

$$+2\lambda_n\left\{(Ay_{n-1}, x_{n+1}-y_n)+(Ay_n, x_{\alpha_n}-x_{n+1})\right\}+$$
$$+2\alpha_n\lambda_n(x_n-z, x_{\alpha_n}-y_n)=$$
$$=\|x_n-x_{\alpha_n}\|^2-\|x_n-y_n\|^2-\|y_n-x_{n+1}\|^2+$$
$$+2\lambda_n\left\{(Ay_n, x_{\alpha_n}-y_n)+(Ay_{n-1}, x_{n+1}-y_n)+(Ay_n, y_n-x_{n+1})\right\}+$$
$$+2\alpha_n\lambda_n(x_n-z, x_{\alpha_n}-y_n). \quad (2.42)$$

Ліпшицевість оператора A гарантує виконання нерівності

$$(Ay_{n-1}-Ay_n, x_{n+1}-y_n)\leqslant L\|y_{n-1}-y_n\|\,\|x_{n+1}-y_n\|\leqslant$$
$$\leqslant\frac{L}{2}\|y_{n-1}-y_n\|^2+\frac{L}{2}\|y_n-x_{n+1}\|^2.$$

Використавши вищенаведену оцінку в (2.42), отримуємо

$$\|x_{n+1}-x_{\alpha_n}\|^2\leqslant\|x_n-x_{\alpha_n}\|^2-\|x_n-y_n\|^2-\|y_n-x_{n+1}\|^2+$$
$$+\lambda_nL\|y_{n-1}-y_n\|^2+\lambda_nL\|y_n-x_{n+1}\|^2+$$
$$+2\alpha_n\lambda_n(x_n-z, x_{\alpha_n}-y_n)+2\lambda_n(Ay_n, x_{\alpha_n}-y_n). \quad (2.43)$$

Член $\|y_{n-1}-y_n\|^2$ оцінимо таким чином

$$\|y_{n-1}-y_n\|^2\leqslant 2\|y_{n-1}-x_n\|^2+2\|y_n-x_n\|^2.$$

Маємо

$$\|x_{n+1}-x_{\alpha_n}\|^2\leqslant\|x_n-x_{\alpha_n}\|^2-\|x_n-y_n\|^2-\|y_n-x_{n+1}\|^2+$$
$$+2\lambda_nL\|y_{n-1}-x_n\|^2+2\lambda_nL\|y_n-x_n\|^2+\lambda_nL\|y_n-x_{n+1}\|^2+$$
$$+2\alpha_n\lambda_n(x_n-z, x_{\alpha_n}-y_n)+2\lambda_n(Ay_n, x_{\alpha_n}-y_n). \quad (2.44)$$

З монотонності оператора A випливає

$$(Ay_n, x_{\alpha_n}-y_n)\leqslant-(Ax_{\alpha_n}, y_n-x_{\alpha_n}),$$

звідки

$$(Ay_n, x_{\alpha_n}-y_n)-\alpha_n(x_{\alpha_n}-z, y_n-x_{\alpha_n})\leqslant-(Ax_{\alpha_n}, y_n-x_{\alpha_n})-\alpha_n(x_{\alpha_n}-z, y_n-x_{\alpha_n}).$$

Оскільки

$$(Ax_{\alpha_n}, y_n - x_{\alpha_n}) + \alpha_n(x_{\alpha_n} - z, y_n - x_{\alpha_n}) \geqslant 0,$$

то

$$(Ay_n, x_{\alpha_n} - y_n) \leqslant \alpha_n(x_{\alpha_n} - z, y_n - x_{\alpha_n}).$$

Урахувавши останню оцінку в (2.44), приходимо до нерівності

$$\begin{aligned} \|x_{n+1} - x_{\alpha_n}\|^2 \leqslant \|x_n - x_{\alpha_n}\|^2 - \|x_n - y_n\|^2 - \|y_n - x_{n+1}\|^2 + \\ + 2\lambda_n L \|y_{n-1} - x_n\|^2 + 2\lambda_n L \|y_n - x_n\|^2 + \lambda_n L \|y_n - x_{n+1}\|^2 + \\ + 2\alpha_n \lambda_n (x_n - x_{\alpha_n}, x_{\alpha_n} - y_n). \end{aligned} \quad (2.45)$$

Оцінимо зверху член $(x_n - x_{\alpha_n}, x_{\alpha_n} - y_n)$. Маємо

$$\begin{aligned} (x_n - x_{\alpha_n}, x_{\alpha_n} - y_n) = (x_n - x_{\alpha_n}, x_{\alpha_n} - x_n) + \\ + (x_n - x_{\alpha_n}, x_n - y_n) \leqslant -\|x_{\alpha_n} - x_n\|^2 + \|x_n - x_{\alpha_n}\| \|x_n - y_n\| \leqslant \\ \leqslant -\|x_{\alpha_n} - x_n\|^2 + \frac{1}{2}\|x_n - x_{\alpha_n}\|^2 + \frac{1}{2}\|x_n - y_n\|^2 = \\ = -\frac{1}{2}\|x_n - x_{\alpha_n}\|^2 + \frac{1}{2}\|x_n - y_n\|^2. \end{aligned} \quad (2.46)$$

З нерівностей (2.45) та (2.46) отримуємо

$$\begin{aligned} \|x_{n+1} - x_{\alpha_n}\|^2 \leqslant (1 - \alpha_n \lambda_n \mu) \|x_n - x_{\alpha_n}\|^2 - \\ - (1 - 2\lambda_n L - \alpha_n \lambda_n) \|x_n - y_n\|^2 - (1 - \lambda_n L) \|y_n - x_{n+1}\|^2 + \\ + 2\lambda_n L \|y_{n-1} - x_n\|^2. \end{aligned}$$

що й було потрібно. □

Доведемо оцінку з якої випливає збіжність до нуля послідовностей $(\|x_n - x_{\alpha_n}\|)$ та $(\|y_{n-1} - x_n\|)$.

Лема 2.12. *Для породжених алгоритмом 2.4 послідовностей (x_n), (y_n) та елементів x_{α_n} при великих $n \in \mathbb{N}$ виконується нерівність*

$$\left\|x_{n+1}-x_{\alpha_{n+1}}\right\|^2+\frac{2\lambda_{n+1}L}{1-\alpha_{n+1}\lambda_{n+1}}\left\|y_n-x_{n+1}\right\|^2\leqslant$$
$$\leqslant\left(1-\frac{\alpha_n\lambda_n}{2}\right)\left(\left\|x_n-x_{\alpha_n}\right\|^2+\frac{2\lambda_nL}{1-\alpha_n\lambda_n}\left\|y_{n-1}-x_n\right\|^2\right)+$$
$$+\frac{8\|x^*-z\|^2}{\lambda_n}\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^3}. \quad (2.47)$$

Доведення. Маємо

$$\left\|x_{n+1}-x_{\alpha_n}\right\|^2=\left\|x_{n+1}-x_{\alpha_{n+1}}\right\|^2+\left\|x_{\alpha_{n+1}}-x_{\alpha_n}\right\|^2+$$
$$+2\left(x_{n+1}-x_{\alpha_{n+1}},x_{\alpha_{n+1}}-x_{\alpha_n}\right)\geqslant\left\|x_{n+1}-x_{\alpha_{n+1}}\right\|^2+$$
$$+\left\|x_{\alpha_{n+1}}-x_{\alpha_n}\right\|^2-2\left\|x_{n+1}-x_{\alpha_{n+1}}\right\|\left\|x_{\alpha_{n+1}}-x_{\alpha_n}\right\|\geqslant$$
$$\geqslant(1-\varepsilon)\left\|x_{n+1}-x_{\alpha_{n+1}}\right\|^2+\frac{(\varepsilon-1)}{\varepsilon}\left\|x_{\alpha_{n+1}}-x_{\alpha_n}\right\|^2, \quad (2.48)$$

де $\varepsilon>0$. Покладемо в (2.48) $\varepsilon=\frac{1}{2}\alpha_n\lambda_n$. Отримаємо

$$\left\|x_{n+1}-x_{\alpha_n}\right\|^2\geqslant\frac{2-\alpha_n\lambda_n}{2}\left\|x_{n+1}-x_{\alpha_{n+1}}\right\|^2-$$
$$-\frac{2-\alpha_n\lambda_n}{\alpha_n\lambda_n}\left\|x_{\alpha_{n+1}}-x_{\alpha_n}\right\|^2. \quad (2.49)$$

В силу правил узгодження значень параметрів α_n, λ_n при великих n маємо $1-\alpha_n\lambda_n>0$. З урахуванням другої нерівності леми 2.9 з (2.49) виводимо

$$\left\|x_{n+1}-x_{\alpha_n}\right\|^2\geqslant\frac{2-\alpha_n\lambda_n}{2}\left\|x_{n+1}-x_{\alpha_{n+1}}\right\|^2-$$
$$-\frac{(2-\alpha_n\lambda_n)}{\alpha_n\lambda_n}\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^2}4\|x^*-z\|^2, \quad (2.50)$$

для всіх $n\geqslant n_0$. Використавши (2.50) в (2.38), отримаємо (для $n\geqslant n_0$)

$$\frac{2-\alpha_n\lambda_n}{2}\left\|x_{n+1}-x_{\alpha_{n+1}}\right\|^2\leqslant(1-\alpha_n\lambda_n)\left\|x_n-x_{\alpha_n}\right\|^2-$$
$$-(1-2\lambda_nL-\alpha_n\lambda_n)\left\|x_n-y_n\right\|^2-(1-\lambda_nL)\left\|y_n-x_{n+1}\right\|^2+$$
$$+2\lambda_nL\left\|y_{n-1}-x_n\right\|^2+\frac{(2-\alpha_n\lambda_n)}{\alpha_n\lambda_n}\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^2}4\|x^*-z\|^2.$$

Звідки випливає нерівність

$$\left\|x_{n+1}-x_{\alpha_{n+1}}\right\|^2 \leqslant \frac{2-2\alpha_n\lambda_n}{2-\alpha_n\lambda_n}\left\|x_n-x_{\alpha_n}\right\|^2 - \\ -\frac{2\left(1-2\lambda_n L-\alpha_n\lambda_n\right)}{2-\alpha_n\lambda_n}\left\|x_n-y_n\right\|^2 - \frac{2\left(1-\lambda_n L\right)}{2-\alpha_n\lambda_n}\left\|y_n-x_{n+1}\right\|^2 + \\ +\frac{4\lambda_n L}{2-\alpha_n\lambda_n}\left\|y_{n-1}-x_n\right\|^2 + \frac{8\|x^*-z\|^2}{\lambda_n}\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^3}. \quad (2.51)$$

Перегрупувавши члени в (2.51), отримаємо

$$\left\|x_{n+1}-x_{\alpha_{n+1}}\right\|^2 + \frac{2\lambda_{n+1}L}{1-\alpha_{n+1}\lambda_{n+1}}\left\|y_n-x_{n+1}\right\|^2 \leqslant \\ \leqslant \frac{2-2\alpha_n\lambda_n}{2-\alpha_n\lambda_n}\left(\left\|x_n-x_{\alpha_n}\right\|^2 + \frac{2\lambda_n L}{1-\alpha_n\lambda_n}\left\|y_{n-1}-x_n\right\|^2\right) - \\ -\frac{2\left(1-2\lambda_n L-\alpha_n\lambda_n\right)}{2-\alpha_n\lambda_n}\left\|x_n-y_n\right\|^2 - \\ -\left(\frac{1-\lambda_n L}{1-\frac{\alpha_n\lambda_n}{2}} - \frac{2\lambda_{n+1}L}{1-\alpha_{n+1}\lambda_{n+1}}\right)\left\|y_n-x_{n+1}\right\|^2 + \\ +\frac{8\|x^*-z\|^2}{\lambda_n}\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^3}. \quad (2.52)$$

Оскільки

$$\lambda_n \in \left[\underline{\lambda}, \overline{\lambda}\right] \subseteq \left(0, \frac{1}{3L}\right) \quad \text{та} \quad \alpha_n > 0, \quad \lim_{n\to\infty}\alpha_n = 0,$$

то починаючи з деякого номера n_1 будуть виконуватися нерівності

$$\frac{1-2\lambda_n L-\alpha_n\lambda_n}{2-\alpha_n\lambda_n} > 0, \quad \frac{1-\lambda_n L}{1-\frac{\alpha_n\lambda_n}{2}} - \frac{2\lambda_{n+1}L}{1-\alpha_{n+1}\lambda_{n+1}} > 0,$$

та

$$\frac{2-2\alpha_n\lambda_n}{2-\alpha_n\lambda_n} < 1-\frac{\alpha_n\lambda_n}{2}.$$

Таким чином, для $n \geqslant N = \max\{n_0, n_1\}$ з (2.52) випливає

$$\left\|x_{n+1}-x_{\alpha_{n+1}}\right\|^2 + \frac{2\lambda_{n+1}L}{1-\alpha_{n+1}\lambda_{n+1}}\left\|y_n-x_{n+1}\right\|^2 \leqslant \\ \leqslant \left(1-\frac{\alpha_n\lambda_n}{2}\right)\left(\left\|x_n-x_{\alpha_n}\right\|^2 + \frac{2\lambda_n L}{1-\alpha_n\lambda_n}\left\|y_{n-1}-x_n\right\|^2\right) + \\ +\frac{8\|x^*-z\|^2}{\lambda_n}\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^3},$$

що й потрібно було довести. □

Сформулюємо теорему про збіжність алгоритму 2.4.

Теорема 2.6. *Нехай H — гільбертовий простір, $C \subseteq H$ — непорожня опукла замкнена множина, $A : H \to H$ — монотонний та L-ліпшицевий на множині C оператор та $S \neq \varnothing$. Нехай виконуються умови (A1)–(A4). Тоді для породжених алгоритмом 2.4 послідовностей (x_n), (y_n) має місце*

$$\lim_{n\to\infty} \|x_n - x^*\| = \lim_{n\to\infty} \|y_n - x^*\| = 0, \tag{2.53}$$

де $x^* \in H$ — єдиний розв'язок задачі (2.36).

Доведення. В силу леми 2.4 та нерівності (2.47) маємо

$$\lim_{n\to\infty} \left(\|x_n - x_{\alpha_n}\|^2 + \frac{2\lambda_n L}{1-\alpha_n\lambda_n} \|y_{n-1} - x_n\|^2 \right) = 0.$$

Звідки

$$\lim_{n\to\infty} \|x_n - x_{\alpha_n}\| = \lim_{n\to\infty} \|y_{n-1} - x_n\| = 0. \tag{2.54}$$

З нерівності

$$\|x_n - x^*\| \leqslant \|x_n - x_{\alpha_n}\| + \|x_{\alpha_n} - x^*\|,$$

леми 2.10 та (2.54) отримуємо шукані співвідношення (2.53). □

З метою позбутися явного використання в алгоритмі 2.4 інформації про значення константи L при заданні меж для λ_n розглянемо такий алгоритм з адаптивним вибором λ_n.

Алгоритм 2.5 (Адаптивний варіант). Обираємо елементи $z \in H$, $x_1 = y_0 \in C$, $\tau \in \left(0, \frac{1}{3}\right)$, $\lambda_1 \in (0, +\infty)$. Покладаємо $n = 1$.

1. Обчислити

$$y_n = P_C\left(\alpha_n\lambda_n z + (1-\alpha_n\lambda_n)x_n - \lambda_n A y_{n-1}\right).$$

2. Обчислити

$$x_{n+1} = P_C\left(\alpha_n\lambda_n z + (1-\alpha_n\lambda_n)x_n - \lambda_n A y_n\right).$$

3. Обчислити

$$\lambda_{n+1} = \begin{cases} \lambda_n, \ \text{якщо} \ \ (Ay_{n-1} - Ay_n, x_{n+1} - y_n) \leqslant 0, \\ \min\left\{\lambda_n, \frac{\tau}{2}\frac{\|y_{n-1}-y_n\|^2 + \|y_n - x_{n+1}\|^2}{(Ay_{n-1}-Ay_n, x_{n+1}-y_n)}\right\}, \quad \text{інакше.} \end{cases}$$

Покласти $n := n + 1$ та перейти на **1**.

Зауваження 2.15. Алгоритм 2.5 є поєднанням алгоритму 2.4 з адаптивним перерахунком λ_n з алгоритму 2.3.

У пропонованому алгоритмі параметр λ_{n+1} залежить від розташування точок y_{n-1}, y_n, x_{n+1}, значень Ay_{n-1}, Ay_n. Ніяка інформація про константу L не використовується. Очевидно, що послідовність (λ_n) незростаюча. Також вона обмежена знизу числом

$$\min\left\{\lambda_1, \frac{\tau}{L}\right\}.$$

Дійсно, маємо

$$(Ay_{n-1} - Ay_n, x_{n+1} - y_n) \leqslant \frac{L}{2}\left(\|y_{n-1} - y_n\|^2 + \|y_n - x_{n+1}\|^2\right).$$

Перейдемо до обґрунтування алгоритму 2.5. А саме, покажемо, що

$$\lim_{n\to\infty} \|x_n - x_{\alpha_n}\| = 0,$$

де x_{α_n} — розв'язок задачі (2.37) при $\varepsilon = \alpha_n$.

Спочатку доведемо важливу нерівність для послідовностей (x_n), (y_n) та елементів x_{α_n}.

Лема 2.13. *Для породжених алгоритмом 2.5 послідовностей* (x_n), (y_n) *та елементів* x_{α_n} *виконується нерівність*

$$\begin{aligned}\|x_{n+1} - x_{\alpha_n}\|^2 &\leqslant (1 - \alpha_n\lambda_n)\|x_n - x_{\alpha_n}\|^2 - \\ &- \left(1 - \tau\frac{\lambda_n}{\lambda_{n+1}}\right)\|x_{n+1} - y_n\|^2 - \left(1 - 2\tau\frac{\lambda_n}{\lambda_{n+1}} - \alpha_n\lambda_n\right)\|y_n - x_n\|^2 + \\ &+ 2\tau\frac{\lambda_n}{\lambda_{n+1}}\|x_n - y_{n-1}\|^2. \quad (2.55)\end{aligned}$$

Доведення. Маємо

$$\begin{aligned}&\|x_{n+1} - x_{\alpha_n}\|^2 = \\ &= \|x_n - x_{\alpha_n}\|^2 - \|x_n - x_{n+1}\|^2 + 2\big(x_{n+1} - x_n, x_{n+1} - x_{\alpha_n}\big) =\end{aligned}$$

$$= \|x_n - x_{\alpha_n}\|^2 - \|x_n - y_n\|^2 - \|y_n - x_{n+1}\|^2 -$$
$$- 2\big(x_n - y_n, y_n - x_{n+1}\big) + 2\big(x_{n+1} - x_n, x_{n+1} - x_{\alpha_n}\big). \quad (2.56)$$

З визначення точок x_{n+1} і y_n випливає

$$\lambda_n(Ay_n, x_{\alpha_n} - x_{n+1}) \geqslant \big(x_{n+1} - x_n + \alpha_n\lambda_n(x_n - z), x_{n+1} - x_{\alpha_n}\big), \quad (2.57)$$

$$\lambda_n(Ay_{n-1}, x_{n+1} - y_n) \geqslant -\big(x_n + \alpha_n\lambda_n(x_n - z) - y_n, y_n - x_{n+1}\big). \quad (2.58)$$

Використавши нерівності (2.57), (2.58) для оцінки скалярних добутків в (2.56), отримуємо

$$\|x_{n+1} - x_{\alpha_n}\|^2 \leqslant \|x_n - x_{\alpha_n}\|^2 - \|x_n - y_n\|^2 - \|y_n - x_{n+1}\|^2 +$$
$$+ 2\lambda_n \left((Ay_n, x_{\alpha_n} - y_n) + (Ay_{n-1} - Ay_n, x_{n+1} - y_n)\right) +$$
$$+ 2\alpha_n\lambda_n \left(x_n - z, x_{\alpha_n} - y_n\right). \quad (2.59)$$

З правила обчислення λ_{n+1} випливає нерівність

$$(Ay_{n-1} - Ay_n, x_{n+1} - y_n) \leqslant \frac{\tau}{2\lambda_{n+1}} \left(\|y_{n-1} - y_n\|^2 - \|x_{n+1} - y_n\|^2\right). \quad (2.60)$$

Для оцінки виразу $(Ay_{n-1} - Ay_n, x_{n+1} - y_n)$ в (2.59) скористаємося (2.60). Отримаємо

$$\|x_{n+1} - x_{\alpha_n}\|^2 \leqslant \|x_n - x_{\alpha_n}\|^2 - \|x_n - y_n\|^2 - \|y_n - x_{n+1}\|^2 +$$
$$+ \tau\frac{\lambda_n}{\lambda_{n+1}}\|y_{n-1} - y_n\|^2 + \tau\frac{\lambda_n}{\lambda_{n+1}}\|x_{n+1} - y_n\|^2 +$$
$$+ 2\lambda_n(Ay_n, x_{\alpha_n} - y_n) + 2\alpha_n\lambda_n\big(x_n - z, x_{\alpha_n} - y_n\big).$$

Член $\|y_{n-1} - y_n\|^2$ оцінимо наступним чином:

$$\|y_{n-1} - y_n\|^2 \leqslant 2\|y_{n-1} - x_n\|^2 + 2\|x_n - y_n\|^2.$$

Маємо

$$\|x_{n+1} - x_{\alpha_n}\|^2 \leqslant \|x_n - x_{\alpha_n}\|^2 - \|x_n - y_n\|^2 - \|y_n - x_{n+1}\|^2 +$$

$$+\,2\tau\frac{\lambda_n}{\lambda_{n+1}}\|y_{n-1}-x_n\|^2+2\tau\frac{\lambda_n}{\lambda_{n+1}}\|x_n-y_n\|^2+\tau\frac{\lambda_n}{\lambda_{n+1}}\|x_{n+1}-y_n\|^2+$$
$$+\,2\lambda_n(Ay_n,x_{\alpha_n}-y_n)+2\alpha_n\lambda_n\big(x_n-z,x_{\alpha_n}-y_n\big).\quad(2.61)$$

Із монотонності оператора A випливає

$$(Ay_n,x_{\alpha_n}-y_n)\leqslant(Ax_{\alpha_n},x_{\alpha_n}-y_n),$$

звідки

$$(Ay_n,x_{\alpha_n}-y_n)-\alpha_n(x_{\alpha_n}-z,y_n-x_{\alpha_n})\leqslant-(Ax_{\alpha_n},y_n-x_{\alpha_n})-\alpha_n(x_{\alpha_n}-z,y_n-x_{\alpha_n}).$$

Оскільки

$$(Ax_{\alpha_n},y_n-x_{\alpha_n})+\alpha_n(x_{\alpha_n}-z,y_n-x_{\alpha_n})\geqslant 0,$$

то

$$(Ay_n,x_{\alpha_n}-y_n)\leqslant\alpha_n(x_{\alpha_n}-z,y_n-x_{\alpha_n}).$$

Врахувавши останню оцінку в (2.61), приходимо до нерівності

$$\|x_{n+1}-x_{\alpha_n}\|^2\leqslant\|x_n-x_{\alpha_n}\|^2-\|x_n-y_n\|^2-\|y_n-x_{n+1}\|^2+$$
$$+\,2\tau\frac{\lambda_n}{\lambda_{n+1}}\|y_{n-1}-x_n\|^2+2\tau\frac{\lambda_n}{\lambda_{n+1}}\|x_n-y_n\|^2+$$
$$+\,\tau\frac{\lambda_n}{\lambda_{n+1}}\|x_{n+1}-y_n\|^2+2\alpha_n\lambda_n\left(x_n-x_{\alpha_n},x_{\alpha_n}-y_n\right).\quad(2.62)$$

Оцінимо зверху член $(x_n-x_{\alpha_n},x_{\alpha_n}-y_n)$. Маємо

$$(x_n-x_{\alpha_n},x_{\alpha_n}-y_n)=$$
$$=(x_n-x_{\alpha_n},x_{\alpha_n}-x_n)+(x_n-x_{\alpha_n},x_n-y_n)\leqslant$$
$$\leqslant-\|x_{\alpha_n}-x_n\|^2+\|x_n-x_{\alpha_n}\|\,\|x_n-y_n\|\leqslant$$
$$\leqslant-\|x_{\alpha_n}-x_n\|^2+\frac{1}{2}\|x_n-x_{\alpha_n}\|^2+\frac{1}{2}\|x_n-y_n\|^2=$$
$$=-\frac{1}{2}\|x_n-x_{\alpha_n}\|^2+\frac{1}{2}\|x_n-y_n\|^2.\quad(2.63)$$

З нерівностей (2.62) і (2.63) отримуємо

$$\|x_{n+1}-x_{\alpha_n}\|^2 \leqslant (1-\alpha_n\lambda_n)\|x_n-x_{\alpha_n}\|^2-$$
$$-\left(1-\tau\frac{\lambda_n}{\lambda_{n+1}}\right)\|x_{n+1}-y_n\|^2-\left(1-2\tau\frac{\lambda_n}{\lambda_{n+1}}-\alpha_n\lambda_n\right)\|y_n-x_n\|^2+$$
$$+2\tau\lambda_n\frac{\lambda_n}{\lambda_{n+1}}\|x_n-y_{n-1}\|^2,$$

що і треба було довести. □

Доведемо тепер оцінку, з якої буде випливати збіжність до нуля послідовностей $(\|x_n-x_{\alpha_n}\|)$ і $(\|x_n-y_{n-1}\|)$.

Лема 2.14. *Для породжених алгоритмом 2.5 послідовностей* (x_n), (y_n) *і елементів* x_{α_n} *при великих* $n\in\mathbb{N}$ *виконується нерівність*

$$\|x_{n+1}-x_{\alpha_{n+1}}\|^2+\frac{2\tau\lambda_{n+1}\lambda_{n+2}^{-1}}{1-\alpha_{n+1}\lambda_{n+1}}\|x_{n+1}-y_n\|^2 \leqslant$$
$$\leqslant\left(1-\frac{\alpha_n\lambda_n}{2}\right)\left(\|x_n-x_{\alpha_n}\|^2+\frac{2\tau\lambda_n\lambda_{n+1}^{-1}}{1-\alpha_n\lambda_n}\|x_n-y_{n-1}\|^2\right)+$$
$$+\frac{8\|x^*-z\|^2}{\lambda_n}\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^3}. \quad (2.64)$$

Доведення. Маємо

$$\|x_{n+1}-x_{\alpha_n}\|^2=$$
$$=\|x_{n+1}-x_{\alpha_{n+1}}\|^2+\|x_{\alpha_{n+1}}-x_{\alpha_n}\|^2+2(x_{n+1}-x_{\alpha_{n+1}},x_{\alpha_{n+1}}-x_{\alpha_n})\geqslant$$
$$\geqslant\|x_{n+1}-x_{\alpha_{n+1}}\|^2+\|x_{\alpha_{n+1}}-x_{\alpha_n}\|^2-2\|x_{n+1}-x_{\alpha_{n+1}}\|\,\|x_{\alpha_{n+1}}-x_{\alpha_n}\|\geqslant$$
$$\geqslant(1-\varepsilon)\|x_{n+1}-x_{\alpha_{n+1}}\|^2+\frac{\varepsilon-1}{\varepsilon}\|x_{\alpha_{n+1}}-x_{\alpha_n}\|^2, \quad (2.65)$$

де $\varepsilon>0$. Покладемо в (2.65) $\varepsilon=\frac{1}{2}\alpha_n\lambda_n$. Отримаємо

$$\|x_{n+1}-x_{\alpha_n}\|^2\geqslant\frac{2-\alpha_n\lambda_n}{2}\|x_{n+1}-x_{\alpha_{n+1}}\|^2-$$
$$-\frac{2-\alpha_n\lambda_n}{\alpha_n\lambda_n}\|x_{\alpha_{n+1}}-x_{\alpha_n}\|^2. \quad (2.66)$$

В силу правил узгодження значень параметрів α_n, λ_n при великих номерах $n\in\mathbb{N}$ маємо $1-\alpha_n\lambda_n>0$. З урахуванням другої нерівності леми 2.9 з (2.66) виводимо

$$\|x_{n+1} - x_{\alpha_n}\|^2 \geqslant \frac{2-\alpha_n\lambda_n}{2}\|x_{n+1} - x_{\alpha_{n+1}}\|^2 - \\ - \frac{(2-\alpha_n\lambda_n)}{\alpha_n\lambda_n}\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^2}4\|x^* - z\|^2. \quad (2.67)$$

для всіх $n \geqslant n_0$. Використавши (2.67) в (2.55), отримаємо (для $n \geqslant n_0$)

$$\frac{2-\alpha_n\lambda_n}{2}\|x_{n+1} - x_{\alpha_{n+1}}\|^2 \leqslant (1-\alpha_n\lambda_n)\|x_n - x_{\alpha_n}\|^2 - \\ - \left(1 - \tau\frac{\lambda_n}{\lambda_{n+1}}\right)\|x_{n+1} - y_n\|^2 - \left(1 - 2\tau\frac{\lambda_n}{\lambda_{n+1}} - \alpha_n\lambda_n\right)\|y_n - x_n\|^2 + \\ + 2\tau\frac{\lambda_n}{\lambda_{n+1}}\|x_n - y_{n-1}\|^2 + \frac{(2-\alpha_n\lambda_n)}{\alpha_n\lambda_n}\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^2}4\|x^* - z\|^2.$$

Звідки випливає нерівність

$$\|x_{n+1} - x_{\alpha_{n+1}}\|^2 \geqslant \frac{2-2\alpha_n\lambda_n}{2-\alpha_n\lambda_n}\|x_n - x_{\alpha_n}\|^2 - \\ - \frac{2(1-\tau\lambda_n\lambda_{n+1}^{-1})}{2-\alpha_n\lambda_n}\|x_{n+1} - y_n\|^2 - \\ - \frac{2(1-2\tau\lambda_n\lambda_{n+1}^{-1} - \alpha_n\lambda_n)}{2-\alpha_n\lambda_n}\|y_n - x_n\|^2 + \\ + \frac{4\tau\lambda_n\lambda_{n+1}^{-1}}{2-\alpha_n\lambda_n}\|x_n - y_{n-1}\|^2 + \frac{8\|x^* - z\|^2}{\lambda_n}\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^3}. \quad (2.68)$$

Перегрупувавши члени в (2.68), отримаємо

$$\|x_{n+1} - x_{\alpha_{n+1}}\|^2 + \frac{2\tau\lambda_{n+1}\lambda_{n+2}^{-1}}{1-\alpha_{n+1}\lambda_{n+1}}\|x_{n+1} - y_n\|^2 \leqslant \\ \leqslant \frac{2-2\alpha_n\lambda_n}{2-\alpha_n\lambda_n}\left(\|x_n - x_{\alpha_n}\|^2 + \frac{2\tau\lambda_n\lambda_{n+1}^{-1}}{1-\alpha_n\lambda_n}\|x_n - y_{n-1}\|^2\right) - \\ - \frac{2(1-2\tau\lambda_n\lambda_{n+1}^{-1} - \alpha_n\lambda_n)}{2-\alpha_n\lambda_n}\|y_n - x_n\|^2 - \\ - \left(\frac{1-\tau\lambda_n\lambda_{n+1}^{-1}}{1-\frac{1}{2}\alpha_n\lambda_n} - \frac{2\tau\lambda_{n+1}\lambda_{n+2}^{-1}}{1-\alpha_{n+1}\lambda_{n+1}}\right)\|x_{n+1} - y_n\|^2 + \\ + \frac{8\|x^* - z\|^2}{\lambda_n}\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^3}. \quad (2.69)$$

Оскільки $\tau \in \left(0, \frac{1}{3}\right)$, існує $\lim\limits_{n\to\infty} \lambda_n > 0$ і $\lim\limits_{n\to\infty} \alpha_n = 0$, то починаючи з деякого номера $n_1 \in \mathbb{N}$ будуть виконуватися нерівності

$$\frac{1-2\tau\lambda_n\lambda_{n+1}^{-1}-\alpha_n\lambda_n}{2-\alpha_n\lambda_n} > 0, \quad \frac{1-\tau\lambda_n\lambda_{n+1}^{-1}}{1-\frac{1}{2}\alpha_n\lambda_n} - \frac{2\tau\lambda_{n+1}\lambda_{n+2}^{-1}}{1-\alpha_{n+1}\lambda_{n+1}} > 0,$$

$$\frac{2-2\alpha_n\lambda_n}{2-\alpha_n\lambda_n} < 1-\frac{\alpha_n\lambda_n}{2}.$$

Таким чином, для $n \geqslant N = \max\{n_0, n_1\}$ з (2.69) випливає нерівність

$$\|x_{n+1}-x_{\alpha_{n+1}}\|^2 + \frac{2\tau\lambda_{n+1}\lambda_{n+2}^{-1}}{1-\alpha_{n+1}\lambda_{n+1}}\|x_{n+1}-y_n\|^2 \leqslant$$
$$\leqslant \left(1-\frac{\alpha_n\lambda_n}{2}\right)\left(\|x_n-x_{\alpha_n}\|^2 - \frac{2\tau\lambda_n\lambda_{n+1}^{-1}}{1-\alpha_n\lambda_n}\|x_n-y_{n-1}\|^2\right) +$$
$$+\frac{8\|x^*-z\|^2}{\lambda_n}\frac{(\alpha_{n+1}-\alpha_n)^2}{\alpha_n^3},$$

що і потрібно було довести. □

Сформулюємо основний результат.

Теорема 2.7. *Нехай H — гільбертовий простір, $C \subseteq H$ — непорожня опукла замкнена множина, $A : H \to H$ — монотонний та L-ліпшицевий на множині C оператор та $S \neq \varnothing$. Нехай виконуються умови (A2)–(A4). Тоді для породжених алгоритмом 2.5 послідовностей (x_n), (y_n) має місце*

$$\lim_{n\to\infty}\|x_n-x^*\| = \lim_{n\to\infty}\|y_n-x^*\| = 0, \tag{2.70}$$

де $x^* \in H$ — єдиний розв'язок задачі (2.36).

Доведення. В силу леми 2.4 і нерівності (2.64) маємо

$$\|x_n-x_{\alpha_n}\|^2 + 2\tau\frac{\lambda_n\lambda_{n+1}^{-1}}{1-\alpha_n\lambda_n\mu}\|x_n-y_{n-1}\|^2 \to 0 \quad \text{при} \quad n\to\infty.$$

Звідки

$$\lim_{n\to\infty}\|x_n-x_{\alpha_n}\| = \lim_{n\to\infty}\|y_{n-1}-x_n\| = 0. \tag{2.71}$$

З нерівності

$$\|x_n-x^*\| \leqslant \|x_n-x_{\alpha_n}\| + \|x_{\alpha_n}-x^*\|,$$

леми 2.10 і (2.71) отримуємо рівності (2.70). □

Зауваження 2.16. Аналогічний результат має місце для модифікації алгоритму 2.5 з заміною інструкції перерахунку λ_n на наступну

$$\lambda_{n+1} = \begin{cases} \min\left\{\lambda_n, \tau\frac{\|y_{n-1}-y_n\|}{\|Ay_{n-1}-Ay_n\|}\right\}, & \text{якщо} \quad Ay_{n-1} \neq Ay_n, \\ \lambda_n, & \text{інакше.} \end{cases}$$

де $\tau \in \left(0, \frac{1}{3}\right)$.

2.8. Висновки до розділу 2

У даному розділі розглядались варіаційні нерівності з операторами, що діють в гільбертовому просторі. Для цих задач запропоновано та обгрунтовано нові алгоритми екстраградієнтного типу, які є модифікаціями методу Л. Д. Попова пошуку сідлових точок опукло-угнутих функцій. Також отримано нові результати для відомих варіантів методу.

А саме, отримано сублінійну оцінку ефективності для функції зазору. Доведено сильну збіжність алгоритму екстраполяції з минулого для варіаційних нерівностей з рівномірно монотонними операторами. Доведено лінійну швидкість збіжності алгоритму екстраполяції з минулого для варіаційних нерівностей з операторами, що задовольняють умову типу узагальненої сильної монотонності. Запропоновано та обгрунтовано адаптивний варіант алгоритму. Запропоновано регуляризовані варіанти алгоритму та доведено теореми про їх сильну збіжність.

Основні результати даного розділу опубліковано в [3, 9, 10, 12, 15].

РОЗДІЛ 3
АЛГОРИТМ ОПЕРАТОРНОЇ ЕКСТРАПОЛЯЦІЇ ДЛЯ ВАРІАЦІЙНИХ НЕРІВНОСТЕЙ В ГІЛЬБЕРТОВОМУ ПРОСТОРІ

Розділ присвячено дослідженню нових ітераційних алгоритмів для розв'язання варіаційних нерівностей в гільбертових просторах. Дані алгоритми є модификаціями відомого «forward-reflected-backward algorithm», що запропонований в [115]. За об'ємом необхідних для здійснення ітераційного кроку обчислень алгоритми має перевагу над екстаградієнтним методом Корпелевич та методом екстраполяції з минулого. Схеми даного типу ще відомі під назвою «optimistic gradient descent ascent» [55, 117, 118].

Розділ побудовано таким чином. В підрозділі 3.1 розглянуто алгоритм операторної екстраполяції та прокоментовано літературу, що присвячено дослідженню його збіжності. В підрозділі 3.2 доведено слабку збіжність алгоритму операторної екстраполяції та отримано сублінійну оцінку ефективності для функції зазору. Сильну збіжність алгоритму операторної екстраполяції для варіаційних нерівностей з рівномірно монотонними операторами доведено в підрозділі 3.3. Лінійну швидкість збіжності алгоритму операторної екстраполяції для варіаційних нерівностей з операторами, що задовольняють умову типу узагальненої сильної монотонності, доведено в підрозділі 3.4. В підрозділі 3.5 запропоновано та досліджено адаптивний варіант алгоритму. В підрозділі 3.6 запропоновано регуляризований варіант алгоритму та доведено теорему про його сильну збіжність. Для регуляризації використано схему Гальперна.

3.1. Алгоритм операторної екстраполяції

Для розв'язання варіаційної нерівності (2.1) розглянемо ітераційний

Алгоритм 3.1. Обираємо $x_0 = x_1 \in C$, $\lambda_n, \mu_n > 0$. Покладаємо $n = 1$.

1. Обчислити

$$m_n = \mu_n \left(Ax_n - Ax_{n-1} \right),$$

$$x_{n+1} = P_C \left(x_n - \lambda_n Ax_n - m_n \right).$$

2. Якщо $x_{n+1} = x_n = x_{n-1}$, то СТОП, інакше покласти $n := n + 1$ та перейти до **1**.

Правило зупинки в алгоритмі обґрунтовується так: при виконанні

$$x_{n+1} = x_n = x_{n-1}$$

маємо

$$x_n = P_C \left(x_n - \lambda_n Ax_n \right),$$

звідки $x_n \in S$.

Зауваження 3.1. Даний алгоритм при $\mu_n = \lambda_{n-1}$ запропонований в [115] під назвою «forward-reflected-backward algorithm». За об'ємом необхідних для здійснення ітераційного кроку обчислень він має перевагу над екстаградієнтним методом Корпелевич та методом екстраполяції з минулого. Схеми даного типу ще відомі під назвою «optimistic gradient descent ascent» [55, 117, 118]. Дослідженню та розробці модифікацій цього методу присвячено роботи [1, 4–9, 116, 131]. Модифікації з брегманівською відстанню досліджено в [7–9], варіанти для задач в просторах Банаха запропоновані в [1, 5, 6]. Регуляризовані методи досліджено в [4]. У дисертації уточнюються та узагальнюються результати вказаних робіт.

Зауваження 3.2. Якщо $C = H$, то алгоритм 3.1 з $\mu_n = \lambda_n = \lambda$ можна записати так:

$$x_{n+1} = x_n - 2\lambda A x_n + \lambda A x_{n-1} \quad \text{або} \quad \begin{cases} x_n = z_n - \lambda A x_{n-1}, \\ z_{n+1} = z_n - \lambda A x_n. \end{cases}$$

Тобто, у цій ситуації алгоритм 3.1 співпадає з агоритмом екстраполяції з минулого.

3.2. Слабка збіжність та сублінійна оцінка ефективності

Припустимо додатково, що оператор A є псевдомонотонним.

Лема 3.1. *Для породженої алгоритмом 3.1 з $\mu_n = \lambda_{n-1}$ послідовності (x_n) виконується нерівність*

$$\begin{aligned} \|x_{n+1} - z\|^2 + 2\lambda_n (Ax_n - Ax_{n+1}, x_{n+1} - z) + \lambda_n L \|x_{n+1} - x_n\|^2 \leqslant \\ \leqslant \|x_n - z\|^2 + 2\lambda_{n-1} (Ax_{n-1} - Ax_n, x_n - z) + \lambda_{n-1} L \|x_n - x_{n-1}\|^2 - \\ - (1 - \lambda_{n-1} L - \lambda_n L) \|x_{n+1} - x_n\|^2, \end{aligned} \tag{3.1}$$

де $z \in S$.

Доведення. Нехай $z \in S$. Маємо

$$\begin{aligned} \|x_{n+1} - z\|^2 \leqslant \|x_n - z\|^2 - \|x_{n+1} - x_n\|^2 + \\ + 2 (\lambda_n A x_n + \lambda_{n-1} (Ax_n - Ax_{n-1}), z - x_{n+1}). \end{aligned} \tag{3.2}$$

З псевдомонотонності оператора A випливає

$$\begin{aligned} (\lambda_n A x_n + \lambda_{n-1} (Ax_n - Ax_{n-1}), z - x_{n+1}) = \lambda_n (Ax_n - Ax_{n+1}, z - x_{n+1}) + \\ + \lambda_{n-1} (Ax_n - Ax_{n-1}, z - x_{n+1}) + \underbrace{\lambda_n (Ax_{n+1}, z - x_{n+1})}_{\leqslant 0} \leqslant \\ \leqslant \lambda_n (Ax_n - Ax_{n+1}, z - x_{n+1}) + \lambda_{n-1} (Ax_n - Ax_{n-1}, z - x_n) + \\ + \lambda_{n-1} (Ax_n - Ax_{n-1}, x_n - x_{n+1}). \end{aligned} \tag{3.3}$$

Застосуємо (3.3) для оцінки правої частини (3.2) та отримаємо

$$\|x_{n+1}-z\|^2 \leqslant \|x_n-z\|^2 - \|x_{n+1}-x_n\|^2 + \\ + 2\lambda_n\left(Ax_n - Ax_{n+1}, z - x_{n+1}\right) + 2\lambda_{n-1}\left(Ax_n - Ax_{n-1}, z - x_n\right) + \\ + 2\lambda_{n-1}\left(Ax_n - Ax_{n-1}, x_n - x_{n+1}\right). \quad (3.4)$$

Оцінимо зверху вираз $2\lambda_{n-1}\left(Ax_n - Ax_{n-1}, x_n - x_{n+1}\right)$ в (3.4). Маємо

$$2\lambda_{n-1}\left(Ax_n - Ax_{n-1}, x_n - x_{n+1}\right) \leqslant 2\lambda_{n-1}\left\|Ax_n - Ax_{n-1}\right\|\left\|x_n - x_{n+1}\right\| \leqslant \\ \leqslant 2\lambda_{n-1}L\left\|x_n - x_{n-1}\right\|\left\|x_{n+1} - x_n\right\| \leqslant \\ \leqslant \lambda_{n-1}L\left\|x_n - x_{n-1}\right\|^2 + \lambda_{n-1}L\left\|x_n - x_{n+1}\right\|^2.$$

Отримаємо

$$\|x_{n+1}-z\|^2 + 2\lambda_n\left(Ax_n - Ax_{n+1}, x_{n+1} - z\right) + \lambda_n L\|x_{n+1}-x_n\|^2 \leqslant \\ \leqslant \|x_n - z\|^2 + 2\lambda_{n-1}\left(Ax_{n-1} - Ax_n, x_n - z\right) + \lambda_{n-1}L\|x_n - x_{n-1}\|^2 - \\ - \left(1 - \lambda_{n-1}L - \lambda_n L\right)\|x_{n+1} - x_n\|^2,$$

що і треба було довести. □

Зауваження 3.3. Для монотонних операторів нерівність леми 3.1 отримана в [1, 115].

Перейдемо до доведення слабкої збіжності алгоритму 3.1 з $\mu_n = \lambda_{n-1}$.

Припустимо, що

$$0 < \inf_n \lambda_n \leqslant \sup_n \lambda_n < \tfrac{1}{2L}.$$

В якості функції Ляпунова оберемо

$$V_n = \|z - x_n\|^2 + 2\lambda_{n-1}(Ax_{n-1} - Ax_n, x_n - z) + \lambda_{n-1}L\|x_n - x_{n-1}\|^2,$$

де $z \in S$.

Нехай $z' \in S$. Беремо таке $\delta > 0$, що $1 - \lambda_{n-1}L - \lambda_n L \geqslant \delta$ для всіх $n \in \mathbb{N}$. З нерівності (3.1) випливає

$$V_{n+1} \leqslant V_n - \delta\|x_{n+1} - x_n\|^2,$$

де $V_n = \|z' - x_n\|^2 + 2\lambda_{n-1}(Ax_{n-1} - Ax_n, x_n - z') + \lambda_{n-1}L\|x_n - x_{n-1}\|^2$.

Покажемо, що $L_n \geqslant 0$ для всіх $n \in \mathbb{N}$. Маємо

$$\begin{aligned} V_n = \|z' - x_n\|^2 + 2\lambda_{n-1}(Ax_{n-1} - Ax_n, x_n - z') + \lambda_{n-1}L\|x_n - x_{n-1}\|^2 \geqslant \\ \geqslant \|x_n - z'\|^2 - 2\lambda_{n-1}\|Ax_{n-1} - Ax_n\|\,\|x_n - z'\| + \lambda_{n-1}L\,\|x_{n-1} - x_n\|^2 \geqslant \\ \geqslant \|x_n - z'\|^2 - 2\lambda_{n-1}L\,\|x_n - x_{n-1}\|\,\|x_n - z'\| + \lambda_{n-1}L\,\|x_{n-1} - x_n\|^2 \geqslant \\ \geqslant (1 - \lambda_{n-1}L)\,\|x_n - z'\|^2 \geqslant 0. \end{aligned}$$

З леми 2.2 випливає існування границі

$$\lim_{n\to\infty}\left(\|z' - x_n\|^2 + 2\lambda_{n-1}(Ax_{n-1} - Ax_n, x_n - z') + \lambda_{n-1}L\|x_n - x_{n-1}\|^2\right)$$

та

$$\sum_{n=1}^{\infty}\|x_{n+1} - x_n\|^2 < +\infty.$$

Звідки отримуємо обмеженість послідовності (x_n) та $\lim_{n\to\infty}\|x_{n+1} - x_n\| = 0$. Оскільки

$$\lim_{n\to\infty}\left(2\lambda_{n-1}(Ax_{n-1} - Ax_n, x_n - z') + \lambda_{n-1}L\|x_n - x_{n-1}\|^2\right) = 0,$$

то послідовності $\left(\|z' - x_n\|^2\right)$ збігаються для всіх $z' \in S$.

Покажемо, що всі слабкі часткові границі послідовності (x_n) належать множині S. Розглянемо підпослідовність (x_{n_k}), що слабко збігається до деякої точки $z \in E$. Ясно, що $z \in C$. Покажемо, що $z \in S$. Маємо

$$(x_{n+1} - x_n + \lambda_n Ax_n + \lambda_{n-1}(Ax_n - Ax_{n-1}), y - x_{n+1}) \geqslant 0 \quad \forall y \in C.$$

Припустимо, що оператор A монотонний. Виводимо оцінку

$$\begin{aligned} (Ay, y - x_n) + (Ax_n, x_n - x_{n+1}) \geqslant (Ax_n, y - x_{n+1}) \geqslant \\ \geqslant \frac{1}{\lambda_n}(x_n - x_{n+1}, y - x_{n+1}) - \frac{\lambda_{n-1}}{\lambda_n}(Ax_n - Ax_{n-1}, y - x_{n+1}) \quad \forall y \in C. \end{aligned}$$

З $\lim_{n\to\infty}\|x_n - x_{n-1}\| = 0$ та ліпшицевості оператора A випливає

$$\lim_{n\to\infty}\|Ax_n - Ax_{n-1}\| = 0.$$

Таким чином,

$$\liminf_{n\to\infty} (Ay, y - x_n) \geqslant 0 \quad \forall y \in C.$$

З іншого боку

$$(Ay, y - z) = \lim_{k\to\infty} (Ay, y - x_{n_k}) \geqslant \liminf_{n\to\infty} (Ay, y - x_n) \geqslant 0 \quad \forall y \in C.$$

Отже, $z \in S$.

Покажемо, що послідовність (x_n) слабко збігається до z. Міркуємо від супротивного. Нехай існує підпослідовність (x_{m_k}) така, що $x_{m_k} \to z'$ слакбо та $z \neq z'$. Ясно, що $z' \in S$. Маємо

$$2(x_n, z - z') = \|z' - x_n\|^2 - \|z - x_n\|^2 + \|z\|^2 - \|z'\|^2 .$$

Звідки випливає існування границі $\lim\limits_{n\to\infty} (x_n, z - z')$. Отримаємо

$$(z, z - z') = \lim_{k\to\infty} (x_{n_k}, z - z') = \lim_{k\to\infty} (x_{m_k}, z - z') = (z', z - z'),$$

тобто, $\|z - z'\| = 0$. Звідки випливає $z = z'$, що і необхідно було довести.

Таким чином, має місце

Теорема 3.1. *Нехай H — гільбертовий простір, $C \subseteq H$ — непорожня опукла замкнена множина, $A : H \to H$ — монотонний та L-ліпшицевий на множині C оператор та $S \neq \varnothing$. Припустимо, що*

$$0 < \inf_n \lambda_n \leqslant \sup_n \lambda_n < \tfrac{1}{2L}.$$

Тоді породжена алгоритмом 3.1 з $\mu_n = \lambda_{n-1}$ послідовність (x_n) слабко збігається до розв'язку варіаційної нерівності (2.1).

У випадку обмеженості допустимої множини C доведемо, що алгоритму 3.1 з $\mu_n = \lambda_n = \frac{1}{2L}$ необхідно зробити $O\left(\frac{LD^2}{\varepsilon}\right)$ ітерацій для отримання точки $x \in C$ з

$$\operatorname{gap}(x) \leqslant \varepsilon,$$

де $\varepsilon > 0$, $D = \sup_{a,b\in C} \|a - b\| < +\infty$.

Нехай $\mu_n = \lambda_{n-1}$. Для послідовності (x_n) має місце нерівність

$$- 2 (\lambda_n A x_n + m_n, y - x_{n+1}) \leqslant \\ \leqslant \|y - x_n\|^2 - \|x_{n+1} - x_n\|^2 - \|y - x_{n+1}\|^2 \quad \forall y \in C. \quad (3.5)$$

Перепишемо (3.5) таким чином

$$\|y - x_n\|^2 - \|y - x_{n+1}\|^2 \geqslant \\ \geqslant 2\lambda_n (A x_{n+1}, x_{n+1} - y) - 2\lambda_n (A x_{n+1} - A x_n, x_{n+1} - y) + \\ + 2\lambda_{n-1} (A x_n - A x_{n-1}, x_n - y) + 2\lambda_{n-1} (A x_n - A x_{n-1}, x_{n+1} - x_n) + \\ + \|x_{n+1} - x_n\|^2. \quad (3.6)$$

Сумуючи (3.6) по n від 1 до N, отримуємо

$$\|y - x_1\|^2 - \|y - x_{N+1}\|^2 \geqslant \\ \geqslant 2 \sum_{n=1}^{N} \lambda_n (A x_{n+1}, x_{n+1} - y) - 2\lambda_N (A x_{N+1} - A x_N, x_{N+1} - y) + \\ + \sum_{n=1}^{N} \left(2\lambda_{n-1} (A x_n - A x_{n-1}, x_{n+1} - x_n) + \|x_{n+1} - x_n\|^2\right). \quad (3.7)$$

Ліпшицевість оператора A та умова $\lambda_n \in \left(0, \frac{1}{2L}\right]$ дають

$$\sum_{n=1}^{N} \left(2\lambda_{n-1} (A x_n - A x_{n-1}, x_{n+1} - x_n) + \|x_{n+1} - x_n\|^2\right) \geqslant \\ \geqslant \sum_{n=1}^{N} \left(-2\lambda_{n-1} L \|x_n - x_{n-1}\| \|x_{n+1} - x_n\| + \|x_{n+1} - x_n\|^2\right) = \\ = \sum_{n=1}^{N} \left(-2\lambda_{n-1} L \|x_n - x_{n-1}\| \|x_{n+1} - x_n\| + \frac{1}{2} \|x_{n+1} - x_n\|^2 + \right. \\ \left. + \frac{1}{2} \|x_n - x_{n-1}\|^2\right) + \\ + \frac{1}{2} \|x_{N+1} - x_N\|^2 \geqslant \frac{1}{2} \|x_{N+1} - x_N\|^2.$$

Використовуючи останню оцінку в (3.7), отримуємо

$$\|y-x_1\|^2 - \|y-x_{N+1}\|^2 \geqslant 2\sum_{n=1}^{N} \lambda_n \left(Ax_{n+1}, x_{n+1}-y\right) -$$
$$- 2\lambda_N \left(Ax_{N+1} - Ax_N, x_{N+1}-y\right) + \frac{1}{2}\left\|x_{N+1}-x_N\right\|^2 \geqslant$$
$$\geqslant 2\sum_{n=1}^{N} \lambda_n \left(Ax_{n+1}, x_{n+1}-y\right) - 2\lambda_N L \left\|x_{N+1}-x_N\right\| \left\|x_{N+1}-y\right\| +$$
$$+ \frac{1}{2}\left\|x_{N+1}-x_N\right\|^2 \geqslant 2\sum_{n=1}^{N} \lambda_n \left(Ax_{n+1}, x_{n+1}-y\right) - \lambda_N L \left\|x_{N+1}-y\right\|^2.$$

Приходимо до нерівності

$$2\sum_{n=1}^{N} \lambda_n \left(Ax_{n+1}, x_{n+1}-y\right) - \lambda_N L \left\|x_{N+1}-y\right\|^2 + \\ + \left\|y-x_{N+1}\right\|^2 \leqslant \left\|y-x_1\right\|^2 \quad \forall y \in C. \quad (3.8)$$

Використовуючи монотонність оператора A, отримуємо

$$\sum_{n=1}^{N} \lambda_n \left(Ax_{n+1}, x_{n+1}-y\right) \geqslant \sum_{n=1}^{N} \lambda_n \left(Ay, x_{n+1}-y\right) = \\ = \left(\sum_{n=1}^{N} \lambda_n\right) \left(Ay, z_{N+1}-y\right), \quad (3.9)$$

де $z_{N+1} = \frac{\sum_{n=1}^{N} \lambda_n x_{n+1}}{\sum_{n=1}^{N} \lambda_n}$. Враховуючи оцінку (3.9) в (3.8), приходимо до нерівності

$$2\left(\sum_{n=1}^{N} \lambda_n\right) \left(Ay, z_{N+1}-y\right) + \left(1-\lambda_N L\right) \left\|x_{N+1}-y\right\|^2 \leqslant \left\|y-x_1\right\|^2 \quad \forall y \in C,$$

звідки випливає

$$\operatorname{gap}\left(z_{N+1}\right) = \sup_{y\in C} \left(Ay, z_{N+1}-y\right) \leqslant \frac{\sup_{y\in C}\left\|y-x_1\right\|^2}{2\sum_{n=1}^{N} \lambda_n}.$$

Таким чином, має місце

Теорема 3.2. *Нехай H — гільбертовий простір, $C \subseteq H$ — непорожня опукла замкнена обмежена множина, $A : H \to H$ — монотонний та L-ліпшицевий на множині C оператор. Нехай (x_n) — послідовність, що породжена алгоритмом 3.1 з $\lambda_n = \mu_n = \frac{1}{2L}$, тобто,*

$$x_{n+1} = P_C\left(x_n - \tfrac{1}{2L}Ax_n - \tfrac{1}{2L}\left(Ax_n - Ax_{n-1}\right)\right) = P_C\left(x_n - \tfrac{1}{2L}\left(2Ax_n - Ax_{n-1}\right)\right).$$

Тоді має місце оцінка

$$\operatorname{gap}\left(z_{N+1}\right) \leqslant \frac{L \sup_{y \in C} \|x_1 - y\|^2}{N},$$

де $z_{N+1} = \frac{1}{N}\sum_{n=1}^{N} x_{n+1}$.

3.3. Сильна збіжність

Припустимо, що оператор $A : H \to H$ є рівномірно монотонним на обмежених підмножинах множини $C \subseteq H$. Тоді варіаційна нерівність (2.1) має єдиний розв'язок $\bar{z} \in C$.

Покажемо, що породжена алгоритмом 3.1 з $\mu_n = \lambda_{n-1}$ послідовність (x_n) сильно збігається до $\bar{z}$.

Осклільки множина $\{x_n\} \cup \{\bar{z}\} \subseteq C$ обмежена, то якщо існує така зростаюча функція $\phi : [0, +\infty) \to [0, +\infty)$, $\phi(0) = 0$ та $\phi(t) > 0$ для $t > 0$, що

$$\left(Ax_{n+1}, \bar{z} - x_{n+1}\right) \leqslant -\phi\left(\|\bar{z} - x_{n+1}\|\right).$$

Замість нерівності леми 3.1 запишемо її уточнену версію

$$\begin{aligned} 2\lambda_n \phi\left(\|\bar{z} - x_{n+1}\|\right) + {} \\ + \|x_{n+1} - \bar{z}\|^2 + 2\lambda_n\left(Ax_n - Ax_{n+1}, x_{n+1} - \bar{z}\right) + \lambda_n L \|x_{n+1} - x_n\|^2 \leqslant \\ \leqslant \|x_n - \bar{z}\|^2 + 2\lambda_{n-1}\left(Ax_{n-1} - Ax_n, x_n - \bar{z}\right) + \lambda_{n-1} L \|x_n - x_{n-1}\|^2 - \\ - \left(1 - \lambda_{n-1}L - \lambda_n L\right)\|x_{n+1} - x_n\|^2. \end{aligned} \quad (3.10)$$

Перепишемо (3.10) у вигляді

$$\begin{aligned} &2\lambda_n \phi\left(\|\bar{z}-x_{n+1}\|\right) \leqslant \\ &\quad \leqslant \left(\|x_n-\bar{z}\|^2+2\lambda_{n-1}\left(Ax_{n-1}-Ax_n, x_n-\bar{z}\right)+\lambda_{n-1}L\|x_n-x_{n-1}\|^2\right)- \\ &\quad -\left(\|x_{n+1}-\bar{z}\|^2+2\lambda_n\left(Ax_n-Ax_{n+1}, x_{n+1}-\bar{z}\right)+\lambda_n L\|x_{n+1}-x_n\|^2\right)- \\ &\qquad\qquad -\left(1-\lambda_{n-1}L-\lambda_n L\right)\|x_{n+1}-x_n\|^2. \quad (3.11) \end{aligned}$$

Нерівність (3.11) та припущення $0<\inf_n \lambda_n \leqslant \sup_n \lambda_n < \frac{1}{2L}$ дають

$$\sum_{n=1}^{\infty} \lambda_n \phi\left(\|\bar{z}-x_{n+1}\|\right) < +\infty \quad \text{та} \quad \lim_{n\to\infty} \phi\left(\|\bar{z}-x_{n+1}\|\right)=0.$$

Отже, $\|\bar{z}-x_{n+1}\| \to 0$ та $\|\bar{z}-x_{n+1}\| \to 0$.

Таким чином, має місце

Теорема 3.3. *Нехай H — гільбертовий простір, $C \subseteq H$ — непорожня опукла замкнена множина, $A : H \to H$ — рівномірно монотонний на обмежених підмножинах множини $C \subseteq H$ та L-ліпшицевий на множині C оператор, $\bar{z} \in C$ — єдиний розв'язок варіаційної нерівності (2.1). Припустимо, що*

$$0<\inf_n \lambda_n \leqslant \sup_n \lambda_n < \tfrac{1}{2L}.$$

Тоді породжена алгоритмом 3.1 з $\mu_n=\lambda_{n-1}$ послідовність (x_n) сильно збігається до точки $\bar{z}$.

3.4. Лінійна швидкість збіжності

Припустимо, що існує єдиний розв'язок $z \in C$ варіаційної нерівності (2.1), а оператор A задовольняє умову

$$(Ax, x-z) \geqslant \mu\|x-z\|^2 \quad \forall x \in C \tag{3.12}$$

для деякого $\mu>0$.

Розглянемо варіант алгоритму 3.1 з лінійною швидкістю збіжності для варіаціної нерівності (2.1), де ліпшицевий оператор A задовольняє (3.12).

Алгоритм 3.2. Для $x_1 = x_0 \in C$ генеруємо послідовність елементів $x_n \in C$ за допомогою ітераційної схеми

$$x_{n+1} = P_C\left(x_n - \tfrac{1}{2L}Ax_n - \tfrac{1}{2(L+\mu)}\left(Ax_n - Ax_{n-1}\right)\right).$$

Покажемо, що

$$\|x_{n+1} - z\|^2 = O\left(\left(1 - \frac{\mu}{L+\mu}\right)^n\right).$$

Розглянемо схему

$$x_{n+1} = P_C\left(x_n - \alpha Ax_n - \beta\left(Ax_n - Ax_{n-1}\right)\right),$$

де $x_1 = x_0 \in C$, $\alpha > 0$, $\beta > 0$.

Нехай z — розв'язок варіаційної нерівності (2.1). Маємо

$$\begin{aligned} \|x_{n+1} - z\|^2 \leqslant \|x_n - z\|^2 - \|x_{n+1} - x_n\|^2 + \\ + 2\left(\alpha Ax_n + \beta\left(Ax_n - Ax_{n-1}\right), z - x_{n+1}\right). \end{aligned} \tag{3.13}$$

З умови (3.12) випливає

$$\begin{aligned} \left(\alpha Ax_n + \beta\left(Ax_n - Ax_{n-1}\right), z - x_{n+1}\right) = \alpha\left(Ax_n - Ax_{n+1}, z - x_{n+1}\right) + \\ + \beta\left(Ax_n - Ax_{n-1}, z - x_{n+1}\right) + \alpha\left(Ax_{n+1}, z - x_{n+1}\right) \leqslant \\ \leqslant \alpha\left(Ax_n - Ax_{n+1}, z - x_{n+1}\right) + \beta\left(Ax_n - Ax_{n-1}, z - x_n\right) + \\ + \beta\left(Ax_n - Ax_{n-1}, x_n - x_{n+1}\right) - \alpha\mu\|x_{n+1} - z\|^2. \end{aligned} \tag{3.14}$$

Застосуємо (3.14) для оцінки правої частини (3.13) та отримаємо

$$\begin{aligned} (1 + 2\alpha\mu)\|x_{n+1} - z\|^2 \leqslant \|x_n - z\|^2 - \|x_{n+1} - x_n\|^2 + \\ + 2\alpha\left(Ax_n - Ax_{n+1}, z - x_{n+1}\right) + 2\beta\left(Ax_n - Ax_{n-1}, z - x_n\right) + \\ + 2\beta\left(Ax_n - Ax_{n-1}, x_n - x_{n+1}\right). \end{aligned} \tag{3.15}$$

Оцінимо зверху вираз $2\beta\left(Ax_n - Ax_{n-1}, x_n - x_{n+1}\right)$ в (3.15). Маємо

$$2\beta\left(Ax_n - Ax_{n-1}, x_n - x_{n+1}\right) \leqslant 2\beta\left\|Ax_n - Ax_{n-1}\right\|\left\|x_n - x_{n+1}\right\| \leqslant$$
$$\leqslant 2\beta L\left\|x_n - x_{n-1}\right\|\left\|x_{n+1} - x_n\right\| \leqslant$$
$$\leqslant \beta L\left\|x_n - x_{n-1}\right\|^2 + \beta L\left\|x_n - x_{n+1}\right\|^2.$$

Отримаємо

$$(1+2\alpha\mu)\|x_{n+1}-z\|^2 + 2\alpha\left(Ax_n - Ax_{n+1}, x_{n+1}-z\right) + \alpha L\|x_{n+1}-x_n\|^2 \leqslant$$
$$\leqslant \|x_n - z\|^2 + 2\beta\left(Ax_{n-1} - Ax_n, x_n - z\right) + \beta L\|x_n - x_{n-1}\|^2 -$$
$$-\left(1 - \alpha L - \beta L\right)\|x_{n+1} - x_n\|^2.$$

Покладемо $\alpha = \frac{1}{2L}$:

$$\left(1+\frac{\mu}{L}\right)\|x_{n+1}-z\|^2 + \frac{1}{L}\left(Ax_n - Ax_{n+1}, x_{n+1}-z\right) + \frac{1}{2}\|x_{n+1}-x_n\|^2 =$$
$$= \left(1+\frac{\mu}{L}\right)\left(\|x_{n+1}-z\|^2 + \frac{1}{L\left(1+\frac{\mu}{L}\right)}\left(Ax_n - Ax_{n+1}, x_{n+1}-z\right) + \right.$$
$$\left. + \frac{1}{2\left(1+\frac{\mu}{L}\right)}\|x_{n+1}-x_n\|^2\right) \leqslant$$
$$\leqslant \|x_n - z\|^2 + 2\beta\left(Ax_{n-1} - Ax_n, x_n - z\right) + \beta L\|x_n - x_{n-1}\|^2 -$$
$$-\left(\frac{1}{2} - \beta L\right)\|x_{n+1} - x_n\|^2.$$

Покладемо $\beta = \frac{1}{2L\left(1+\frac{\mu}{L}\right)}$. Тоді для

$$W_n = \|x_n - z\|^2 + \frac{1}{L\left(1+\frac{\mu}{L}\right)}\left(Ax_{n-1} - Ax_n, x_n - z\right) + \frac{1}{2\left(1+\frac{\mu}{L}\right)}\|x_n - x_{n-1}\|^2$$

маємо

$$\left(1+\frac{\mu}{L}\right)W_{n+1} \leqslant W_n - \left(\frac{1}{2} - \frac{1}{2\left(1+\frac{\mu}{L}\right)}\right)\|x_{n+1}-x_n\|^2 \leqslant W_n.$$

Звідси

$$W_{n+1} \leqslant \left(1+\frac{\mu}{L}\right)^{-1} W_n \leqslant ... \leqslant \left(1+\frac{\mu}{L}\right)^{-n}\|x_1 - z\|^2. \qquad (3.16)$$

Окільки

$$
\begin{aligned}
W_n = \|x_n - z\|^2 + \\
+ \frac{1}{L\left(1+\frac{\mu}{L}\right)}\left(Ax_{n-1} - Ax_n, x_n - z\right) + \frac{1}{2\left(1+\frac{\mu}{L}\right)}\|x_n - x_{n-1}\|^2 \geqslant \\
\geqslant \|x_n - z\|^2 - \\
- \frac{1}{L\left(1+\frac{\mu}{L}\right)}\|Ax_{n-1} - Ax_n\|\,\|x_n - z\| + \frac{1}{2\left(1+\frac{\mu}{L}\right)}\|x_{n-1} - x_n\|^2 \geqslant \\
\geqslant \|x_n - z\|^2 - \\
- \frac{1}{\left(1+\frac{\mu}{L}\right)}\|x_n - x_{n-1}\|\,\|x_n - z\| + \frac{1}{2\left(1+\frac{\mu}{L}\right)}\|x_{n-1} - x_n\|^2 \geqslant \\
\geqslant \left(1 - \frac{1}{2\left(1+\frac{\mu}{L}\right)}\right)\|x_n - z\|^2 \geqslant \frac{1}{2}\|x_n - z\|^2,
\end{aligned}
$$

то з (3.16) випливає

$$\|x_{n+1} - z\|^2 \leqslant \left(1 - \frac{\mu}{L+\mu}\right)^n 2\|x_1 - z\|^2.$$

Таким чином, має місце

Теорема 3.4. *Нехай H — гільбертовий простір, $C \subseteq H$ — непорожня опукла замкнена множина, $A : H \to H$ — L-ліпшицевий на множині C оператор, існує єдиний розв'язок $z \in C$ варіаційної нерівності (2.1) та виконується умова (3.12). Тоді для породженої алгоритмом 3.2 послідовності (x_n) виконується оцінка*

$$\|x_{n+1} - z\|^2 \leqslant \left(1 - \frac{\mu}{L+\mu}\right)^n 2\|x_1 - z\|^2, \quad n \geqslant 1.$$

Зауваження 3.4. З теореми 3.4 випливає оцінка

$$\|x_{n+1} - z\|^2 \leqslant e^{-\frac{\mu}{2L}n} 2\|x_1 - z\|^2.$$

Зауваження 3.5. Для задачі пошуку нуля μ-сильно монотонного та L-ліпшицевого оператора A (випадок нерівності (2.1) без обмежень, $C = H$) в роботі [117] для алгоритму

$$x_{n+1} = x_n - \tfrac{1}{4L}Ax_n - \tfrac{1}{4L}\left(Ax_n - Ax_{n-1}\right)$$

отримана оцінка

$$\|x_{n+1}-z\|^2 \leqslant \left(1-\frac{\mu}{4L}\right)^n 1024\frac{L^2}{\mu^2}\|x_1-z\|^2, \quad n \geqslant 1.$$

3.5. Адаптивний алгоритм операторної екстраполяції

Розглянемо алгоритм операторної екстраполяції з адаптивним вибором параметрів λ_n.

Алгоритм 3.3 (Адаптивний алгоритм). Обираємо $x_0 = x_1 \in C$, $\tau \in \left(0, \frac{1}{2}\right)$ та число $\lambda_1 = \lambda_0 > 0$. Покладаємо $n = 1$.

1. Обчислити

$$m_n = \lambda_{n-1}\left(Ax_n - Ax_{n-1}\right),$$

$$x_{n+1} = P_C\left(x_n - \lambda_n Ax_n - m_n\right).$$

2. Якщо $x_{n+1} = x_n = x_{n-1}$, то СТОП, інакше перейти до **3**.

3. Обчислити

$$\lambda_{n+1} = \begin{cases} \min\left\{\lambda_n, \tau\frac{\|x_{n+1}-x_n\|}{\|Ax_{n+1}-Ax_n\|}\right\}, & \text{якщо} \quad Ax_{n+1} \neq Ax_n, \\ \lambda_n, & \text{інакше}. \end{cases}$$

Покласти $n := n+1$ та перейти до 1.

Зауваження 3.6. Послідовність (λ_n), що задається на третьому етапі ітераційного кроку в алгоритмі 2, незростаюча та обмежена знизу числом $\min\left\{\lambda_1, \frac{\tau}{L}\right\}$. Отже, існує границя $\lim\limits_{n\to\infty}\lambda_n > 0$.

В якості функції Ляпунова оберемо

$$W_n = \|z-x_n\|^2 + 2\lambda_{n-1}\left(Ax_{n-1}-Ax_n, x_n-z\right) + \tau\frac{\lambda_{n-1}}{\lambda_n}\|x_n-x_{n-1}\|^2,$$

де $z \in S$.

Має місце

Лема 3.2. *Для послідовності* (x_n)*, що породжена алгоритмом 3.3, виконується нерівність*

$$W_{n+1} \leqslant W_n - \left(1-\tau\frac{\lambda_{n-1}}{\lambda_n}-\tau\frac{\lambda_n}{\lambda_{n+1}}\right)\|x_{n+1}-x_n\|^2.$$

Доведення. Нехай $z \in S$. Як і в доведенні леми 3.1 приходимо до нерівності

$$\|z - x_{n+1}\|^2 \leqslant \|z - x_n\|^2 - \|x_{n+1} - x_n\|^2 + \\ + 2\lambda_n \left(Ax_n - Ax_{n+1}, z - x_{n+1}\right) + 2\lambda_{n-1}\left(Ax_n - Ax_{n-1}, z - x_n\right) + \\ + 2\lambda_{n-1}\left(Ax_n - Ax_{n-1}, x_n - x_{n+1}\right). \quad (3.17)$$

Використовуючи правило обчислення λ_{n+1}, оцінимо зверху доданок

$$2\lambda_{n-1}\left(Ax_n - Ax_{n-1}, x_n - x_{n+1}\right)$$

в нерівності (3.17). Маємо

$$2\lambda_{n-1}\left(Ax_n - Ax_{n-1}, x_n - x_{n+1}\right) \leqslant 2\lambda_{n-1}\|Ax_n - Ax_{n-1}\|\,\|x_n - x_{n+1}\| \leqslant \\ \leqslant 2\tau\frac{\lambda_{n-1}}{\lambda_n}\|x_n - x_{n-1}\|\,\|x_{n+1} - x_n\| \leqslant \\ \leqslant \tau\frac{\lambda_{n-1}}{\lambda_n}\|x_n - x_{-1}\|^2 + \tau\frac{\lambda_{n-1}}{\lambda_n}\|x_n - x_{n+1}\|^2.$$

Приходимо до нерівності

$$\|z - x_{n+1}\|^2 + 2\lambda_n\left(Ax_n - Ax_{n+1}, x_{n+1} - z\right) + \tau\frac{\lambda_n}{\lambda_{n+1}}\|x_{n+1} - x_n\|^2 \leqslant \\ \leqslant \|z - x_n\|^2 + 2\lambda_{n-1}\left(Ax_{n-1} - Ax_n, x_n - z\right) + \tau\frac{\lambda_{n-1}}{\lambda_n}\|x_n - x_{n-1}\|^2 - \\ - \left(1 - \tau\frac{\lambda_{n-1}}{\lambda_n} - \tau\frac{\lambda_n}{\lambda_{n+1}}\right)\|x_{n+1} - x_n\|^2,$$

що і потрібно було довести. □

Сформулюємо основний результат про збіжність алгоритму 3.3.

Теорема 3.5. *Нехай C — непорожня опукла та замкнена підмножина гільбертового простору H, $A : H \to H$ — монотонний та ліпшицевий оператор, $S \neq \varnothing$. Тоді послідовність (x_n), що породжена алгоритмом 3.3, слабко збігається до деякої точки $z \in S$.*

Доведення. Нехай $z' \in S$. Для функції Ляпунова

$$W_n = \|z' - x_n\|^2 + 2\lambda_{n-1}\left(Ax_{n-1} - Ax_n, x_n - z'\right) + \tau\frac{\lambda_{n-1}}{\lambda_n}\|x_n - x_{n-1}\|^2,$$

має місце нерівність леми 3.2

$$W_{n+1} \leqslant W_n - \left(1 - \tau\frac{\lambda_{n-1}}{\lambda_n} - \tau\frac{\lambda_n}{\lambda_{n+1}}\right)\|x_{n+1} - x_n\|^2.$$

Оскільки існує границя $\lim\limits_{n\to\infty}\lambda_n > 0$, то

$$1 - \tau\mu\frac{\lambda_{n-1}}{\lambda_n} - \tau\frac{\lambda_n}{\lambda_{n+1}} \to 1 - 2\tau\mu \in (0,1) \quad \text{при} \quad n \to \infty.$$

Покажемо, що $W_n \geqslant 0$ для всіх достатньо великих $n \in \mathbb{N}$. Маємо

$$\begin{aligned} W_n &= \|x_n - z'\|^2 + 2\lambda_{n-1}\left(Ax_{n-1} - Ax_n, x_n - z'\right) + \tau\frac{\lambda_{n-1}}{\lambda_n}\|x_{n-1} - x_n\|^2 \geqslant \\ &\geqslant \|x_n - z'\|^2 - 2\lambda_{n-1}\|Ax_{n-1} - Ax_n\|\,\|x_n - z'\| + \tau\frac{\lambda_{n-1}}{\lambda_n}\|x_{n-1} - x_n\|^2 \geqslant \\ &\geqslant \|x_n - z'\|^2 - 2\tau\frac{\lambda_{n-1}}{\lambda_n}\|x_n - x_{n-1}\|\,\|x_n - z'\| + \tau\frac{\lambda_{n-1}}{\lambda_n}\|x_{n-1} - x_n\|^2 \geqslant \\ &\geqslant \left(1 - \tau\frac{\lambda_{n-1}}{\lambda_n}\right)\|x_n - z'\|^2. \end{aligned}$$

Оскільки існує таке $n_0 \in \mathbb{N}$, що $1 - \tau\frac{\lambda_{n-1}}{\lambda_n} > 0$ для всіх $n \geqslant n_0$, то $W_n \geqslant 0$ починаючи з n_0.

Тепер з леми 2.2 можемо зробити висновок, що існує границя

$$\lim_{n\to\infty}\left(\|z' - x_n\|^2 + 2\lambda_{n-1}\left(Ax_{n-1} - Ax_n, x_n - z'\right) + \tau\frac{\lambda_{n-1}}{\lambda_n}\|x_n - x_{n-1}\|^2\right)$$

та

$$\sum_{n=1}^{\infty}\left(1 - \tau\frac{\lambda_{n-1}}{\lambda_n} - \tau\frac{\lambda_n}{\lambda_{n+1}}\right)\|x_{n+1} - x_n\|^2 < +\infty.$$

Звідки отримуємо обмеженість послідовності (x_n) та $\lim\limits_{n\to\infty}\|x_{n+1} - x_n\| = 0$. Оскільки

$$\lim_{n\to\infty}\left(2\lambda_{n-1}\left(Ax_{n-1} - Ax_n, x_n - z'\right) + \tau\frac{\lambda_{n-1}}{\lambda_n}\|x_n - x_{n-1}\|^2\right) = 0,$$

то збігаються послідовності $(\|z' - x_n\|)$ для всіх $z' \in S$.

Далі, міркуваннями з доведення теореми 3.1, прийдемо до потрібного результату. $\square$

Зауваження 3.7. Для операторного рівняння

$$\text{знайти } x \in H: \ Ax = 0$$

алгоритм 3.3 дає такий ітераційний процес

$$\begin{cases} x_{n+1} = x_n - \lambda_n A x_n - \lambda_{n-1} \left(A x_n - A x_{n-1} \right), \\ \lambda_{n+1} = \begin{cases} \min \left\{ \lambda_n, \tau \frac{\|x_{n+1} - x_n\|}{\|A x_{n+1} - A x_n\|} \right\}, & \text{якщо } \ A x_{n+1} \neq A x_n, \\ \lambda_n, & \text{інакше.} \end{cases} \end{cases}$$

3.6. Регуляризований алгоритм операторної екстраполяції

Спираючись на відомий метод Гальперна апроксимації нерухомих точок нерозтягуючих операторів [24, 67] побудуємо такий регуляризований варіант алгоритму (3.1).

Алгоритм 3.4 (Регуляризований алгоритм). Задаємо елементи $y \in H$, $x_0 = x_1 \in C$, послідовність додатніх чисел (λ_n) та таку послідовність (α_n), що

$$\alpha_n \in (0, 1), \quad \lim_{n \to \infty} \alpha_n = 0, \quad \sum_{n=1}^{\infty} \alpha_n = +\infty.$$

Ітерації. Генеруємо послідовність (x_n) за допомогою ітераційної схеми

$$x_{n+1} = P_C \left(\alpha_n y + (1 - \alpha_n) x_n - \lambda_n A x_n - \right.$$
$$\left. - (1 - \alpha_n) \lambda_{n-1} \left(A x_n - A x_{n-1} \right) \right).$$

Відносно додатніх параметрів λ_n припустимо виконання такої умови:

$$0 < \inf_n \lambda_n \leqslant \sup_n \lambda_n < \tfrac{1}{2L}. \tag{3.18}$$

Доведемо, що послідовність (x_n), що породжена алгоритмом 3.4, сильно збігається до проєкції точки y на множину S.

Зауваження 3.8. Для пошуку нормального розв'язку операторного рівняння $Ax = 0$ можна використати схему

$$x_{n+1} = (1 - \alpha_n)\, x_n - \lambda_n A x_n - (1 - \alpha_n)\, \lambda_{n-1} \left(A x_n - A x_{n-1} \right).$$

Зауваження 3.9. Для гладкої сідлової задачі

$$\min_{x \in C} \max_{y \in D} L\,(x, y)$$

алгоритм 3.4 має вигляд

$$\begin{cases} x_{n+1} = P_C \left(\alpha_n x + (1 - \alpha_n)\, x_n - \lambda_n \nabla_1 L\,(x_n, y_n) - \right. \\ \qquad \left. - (1 - \alpha_n)\, \lambda_{n-1} \left(\nabla_1 L\,(x_n, y_n) - \nabla_1 L\,(x_{n-1}, y_{n-1}) \right) \right), \\ y_{n+1} = P_D \left(\alpha_n y + (1 - \alpha_n)\, y_n + \lambda_n \nabla_2 L\,(x_n, y_n) + \right. \\ \qquad \left. + (1 - \alpha_n)\, \lambda_{n-1} \left(\nabla_2 L\,(x_n, y_n) - \nabla_2 L\,(x_{n-1}, y_{n-1}) \right) \right). \end{cases}$$

Перейдемо до доведення сильної збіжності алгоритму 3.4.

Має місце

Лема 3.3. *Для послідовності* (x_n)*, що породжена алгоритмом 3.4, виконується нерівність*

$$\begin{aligned} &\|x_{n+1} - z\|^2 + 2\lambda_n \left(A x_n - A x_{n+1}, x_{n+1} - z \right) + \frac{1}{2} \|x_{n+1} - x_n\|^2 \leqslant \\ &\leqslant (1 - \alpha_n) \left(\|x_n - z\|^2 + 2\lambda_{n-1} \left(A x_{n-1} - A x_n, x_n - z \right) + \frac{1}{2} \|x_n - x_{n-1}\|^2 \right) + \\ &\qquad + \alpha_n \|y - z\|^2 - \alpha_n \|y - x_{n+1}\|^2 - \\ &\qquad - \left(\frac{1}{2} - \alpha_n - (1 - \alpha_n)\, \lambda_{n-1} L \right) \|x_{n+1} - x_n\|^2 - \\ &\qquad - (1 - \alpha_n) \left(\frac{1}{2} - \lambda_{n-1} L \right) \|x_n - x_{n-1}\|^2, \end{aligned} \tag{3.19}$$

де $z \in S$.

Доведення. Нехай $z \in S$. Маємо

$$(x_{n+1} - \alpha_n y - (1 - \alpha_n)\, x_n + \lambda_n A x_n +$$

$$+ (1-\alpha_n)\,\lambda_{n-1}\,(Ax_n - Ax_{n-1})\,, z - x_{n+1}) \geqslant 0. \quad (3.20)$$

Монотонність оператора A та включення $z \in S$ дає

$$\begin{gathered}(\lambda_n Ax_n + (1-\alpha_n)\,\lambda_{n-1}\,(Ax_n - Ax_{n-1})\,, z - x_{n+1}) = \\ = \lambda_n\,(Ax_n - Ax_{n+1}, z - x_{n+1}) + \\ + (1-\alpha_n)\,\lambda_{n-1}\,(Ax_n - Ax_{n-1}, z - x_{n+1}) + \underbrace{\lambda_n\,(Ax_{n+1}, z - x_{n+1})}_{\leqslant 0} \leqslant \\ \leqslant \lambda_n\,(Ax_n - Ax_{n+1}, z - x_{n+1}) + \\ + (1-\alpha_n)\,\lambda_{n-1}\,(Ax_n - Ax_{n-1}, z - x_n) + \\ + (1-\alpha_n)\,\lambda_{n-1}\,(Ax_n - Ax_{n-1}, x_n - x_{n+1})\,. \quad (3.21)\end{gathered}$$

Використавши (3.21) в (3.20), отримаємо

$$\begin{gathered}0 \leqslant 2\,(x_{n+1} - \alpha_n y - (1-\alpha_n)\,x_n, z - x_{n+1}) + \\ + 2\lambda_n\,(Ax_n - Ax_{n+1}, z - x_{n+1}) + \\ + 2\,(1-\alpha_n)\,\lambda_{n-1}\,(Ax_n - Ax_{n-1}, z - x_n) + \\ + 2\,(1-\alpha_n)\,\lambda_{n-1}\,(Ax_n - Ax_{n-1}, x_n - x_{n+1})\,. \quad (3.22)\end{gathered}$$

Оцінимо зверху доданок $2\lambda_{n-1}\,(Ax_n - Ax_{n-1}, x_n - x_{n+1})$ в (3.22). Маємо

$$\begin{gathered}2\lambda_{n-1}\,(Ax_n - Ax_{n-1}, x_n - x_{n+1}) \leqslant 2\lambda_{n-1}\,\|Ax_n - Ax_{n-1}\| \cdot \|x_n - x_{n+1}\| \leqslant \\ \leqslant 2\lambda_{n-1} L\,\|x_n - x_{n-1}\| \cdot \|x_{n+1} - x_n\| \leqslant \\ \leqslant \lambda_{n-1} L\,\|x_n - x_{n-1}\|^2 + \lambda_{n-1} L\,\|x_n - x_{n+1}\|^2\,.\end{gathered}$$

Перетворимо доданок $2\,(x_{n+1} - \alpha_n y - (1-\alpha_n)\,x_n, z - x_{n+1})$ в (3.22). Маємо

$$\begin{gathered}2\,(x_{n+1} - \alpha_n y - (1-\alpha_n)\,x_n, z - x_{n+1}) = \|\alpha_n y + (1-\alpha_n)\,x_n - z\|^2 - \\ - \|x_{n+1} - z\|^2 - \|x_{n+1} - \alpha_n y - (1-\alpha_n)\,x_n\|^2\,. \quad (3.23)\end{gathered}$$

Для перетворення різниці

$$\|\alpha_n y + (1-\alpha_n)\,x_n - z\|^2 - \|\alpha_n y + (1-\alpha_n)\,x_n - x_{n+1}\|^2$$

в (3.23) використаємо таку тотожність

$$\|\alpha u+(1-\alpha)\,v-w\|^2=\|v-w-\alpha\,(v-u)\|^2=$$
$$=\|v-w\|^2-2\alpha\,(v-w,v-u)+\alpha^2\,\|v-u\|^2=$$
$$=\|v-w\|^2-\alpha\,\|v-u\|^2-\alpha\,\|v-w\|^2+\alpha\,\|u-w\|^2+\alpha^2\,\|v-u\|^2,$$

де u, v, $w\in H$, $\alpha>0$. Отримаємо рівність

$$\|\alpha_n y+(1-\alpha_n)\,x_n-z\|^2-\|\alpha_n y+(1-\alpha_n)\,x_n-x_{n+1}\|^2=$$
$$=\|x_n-z\|^2-\alpha_n\,\|x_n-y\|^2-\alpha_n\,\|x_n-z\|^2+$$
$$+\,\alpha_n\,\|y-z\|^2+\alpha_n^2\,\|x_n-y\|^2-$$
$$-\,\|x_n-x_{n+1}\|^2+\alpha_n\,\|x_n-y\|^2+\alpha_n\,\|x_n-x_{n+1}\|^2-$$
$$-\,\alpha_n\,\|y-x_{n+1}\|^2-\alpha_n^2\,\|x_n-y\|^2=$$
$$=(1-\alpha_n)\,\|x_n-z\|^2-(1-\alpha_n)\,\|x_n-x_{n+1}\|^2+$$
$$+\,\alpha_n\,\|y-z\|^2-\alpha_n\,\|y-x_{n+1}\|^2.$$

Прийшли до нерівності

$$0\leqslant(1-\alpha_n)\,\|x_n-z\|^2-(1-\alpha_n)\,\|x_n-x_{n+1}\|^2+$$
$$+\,\alpha_n\,\|y-z\|^2-\alpha_n\,\|y-x_{n+1}\|^2-\|x_{n+1}-z\|^2+$$
$$+\,2\lambda_n\,(Ax_n-Ax_{n+1},z-x_{n+1})+$$
$$+\,2\,(1-\alpha_n)\,\lambda_{n-1}\,(Ax_n-Ax_{n-1},z-x_n)+$$
$$+\,(1-\alpha_n)\,\lambda_{n-1}L\,\|x_n-x_{n-1}\|^2+(1-\alpha_n)\,\lambda_{n-1}L\,\|x_n-x_{n+1}\|^2. \quad (3.24)$$

Перегрупуємо члени в (3.24) та отримаємо

$$\|x_{n+1}-z\|^2+2\lambda_n\,(Ax_n-Ax_{n+1},x_{n+1}-z)+\frac{1}{2}\,\|x_{n+1}-x_n\|^2\leqslant$$
$$\leqslant(1-\alpha_n)\left(\|x_n-z\|^2+2\lambda_{n-1}\,(Ax_{n-1}-Ax_n,x_n-z)+\frac{1}{2}\,\|x_n-x_{n-1}\|^2\right)+$$
$$+\alpha_n\,\|y-z\|^2-\alpha_n\,\|y-x_{n+1}\|^2-\left(\frac{1}{2}-\alpha_n-(1-\alpha_n)\,\lambda_{n-1}L\right)\|x_{n+1}-x_n\|^2-$$

$$- \left(1-\alpha_n\right)\left(\frac{1}{2}-\lambda_{n-1}L\right)\left\|x_n-x_{n-1}\right\|^2,$$

що й потрібно було довести. □

Лема 3.4. *Для послідовності* (x_n)*, що породжена алгоритмом 3.4, виконується нерівність*

$$\begin{gathered}\left\|x_{n+1}-z\right\|^2+2\lambda_n\left(Ax_n-Ax_{n+1},x_{n+1}-z\right)+\frac{1}{2}\left\|x_{n+1}-x_n\right\|^2\leqslant\\ \leqslant\left(1-\alpha_n\right)\left(\left\|x_n-z\right\|^2+2\lambda_{n-1}\left(Ax_{n-1}-Ax_n,x_n-z\right)+\frac{1}{2}\left\|x_n-x_{n-1}\right\|^2\right)+\\ +2\alpha_n\left(y-z,x_{n+1}-z\right)-\left(\frac{1}{2}-\alpha_n-\left(1-\alpha_n\right)\lambda_{n-1}L\right)\left\|x_{n+1}-x_n\right\|^2-\\ -\left(1-\alpha_n\right)\left(\frac{1}{2}-\lambda_{n-1}L\right)\left\|x_n-x_{n-1}\right\|^2,\end{gathered} \tag{3.25}$$

де $z\in S$*.*

Доведення. Застосуємо елементарну нерівність

$$\|a+b\|^2\leqslant\|a\|^2+2\left(b,a+b\right).$$

Маємо

$$\begin{gathered}\left\|y-z\right\|^2=\left\|y-x_{n+1}+x_{n+1}-z\right\|^2\leqslant\\ \leqslant\left\|y-x_{n+1}\right\|^2+2\left(y-z,x_{n+1}-z\right).\end{gathered} \tag{3.26}$$

Використавши (3.26) в нерівності (3.19), отримаємо (3.25), що і потрібно було довести. □

Доведемо обмеженість послідовності (x_n).

Лема 3.5. *Нехай виконана умова* (3.18)*. Тоді породжена алгоритмом 3.4 послідовність* (x_n) *є обмеженою.*

Доведення. Оскільки

$$0<\inf_n\lambda_n\leqslant\sup_n\lambda_n<\tfrac{1}{2L},\quad \lim_{n\to\infty}\alpha_n=0,$$

то існує такий номер $n_0 \geqslant 1$, що

$$\frac{1}{2} - \alpha_n - (1 - \alpha_n)\lambda_{n-1}L = \frac{1}{2} - \lambda_{n-1}L - \alpha_n(1 - \lambda_{n-1}L) > 0 \tag{3.27}$$

та

$$(1 - \alpha_n)\left(\frac{1}{2} - \lambda_{n-1}L\right) > 0. \tag{3.28}$$

З (3.19) та (3.27), (3.28) випливає, що для $n \geqslant n_0$ виконується нерівність

$$\xi_{n+1} \leqslant (1 - \alpha_n)\xi_n + \alpha_n \|y - z\|^2, \tag{3.29}$$

де

$$\xi_n = \|x_n - z\|^2 + 2\lambda_{n-1}(Ax_{n-1} - Ax_n, x_n - z) + \frac{1}{2}\|x_n - x_{n-1}\|^2, \quad z \in S.$$

Оцінимо знизу ξ_n. Маємо

$$\begin{aligned}
\xi_n &= \|x_n - z\|^2 + 2\lambda_{n-1}(Ax_{n-1} - Ax_n, x_n - z) + \frac{1}{2}\|x_n - x_{n-1}\|^2 \geqslant \\
&\geqslant \|x_n - z\|^2 - 2\lambda_{n-1}\|Ax_{n-1} - Ax_n\|\|x_n - z\| + \frac{1}{2}\|x_{n-1} - x_n\|^2 \geqslant \\
&\geqslant \|x_n - z\|^2 - 2\lambda_{n-1}L\|x_{n-1} - x_n\|\|x_n - z\| + \frac{1}{2}\|x_{n-1} - x_n\|^2 \geqslant \\
&\geqslant (1 - \lambda_{n-1}L)\|x_n - z\|^2 + \left(\frac{1}{2} - \lambda_{n-1}L\right)\|x_n - x_{n-1}\|^2 \geqslant 0.
\end{aligned} \tag{3.30}$$

З нерівностей (3.29), (3.30) та леми 2.3 випливає обмеженість послідовностей (ξ_n) та (x_n), що і потрібно було довести. □

Сформулюємо основний результат.

Теорема 3.6. *Нехай C — непорожня опукла та замкнена підмножина гільбертового простору H, $A : H \to H$ — монотонний та ліпшицевий на множині C оператор, $S \neq \varnothing$, $y \in H$, виконується умова (3.18). Тоді послідовність (x_n), що породжена алгоритмом 3.4, сильно збігається до точки $z = P_S y$.*

Доведення. Нехай $z = P_S y$. З леми 3.5 випливає існування такого числа $M \geqslant 0$, що

$$|(y - z, x_{n+1} - z)| \leqslant M$$

для всіх $n \geqslant 1$. Тоді з леми 3.4 випливає нерівність

$$\xi_{n+1} - \xi_n + \alpha_n \xi_n + \left(\frac{1}{2} - \alpha_n - (1-\alpha_n)\lambda_{n-1} L\right) \|x_{n+1} - x_n\|^2 + \\ + (1-\alpha_n)\left(\frac{1}{2} - \lambda_{n-1} L\right) \|x_n - x_{n-1}\|^2 \leqslant 2\alpha_n M, \quad (3.31)$$

де $\xi_n = \|x_n - z\|^2 + 2\lambda_{n-1}(Ax_{n-1} - Ax_n, x_n - z) + \frac{1}{2}\|x_n - x_{n-1}\|^2$.

Розглянемо числову послідовність (ξ_n). Можливі два варіанти:

a) існує номер $\bar{n} \geqslant 1$ такий, що $\xi_{n+1} \leqslant \xi_n$ для всіх $n \geqslant \bar{n}$;

b) існує зростаюча послідовність номерів (n_k) така, що $\xi_{n_k+1} > \xi_{n_k}$ для всіх $k \geqslant 1$.

Спочатку розглянемо варіант a). В цьому випадку існує $\lim_{n\to\infty} \xi_n$. Оскільки $\xi_{n+1} - \xi_n \to 0$, $\alpha_n \to 0$ та виконуються нерівності (3.31), то при $n \to \infty$ маємо

$$\|x_{n+1} - x_n\| \to 0. \quad (3.32)$$

Покажемо, що всі слабкі часткові границі послідовності (x_n) належать множині S. Розглянемо підпослідовність (x_{n_k}), що слабко збігається до деякої точки $w \in H$. Ясно, що $w \in C$. Покажемо, що $w \in S$. Маємо

$$(x_{n_k+1} - \alpha_{n_k} y - (1-\alpha_{n_k}) x_{n_k} + \lambda_{n_k} A x_{n_k} + \\ + (1-\alpha_{n_k}) \lambda_{n_k-1} (Ax_{n_k} - Ax_{n_k-1}), y - x_{n_k+1}) \geqslant 0 \quad \forall y \in C.$$

Звідки, використовуючи монотонність оператора A, виводимо оцінку

$$(Ay, y - x_{n_k}) + (Ax_{n_k}, x_{n_k} - x_{n_k+1}) \geqslant (Ax_{n_k}, y - x_{n_k+1}) \geqslant \\ \geqslant \frac{1}{\lambda_{n_k}} (\alpha_{n_k}(y - x_{n_k}) + x_{n_k} - x_{n_k+1}, y - x_{n_k+1}) - \\ - (1-\alpha_{n_k}) \frac{\lambda_{n_k-1}}{\lambda_{n_k}} (Ax_{n_k} - Ax_{n_k-1}, y - x_{n_k+1}) \quad \forall y \in C.$$

З $\lim_{n\to\infty} \alpha_n = 0$, обмеженості послідовності (x_n), (3.32) та ліпшицевості оператора A випливає

$$\liminf_{k\to\infty} (Ay, y - x_{n_k}) \geqslant 0 \quad \forall y \in C.$$

З іншого боку

$$(Ay, y - w) = \lim_{k\to\infty} (Ay, y - x_{n_k}) = \liminf_{k\to\infty} (Ay, y - x_{n_k}) \geqslant 0 \quad \forall y \in C.$$

Отже, $w \in S$.

Доведемо, що

$$\limsup_{n\to\infty} (y - z, x_{n+1} - z) \leqslant 0. \tag{3.33}$$

Розглянемо таку підпослідовність (x_{n_k}), що

$$\lim_{k\to\infty} (y - z, x_{n_k} - z) = \limsup_{n\to\infty} (y - z, x_{n+1} - z)\,.$$

Можна вважати, що $x_{n_k} \rightharpoonup w \in S$. Тоді отримуємо

$$\lim_{k\to\infty} (y - z, x_{n_k} - z) = (y - z, w - z) = (y - P_S y, w - P_S y) \leqslant 0,$$

чим доводимо (3.33).

Тепер з (3.33), нерівності

$$\xi_{n+1} \leqslant (1 - \alpha_n)\, \xi_n + 2\alpha_n\, (y - z, x_{n+1} - z)\,,$$

що виконується для достатньо великих n, та леми 2.4 робимо висновок, що

$$\xi_n = \|x_n - z\|^2 + 2\lambda_{n-1}\, (Ax_{n-1} - Ax_n, x_n - z) + \frac{1}{2}\, \|x_n - x_{n-1}\|^2 \to 0.$$

З (3.30) отримуємо

$$\lim_{n\to\infty} \|x_n - z\| = 0.$$

Розглянемо варіант b). В цьому випадку існує неспадна послідовність номерів (m_k) з властивостями (лема 2.5):

i) $m_k \to +\infty$;

ii) $\xi_{m_k+1} \geqslant \xi_{m_k}$ для всіх $k \geqslant n_1$;

iii) $\xi_{m_k+1} \geqslant \xi_k$ для всіх $k \geqslant n_1$.

З нерівності леми 3.4, (3.30) та ii) випливає

$$\left(\frac{1}{2} - \alpha_{m_k} - (1 - \alpha_{m_k})\, \lambda_{m_k - 1} L\right) \|x_{m_k+1} - x_{m_k}\|^2 + \\ + (1 - \alpha_{m_k}) \left(\frac{1}{2} - \lambda_{m_k-1} L\right) \|x_{m_k} - x_{m_k-1}\|^2 \leqslant 2\alpha_{m_k} M.$$

Звідки

$$\lim_{k\to\infty} \|x_{m_k+1} - x_{m_k}\| = \lim_{k\to\infty} \|x_{m_k} - x_{m_k-1}\| = 0.$$

Міркуваннями, що подібні вищевикладеним, доводимо, що слабкі часткові границі послідовності (x_{m_k}) належать множині S. З тотожності

$$(y - z, x_{m_k+1} - z) = (y - z, x_{m_k} - z) + (y - z, x_{m_k+1} - x_{m_k})$$

отримуємо

$$\limsup_{k\to\infty} (y - z, x_{m_k+1} - z) = \limsup_{k\to\infty} (y - z, x_{m_k} - z)\,.$$

Як і раніше отримуємо нерівність

$$\limsup_{k\to\infty} (y - z, x_{m_k+1} - z) \leqslant 0.$$

Далі маємо

$$\xi_{m_k+1} \leqslant (1 - \alpha_{m_k})\, \xi_{m_k} + 2\alpha_{m_k} (y - z, x_{m_k+1} - z) \leqslant \\ \leqslant (1 - \alpha_{m_k})\, \xi_{m_k+1} + 2\alpha_{m_k} (y - z, x_{m_k+1} - z)\,.$$

Звідки, ураховуючи умову iii), отримуємо

$$\xi_k \leqslant \xi_{m_k+1} \leqslant 2\,(y - z, x_{m_k+1} - z)\,.$$

Таким чином,

$$\limsup_{k\to\infty} \xi_k \leqslant 2 \limsup_{k\to\infty} (y - z, x_{m_k+1} - z) \leqslant 0.$$

Отже, $\lim_{n\to\infty} \xi_n = 0$ та, в свою чергу, $\lim_{n\to\infty} \|x_n - z\| = 0$, що і потрібно було довести. $\square$

Зауваження 3.10. Важливим питанням є дослідження асимптотичної поведінки алгоритму 3.4 у ситуації $C = H$:

$$x_{n+1} = \alpha_n y + (1 - \alpha_n) x_n - \lambda_n A x_n - (1 - \alpha_n) \lambda_{n-1} (A x_n - A x_{n-1}).$$

А саме, питання про поведінку норми $\|Ax_n\|$. На нашу думку повинна виконуватись оцінка

$$\|Ax_n\| = O\left(\frac{1}{n}\right).$$

Зауважимо, що в роботі [180] для екстраградієнтного методу отримана оцінка

$$\|Ax_n\| = O\left(\frac{1}{\sqrt{n}}\right),$$

а в роботі [181] отримана оцінка

$$\|Ax_n\| = O\left(\frac{1}{n}\right)$$

для екстраградієнтного методу з регуляризацією Гальперна

$$\begin{cases} y_n = x_n + \frac{1}{n+2} (x_0 - x_n) - \frac{1}{8L} A x_n, \\ x_{n+1} = x_n + \frac{1}{n+2} (x_0 - x_n) - \frac{1}{8L} A y_n. \end{cases}$$

Зауваження 3.11. Параметри λ_n алгоритму 3.1 задавалась виходячи з умови (3.18). Тобто, явно використовувалась інформація про константи ліпшицевості оператора A. Схема з статей $[1, 131]$ дозволяє побудувати модифікацію алгоритму 3.4 з адаптивним вибором величини λ_n, що не вимагає знання ліпшицевих констант операторів та процедур типу лінійного пошуку.

Зауваження 3.12. Відштовхуючись від результатів робіт $[1, 3, 131]$ можна отримати аналог алгоритму 3.4 з узагальненою проєкцію Альбера для розв'язання варіаційних нерівностей в рівномірно опуклих та рівномірно гладких банахових просторах. Дане питання докладно розглянемо в розділі 4.

3.7. Висновки до розділу 3

У розділі досліджено нові ітераційні алгоритми для розв'язання варіаційних нерівностей в гільбертових просторах. Дані алгоритми є модификаціями відомого «forward-reflected-backward algorithm», що запропонований в [115].

А саме, доведено слабку збіжність алгоритму операторної екстраполяції та отримано сублінійну оцінку ефективності для функції зазору. Доведено сильну збіжність алгоритму операторної екстраполяції для варіаційних нерівностей з рівномірно монотонними операторами. Доведено лінійну швидкість збіжності алгоритму операторної екстраполяції для варіаційних нерівностей з операторами, що задовольняють умову типу узагальненої сильної монотонності. Запропоновано та обгрунтовано адаптивний варіант алгоритму. Запропоновано регуляризований варіант алгоритму та доведено теорему про його сильну збіжність. Для регуляризації використано схему Гальперна.

Основні результати даного розділу опубліковано в [4–9, 11–14].

РОЗДІЛ 4
МЕТОДИ З ДИВЕРГЕНЦІЄЮ БРЕГМАНА ТА В БАНАХОВИХ ПРОСТОРАХ

У підрозділі 4.1 досліджено варіанти алгоритмів екстраполяції з минулого та операторної екстраполяції з дивергенцією Брегмана для розв'язання варіаційних нерівностей з монотонними та ліпшицевими операторами, що діють в скінченновимірному дійсному лінійному просторі. Основні результати: оцінки ефективності в термінах функції зазору.

У підрозділі 4.2 досліджено нові алгоритми для розв'язання варіаційних нерівностей з монотонними, ліпшицевими операторами, що діють в 2-рівномірно опуклих та рівномірно гладких банахових просторах. Перший алгоритм — варіант методу операторної екстраполяції, що використовує узагальнену проєкцію Альбера замість метричної. Другий алгоритм є адаптивним варіантом першого, де використовується правило поновлення величини кроку, що не вимагає знання ліпшицевих констант та лінійного пошуку. Основні результати: оцінка ефективності в термінах функції зазору та теорема про слабку збіжність методу. Також встановлена слабка збіжність адаптивного варіанту методу, де використовується правило поновлення величини кроку, що не вимагає знання ліпшицевих констант та лінійного пошуку.

4.1. Алгоритми з дивергенцією Брегмана

Розглянемо варіанти алгоритмів з дивергенцією Брегмана для розв'язання варіаційних нерівностей з монотонними та ліпшицевими операторами, що діють в скінченновимірному дійсному лінійному просторі.

Нехай E — скінченновимірний дійсний лінійний простір. Введемо у цьому просторі норму $\|\cdot\|$ (не обов'язково евклідову). E^* — спряжений простір. Для $a \in E^*$ та $b \in E$ через $\langle a, b\rangle$ позначимо значення линійної функції a в точці b. Спряжена норма на E^* позначається $\|\cdot\|_*$.

Нехай $C \subseteq E$ — непорожня замкнена та опукла множина, A — оператор, що діє з E в E^*. Розглянемо варіаційну нерівність:

$$\text{знайти } x \in C: \ \langle Ax, y - x\rangle \geqslant 0 \quad \forall y \in C, \tag{4.1}$$

множину розв'язків якої позначимо S.

Перейдемо до опису результатів.

4.1.1. Допоміжні відомості та алгоритми. Введемо необхідні для алгоритмів конструкції.

Нехай $\varphi : E \to \mathbb{R} \cup \{+\infty\}$ — власна замкнена опукла функція, що диференційовна на $\operatorname{dom}(\partial\varphi)$. Дивергенція Брегмана (відстань Брегмана), що відповідає функції φ, задається формулою [169]

$$V(a,b) = \varphi(a) - \varphi(b) - \langle \nabla\varphi(b), a - b\rangle \quad \forall a \in \operatorname{dom}(\varphi),\ b \in \operatorname{dom}(\partial\varphi).$$

Припустимо, що виконані такі умови:

- функція φ власна замкнена та опукла;
- функція φ диференційовна на $\operatorname{dom}(\partial\varphi)$;
- $C \subseteq \operatorname{dom}(\varphi)$;
- на множині C функція φ є сильно опуклою відносно норми $\|\cdot\|$ з константою сильної опуклості $\sigma > 0$.

Зрозуміло, що

$$V(a,b) \geqslant \tfrac{\sigma}{2}\|a - b\|^2 \quad \forall a \in C \quad \forall b \in C \cap \operatorname{dom}(\partial\varphi).$$

та

$$V(a,b) = 0 \quad \Leftrightarrow \quad a = b.$$

Має місце корисна 3-точкова тотожність [169]:

$$V\left(a,c\right)=V\left(a,b\right)+V\left(b,c\right)+\left\langle \nabla\varphi\left(b\right)-\nabla\varphi\left(c\right),a-b\right\rangle ,$$

де $a\in\mathrm{dom}(\varphi)$, $b,c\in\mathrm{dom}(\partial\varphi)$.

Розглянемо два основних приклади дивергенцій Брегмана. Для

$$\varphi\left(\cdot\right)=\tfrac{1}{2}\left\|\cdot\right\|_2^2,$$

де $\left\|\cdot\right\|_2$ — евклідова норма, маємо

$$V\left(x,y\right)=\tfrac{1}{2}\left\|x-y\right\|_2^2.$$

Для ймовірносного симплексу

$$\Delta^m=\left\{x\in\mathbb{R}^m:\quad x_i\geqslant0,\quad\sum_{i=1}^m x_i=1\right\}$$

та функції від'ємної ентропії

$$\varphi\left(x\right)=\sum_{i=1}^m x_i\ln x_i$$

(за нерівністю Пінскера [182] вона сильно опукла з константою 1 відносно ℓ_1-норми на симплексі Δ^m) отримуємо дивергенцію Кульбака–Лейблера [169]

$$V\left(x,y\right)=\sum_{i=1}^m x_i\ln\left(\frac{x_i}{y_i}\right),$$

де $x\in\Delta^m$, $y\in\Delta_+^m=\{x\in\mathbb{R}^m:\quad x_i>0,\quad\sum_{i=1}^m x_i=1\}$.

Розглянемо сильно опуклі задачі мінімізації вигляду

$$P_x^C\left(a\right)=\mathrm{argmin}_{y\in C}\left\{-\left\langle a,y-x\right\rangle+V\left(y,x\right)\right\},\tag{4.2}$$

де $a\in E^*$, $x\in\mathrm{dom}(\partial\varphi)$. Відомо [169], що задача (4.2) має єдиний розв'язок $z\in C\cap\mathrm{dom}(\partial\varphi)$, причому

$$-\left\langle a,y-z\right\rangle+\left\langle\nabla\varphi\left(z\right)-\nabla\varphi\left(x\right),y-z\right\rangle\geqslant0\quad\forall y\in C.$$

Останню нерівність, ураховуючи 3-точкову тотожність, можна записати у вигляді

$$V(y,z) \leqslant V(y,x) - V(y,z) - \langle a, y-z \rangle \quad \forall y \in C.$$

Зауваження 4.1. Точка $P_x^C(a)$ в евклідовому випадку співпадає з евклідовою метричною проєкцією

$$P_C(x+a) = \operatorname{argmin}_{y \in C} \|y - (x+a)\|_2 .$$

Зауваження 4.2. Для ймовірносного симплексу Δ^m та дивергенції Кульбака–Лейблера маємо [169]

$$P_x^{\Delta^m}(a) = \left(\frac{x_1 e^{a_1}}{\sum_{j=1}^m x_j e^{a_j}}, \frac{x_2 e^{a_2}}{\sum_{j=1}^m x_j e^{a_j}}, \ldots, \frac{x_m e^{a_m}}{\sum_{j=1}^m x_j e^{a_j}} \right), \quad a \in \mathbb{R}^m,\ x \in \Delta_+^m.$$

Зауваження 4.3. Часто відображення $P_x^C : E^* \to C \cap \operatorname{dom}(\partial\varphi)$ називають прокс-відображенням [183] та позначають $P_x^C(a)$ через $\operatorname{Mirr}_x(a)$ [184].

Припустимо, що виконані такі умови:

- множина $C \subseteq E$ – опукла та замкнена;
- оператор $A : E \to E^*$ – монотонний на C, тобто

  $$\langle Ax - Ay, x - y \rangle \geqslant 0 \quad \forall x, y \in C,$$

  та ліпшицевий на C (з константою $L > 0$), тобто

  $$\|Ax - Ay\|_* \leqslant \|x - y\| \quad \forall x, y \in C;$$

- множина S непорожня.

Розглянемо дуальну варіаційну нерівність:

$$x \in C : \ \langle Ay, x - y \rangle \leqslant 0 \quad \forall y \in C. \tag{4.3}$$

Множину розв'язків задачі (4.3) позначимо S^d. Множина S^d опукла та замкнена. Для монотонних операторів A завжди маємо $S \subseteq S^d$. В наших умовах маємо $S^d = S$.

Перший з розглянутих алгоритмів має вигляд.

Алгоритм 4.1 (Екстраполяція з минулого). Для $x_1 = y_0 \in C \cap \operatorname{dom}(\partial\varphi)$ генерируємо послідовність елементів x_n, $y_n \in C \cap \operatorname{dom}(\partial\varphi)$ за допомогою ітераційної схеми

$$\begin{cases} y_n = P_{x_n}^C\left(-\lambda_n A y_{n-1}\right), \\ x_{n+1} = P_{x_n}^C\left(-\lambda_n A y_n\right), \end{cases}$$

де $\lambda_n > 0$.

Зауваження 4.4. Алгоритм 4.1 запропоновано в [155]. В роботах [9,104–109] досліджувалась його збіжність та прононувались деякі модифікації.

При виконанні для деякого $n \in \mathbb{N}$ в алгоритмі 4.1 рівностей

$$y_n = y_{n-1} = x_n \quad \text{або} \quad x_{n+1} = x_n = y_n \tag{4.4}$$

має місце включення $y_n \in S$. Дійсно, рівність

$$x_{n+1} = P_{x_n}^C\left(-\lambda_n A y_n\right)$$

рівносильна нерівності

$$\langle Ay_n, y - x_{n+1}\rangle + \frac{\langle \nabla\varphi(x_{n+1}) - \nabla\varphi(x_n), y - x_{n+1}\rangle}{\lambda_n} \geqslant 0 \quad \forall y \in C.$$

З другої рівності (4.4) випливає

$$\langle Ay_n, y - y_n\rangle \geqslant 0 \;\; \forall y \in C,$$

тобто, $y_n \in S$.

Аналогічно, з

$$\langle Ay_{n-1}, y - y_n\rangle + \frac{\langle \nabla\varphi(y_n) - \nabla\varphi(x_n), y - y_n\rangle}{\lambda_n} \geqslant 0 \quad \forall y \in C$$

при першій рівності в (4.4) отримуємо $y_n \in S$.

Далі припустимо, що для всіх номерів $n \in \mathbb{N}$ умова (4.4) не має місця.

Другий з розглянутих алгоритмів має вигляд.

Алгоритм 4.2 (Операторна екстраполяція). Обираємо $x_0 = x_1 \in C \cap \operatorname{dom}(\partial\varphi)$, $\lambda_n, \mu_n > 0$. Покладаємо $n = 1$.

1. Обчислити

$$x_{n+1} = P_{x_n}^C \left(-\lambda_n A x_n - \mu_n (A x_n - A x_{n-1})\right).$$

2. Якщо $x_{n+1} = x_n = x_{n-1}$, то СТОП, інакше покласти $n := n + 1$ та перейти до **1**.

Зауваження 4.5. Алгоритм 4.2 запропоновано в [9]. Алгоритм 4.2 є модифікацією методу з попереднього розділу, що використовує дивергенцію Брегмана замість квадрату евклідової норми.

Правило зупинки в алгоритмі 4.2 обґрунтовується так: при виконанні

$$x_{n+1} = x_n = x_{n-1}$$

маємо

$$x_n = P_{x_n}^C \left(-\lambda_n A x_n\right),$$

звідки $x_n \in S$.

Наведемо дві версії алгоритму 4.2.

Розглянемо варіаційну нерівність на стандартному симплексі:

$$\text{знайти} \ \ x \in \Delta^m : \ \ (Ax, y - x) \geqslant 0 \ \ \forall y \in \Delta_m.$$

Обираючи дивергенцію Кульбака–Лейблера та $\mu_n = \lambda_n = \lambda > 0$, одержуємо таку версію:

$$x_i^{n+1} = \frac{x_i^n e^{-\lambda(2Ax_n - Ax_{n-1})_i}}{\sum_{j=1}^d x_j^n e^{-\lambda(2Ax_n - Ax_{n-1})_j}}, \ \ i = 1, ..., m,$$

де $(a)_i \in \mathbb{R}$ — i-та координата вектора $a \in \mathbb{R}^m$.

Розглянемо варіаційну нерівність на добутку стандартних симплексів:

$$\text{знайти} \ \ x \in \Delta^{m_1} \times \Delta^{m_2} : \ \ (Ax, y - x) \geqslant 0 \ \ \forall y \in \Delta^{m_1} \times \Delta^{m_2}. \tag{4.5}$$

За сепарабельною функцією

$$\varphi(x) = \varphi_1(x_1) + \varphi_2(x_2) = \sum_{i=1}^{m_1} x_{1,i} \ln x_{1,i} + \sum_{i=1}^{m_2} x_{2,i} \ln x_{2,i},$$

де $x = (x_1, x_2) = (\underbrace{x_{1,1}, x_{1,2}, \ldots x_{1,m_1}}_{x_1}, \underbrace{x_{2,1}, x_{2,2}, \ldots x_{2,m_2}}_{x_2}) \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$, побудуємо дивергенцію Брегмана на $C = \Delta^{m_1} \times \Delta^{m_2}$:

$$V(x, y) = V_1(x_1, y_1) + V_2(x_2, y_2) = \sum_{i=1}^{m_1} x_{1,i} \ln \frac{x_{1,i}}{y_{1,i}} + \sum_{i=1}^{m_2} x_{2,i} \ln \frac{x_{2,i}}{y_{2,i}}.$$

Алгоритм 4.2 для нерівності (4.5) з таким вибором дивергенції та $\mu_n = \lambda_n = \lambda > 0$ приймає вигляд:

$$x_{k,i}^{n+1} = \frac{x_{k,i}^n \exp\left(-\lambda\left(2Ax_n - Ax_{n-1}\right)_{k,i}\right)}{\sum_{j=1}^{m_k} x_{k,j}^n \exp\left(-\lambda\left(2Ax_n - Ax_{n-1}\right)_{k,j}\right)}, \quad k = 1, 2, \; i = 1, \ldots, m_k,$$

де $(a)_{k,i}$ — $\left(\sum_{t=1}^{k-1} m_t + i\right)$-та координата вектора $a \in \mathbb{R}^{m_1} \times \mathbb{R}^{m_2}$.

4.1.2. Сублінійні оцінки ефективності. У випадку обмеженості множини C доведемо, що алгоритмам необхідно зробити $O\left(\frac{1}{\varepsilon}\right)$ ітерацій для отримання точки $x \in C$ з

$$\operatorname{gap}(x) = \sup_{y \in C} \langle Ay, x - y \rangle \leqslant \varepsilon.$$

Почнемо з аналізу алгоритму 4.1.

Для породжених алгоритмом 4.1 послідовностей (x_n) та (y_n) мають місце нерівності

$$-\lambda_n \langle Ay_{n-1}, y - y_n \rangle \leqslant V(y, x_n) - V(y_n, x_n) - V(y, y_n) \quad \forall y \in C. \tag{4.6}$$

$$-\lambda_n \langle Ay_n, y - x_{n+1} \rangle \leqslant V(y, x_n) - V(x_{n+1}, x_n) - V(y, x_{n+1}) \quad \forall y \in C. \tag{4.7}$$

З (4.7) випливає

$$V\left(y,x_{n+1}\right)\leqslant V\left(y,x_n\right)-V\left(x_{n+1},x_n\right)+\lambda_n\left\langle Ay_n,y-x_{n+1}\right\rangle= \\ =V\left(y,x_n\right)-V\left(x_{n+1},x_n\right)+\lambda_n\left\langle Ay_n,y_n-x_{n+1}\right\rangle+ \\ +\lambda_n\left\langle Ay_n,y-y_n\right\rangle.$$

Урахувавши монотонность оператора A, отримаємо

$$V\left(y,x_{n+1}\right)\leqslant V\left(y,x_n\right)-V\left(x_{n+1},x_n\right)+\lambda_n\left\langle Ay_n,y_n-x_{n+1}\right\rangle+ \\ +\lambda_n\left\langle Ay,y-y_n\right\rangle. \quad (4.8)$$

З (4.6) випливає

$$-V\left(x_{n+1},x_n\right)\leqslant\lambda_n\left\langle Ay_{n-1},x_{n+1}-y_n\right\rangle-V\left(y_n,x_n\right)-V\left(x_{n+1},y_n\right). \quad (4.9)$$

Оцінимо зверху $-V\left(x_{n+1},x_n\right)$ в (4.8) за допомогою (4.9). Отримаємо

$$V\left(y,x_{n+1}\right)\leqslant V\left(y,x_n\right)-V\left(x_{n+1},y_n\right)-V\left(y_n,x_n\right)+ \\ +\lambda_n\left\langle Ay_n-Ay_{n-1},y_n-x_{n+1}\right\rangle+ \\ +\lambda_n\left\langle Ay,y-y_n\right\rangle. \quad (4.10)$$

Оцінимо зверху доданок $\lambda_n\left\langle Ay_n-Ay_{n-1},y_n-x_{n+1}\right\rangle$ в (4.10). Маємо

$$\lambda_n\left\langle Ay_n-Ay_{n-1},y_n-x_{n+1}\right\rangle\leqslant\lambda_nL\left\|y_{n-1}-y_n\right\|\left\|x_{n+1}-y_n\right\|\leqslant \\ \leqslant\frac{\lambda_nL}{2}\left\|y_{n-1}-y_n\right\|^2+\frac{\lambda_nL}{2}\left\|y_n-x_{n+1}\right\|^2\leqslant \\ \leqslant\lambda_nL\left\|y_{n-1}-x_n\right\|^2+\lambda_nL\left\|y_n-x_n\right\|^2+\frac{\lambda_nL}{2}\left\|y_n-x_{n+1}\right\|^2\leqslant \\ \leqslant\frac{2\lambda_nL}{\sigma}V(x_n,y_{n-1})+\frac{2\lambda_nL}{\sigma}V(y_n,x_n)+\frac{\lambda_nL}{\sigma}V(x_{n+1},y_n).$$

Приходимо до нерівності

$$V\left(y,x_{n+1}\right)\leqslant V\left(y,x_n\right)-V\left(y_n,x_n\right)-V\left(x_{n+1},y_n\right)+ \\ +\frac{2\lambda_nL}{\sigma}V(x_n,y_{n-1})+\frac{2\lambda_nL}{\sigma}V(y_n,x_n)+\frac{\lambda_nL}{\sigma}V(x_{n+1},y_n)+ \\ +\lambda_n\left\langle Ay,y-y_n\right\rangle. \quad (4.11)$$

Перепишемо (4.11) у вигляді

$$\lambda_n \langle Ay, y_n - y\rangle \leqslant \left(V(y, x_n) + \frac{2\lambda_n L}{\sigma} V(x_n, y_{n-1})\right) -$$
$$- \left(V(y, x_{n+1}) + \frac{2\lambda_{n+1} L}{\sigma} V(x_{n+1}, y_n)\right) -$$
$$- \left(1 - \frac{2\lambda_{n+1} L}{\sigma} - \frac{\lambda_n L}{\sigma}\right) V(x_{n+1}, y_n) -$$
$$- \left(1 - \frac{2\lambda_n L}{\sigma}\right) V(y_n, x_n). \quad (4.12)$$

Припустимо, що $\lambda_n \in \left(0, \frac{\sigma}{3L}\right]$. Тоді з (4.12) випливає

$$2\lambda_n \langle Ay, y_n - y\rangle \leqslant \left(V(y, x_n) + \frac{2\lambda_n L}{\sigma} V(x_n, y_{n-1})\right) -$$
$$- \left(V(y, x_{n+1}) + \frac{2\lambda_{n+1} L}{\sigma} V(x_{n+1}, y_n)\right). \quad (4.13)$$

Просумувавши (4.13) по n від 1 до N отримаємо

$$2 \sum_{n=1}^{N} \lambda_n \langle Ay, y_n - y\rangle \leqslant V(y, x_1) + \frac{2\lambda_1 L}{\sigma} V(x_1, y_0),$$

та

$$\langle Ay, z_N - y\rangle \leqslant \frac{V(y, x_1)}{2\sum_{n=1}^{N} \lambda_n}, \quad (4.14)$$

де $z_N = \frac{\sum_{n=1}^{N} \lambda_n y_n}{\sum_{n=1}^{N} \lambda_n}$. Переходимо до супремуму по $y \in C$ в (4.14)

$$\operatorname{gap}(z_N) \leqslant \frac{\sup_{y \in C} V(y, x_1)}{2\sum_{n=1}^{N} \lambda_n}.$$

Таким чином, має місце

Теорема 4.1. *Нехай $C \subseteq E$ — непорожня опукла замкнена обмежена множина, $A : E \to E^*$ — монотонний та L-ліпшицевий на множині C оператор. Нехай (y_n) — послідовність, що породжена алгоритмом 4.1 з $\lambda_n = \frac{\sigma}{3L}$, тобто,*

$$\begin{cases} x_1 = y_0 \in C, \\ y_n = P_{x_n}^C\left(-\frac{\sigma}{3L} A y_{n-1}\right), \\ x_{n+1} = P_{x_n}^C\left(-\frac{\sigma}{3L} A y_n\right). \end{cases}$$

Тоді для послідовності середніх $z_N = \frac{1}{N}\sum_{n=1}^N y_n$ *має місце оцінка*

$$\operatorname{gap}(z_N) \leqslant \frac{L\frac{3}{2\sigma}\sup_{y\in C} V(y, x_1)}{N}.$$

Проаналізуємо алгоритм 4.2.

Нехай $\lambda_n \in \left(0, \frac{\sigma}{2L}\right]$, $\mu_n = \lambda_{n-1}$. Для послідовності (x_n) має місце нерівність

$$\begin{aligned} -\langle \lambda_n A x_n + \lambda_{n-1}(Ax_n - Ax_{n-1}), y - x_{n+1}\rangle \leqslant \\ \leqslant V(y, x_n) - V(x_{n+1}, x_n) - V(y, x_{n+1}) \quad \forall y \in C. \end{aligned} \tag{4.15}$$

Перепишемо (4.15) таким чином

$$\begin{aligned} V(y, x_n) - V(y, x_{n+1}) \geqslant \\ \geqslant \lambda_n \langle A x_{n+1}, x_{n+1} - y\rangle - \lambda_n \langle A x_{n+1} - A x_n, x_{n+1} - y\rangle + \\ + \lambda_{n-1}\langle Ax_n - Ax_{n-1}, x_n - y\rangle + \lambda_{n-1}\langle Ax_n - Ax_{n-1}, x_{n+1} - x_n\rangle + \\ + V(x_{n+1}, x_n). \end{aligned} \tag{4.16}$$

Сумуючи (4.16) по n від 1 до N, отримуємо

$$\begin{aligned} V(y, x_1) - V(y, x_{N+1}) \geqslant \\ \geqslant \sum_{n=1}^N \lambda_n \langle A x_{n+1}, x_{n+1} - y\rangle - \lambda_N \langle A x_{N+1} - A x_N, x_{N+1} - y\rangle + \\ + \sum_{n=1}^N \left(\lambda_{n-1}\langle Ax_n - Ax_{n-1}, x_{n+1} - x_n\rangle + V(x_{n+1}, x_n)\right). \end{aligned} \tag{4.17}$$

Ліпшицевість оператора A дає

$$\begin{aligned} \sum_{n=1}^N \left(\lambda_{n-1}\langle Ax_n - Ax_{n-1}, x_{n+1} - x_n\rangle + V(x_{n+1}, x_n)\right) \geqslant \\ \geqslant \sum_{n=1}^N \left(-\frac{\lambda_{n-1}L}{2}\|x_n - x_{n-1}\|^2 - \frac{\lambda_{n-1}L}{2}\|x_n - x_{n+1}\|^2 + \frac{\sigma}{2}\|x_{n+1} - x_n\|^2\right) \geqslant \end{aligned}$$

$$\geqslant \sum_{n=1}^{N} \left(-\frac{\sigma}{4} \|x_n - x_{n-1}\|^2 - \frac{\sigma}{4} \|x_n - x_{n+1}\|^2 + \frac{\sigma}{2} \|x_{n+1} - x_n\|^2 \right) =$$

$$= \sum_{n=1}^{N} \left(-\frac{\sigma}{4} \|x_n - x_{n-1}\|^2 + \frac{\sigma}{4} \|x_{n+1} - x_n\|^2 \right) = \frac{\sigma}{4} \left\| x_{N+1} - x_N \right\|^2 .$$

Використовуючи останню оцінку в (4.17), отримуємо

$$\begin{aligned} V\left(y, x_1\right) - V\left(y, x_{N+1}\right) \geqslant \\ \geqslant \sum_{n=1}^{N} \lambda_n \left\langle A x_{n+1}, x_{n+1} - y \right\rangle - \lambda_N \left\langle A x_{N+1} - A x_N, x_{N+1} - y \right\rangle + \\ + \frac{\sigma}{4} \left\| x_{N+1} - x_N \right\|^2 \geqslant \\ \geqslant \sum_{n=1}^{N} \lambda_n \left\langle A x_{n+1}, x_{n+1} - y \right\rangle - \lambda_N L \left\| x_{N+1} - x_N \right\| \left\| x_{N+1} - y \right\| + \\ + \frac{\sigma}{4} \left\| x_{N+1} - x_N \right\|^2 \geqslant \\ \geqslant \sum_{n=1}^{N} \lambda_n \left\langle A x_{n+1}, x_{n+1} - y \right\rangle - \frac{\lambda_N L}{2} \left\| x_{N+1} - y \right\|^2 . \end{aligned}$$

Приходимо до нерівності

$$\sum_{n=1}^{N} \lambda_n \left\langle A x_{n+1}, x_{n+1} - y \right\rangle - \frac{\lambda_N L}{2} \left\| x_{N+1} - y \right\|^2 + \\ + V\left(y, x_{N+1}\right) \leqslant V\left(y, x_1\right) \quad \forall y \in C. \quad (4.18)$$

Використовуючи монотонність оператора A, отримуємо

$$\sum_{n=1}^{N} \lambda_n \left\langle A x_{n+1}, x_{n+1} - y \right\rangle \geqslant \sum_{n=1}^{N} \lambda_n \left\langle A y, x_{n+1} - y \right\rangle = \\ = \left(\sum_{n=1}^{N} \lambda_n \right) \left\langle A y, z_{N+1} - y \right\rangle , \quad (4.19)$$

де $z_{N+1} = \frac{\sum_{n=1}^{N} \lambda_n x_{n+1}}{\sum_{n=1}^{N} \lambda_n}$. Враховуючи оцінку (4.19) в (4.18), приходимо до нерівності

$$\left(\sum_{n=1}^{N} \lambda_n \right) \left\langle A y, z_{N+1} - y \right\rangle + \left(\frac{\sigma}{2} - \frac{\lambda_N L}{2} \right) \left\| x_{N+1} - y \right\|^2 \leqslant V\left(y, x_1\right) \quad \forall y \in C,$$

звідки випливає

$$\operatorname{gap}(z_{N+1}) = \sup_{y\in C} \langle Ay, z_{N+1} - y\rangle \leqslant \frac{\sup_{y\in C} V(y, x_1)}{\sum_{n=1}^{N} \lambda_n}.$$

Таким чином, має місце

Теорема 4.2. *Нехай* (x_n) *— послідовність, що породжена алгоритмом 4.2 з* $\lambda_n = \mu_n = \frac{\sigma}{2L}$, *тобто,*

$$x_{n+1} = P_{x_n}^{C}\left(-\tfrac{\sigma}{2L}\left(2Ax_n - Ax_{n-1}\right)\right).$$

Тоді має місце оцінка

$$\operatorname{gap}(z_{N+1}) \leqslant \frac{L\frac{2}{\sigma}\sup_{y\in C} V(y, x_1)}{N},$$

де $z_{N+1} = \frac{1}{N}\sum_{n=1}^{N} x_{n+1}$.

Зауваження 4.6. Припустимо, що замість ліпшицевості A виконана така умова:

- оператор $A : E \to E^*$ — відносно ліпшицевий на C (з константою $L > 0$), тобто

$$\langle Ax - Ay, x - z\rangle \leqslant LV(x, y) + LV(z, x),$$

де $x,\ y \in C \cap \operatorname{dom}(\partial\varphi)$, $z \in C$;

Якщо оператор $A : E \to E^*$ — L-ліпшицевий на C, то він є $\frac{L}{\sigma}$-відносно ліпшицевим на C [185]. Дійсно,

$$\langle Ax - Ay, x - z\rangle \leqslant \|Ax - Ay\|_*\|x - z\| \leqslant L\|x - y\|\|x - z\| \leqslant$$
$$\leqslant \frac{L}{2}\|x - y\|^2 + \frac{L}{2}\|x - z\|^2 \leqslant \frac{L}{\sigma}V(x, y) + \frac{L}{\sigma}V(z, x).$$

Актуальною є задача отримання оцінок для алгоритмів 4.1 та 4.2 в класі відносно ліпшицевих монотонних операторів.

Зауваження 4.7. У роботі [9] побудовані адаптивні варіанти алгоритмів 4.1, 4.2.

4.2. Алгоритм операторної екстраполяції в банахових просторах

Розглянемо результати для алгоритму операторної екстраполяції розв'язання варіаційних нерівностей з монотонними, ліпшицевими операторами, що діють в 2-рівномірно опуклих та рівномірно гладких банахових просторах.

4.2.1. Постановка задачі та допоміжні відомості. Спочатку нагадаємо декілька понять та фактів геометрії банахових просторів та нелінійного аналізу, що необхідні для формулювання та доведення результатів [122, 123, 186–188].

Нехай E — дійсний банаховий простір з нормою $\|\cdot\|$, E^* — спряжений до E простір, $\langle x^*, x\rangle$ — значення функціоналу $x^* \in E^*$ на елементі $x \in E$. Норму в E^* будемо позначати $\|\cdot\|_*$.

Нехай

$$S_E = \{x \in E : \ \ \|x\| = 1\}.$$

Банаховий простір E називають строго опуклим, якщо для всіх $x, y \in S_E$ та $x \neq y$ маємо

$$\left\|\frac{x+y}{2}\right\| < 1.$$

Модуль опуклості простору E визначається таким чином

$$\delta_E(\varepsilon) = \inf\left\{1 - \left\|\frac{x+y}{2}\right\| : \ \ x, y \in S_E, \ \ \|x-y\| = \varepsilon\right\} \quad \forall\varepsilon \in (0,2].$$

Банаховий простір E називають рівномірно опуклим, якщо

$$\delta_E(\varepsilon) > 0 \quad \text{для всіх} \quad \varepsilon \in (0,2].$$

Банаховий простір E називають 2-рівномірно опуклим, якщо існує таке $c > 0$, що

$$\delta_E(\varepsilon) \geqslant c\varepsilon^2 \quad \text{для всіх} \quad \varepsilon \in (0,2].$$

Очевидно, що 2-рівномірно опуклий простір є рівномірно опуклим. Відомо, що рівномірно опуклий банаховий простір рефлексивний [186].

Банаховий простір E називають гладким, якщо границя

$$\lim_{t\to 0}\frac{\|x+ty\|-\|x\|}{t} \tag{4.20}$$

існує для всіх x, $y\in S_E$. Банаховий простір E називають рівномірно гладким, якщо границя (4.20) існує рівномірно за x, $y\in S_E$.

Відомо, що гільбертові простори та простори L_p $(1<p\leqslant 2)$ є 2-рівномірно опуклими та рівномірно гладкими (простори L_p рівномірно гладкі для $p\in(1,\infty)$) [186].

Основну інформацію про монотонні оператори та варіаційні нерівності в банахових просторах можна знайти в [29, 31, 122]. Згадаємо лише два мотиваційних приклади монотонних операторів, що діють у банахових просторах [122].

Для $p\geqslant 2$, визначимо оператор A за допомогою формули

$$Au=|u\left(x\right)|^{p-2}u\left(x\right)\int\limits_{\mathbb{R}^3}\frac{|u\left(y\right)|^p}{\|x-y\|_2}dy.$$

Оператор A потенційний, монотонний та діє $L_p\left(\mathbb{R}^3\right)$ з $L_q\left(\mathbb{R}^3\right)$, де $p^{-1}+q^{-1}=1$. Зауважимо, що оператор A є градієнтом опуклого функціоналу

$$F\left(u\right)=\frac{1}{2p}\int\limits_{\mathbb{R}^3}\int\limits_{\mathbb{R}^3}\frac{|u\left(x\right)|^p\,|u\left(y\right)|^p}{\|x-y\|_2}\,dx\,dy.$$

Нехай $G\subseteq\mathbb{R}^n$ — обмежена область. Диференціальний вираз

$$Au=-\sum_{i=1}^{n}\tfrac{\partial}{\partial x_i}\left(a_i\left(x,\left|\tfrac{\partial u}{\partial x_i}\right|^{p-1}\right)\left|\tfrac{\partial u}{\partial x_i}\right|^{p-2}\tfrac{\partial u}{\partial x_i}\right)+a_0\left(x,|u|^{p-1}\right)|u|^{p-2}\,u,$$

де $p>1$, функції $a_i\left(x,s\right)$, $i=0,1,...,n$, вимірні по x для всіх $s\in[0,+\infty)$ та неперервні по s для майже всіх $x\in G$, $|a_i\left(x,s\right)|\leqslant M$ для всіх $s\in[0,+\infty)$ та майже всіх $x\in G$, задає монотонний оператор, що діє з простору Соболєва $W^1_{0,p}\left(G\right)$ в $\left(W^1_{0,p}\left(G\right)\right)^*$.

Розглянемо варіаційну нерівність:

$$x \in C: \ \langle Ax, y - x \rangle \geqslant 0 \quad \forall y \in C, \tag{4.21}$$

де C — непорожня підмножина 2-рівномірно опуклого та рівномірно гладкого банахового простору E, A — оператор, що діє з E в E^*. Множину розв'язків (4.21) позначимо S.

Припустимо, що виконані такі умови:

- множина $C \subseteq E$ – опукла та замкнена;
- оператор $A : E \to E^*$ – монотонний на C, тобто

  $$\langle Ax - Ay, x - y \rangle \geqslant 0 \quad \forall x, y \in C,$$

  та ліпшицевий на C (з константою $L > 0$), тобто

  $$\|Ax - Ay\|_* \leqslant L \|x - y\| \quad \forall x, y \in C;$$

- множина S непорожня.

Розглянемо дуальну варіаційну нерівність:

$$x \in C: \ \langle Ay, x - y \rangle \leqslant 0 \quad \forall y \in C. \tag{4.22}$$

Множину розв'язків задачі (4.22) позначимо S^d. Зауважимо, що множина S^d опукла та замкнена [31, 122]. Нерівність (4.22) називають слабким або дуальним формулюванням (4.21) (або нерівністю типу Мінті), а розв'язки нерівності (4.22) – слабкими розв'язками (4.21). Для монотонних операторів A завжди маємо $S \subseteq S^d$. В наших умовах маємо $S^d = S$ [31].

Багатозначний оператор $J : E \to 2^{E^*}$, що діє за правилом

$$Jx = \left\{ x^* \in E^* : \ \ \langle x^*, x \rangle = \|x\|^2 = \|x^*\|_*^2 \right\},$$

називають нормалізованим дуальным відображенням [186]. Відомо, що [186]:

- якщо простір E гладкий, то відображення J однозначне;

- якщо простір E строго опуклий, то відображення J ін'єктивне та строго монотонне;
- якщо простір E рефлексивний, то відображення J сюр'єктивне;
- якщо простір E рівномірно гладкий, то відображення J рівномірно неперервне на обмежених підмножинах E.

Нехай E — гладкий банаховий простір. Розглянемо введений Я. І. Альбером [122, 123] функціонал

$$\phi(x, y) = \|x\|^2 - 2\langle Jy, x\rangle + \|y\|^2 \quad \forall x, y \in E.$$

З означення ϕ випливає корисна 3-точкова тотожність:

$$\phi(x, y) - \phi(x, z) - \phi(z, y) = 2\langle Jz - Jy, x - z\rangle \quad \forall x, y, z \in E.$$

Якщо простір E строго опуклий, то для $x, y \in E$ маємо

$$\phi(x, y) = 0 \Leftrightarrow x = y.$$

Нехай E — рівномірно опуклий та рівномірно гладкий банаховий простір, (x_n), (y_n) — обмежені послідовності елементів E. Тоді має місце [123]

$$\|x_n - y_n\| \to 0 \quad \Leftrightarrow \quad \|Jx_n - Jy_n\|_* \to 0 \quad \Leftrightarrow \quad \phi(x_n, y_n) \to 0. \tag{4.23}$$

Лема 4.1 (Aoyama K., Kohsaka F., [187]). *Нехай E — 2-рівномірно опуклий та гладкий банаховий простір. Тоді для деякого $\mu \geqslant 1$ виконується нерівність*

$$\phi(x, y) \geqslant \frac{1}{\mu}\|x - y\|^2 \quad \forall x, y \in E. \tag{4.24}$$

Для банахових просторів ℓ_p, L_p та W_p^m $(1 < p \leqslant 2)$ маємо $\mu = \frac{1}{p-1}$ [188]. Для гільбертового простору нерівність (4.24) перетворюється на тотожність з $\mu = 1$.

Використовуючи функціонал ϕ можна визначити новий оператор проєктування. Нехай K — непорожня замкнена та опукла підмножина рефлексивного, строго опуклого та гладкого простору E. Відомо [122, 123], що для кожного $x \in E$ існує єдиний елемент $z \in K$ такий, що

$$\phi(z, x) = \inf_{y \in K} \phi(y, x).$$

Цей елемент z позначають $\Pi_K x$, а відповідний оператор

$$\Pi_K : E \to K$$

називають узагальненою проєкцією E на K (узагальненою проєкцією Альбера) [122, 123].

Зауваження 4.8. Якщо E — гільбертовий простір, то Π_K співпадає з метричною проєкцією на множину K.

Лема 4.2 (Alber Y. I., [123]). *Нехай K — замкнена та опукла підмножина рефлексивного, строго опуклого та гладкого простору E, $x \in E$, $z \in K$. Тоді*

$$z = \Pi_K x \quad \Leftrightarrow \quad \langle Jz - Jx, y - z \rangle \geqslant 0 \quad \forall y \in K.$$

Нерівність леми 4.2 рівносильна такій [123]:

$$\phi(y, \Pi_K x) + \phi(\Pi_K x, x) \leqslant \phi(y, x) \quad \forall y \in K.$$

Варіаційну нерівність (4.21) можна сформулювати як задачу пошуку нерухомої точки [123]:

$$x = \Pi_C J^{-1}(Jx - \lambda Ax), \tag{4.25}$$

де $\lambda > 0$. Формулювання (4.25) веде до ітераційної схеми

$$x_{n+1} = \Pi_C J^{-1}(Jx_n - \lambda Ax_n), \tag{4.26}$$

яка вивчалась в [82] для обернено сильно монотонних операторів $A : E \to E^*$. Але для ліпшицевих монотонних операторів схема (4.26) в загальному випадку не збігається.

Основним елементом алгоритмів є обчислення за даними точками $x \in E$ та $x^* \in E^*$ нової точки

$$x^+ = \Pi_K J^{-1} \left(Jx - x^* \right).$$

З леми 4.2 та 3-точкової тотожності випливає фундаментальна для аналізу алгоритмів нерівність

$$\phi\left(y, x^+\right) \leqslant \phi\left(y, x\right) - \phi\left(x^+, x\right) + 2\left\langle x^*, y - x^+ \right\rangle \quad \forall y \in K.$$

Для розв'язання варіаційної нерівності (4.21) пропонуємо наступний алгоритм.

Алгоритм 4.3 (Операторна екстраполяція). Обираємо $x_0 = x_1 \in E$, $\lambda_n > 0$. Покладаємо $n = 1$.

1. Обчислити

$$m_n = \lambda_{n-1} \left(Ax_n - Ax_{n-1} \right),$$
$$x_{n+1} = \Pi_C J^{-1} \left(Jx_n - \lambda_n Ax_n - m_n \right).$$

2. Якщо $x_{n+1} = x_n = x_{n-1}$, то СТОП, інакше покласти $n := n + 1$ та перейти до **1**.

Зауваження 4.9. Алгоритм 4.3 є модификацією методу з попереднього розділу для задач в банахових просторах, що використовує узагальнену проєкцію Альбера [123] замість метричної.

Правило зупинки обґрунтовується тотожністю (4.25), що рівносильна варіаційній нерівності (4.21). Дійсно, при виконанні $x_{n+1} = x_n = x_{n-1}$ маємо $x_n = \Pi_C J^{-1} \left(Jx_n - \lambda_n Ax_n \right)$, звідки $y_n \in S$.

У випадку $C = E$ варіаційна нерівність (4.21) набуває вигляду операторного рівняння

$$\text{знайти} \quad x \in E : \quad Ax = 0 \tag{4.27}$$

Для (4.27) алгоритм 4.3 дає такий ітераційний процес

$$Jx_{n+1} = Jx_n - \lambda_n Ax_n - \lambda_{n-1}\left(Ax_n - Ax_{n-1}\right), \tag{4.28}$$

який збіжний лише за умови монотонності оператора A.

Метод

$$Jx_{n+1} = Jx_n - \lambda_n Ax_n$$

у цьому випадку може слабко збігатись лише в ергодичному розумінні [123]. А за об'ємом необхідних для здійснення ітераційного кроку обчислень процес (4.28) має переваги над екстаградієнтним методом

$$\begin{cases} Jx_{n+\frac{1}{2}} = Jx_n - \lambda_n Ax_n, \\ Jx_{n+1} = Jx_n - \lambda_n Ax_{n+\frac{1}{2}}, \end{cases}$$

та неявним методом, що використовує резольвенту

$$x_{n+1} = \left(J + \lambda_n A\right)^{-1} Jx_n.$$

Схему (4.28) можна подати у вигляді двоетапного процесу

$$Jx_n = Jy_n - \lambda_{n-1} Ax_{n-1},$$

$$Jy_{n+1} = Jy_n - \lambda_{n-1} Ax_n.$$

4.2.2. Сублінійна оцінка ефективності та слабка збіжність. У випадку обмеженості множини C доведемо, що алгоритму 4.3 необхідно зробити $O\left(\frac{LD^2}{\varepsilon}\right)$ ітерацій для отримання точки $x \in C$ з

$$\operatorname{gap}(x) = \sup_{y \in C} \langle Ay, x - y \rangle \leqslant \varepsilon,$$

де $\varepsilon > 0$, $D^2 = \sup_{a,b \in C} \phi(a,b) < +\infty$.

Теорема 4.3. *Нехай (x_n) — послідовність, що породжена алгоритмом 4.3. Припустимо, що $\lambda_n \in \left(0, \frac{1}{2\mu L}\right]$. Тоді для послідовності чезарівських* *середніх* $z_{N+1}=\frac{\sum_{n=1}^{N}\lambda_n x_{n+1}}{\sum_{n=1}^{N}\lambda_n}$ *має місце нерівність*

$$\operatorname{gap}\left(z_{N+1}\right)\leqslant\frac{\sup_{y\in C}\phi\left(y,x_1\right)}{2\sum_{n=1}^{N}\lambda_n}.$$

Теорема 4.4. *Нехай* (x_n) *— послідовність, що породжена алгоритмом 4.3 з* $\lambda_n=\frac{1}{2\mu L}$*, тобто,*

$$x_{n+1}=\Pi_C J^{-1}\left(Jx_n-\tfrac{1}{2\mu L}\left(2Ax_n-Ax_{n-1}\right)\right).$$

Тоді має місце оцінка

$$\operatorname{gap}\left(z_{N+1}\right)\leqslant\frac{L\mu\sup_{y\in C}\phi\left(y,x_1\right)}{N},$$

де $z_{N+1}=\frac{1}{N}\sum_{n=1}^{N}x_{n+1}$.

Доведення теореми 4.3. Для послідовності (x_n) має місце нерівність

$$-2\left\langle\lambda_n Ax_n+m_n,y-x_{n+1}\right\rangle\leqslant \\ \leqslant\phi\left(y,x_n\right)-\phi\left(x_{n+1},x_n\right)-\phi\left(y,x_{n+1}\right)\quad\forall y\in C. \tag{4.29}$$

Перепишемо (4.29) таким чином

$$\phi\left(y,x_n\right)-\phi\left(y,x_{n+1}\right)\geqslant \\ \geqslant 2\lambda_n\left\langle Ax_{n+1},x_{n+1}-y\right\rangle-2\lambda_n\left\langle Ax_{n+1}-Ax_n,x_{n+1}-y\right\rangle+ \\ +2\lambda_{n-1}\left\langle Ax_n-Ax_{n-1},x_n-y\right\rangle+2\lambda_{n-1}\left\langle Ax_n-Ax_{n-1},x_{n+1}-x_n\right\rangle+ \\ +\phi\left(x_{n+1},x_n\right). \tag{4.30}$$

Сумуючи (4.30) по n від 1 до N, отримуємо

$$\phi\left(y,x_1\right)-\phi\left(y,x_{N+1}\right)\geqslant \\ \geqslant 2\sum_{n=1}^{N}\lambda_n\left\langle Ax_{n+1},x_{n+1}-y\right\rangle-2\lambda_N\left\langle Ax_{N+1}-Ax_N,x_{N+1}-y\right\rangle+ \\ +\sum_{n=1}^{N}\left(2\lambda_{n-1}\left\langle Ax_n-Ax_{n-1},x_{n+1}-x_n\right\rangle+\phi\left(x_{n+1},x_n\right)\right). \tag{4.31}$$

Ліпшицевість оператора A та нерівність (4.24) дає

$$\sum_{n=1}^{N}\left(2\lambda_{n-1}\left\langle Ax_n - Ax_{n-1}, x_{n+1}-x_n\right\rangle + \phi\left(x_{n+1}, x_n\right)\right) \geqslant$$
$$\geqslant \sum_{n=1}^{N}\left(-2\lambda_{n-1}L\left\|x_n - x_{n-1}\right\|\left\|x_{n+1}-x_n\right\| + \frac{1}{\mu}\left\|x_{n+1}-x_n\right\|^2\right) =$$
$$= \sum_{n=1}^{N}\left(-2\lambda_{n-1}L\left\|x_n - x_{n-1}\right\|\left\|x_{n+1}-x_n\right\| + \frac{1}{2\mu}\left\|x_{n+1}-x_n\right\|^2 +\right.$$
$$\left.+\frac{1}{2\mu}\left\|x_n - x_{n-1}\right\|^2\right) + \frac{1}{2\mu}\left\|x_{N+1}-x_N\right\|^2 \geqslant \frac{1}{2\mu}\left\|x_{N+1}-x_N\right\|^2.$$

Використовуючи останню оцінку в (4.31), отримуємо

$$\phi\left(y, x_1\right) - \phi\left(y, x_{N+1}\right) \geqslant$$
$$\geqslant 2\sum_{n=1}^{N}\lambda_n\left\langle Ax_{n+1}, x_{n+1}-y\right\rangle - 2\lambda_N\left\langle Ax_{N+1}-Ax_N, x_{N+1}-y\right\rangle +$$
$$+\frac{1}{2\mu}\left\|x_{N+1}-x_N\right\|^2 \geqslant$$
$$\geqslant 2\sum_{n=1}^{N}\lambda_n\left\langle Ax_{n+1}, x_{n+1}-y\right\rangle - 2\lambda_N L\left\|x_{N+1}-x_N\right\|\left\|x_{N+1}-y\right\| +$$
$$+\frac{1}{2\mu}\left\|x_{N+1}-x_N\right\|^2 \geqslant$$
$$\geqslant 2\sum_{n=1}^{N}\lambda_n\left\langle Ax_{n+1}, x_{n+1}-y\right\rangle - \lambda_N L\left\|x_{N+1}-y\right\|^2.$$

Приходимо до нерівності

$$2\sum_{n=1}^{N}\lambda_n\left\langle Ax_{n+1}, x_{n+1}-y\right\rangle - \lambda_N L\left\|x_{N+1}-y\right\|^2 +$$
$$+\phi\left(y, x_{N+1}\right) \leqslant \phi\left(y, x_1\right) \quad \forall y \in C. \quad (4.32)$$

Використовуючи монотонність оператора A, отримуємо

$$\sum_{n=1}^{N}\lambda_n\left\langle Ax_{n+1}, x_{n+1}-y\right\rangle \geqslant \sum_{n=1}^{N}\lambda_n\left\langle Ay, x_{n+1}-y\right\rangle =$$

$$= \left(\sum_{n=1}^{N} \lambda_n \right) \langle Ay, z_{N+1} - y \rangle, \quad (4.33)$$

де $z_{N+1} = \frac{\sum_{n=1}^{N} \lambda_n x_{n+1}}{\sum_{n=1}^{N} \lambda_n}$. Враховуючи оцінку (4.33) в (4.32), приходимо до нерівності

$$2 \left(\sum_{n=1}^{N} \lambda_n \right) \langle Ay, z_{N+1} - y \rangle + \left(\frac{1}{\mu} - \lambda_N L \right) \|x_{N+1} - y\|^2 \leqslant \phi(y, x_1) \quad \forall y \in C,$$

звідки випливає

$$\operatorname{gap}(z_{N+1}) = \sup_{y \in C} \langle Ay, z_{N+1} - y \rangle \leqslant \frac{\sup_{y \in C} \phi(y, x_1)}{2 \sum_{n=1}^{N} \lambda_n},$$

що і потрібно було довести. □

Перейдемо до доведення слабкої збіжності алгоритму 4.3.

В якості функції Ляпунова оберемо

$$V_n = \phi(z, x_n) - 2 \langle m_n, x_n - z \rangle + \mu \lambda_{n-1} L \phi(x_n, x_{n-1}),$$

де $z \in S$.

Має місце

Лема 4.3. *Для послідовності (x_n), що породжена алгоритмом 4.3, виконується нерівність*

$$V_{n+1} \leqslant V_n - (1 - \mu \lambda_{n-1} L - \mu \lambda_n L) \phi(x_{n+1}, x_n),$$

де $z \in S$.

Доведення. Нехай $z \in S$. Маємо

$$\phi(z, x_{n+1}) \leqslant \phi(z, x_n) - \phi(x_{n+1}, x_n) + 2 \langle \lambda_n A x_n + m_n, z - x_{n+1} \rangle. \quad (4.34)$$

Монотонність оператора A дає

$$\langle \lambda_n A x_n + m_n, z - x_{n+1} \rangle = \langle \lambda_n A x_n + \lambda_{n-1} (A x_n - A x_{n-1}), z - x_{n+1} \rangle =$$

$$= \lambda_n \langle Ax_n - Ax_{n+1}, z - x_{n+1} \rangle + \lambda_{n-1} \langle Ax_n - Ax_{n-1}, z - x_{n+1} \rangle +$$
$$+ \underbrace{\lambda_n \langle Ax_{n+1}, z - x_{n+1} \rangle}_{\leqslant 0} \leqslant \lambda_n \langle Ax_n - Ax_{n+1}, z - x_{n+1} \rangle +$$
$$+ \lambda_{n-1} \langle Ax_n - Ax_{n-1}, z - x_n \rangle + \lambda_{n-1} \langle Ax_n - Ax_{n-1}, x_n - x_{n+1} \rangle . \quad (4.35)$$

Використавши (4.35) в (4.34), отримаємо

$$\phi(z, x_{n+1}) \leqslant \phi(z, x_n) - \phi(x_{n+1}, x_n) + 2\lambda_n \langle Ax_n - Ax_{n+1}, z - x_{n+1} \rangle +$$
$$+ 2\lambda_{n-1} \langle Ax_n - Ax_{n-1}, z - x_n \rangle + 2\lambda_{n-1} \langle Ax_n - Ax_{n-1}, x_n - x_{n+1} \rangle . \quad (4.36)$$

Оцінимо зверху доданок $2\lambda_{n-1} \langle Ax_n - Ax_{n-1}, x_n - x_{n+1} \rangle$ в (4.36). Маємо

$$2\lambda_{n-1} \langle Ax_n - Ax_{n-1}, x_n - x_{n+1} \rangle \leqslant$$
$$\leqslant 2\lambda_{n-1} \|Ax_n - Ax_{n-1}\|_* \|x_n - x_{n+1}\| \leqslant 2\lambda_{n-1} L \|x_n - x_{n-1}\| \|x_{n+1} - x_n\| \leqslant$$
$$\leqslant \lambda_{n-1} L \|x_n - x_{n-1}\|^2 + \lambda_{n-1} L \|x_n - x_{n+1}\|^2 \leqslant \mu\lambda_{n-1} L\phi(x_n, x_{n-1}) +$$
$$+ \mu\lambda_{n-1} L\phi(x_{n+1}, x_n) .$$

Приходимо до нерівності

$$\phi(z, x_{n+1}) + 2\lambda_n \langle Ax_n - Ax_{n+1}, x_{n+1} - z \rangle + \mu\lambda_n L\phi(x_{n+1}, x_n) \leqslant$$
$$\leqslant \phi(z, x_n) + 2\lambda_{n-1} \langle Ax_{n-1} - Ax_n, x_n - z \rangle + \mu\lambda_{n-1} L\phi(x_n, x_{n-1}) -$$
$$- (1 - \mu\lambda_{n-1} L - \mu\lambda_n L) \phi(x_{n+1}, x_n) ,$$

що й потрібно було довести. □

Теорема 4.5. *Нехай C — непорожня опукла та замкнена підмножина 2-рівномірно опуклого та рівномірно гладкого банахового простору E, $A : E \to E^*$ — монотонний та ліпшицевий на множині C оператор, $S \neq \varnothing$. Припустимо, що нормалізоване дуальне відображення J секвенційно слабко неперервне та послідовність (λ_n) така, що*

$$0 < \inf_n \lambda_n \leqslant \sup_n \lambda_n < \frac{1}{2\mu L}.$$

Тоді послідовність (x_n), що породжена алгоритмом 4.3, слабко збігається до деякої точки $z \in S$.

Доведення. Нехай $z' \in S$. Беремо таке $\delta > 0$, що $1 - \mu\lambda_{n-1}L - \mu\lambda_n L \geqslant \delta$ для всіх $n \in N$. З нерівності леми 4.3 випливає

$$V_{n+1} \leqslant V_n - \delta\phi\left(x_{n+1}, x_n\right),$$

де $V_n = \phi\left(z', x_n\right) + 2\left\langle m_n, x_n - z'\right\rangle + \mu\lambda_{n-1}L\phi\left(x_n, x_{n-1}\right)$.

Покажемо, що $V_n \geqslant 0$ для всіх $n \in \mathbb{N}$. Маємо

$$\begin{gathered}
V_n = \phi\left(z', x_n\right) + 2\lambda_{n-1}\left\langle Ax_{n-1} - Ax_n, x_n - z'\right\rangle + \mu\lambda_{n-1}L\phi\left(x_n, x_{n-1}\right) \geqslant \\
\geqslant \frac{1}{\mu}\left\|x_n - z'\right\|^2 - 2\lambda_{n-1}\left\|Ax_{n-1} - Ax_n\right\|_* \left\|x_n - z'\right\| + \\
+ \lambda_{n-1}L\left\|x_{n-1} - x_n\right\|^2 \geqslant \\
\geqslant \frac{1}{\mu}\left\|x_n - z'\right\|^2 - 2\lambda_{n-1}L\left\|x_n - x_{n-1}\right\|\left\|x_n - z'\right\| + \lambda_{n-1}L\left\|x_{n-1} - x_n\right\|^2 \geqslant \\
\geqslant \left(\frac{1}{\mu} - \lambda_{n-1}L\right)\left\|x_n - z'\right\|^2 \geqslant 0.
\end{gathered}$$

З леми 2.2 випливає існування границі

$$\lim_{n\to\infty}\left(\phi\left(z', x_n\right) + 2\lambda_{n-1}\left\langle Ax_{n-1} - Ax_n, x_n - z'\right\rangle + \mu\lambda_{n-1}L\phi\left(x_n, x_{n-1}\right)\right)$$

та

$$\sum_{n=1}^{\infty}\phi\left(x_{n+1}, x_n\right) < +\infty.$$

Звідки отримуємо обмеженість послідовності (x_n) та

$$\lim_{n\to\infty}\phi\left(x_{n+1}, x_n\right) = \lim_{n\to\infty}\left\|x_{n+1} - x_n\right\| = 0.$$

Оскільки

$$\lim_{n\to\infty}\left(2\lambda_{n-1}\left\langle Ax_{n-1} - Ax_n, x_n - z'\right\rangle + \mu\lambda_{n-1}L\phi\left(x_n, x_{n-1}\right)\right) = 0,$$

то послідовності $\left(\phi\left(z', x_n\right)\right)$ збігаються для всіх $z' \in S$.

Покажемо, що всі слабкі часткові границі послідовності (x_n) належать множині S. Розглянемо підпослідовність (x_{n_k}), що слабко збігається до деякої точки $z \in E$. Ясно, що $z \in C$. Покажемо, що $z \in S$. Маємо

$$\langle Jx_{n+1} - Jx_n + \lambda_n Ax_n + \lambda_{n-1}(Ax_n - Ax_{n-1}), y - x_{n+1} \rangle \geqslant 0 \quad \forall y \in C.$$

Звідки, використовуючи монотонність оператора A, виводимо оцінку

$$\begin{aligned} &\langle Ay, y - x_n \rangle + \langle Ax_n, x_n - x_{n+1} \rangle \geqslant \langle Ax_n, y - x_{n+1} \rangle \geqslant \\ &\geqslant \frac{1}{\lambda_n} \langle Jx_n - Jx_{n+1}, y - x_{n+1} \rangle - \frac{\lambda_{n-1}}{\lambda_n} \langle Ax_n - Ax_{n-1}, y - x_{n+1} \rangle \quad \forall y \in C. \end{aligned}$$

З $\lim_{n\to\infty} \|x_n - x_{n-1}\| = 0$ та ліпшицевості оператора A випливає

$$\lim_{n\to\infty} \|Ax_n - Ax_{n-1}\|_* = 0.$$

Завдяки рівномірній неперервності на обмежених множинах нормалізованого дуального відображення J отримуємо $\lim_{n\to\infty} \|Jx_n - Jx_{n+1}\|_* = 0$. Таким чином,

$$\liminf_{n\to\infty} \langle Ay, y - x_n \rangle \geqslant 0 \quad \forall y \in C.$$

З іншого боку

$$\langle Ay, y - z \rangle = \lim_{k\to\infty} \langle Ay, y - x_{n_k} \rangle \geqslant \liminf_{n\to\infty} \langle Ay, y - x_n \rangle \geqslant 0 \quad \forall y \in C.$$

Отже, $z \in S$.

Покажемо, що послідовність (x_n) слабко збігається до z. Міркуємо від супротивного. Нехай існує підпослідовність (x_{m_k}) така, що $x_{m_k} \rightharpoonup z'$ та $z \neq z'$. Ясно, що $z' \in S$. Маємо

$$2\langle Jx_n, z - z' \rangle = \phi(z', x_n) - \phi(z, x_n) + \|z\|^2 - \|z'\|^2.$$

Звідки випливає існування границі $\lim_{n\to\infty} \langle Jx_n, z - z' \rangle$. Завдяки секвенційній слабкій неперервності нормалізованого дуального відображення J отримаємо

$$\langle Jz, z - z' \rangle = \lim_{k\to\infty} \langle Jx_{n_k}, z - z' \rangle = \lim_{k\to\infty} \langle Jx_{m_k}, z - z' \rangle = \langle Jz', z - z' \rangle,$$

тобто, $\langle Jz - Jz', z - z' \rangle = 0$. Звідки випливає $z = z'$, що і необхідно було довести. $\square$

Параметри λ_n алгоритму 4.3 задавалась виходячи з умови

$$0 < \inf_n \lambda_n \leqslant \sup_n \lambda_n < \frac{1}{2\mu L}.$$

Тобто, використовувалась інформація про константи ліпшицевості оператора A.

Відштовхуючись від алгоритму 4.3 та робіт [95,109,153,168] у статті [131] побудовано алгоритм з адаптивним вибором величини λ_n, що не вимагає знання ліпшицевих констант та процедур типу лінійного пошуку. У даній роботі доведемо збіжність даного алгоритму за допомогою ляпуновського аналізу.

Припустимо, що відома лише константа $\mu \geqslant 1$ з нерівності (4.24).

Алгоритм 4.4 (Адаптивна версія). Обираємо $x_0 = x_1 \in E$, $\tau \in \left(0, \frac{1}{2\mu}\right)$ та число $\lambda_1 = \lambda_0 > 0$. Покладаємо $n = 1$.

1. Обчислити

$$m_n = \lambda_{n-1} \left(Ax_n - Ax_{n-1} \right),$$

$$x_{n+1} = \Pi_C J^{-1} \left(Jx_n - \lambda_n Ax_n - m_n \right).$$

2. Якщо $x_{n+1} = x_n = x_{n-1}$, то СТОП, інакше перейти до **3**.

3. Обчислити

$$\lambda_{n+1} = \begin{cases} \min\left\{ \lambda_n, \tau \frac{\|x_{n+1} - x_n\|}{\|Ax_{n+1} - Ax_n\|_*} \right\}, & \text{якщо} \quad Ax_{n+1} \neq Ax_n, \\ \lambda_n, & \text{інакше.} \end{cases}$$

Покласти $n := n + 1$ та перейти до **1**.

Зауваження 4.10. Для операторного рівняння (4.27) алгоритм 4.4 дає такий ітераційний процес

$$\begin{cases} Jx_{n+1} = Jx_n - \lambda_n Ax_n - \lambda_{n-1} \left(Ax_n - Ax_{n-1} \right), \\ \lambda_{n+1} = \begin{cases} \min\left\{ \lambda_n, \tau \frac{\|x_{n+1} - x_n\|}{\|Ax_{n+1} - Ax_n\|_*} \right\}, & \text{якщо} \ Ax_{n+1} \neq Ax_n, \\ \lambda_n, & \text{інакше.} \end{cases} \end{cases}$$

Послідовність (λ_n), що задається на третьому етапі ітераційного кроку в алгоритмі 4.4, незростаюча та обмежена знизу числом $\min\left\{\lambda_1, \frac{\tau}{L}\right\}$. Отже, існує $\lim\limits_{n\to\infty}\lambda_n > 0$.

В якості функції Ляпунова оберемо

$$W_n = \phi(z, x_n) - 2\langle m_n, x_n - z\rangle + \tau\mu\frac{\lambda_{n-1}}{\lambda_n}L\phi(x_n, x_{n-1}),$$

де $z \in S$.

Лема 4.4. *Для послідовності (x_n), що породжена алгоритмом 4.4, виконується нерівність*

$$W_{n+1} \leqslant W_n - \left(1 - \tau\mu\frac{\lambda_{n-1}}{\lambda_n} - \tau\mu\frac{\lambda_n}{\lambda_{n+1}}\right)\phi(x_{n+1}, x_n).$$

Доведення. Нехай $z \in S$. Як і в доведенні леми 4.3 приходимо до нерівності

$$\begin{aligned}\phi(z, x_{n+1}) \leqslant \phi(z, x_n) - \phi(x_{n+1}, x_n) + 2\lambda_n\langle Ax_n - Ax_{n+1}, z - x_{n+1}\rangle + \\ + 2\lambda_{n-1}\langle Ax_n - Ax_{n-1}, z - x_n\rangle + 2\lambda_{n-1}\langle Ax_n - Ax_{n-1}, x_n - x_{n+1}\rangle. \quad (4.37)\end{aligned}$$

Використовуючи правило обчислення λ_{n+1}, оцінимо зверху доданок

$$2\lambda_{n-1}\langle Ax_n - Ax_{n-1}, x_n - x_{n+1}\rangle$$

в нерівності (4.37). Маємо

$$\begin{aligned}2\lambda_{n-1}\langle Ax_n - Ax_{n-1}, x_n - x_{n+1}\rangle \leqslant 2\lambda_{n-1}\|Ax_n - Ax_{n-1}\|_*\|x_n - x_{n+1}\| \leqslant \\ \leqslant 2\tau\frac{\lambda_{n-1}}{\lambda_n}\|x_n - x_{n-1}\|\|x_{n+1} - x_n\| \leqslant \\ \leqslant \tau\frac{\lambda_{n-1}}{\lambda_n}\|x_n - x_{n-1}\|^2 + \tau\frac{\lambda_{n-1}}{\lambda_n}\|x_n - x_{n+1}\|^2 \leqslant \\ \leqslant \tau\mu\frac{\lambda_{n-1}}{\lambda_n}\phi(x_n, x_{n-1}) + \tau\mu\frac{\lambda_{n-1}}{\lambda_n}\phi(x_{n+1}, x_n).\end{aligned}$$

Приходимо до нерівності

$$\phi(z,x_{n+1})+2\lambda_n\langle Ax_n-Ax_{n+1},x_{n+1}-z\rangle+\tau\mu\frac{\lambda_n}{\lambda_{n+1}}\phi(x_{n+1},x_n)\leqslant$$
$$\leqslant\phi(z,x_n)+2\lambda_{n-1}\langle Ax_{n-1}-Ax_n,x_n-z\rangle+\tau\mu\frac{\lambda_{n-1}}{\lambda_n}\phi(x_n,x_{n-1})-$$
$$-\left(1-\tau\mu\frac{\lambda_{n-1}}{\lambda_n}-\tau\mu\frac{\lambda_n}{\lambda_{n+1}}\right)\phi(x_{n+1},x_n),$$

що і потрібно було довести. □

Має місце

Теорема 4.6. *Нехай C — непорожня опукла та замкнена підмножина 2-рівномірно опуклого та рівномірно гладкого банахового простору E, $A : E\to E^*$ — монотонний та ліпшицевий оператор, $S\neq\varnothing$. Припустимо, що нормалізоване дуальне відображення J секвенційно слабко неперервне. Тоді послідовність (x_n), що породжена алгоритмом 4.4, слабко збігається до деякої точки $z\in S$.*

Доведення. Нехай $z'\in S$. Для функції Ляпунова

$$W_n=\phi(z',x_n)+2\lambda_{n-1}\langle Ax_{n-1}-Ax_n,x_n-z'\rangle+\tau\mu\frac{\lambda_{n-1}}{\lambda_n}\phi(x_n,x_{n-1}),$$

має місце нерівність леми 4.4

$$W_{n+1}\leqslant W_n-\left(1-\tau\mu\frac{\lambda_{n-1}}{\lambda_n}-\tau\mu\frac{\lambda_n}{\lambda_{n+1}}\right)\phi(x_{n+1},x_n).$$

Оскільки існує границя $\lim\limits_{n\to\infty}\lambda_n>0$, то

$$1-\tau\mu\frac{\lambda_{n-1}}{\lambda_n}-\tau\mu\frac{\lambda_n}{\lambda_{n+1}}\to 1-2\tau\mu\in(0,1)\quad\text{при}\quad n\to\infty.$$

Покажемо, що $W_n\geqslant 0$ для всіх достатньо великих $n\in\mathbb{N}$. Маємо

$$W_n=\phi(z',x_n)+2\lambda_{n-1}\langle Ax_{n-1}-Ax_n,x_n-z'\rangle+\tau\mu\frac{\lambda_{n-1}}{\lambda_n}\phi(x_n,x_{n-1})\geqslant$$
$$\geqslant\frac{1}{\mu}\|x_n-z'\|^2-2\lambda_{n-1}\|Ax_{n-1}-Ax_n\|_*\|x_n-z'\|+$$
$$+\tau\frac{\lambda_{n-1}}{\lambda_n}\|x_{n-1}-x_n\|^2\geqslant$$

$$\geqslant \frac{1}{\mu}\left\|x_n - z'\right\|^2 - 2\tau\frac{\lambda_{n-1}}{\lambda_n}\left\|x_n - x_{n-1}\right\|\left\|x_n - z'\right\| + \tau\frac{\lambda_{n-1}}{\lambda_n}\left\|x_{n-1} - x_n\right\|^2 \geqslant$$

$$\geqslant \left(\frac{1}{\mu} - \tau\frac{\lambda_{n-1}}{\lambda_n}\right)\left\|x_n - z'\right\|^2.$$

Оскільки існує таке $n_0 \in \mathbb{N}$, що $\frac{1}{\mu} - \tau\frac{\lambda_{n-1}}{\lambda_n} > 0$ для всіх $n \geqslant n_0$, то $W_n \geqslant 0$ починаючи з n_0.

Тепер з леми 2.2 можемо зробити висновок, що існує границя

$$\lim_{n\to\infty}\left(\phi\left(z', x_n\right) + 2\lambda_{n-1}\left\langle Ax_{n-1} - Ax_n, x_n - z'\right\rangle + \tau\mu\frac{\lambda_{n-1}}{\lambda_n}\phi\left(x_n, x_{n-1}\right)\right)$$

та

$$\sum_{n=1}^{\infty}\left(1 - \tau\mu\frac{\lambda_{n-1}}{\lambda_n} - \tau\mu\frac{\lambda_n}{\lambda_{n+1}}\right)\phi\left(x_{n+1}, x_n\right) < +\infty.$$

Звідки отримуємо обмеженість послідовності (x_n) та

$$\lim_{n\to\infty}\phi\left(x_{n+1}, x_n\right) = \lim_{n\to\infty}\left\|x_{n+1} - x_n\right\| = 0.$$

Оскільки

$$\lim_{n\to\infty}\left(2\lambda_{n-1}\left\langle Ax_{n-1} - Ax_n, x_n - z'\right\rangle + \tau\mu\frac{\lambda_{n-1}}{\lambda_n}\phi\left(x_n, x_{n-1}\right)\right) = 0,$$

то збігаються послідовності $(\phi\left(z', x_n\right))$ для всіх $z' \in S$.

Далі, міркуваннями доведення теореми 4.5, прийдемо до результату. □

Зауваження 4.11. У зв'язку з проведеним дослідженням вкажемо на два питання. По-перше, всі результати отримані для класу 2-рівномірно опуклих і рівномірно гладких банахових просторів, який не містить важливих для застосувань просторів L_p і W_p^m $(2 < p < +\infty)$. Дуже бажано позбавитися цього обмеження. По-друге, для можливості ефективного застосування алгоритмів для нелінійних задач у банахових просторах необхідні швидкі та стійкі алгоритми обчислення узагальненої проєкції Альбера для широкого набору множин.

4.3. Висновки до розділу 4

Досліджено варіанти алгоритмів екстраполяції з минулого та операторної екстраполяції з дивергенцією Брегмана для розв'язання варіаційних нерівностей з монотонними та ліпшицевими операторами, що діють в скінченновимірному дійсному лінійному просторі. Основні результати: $O\left(\frac{1}{\varepsilon}\right)$-оцінки ефективності в термінах функції зазору.

Для варіаційних нерівностей з монотонними, ліпшицевими операторами, що діють в 2-рівномірно опуклому та рівномірно гладкому банаховому просторі, доведено теореми про слабку збіжність двох варіантів методу операторної екстраполяції та $O\left(\frac{1}{\varepsilon}\right)$-оцінку ефективності в термінах функції зазору.

Основні результати даного розділу опубліковано в [1, 2, 5–9].

РОЗДІЛ 5
ЕКСПЕРИМЕНТИ

Нижче описано експерименти, проведені для перевірки теорем 2.2, 2.4, 3.2 та 3.4, що містять оцінки швидкості збіжності розглянутих алгоритмів. Розглянуто білінійні та квадратичні сідлові задачі.

Програмна реалізація здійснена на мові Python 3.8.5 з використанням NumPy 1.26. Експерименти проводились на пристрої з 64-розрядною операційною системою Windows 7 Ultimate, процесором Intel(R) Core(TM) i5-3230M, 2.6GHz та оперативною пам'яттю (RAM) 6 ГБ. Результати проведених експериментів підтверджують теоретичні оцінки.

Також проведено експерименти з навчанням простих GANs за допомогою двох рандомізованих варіантів методу з експоненційним усередненням. Для експериментів було створено Deep Convolutional Generative Adversarial Network (DCGAN). Для її побудови використано Keras Sequential API з навчальним циклом tf.GradientTape.

Навчання мереж проводилось за допомогою обчислювальних можливостей Google Colab з використанням GPU типу T4.

5.1. Білінійна сідлова задача

Розглянемо задачу:

$$\min_{x\in\Delta^n}\max_{y\in\Delta^m}\langle Kx, y\rangle \tag{5.1}$$

де K — дійсна матриця $m\times n$, $\Delta^n=\{x\in\mathbb{R}^n:\ \sum_{i=1}^n x_i=1,\ x_i\geqslant 0\}$, $\Delta^m=\{y\in\mathbb{R}^m:\ \sum_{i=1}^n y_i=1,\ y_i\geqslant 0\}$ — стандартні симплекси.

Функція зазору (2.9) для задачі (5.1) обчислюється так

$$\mathrm{gap}\,(x,y) = \max_i (Kx)_i - \min_j (K^*y)_j.$$

Для (5.1) варіанти алгоритмів з теорем 2.2 та 3.2 мають вигляд

$$\begin{cases} x_1 = p_0 \in \Delta^n, \quad y_1 = q_0 \in \Delta^m, \\ p_k = P_{\Delta^n}\left(x_k - \frac{1}{3\|K\|_s} K^* q_{k-1}\right), \\ q_k = P_{\Delta^m}\left(y_k + \frac{1}{3\|K\|_s} K p_{k-1}\right), \\ x_{k+1} = P_{\Delta^n}\left(x_k - \frac{1}{3\|K\|_s} K^* q_k\right), \\ y_{k+1} = P_{\Delta^m}\left(y_k + \frac{1}{3\|K\|_s} K p_k\right), \end{cases} \tag{5.2}$$

та

$$\begin{cases} x_1 = x_0 \in \Delta^n, \quad y_1 = y_0 \in \Delta^m, \\ x_{k+1} = P_{\Delta^n}\left(x_k - \frac{1}{2\|K\|_s}\left(2K^* y_k - K^* y_{k-1}\right)\right), \\ y_{k+1} = P_{\Delta^m}\left(y_k + \frac{1}{2\|K\|_s}\left(2K x_k - K x_{k-1}\right)\right), \end{cases} \tag{5.3}$$

відповідно. Тут $\|K\|_s = \sqrt{\lambda_{max}(K^*K)}$ — спектральна норма матриці K.

Елементи матриці K генерувались випадково. Використали рівномірний розподіл на $\{-5, -4, ..., 4, 5\}$.

В експериментах за стартову точку обирали

$$(x_1, y_1) = \left(\frac{1}{n}, ..., \frac{1}{n}, \frac{1}{m}, ..., \frac{1}{m}\right) \in \Delta^n \times \Delta^m.$$

Для проєктування на симплекс використовувався перший алгоритм роботи [189].

Для (5.2) динаміка зменшення значення $\mathrm{gap}\,(z_N)$, де

$$z_N = \left(\tfrac{1}{N}\sum_{k=1}^{N} p_k, \tfrac{1}{N}\sum_{k=1}^{N} q_k\right),$$

подана на Рис. 5.1–5.4.

Для (5.3) динаміка зменшення значення $\mathrm{gap}\,(z_N)$ де

$$z_N = \left(\tfrac{1}{N-1}\sum_{k=2}^{N} x_k, \tfrac{1}{N-1}\sum_{k=2}^{N} y_k\right),$$

подана на Рис. 5.5–5.8.

Наведені графіки демонструють відповідність результатів розрахунків теоретичним оцінкам.

В Таблиці 5.1 вказано за яких час алгоритмами досягалось виконання нерівності $\operatorname{gap}(z_N) < 0,01$.

Таблиця 5.1

Час, потрібний для $\operatorname{gap}(z_N) < 0,01$, секунди

Розмірність K	Алгоритм (5.2)	Алгоритм (5.3)
100×100	5.64	1.85
200×200	22.48	6.38
100×300	28.08	7.46
500×500	131.82	44.16

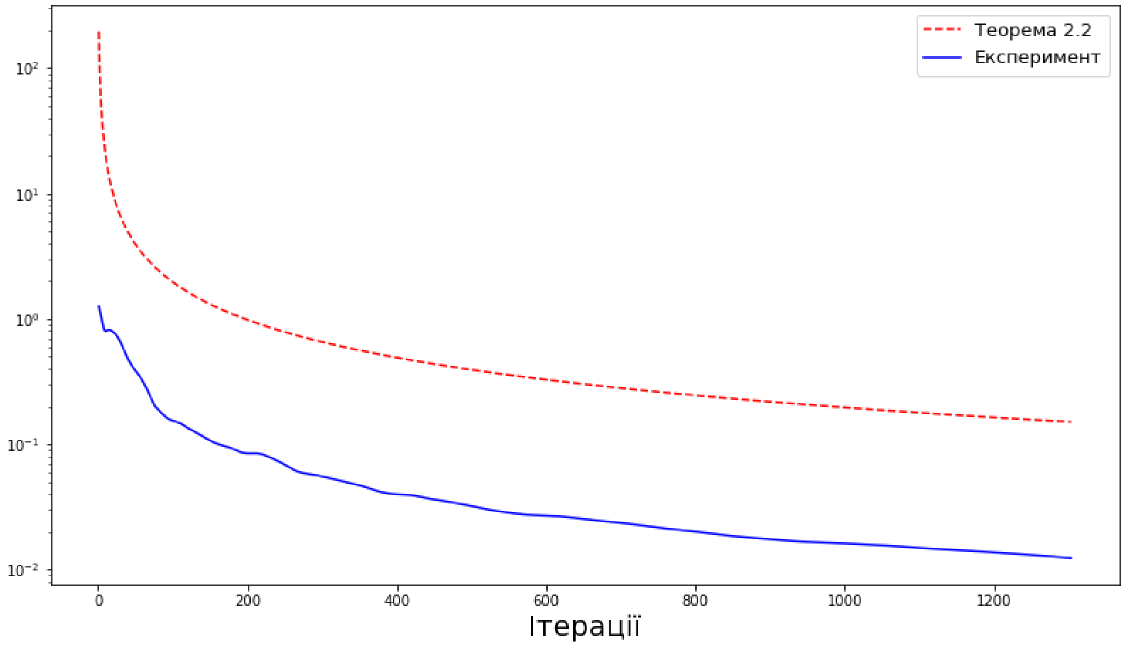


Рис. 5.1. Алгоритм (5.2), розмірність K — 100×100.

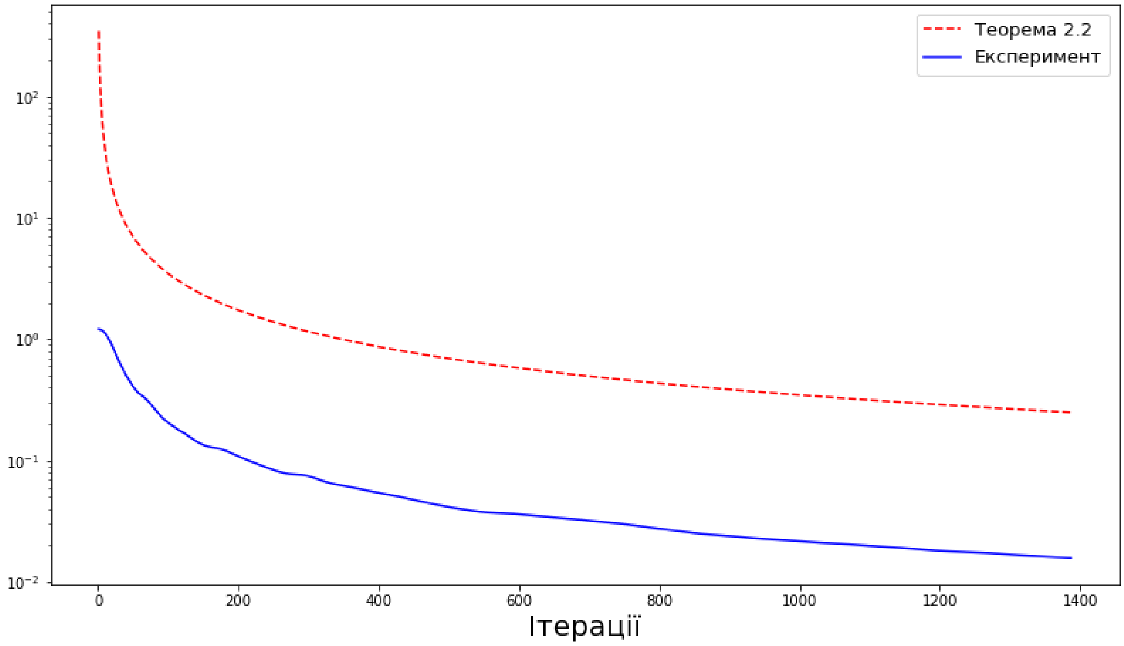


Рис. 5.2. Алгоритм (5.2), розмірність K — 200×200.

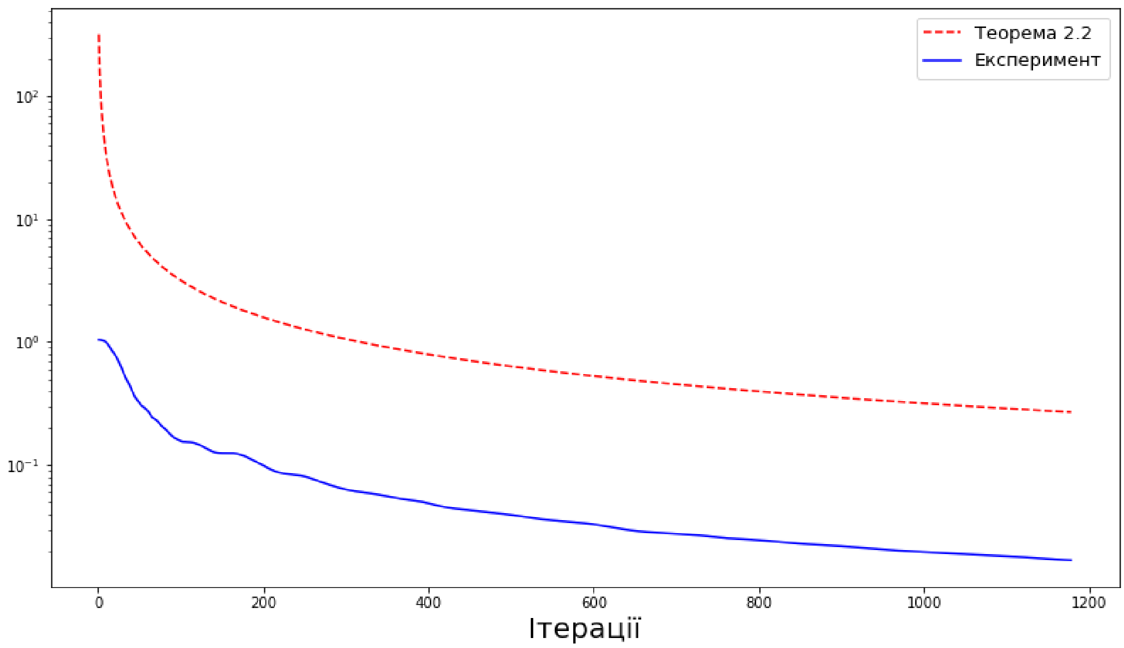


Рис. 5.3. Алгоритм (5.2), розмірність K — 100×300.

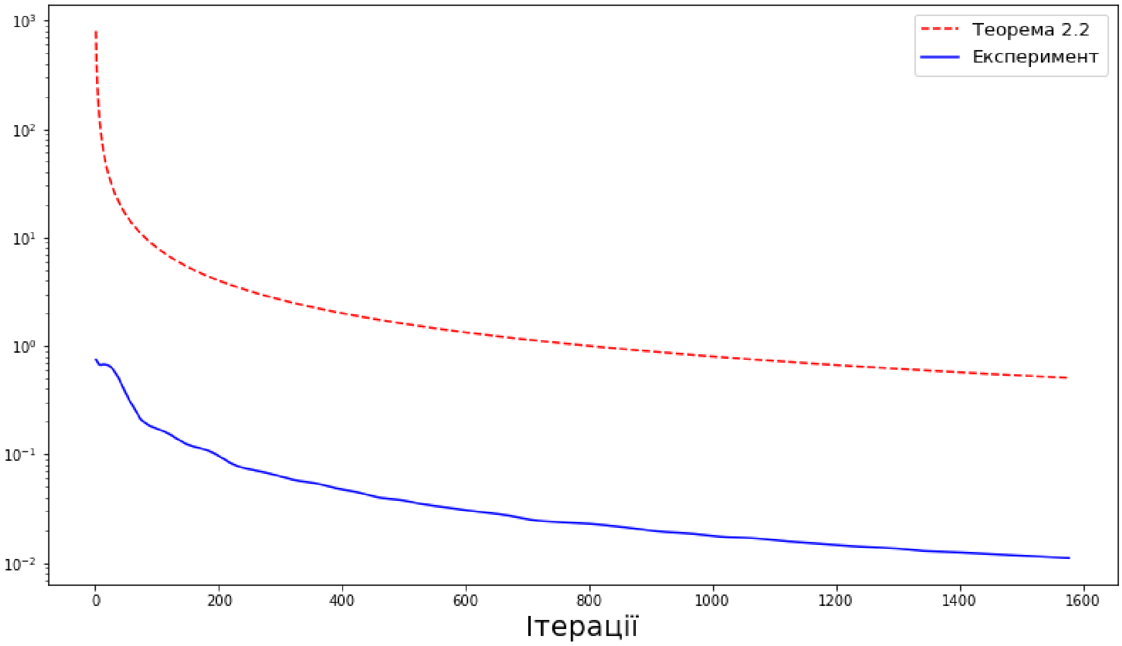


Рис. 5.4. Алгоритм (5.2), розмірність K — 500×500.

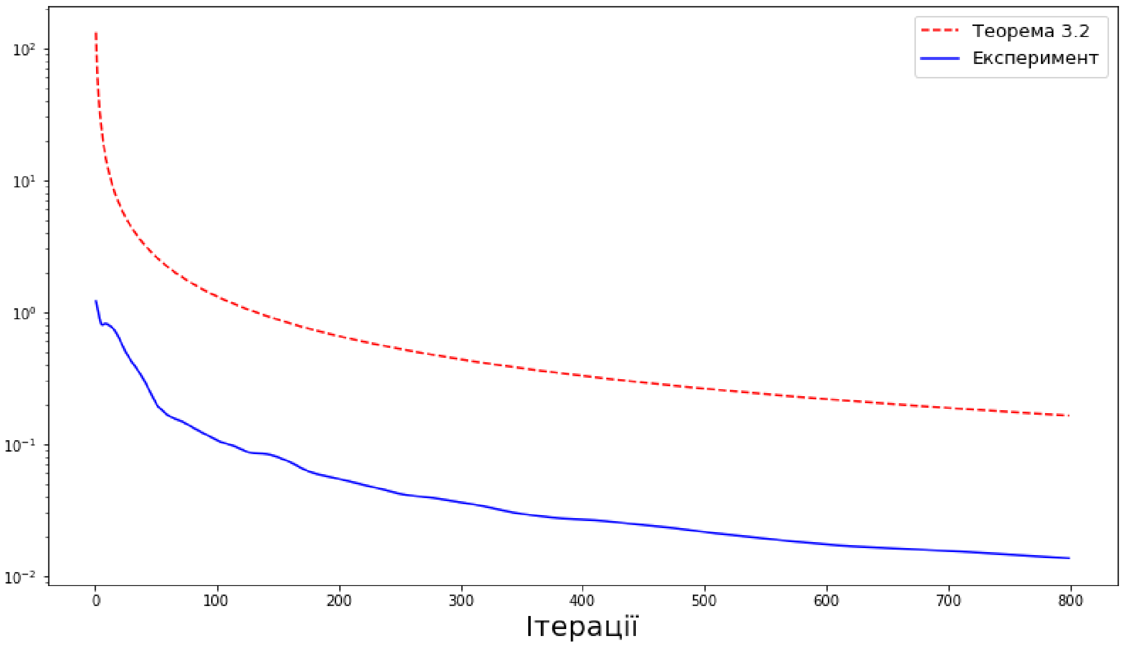


Рис. 5.5. Алгоритм (5.3), розмірність K — 100×100.

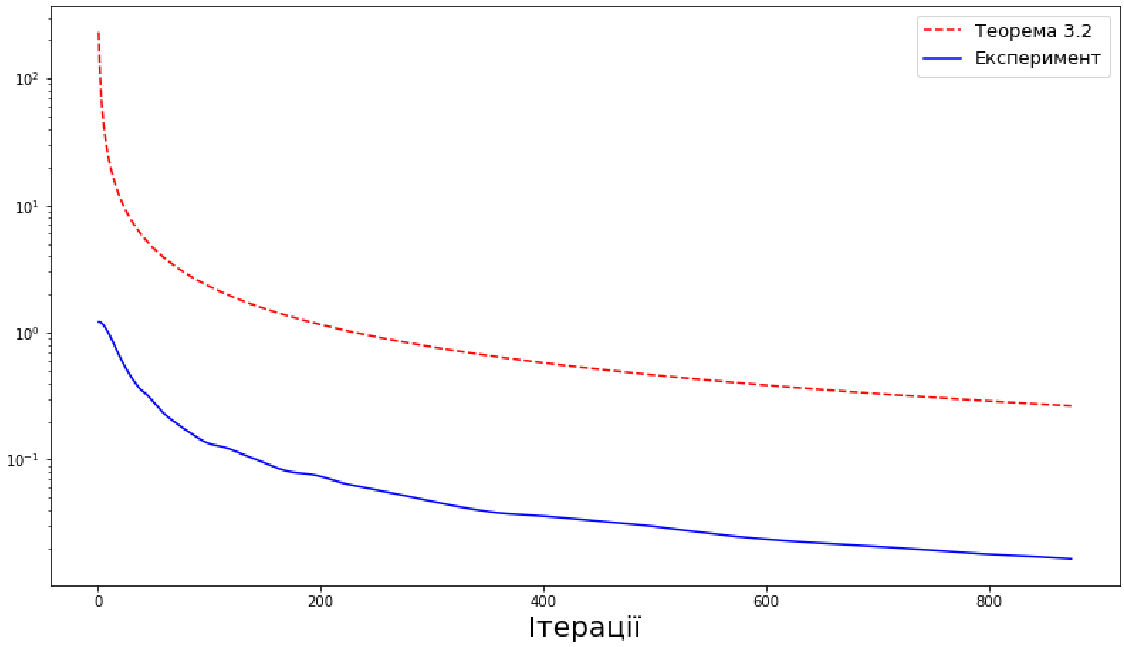


Рис. 5.6. Алгоритм (5.3), розмірність K — 200×200.

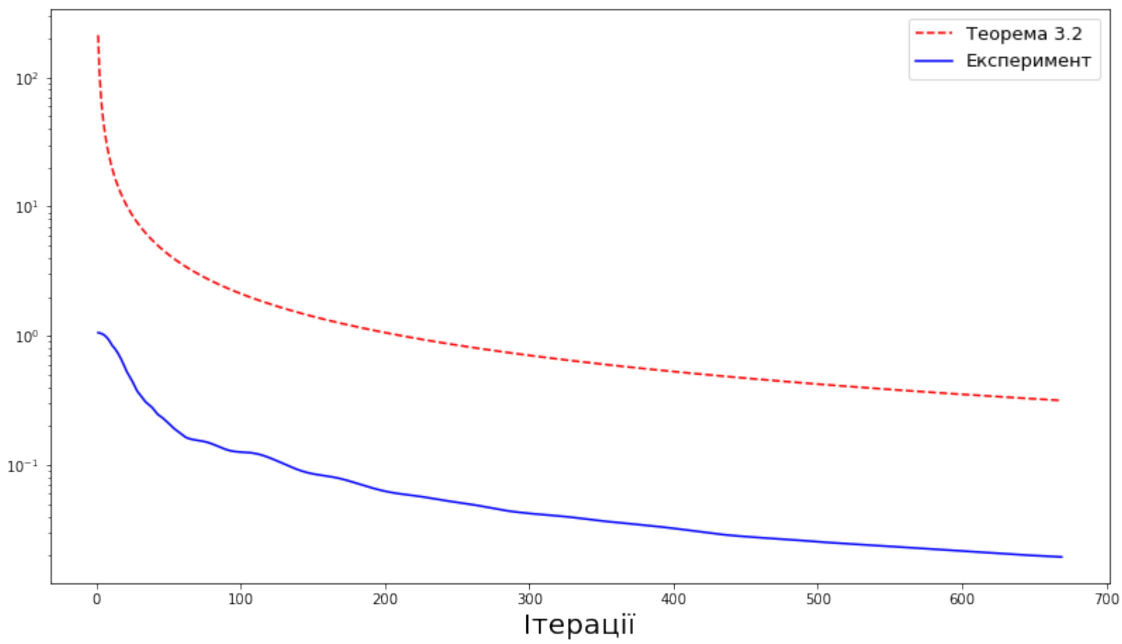


Рис. 5.7. Алгоритм (5.3), розмірність K — 100×300.

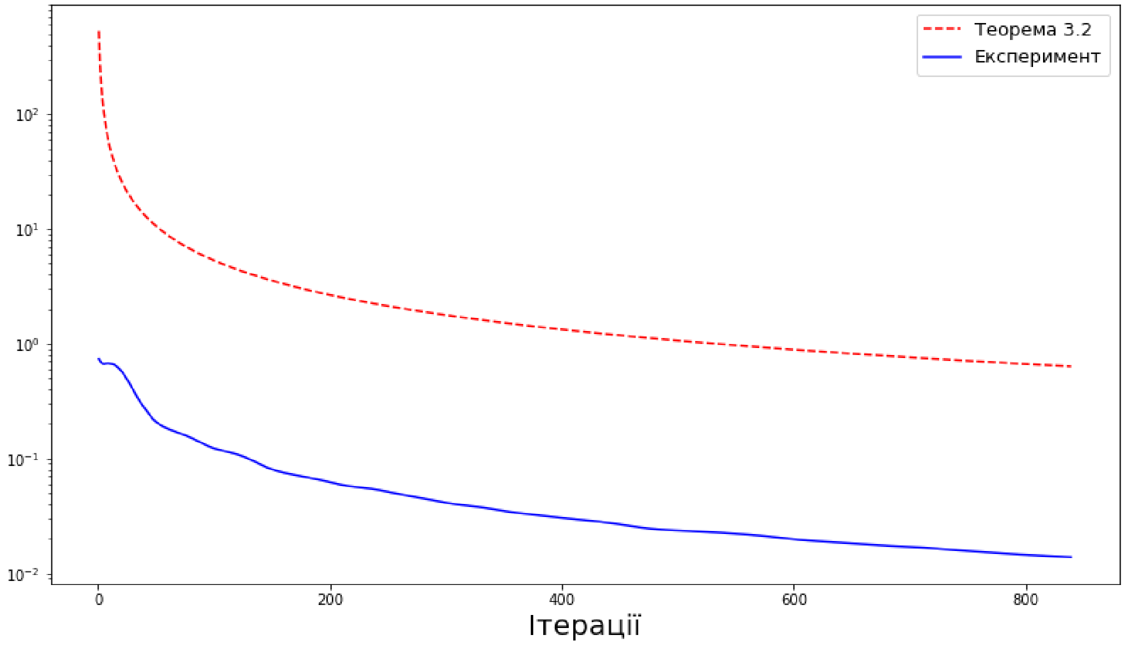


Рис. 5.8. Алгоритм (5.3), розмірність K — 500×500.

5.2. Квадратична сідлова задача

Розглянемо задачу квадратичну сідлову задачу:

$$\min_{x \in \mathbb{R}^n} \max_{y \in \mathbb{R}^m} \left(\frac{\alpha}{2} \|x\|_2^2 + \langle a, x \rangle + \langle Kx, y \rangle - \langle b, y \rangle - \frac{\alpha}{2} \|y\|_2^2 \right), \tag{5.4}$$

де K — дійсна матриця $m \times n$, $a \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $\alpha > 0$.

Задача (5.4) має єдиний розв'язок та рівносильна системі лінійних алгебраїчних рівнянь:

$$\begin{pmatrix} \alpha I_{n \times n} & K^* \\ -K & \alpha I_{m \times m} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = -\begin{pmatrix} a \\ b \end{pmatrix},$$

де $I_{n \times n}$ — одинична матриця $n \times n$.

Зауважимо, що до задачі вигляду (5.4) зводиться задача ℓ_2-регуляризованої ℓ_2-регресії. Дійсно, задачу

$$\frac{1}{2} \|Kx - b\|_2^2 + \frac{\lambda}{2} \|x\|_2^2 \to \min_{x \in \mathbb{R}^n},$$

де $b \in \mathbb{R}^m$, $K : \mathbb{R}^n \to \mathbb{R}^m$ — лінійний оператор, $\lambda > 0$, можна записати у сідловому вигляді

$$\min_{x\in\mathbb{R}^n}\max_{y\in\mathbb{R}^m}\left(\langle Kx,y\rangle+\frac{\lambda}{2}\|x\|_2^2-\frac{1}{2}\|y\|_2^2-\langle b,y\rangle\right).$$

Для (5.4) варіанти алгоритмів з теорем 2.4 та 3.4 мають вигляд

$$\begin{cases} x_1=p_0\in\mathbb{R}^n,\quad y_1=q_0\in\mathbb{R}^m,\\ p_k=x_k-\frac{1}{3\sqrt{\|K\|_s^2+\alpha^2}}\left(K^*q_{k-1}+a+\alpha p_{k-1}\right),\\ q_k=y_k+\frac{1}{3\sqrt{\|K\|_s^2+\alpha^2}}\left(Kp_{k-1}-b-\alpha q_{k-1}\right),\\ x_{k+1}=x_k-\frac{1}{3\sqrt{\|K\|_s^2+\alpha^2}}\left(K^*q_k+a+\alpha p_k\right),\\ y_{k+1}=y_k+\frac{1}{3\sqrt{\|K\|_s^2+\alpha^2}}\left(Kp_k-b-\alpha q_k\right), \end{cases} \tag{5.5}$$

та

$$\begin{cases} x_1=x_0\in\mathbb{R}^n,\quad y_1=y_0\in\mathbb{R}^m,\\ x_{k+1}=x_k-\frac{1}{2\sqrt{\|K\|_s^2+\alpha^2}}\left(K^*y_k+a+\alpha x_k\right)-\frac{1}{2\left(\sqrt{\|K\|_s^2+\alpha^2}+\alpha\right)}\times\\ \times\left(K^*y_k-K^*y_{k-1}+\alpha x_k-\alpha x_{k-1}\right),\\ y_{k+1}=y_k+\frac{1}{2\sqrt{\|K\|_s^2+\alpha^2}}\left(Kx_k-b-\alpha y_k\right)+\frac{1}{2\left(\sqrt{\|K\|_s^2+\alpha^2}+\alpha\right)}\times\\ \times\left(Kx_k-Kx_{k-1}-\alpha y_k+\alpha y_{k-1}\right), \end{cases} \tag{5.6}$$

відповідно.

Елементи матриці K генерувались випадково. Використали рівномірний розподіл на $\{-5,-4,...,4,5\}$. Розглядали випадок $a=0$, $b=0$ та $\alpha=0,1$. Тоді $x=0$, $y=0$ — єдиний розв'язок (5.4).

В експериментах за стартову точку обирали

$$(x_1,y_1)=(1,...,1,1,...,1)\in\mathbb{R}^n\times\mathbb{R}^m.$$

Для (5.5) динаміка зменшення значення Δ_N, де

$$\Delta_N=\sqrt{\|x_N\|_2^2+\|y_N\|_2^2},$$

подана на Рис. 5.9–5.12. Для (5.6) динаміка зменшення значення Δ_N подана на Рис. 5.13–5.16. Наведені графіки демонструють відповідність результатів розрахунків теоретичним оцінкам теорем 2.4, 3.4.

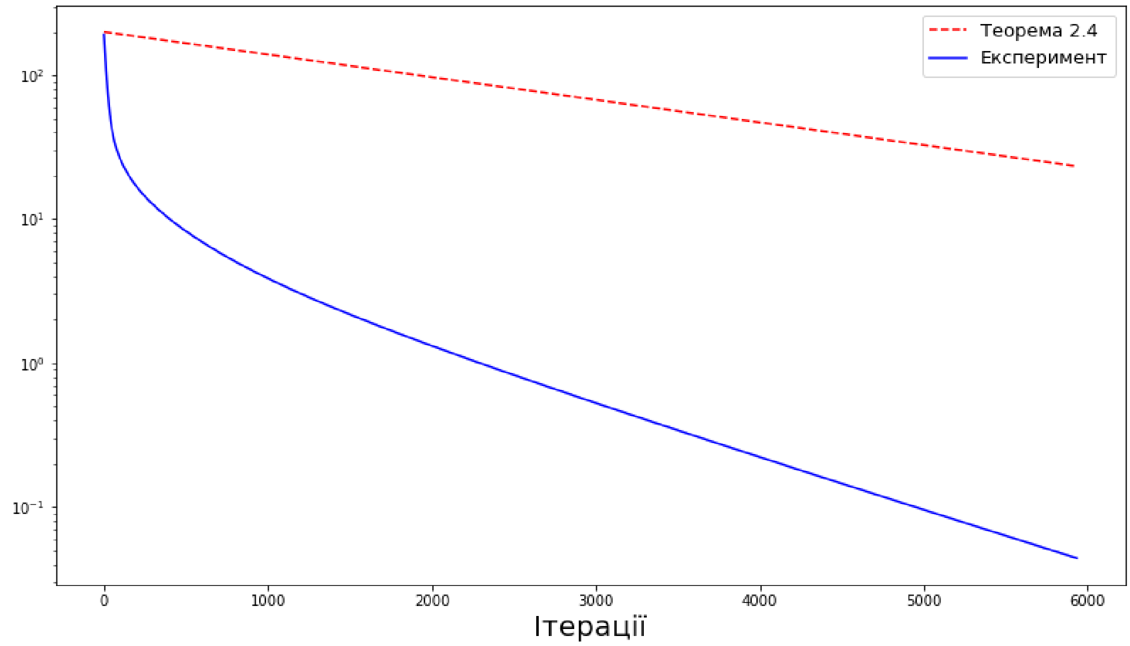


Рис. 5.9. Алгоритм (5.5), розмірність K — 100×100.

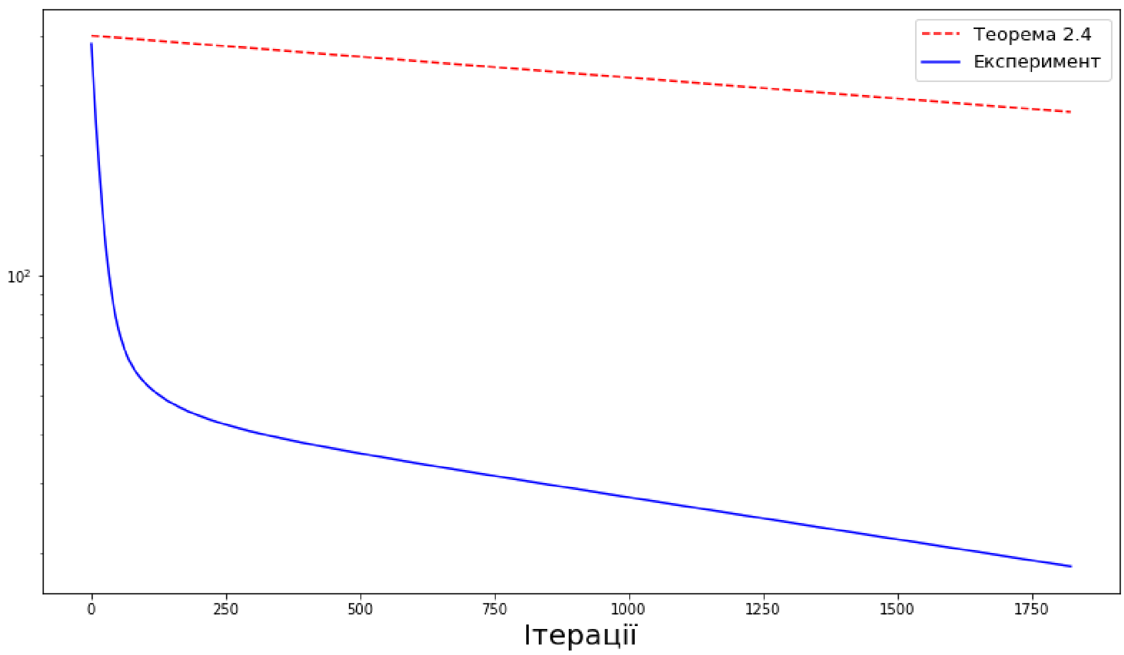


Рис. 5.10. Алгоритм (5.5), розмірність K — 100×300.

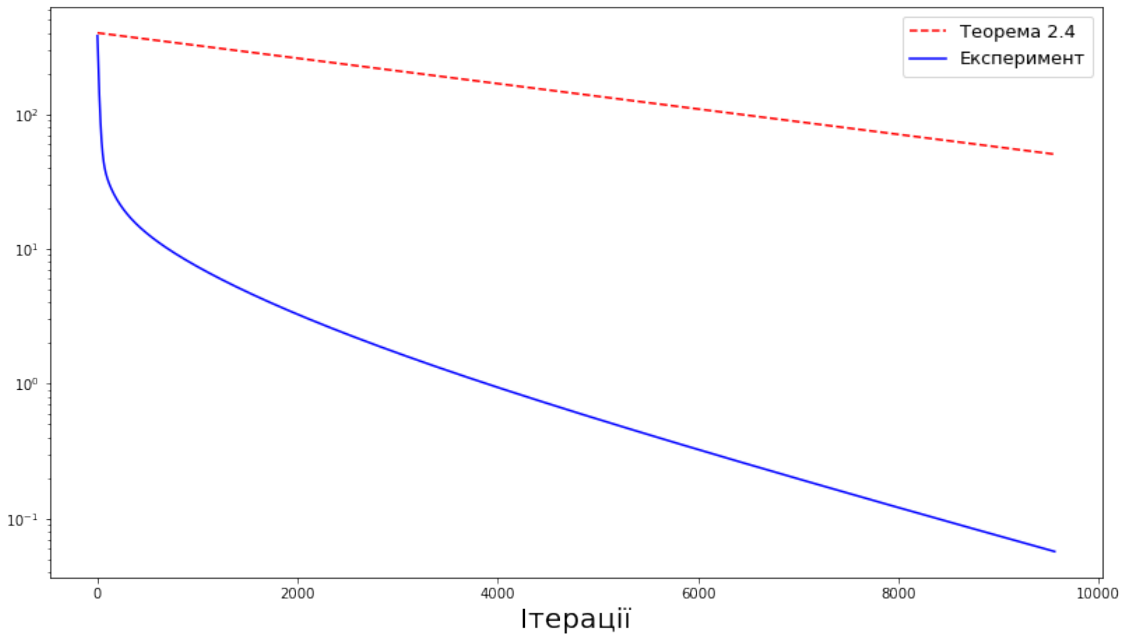


Рис. 5.11. Алгоритм (5.5), розмірність K — 200×200.

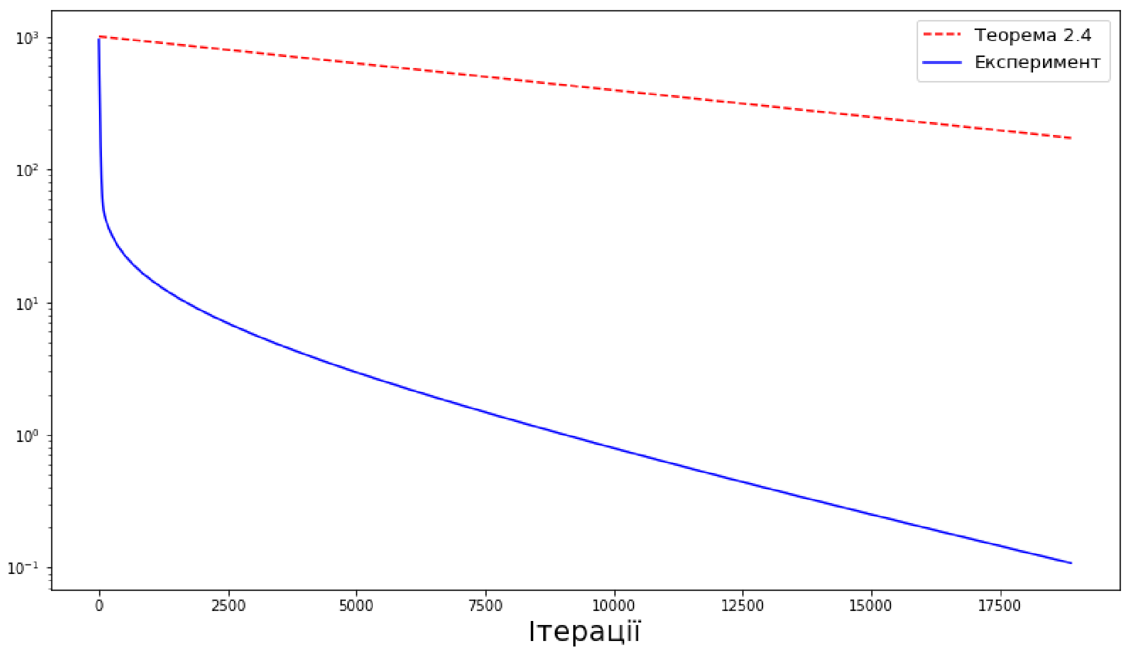


Рис. 5.12. Алгоритм (5.5), розмірність K — 500×500.

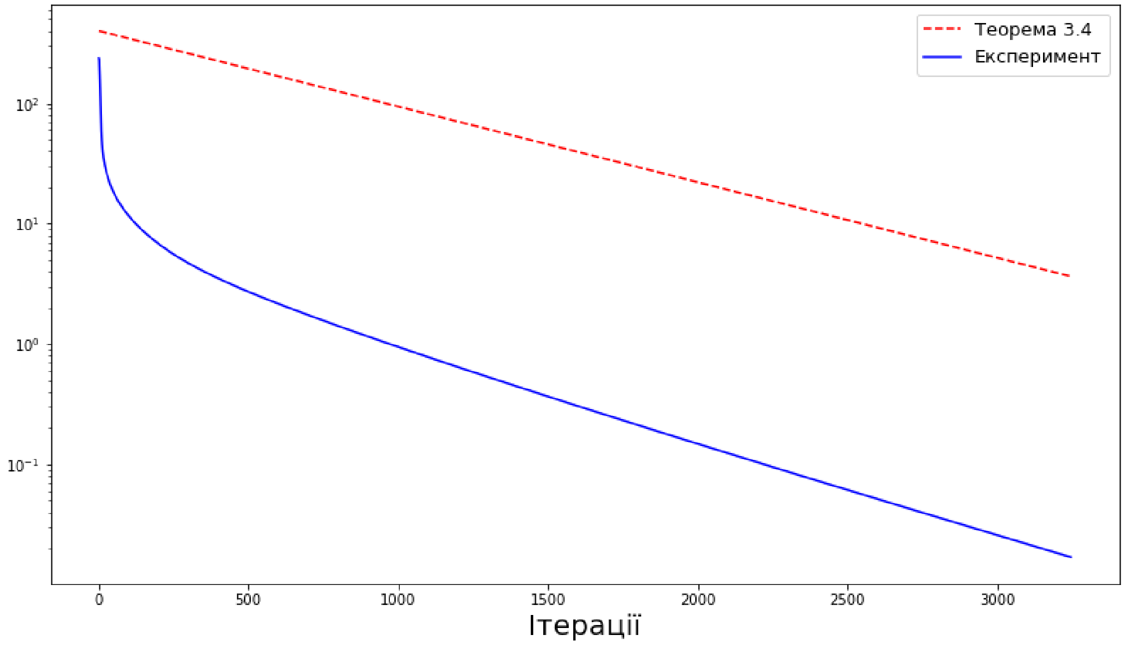


Рис. 5.13. Алгоритм (5.6), розмірність K — 100×100.

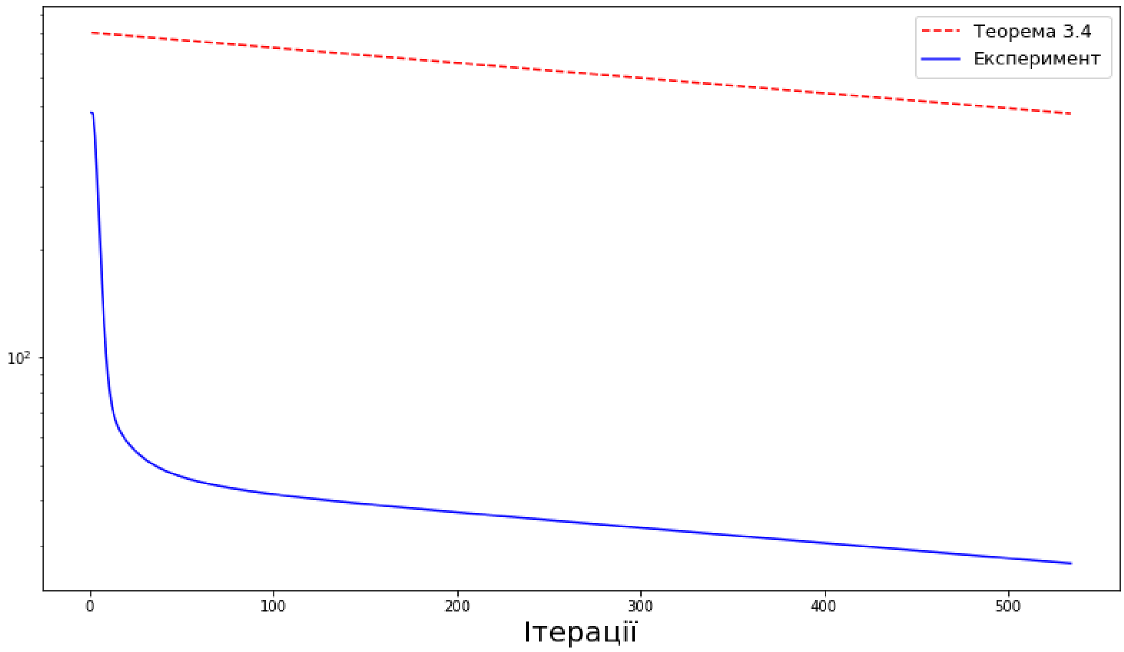


Рис. 5.14. Алгоритм (5.6), розмірність K — 100×300.

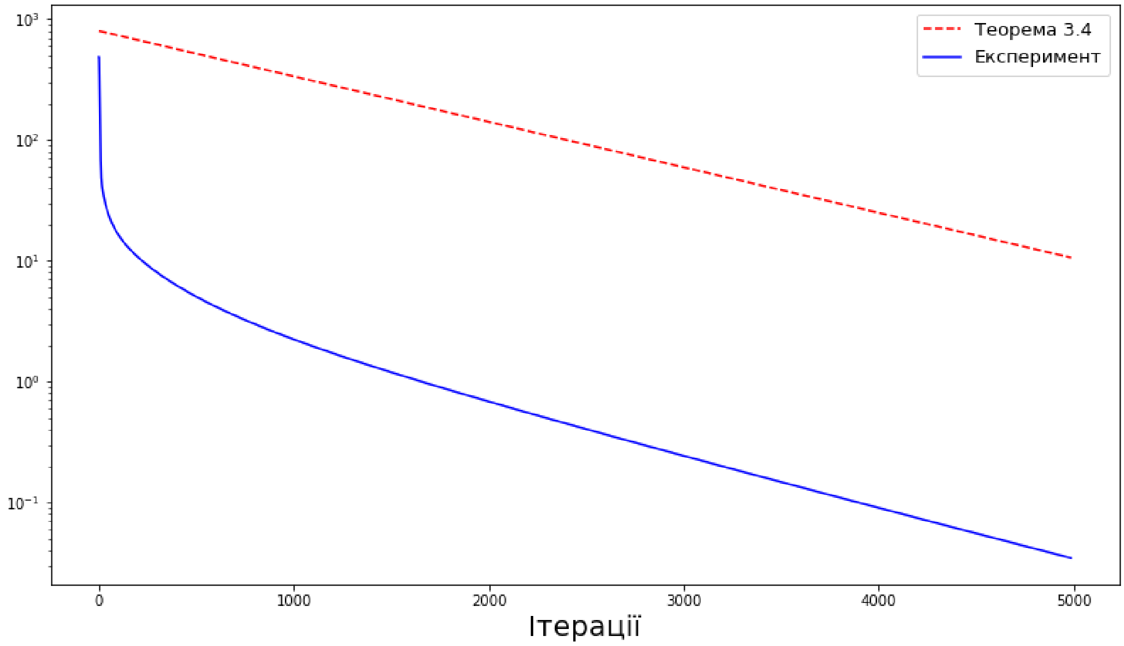


Рис. 5.15. Алгоритм (5.6), розмірність K — 200×200.

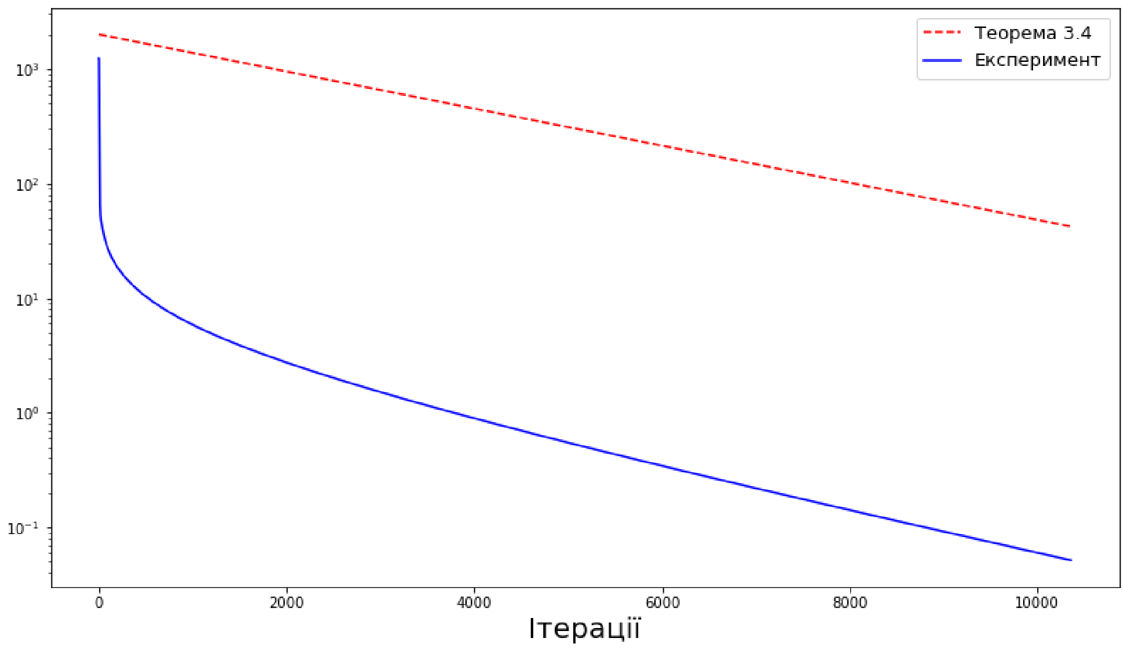


Рис. 5.16. Алгоритм (5.6), розмірність K — 500×500.

В Таблиці 5.2 вказано за яких час алгоритмами досягалось виконання нерівності $\Delta_N < 0,001$.

Таблиця 5.2

Час, потрібний для $\Delta_N < 0,001$, секунди

Розмірність K	Алгоритм (5.5)	Алгоритм (5.6)
100×100	2.95	1.41
200×200	3.99	2.99
500×500	103.55	63.17

5.3. Змагальні породжуючі мережі

Задача навчання GAN [53] формулюється як стохастична сідлова задача

$$\min_{x \in \mathbb{R}^p} \max_{y \in \mathbb{R}^q} L(x, y), \quad L(x, y) = \mathbb{E}_\xi l(x, y, \xi),$$

де ξ — випадкова величина.

Нехай

$$w = (x, y) \in \mathbb{R}^p \times \mathbb{R}^q, \quad F_\xi(w) = (\nabla_x l(x, y, \xi), -\nabla_y l(x, y, \xi)).$$

Розглянемо дві модифікації методу екстраполяції з минулого, що відомі під назвою стохастичний оптимістичний градієнтний метод (stochastic optimistic gradient method, SOG) [190]:

$$w_{n+1} = w_n - \lambda((1 + \alpha) F_{\xi_n}(w_n) - \alpha F_{\xi_{n-1}}(w_{n-1})),$$

$$w_{n+1} = w_n - \lambda((1 + \alpha) F_{\xi_n}(w_n) - \alpha F_{\xi_n}(w_{n-1})),$$

де $\lambda > 0$ — величина кроку, $\alpha > 0$ — параметр, що задає степінь «оптимізму», ξ_n — випадкові величини. Ці схеми обчислювально простіші за стохастичний екстраградієнтний метод [58]. Але вони більш чутливі до стохастичних шумів [57].

Для зменшення впливу шуму можна використати експоненційне усереднення (exponential moving average, EMA) стохастичних градієнтів [191]. Це усереднення має вигляд:

$$F_n = \begin{cases} F_{\xi_0}(w_0), & n = 0 \ , \\ (1-\beta)F_{\xi_n}(w_n) + \beta F_{n-1}, & n \neq 0 \ , \end{cases}$$

де $\beta \in [0, 1]$. Будемо використовувати різні значення β, разом з $\beta = 0.9$ (як рекомендують в [190]).

В роботі [190] запропоновано нові алгоритми з експоненційним усередненням (Optimistic EMA gradients):

— метод Omega:

$$w_{n+1} = w_n - \lambda((1+\alpha)F_{\xi_n}(w_n) - \alpha F_{n-1});$$

— метод Omega M:

$$w_{n+1} = w_n - \lambda((1+\alpha)F_n - \alpha F_{n-1}).$$

Автори [190] не отримали теоретичних гарантій збіжності. Але продемонстрували задовільну роботу алгоритмів на модельних білінійних, квадратичних та лінійно-квадратичних стохастичних іграх.

Продемонструємо поведінку даних методів при навчанні найпростіших генеруючих змагальних нейронних мереж та порівняємо з методом SOG.

Для експериментів було створено Deep Convolutional Generative Adversarial Network (DCGAN), архітектура якої вказана на Рис. 5.17.

Для побудови мереж використано Keras Sequential API з навчальним циклом tf.GradientTape.

Генератор використовує шари tf.keras.layers.Conv2DTranspose (шари підвищення дискретизації) для створення зображення з шуму. І генератор, і дискримінатор визначаються за допомогою Keras Sequential API.

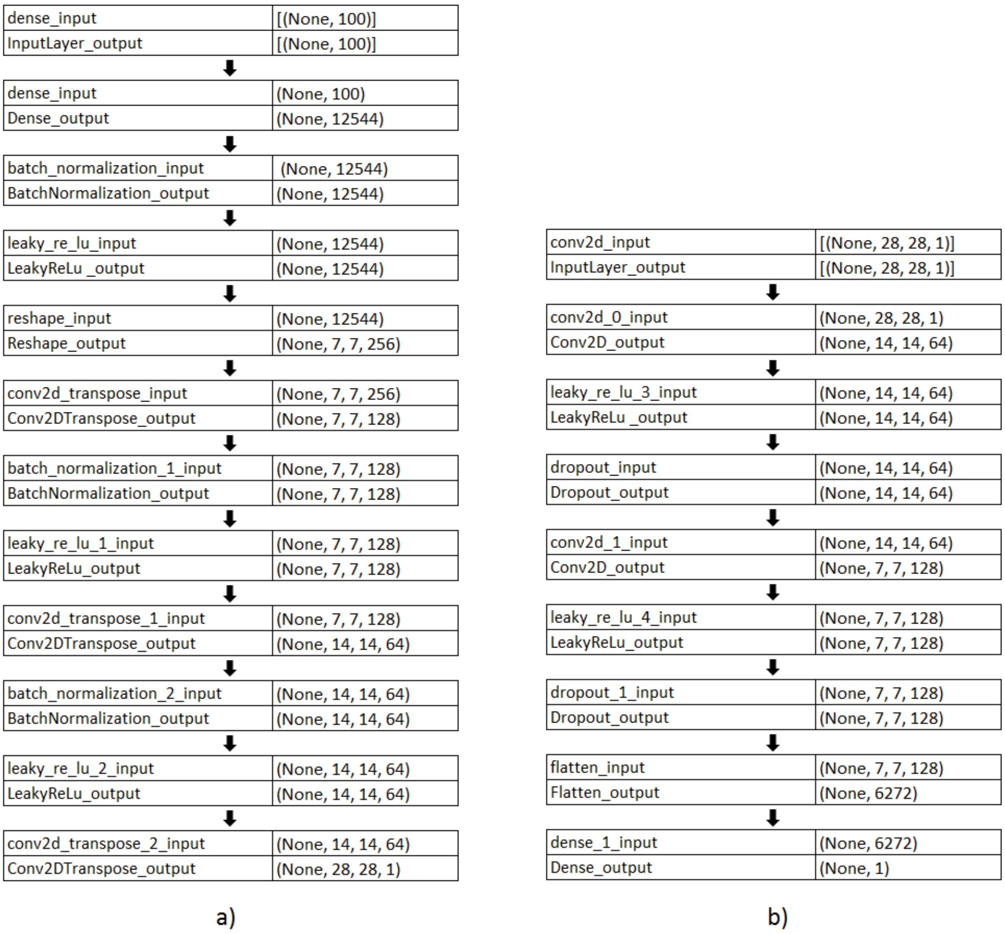


Рис. 5.17. Архітектури: генератор (a), дискримінатор (b).

За допомогою описаних алгоритмів проводилось тренування по генеруванню рукописних цифр за допомогою набору MNIST.

Цілові функції генератора та дискримінатора такі (див. підрозділ 1.3.3):

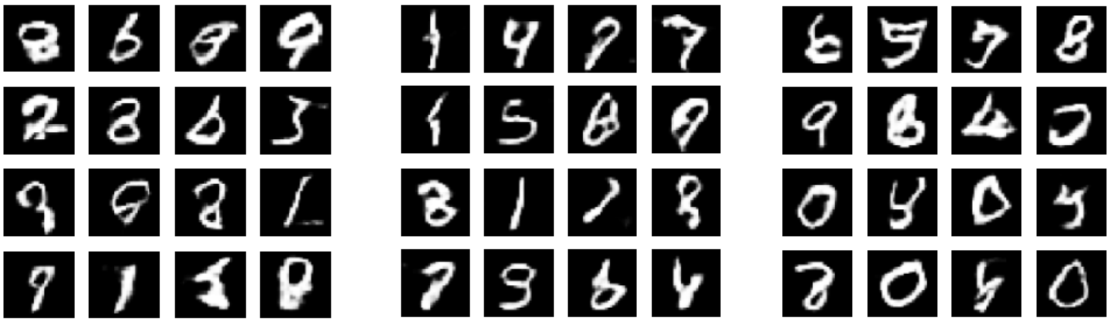
$$L_G(x, y) = \mathbb{E}_{z\sim\mu} \log\left(1 - D_y\left(G_x(z)\right)\right),$$

$$L_D(x, y) = \mathbb{E}_{u\sim p_{data}} \log D_y(u) + \mathbb{E}_{z\sim\mu} \log\left(1 - D_y\left(G_x(z)\right)\right).$$

Розмір міні-батчів — 256 зображень. В алгоритмах обирали $\lambda = 0.001$.

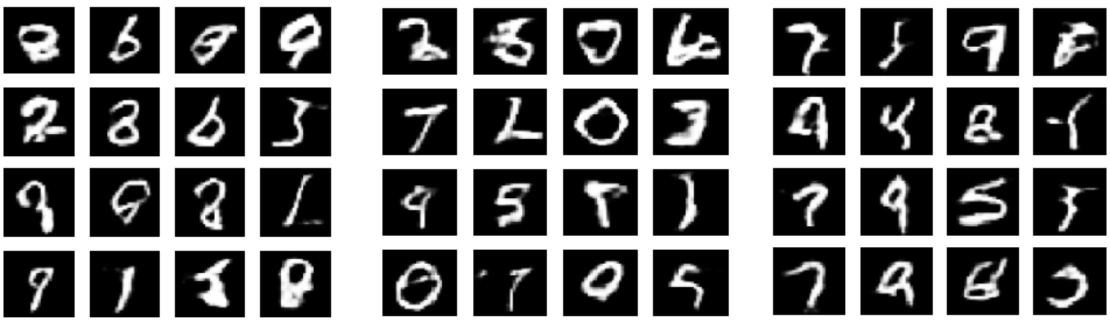
Навчання проводилось за допомогою можливостей Google Colab з використанням GPU типу T4.

Наведемо результати роботи генератора через 500 епох навчання (Рис. 5.18–5.24). Також наведемо графіки зміни цільових функцій генератора та дискримінатора (Рис. 5.25–5.31).



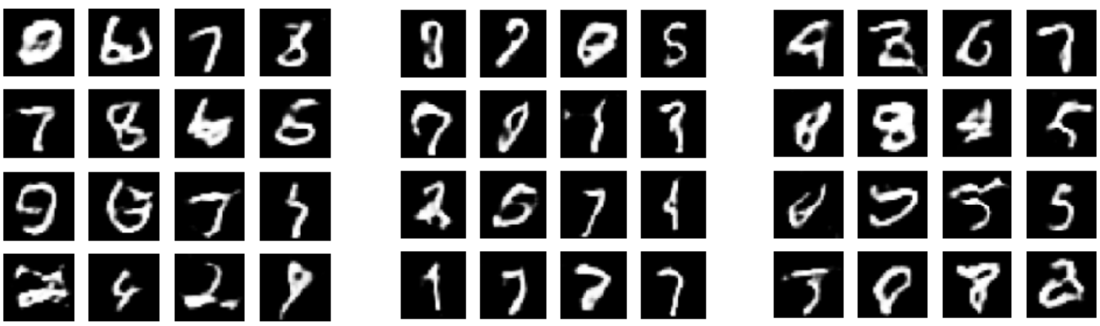
a) b) c)

Рис. 5.18. Синтетичні цифри, згенеровані після 500 епох з $\alpha = 0.1$, $\beta = 0.1$: SOG (a), Omega (b), Omega M (c).



a) b) c)

Рис. 5.19. Синтетичні цифри, згенеровані після 500 епох з $\alpha = 0.1$, $\beta = 0.5$: SOG (a), Omega (b), Omega M (c).



a) b) c)

Рис. 5.20. Синтетичні цифри, згенеровані після 500 епох з $\alpha = 0.5$, $\beta = 0.1$: SOG (a), Omega (b), Omega M (c).

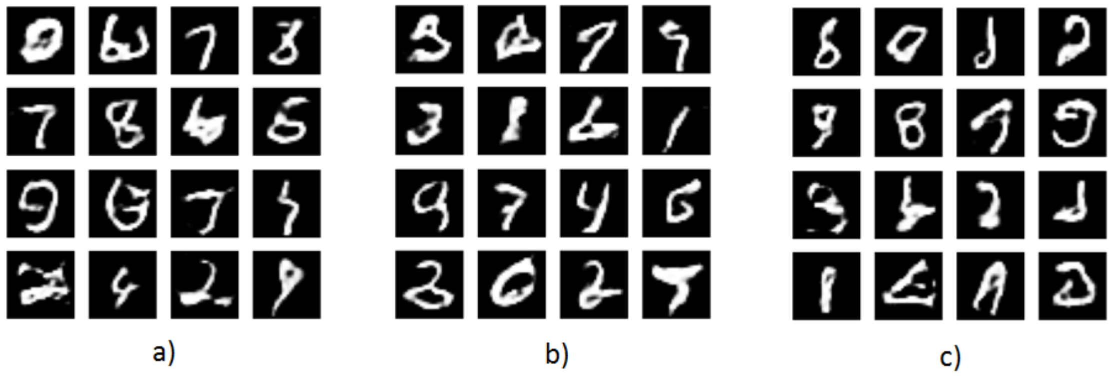


Рис. 5.21. Синтетичні цифри, згенеровані після 500 епох з $\alpha = 0.5$, $\beta = 0.9$: SOG (a), Omega (b), Omega M (c).

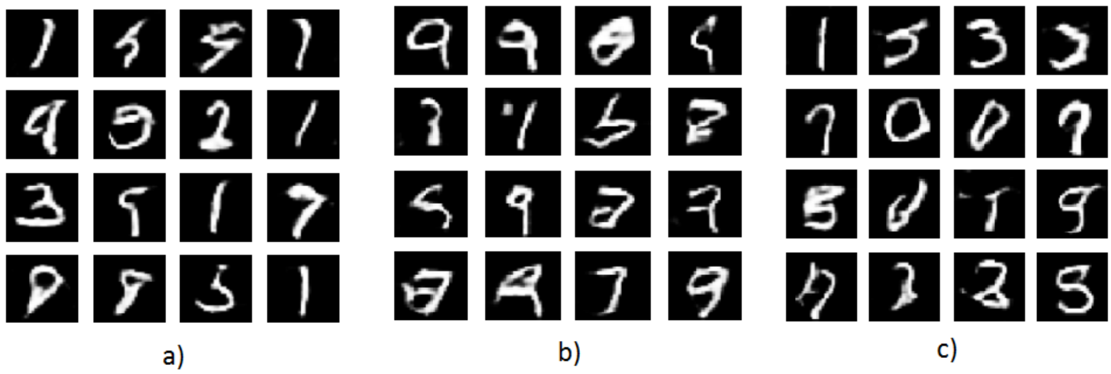


Рис. 5.22. Синтетичні цифри, згенеровані після 500 епох з $\alpha = 1.0$, $\beta = 0.1$: SOG (a), Omega (b), Omega M (c).

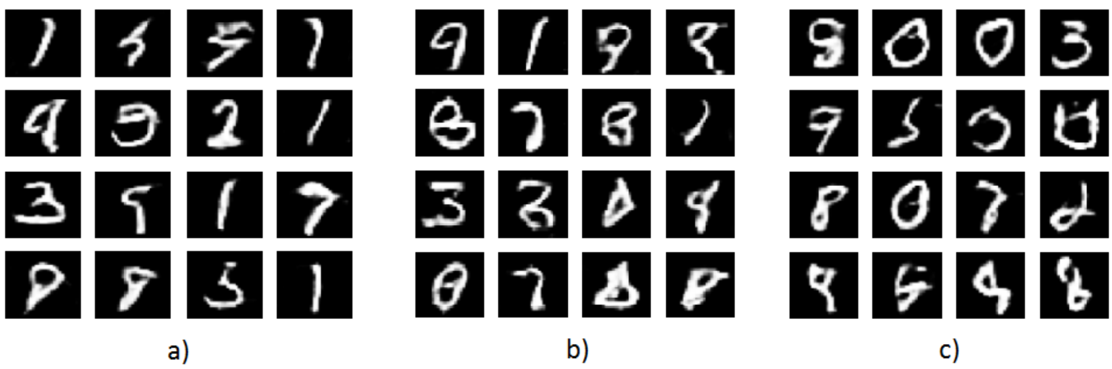


Рис. 5.23. Синтетичні цифри, згенеровані після 500 епох з $\alpha = 1.0$, $\beta = 0.5$: SOG (a), Omega (b), Omega M (c).

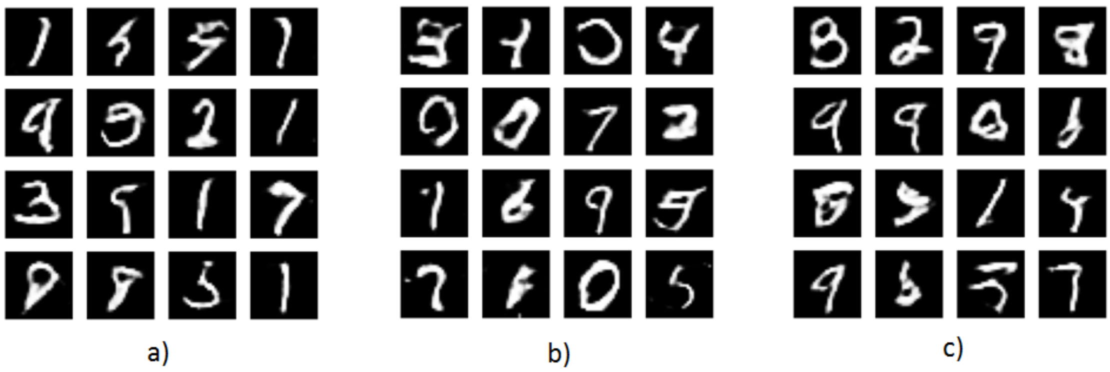


Рис. 5.24. Синтетичні цифри, згенеровані після 500 епох з $\alpha = 1.0$, $\beta = 0.9$: SOG (a), Omega (b), Omega M (c).

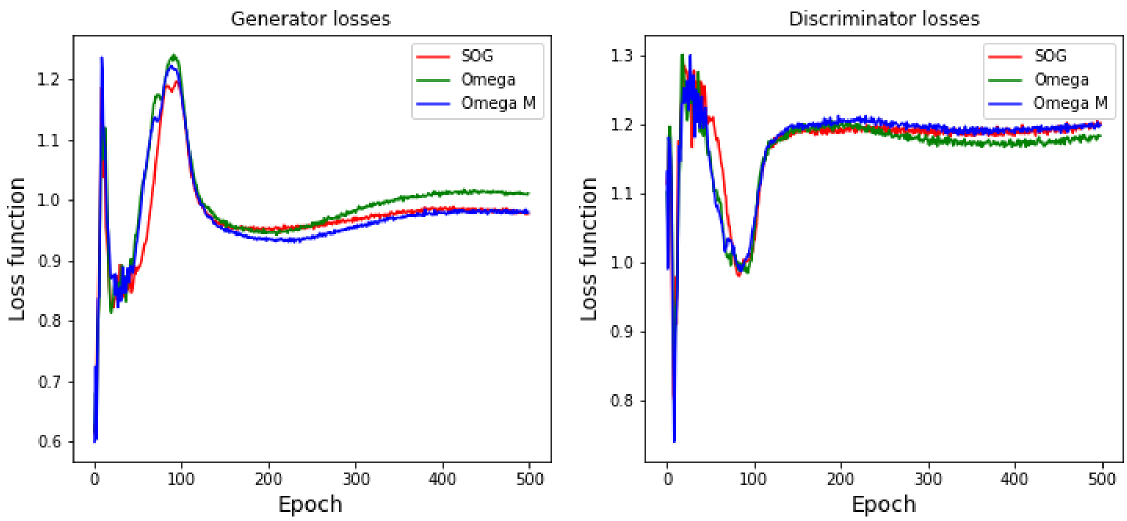


Рис. 5.25. Зміна цільових функцій генератора та дискримінатора за 500 епох з $\alpha = 0.1$, $\beta = 0.1$.

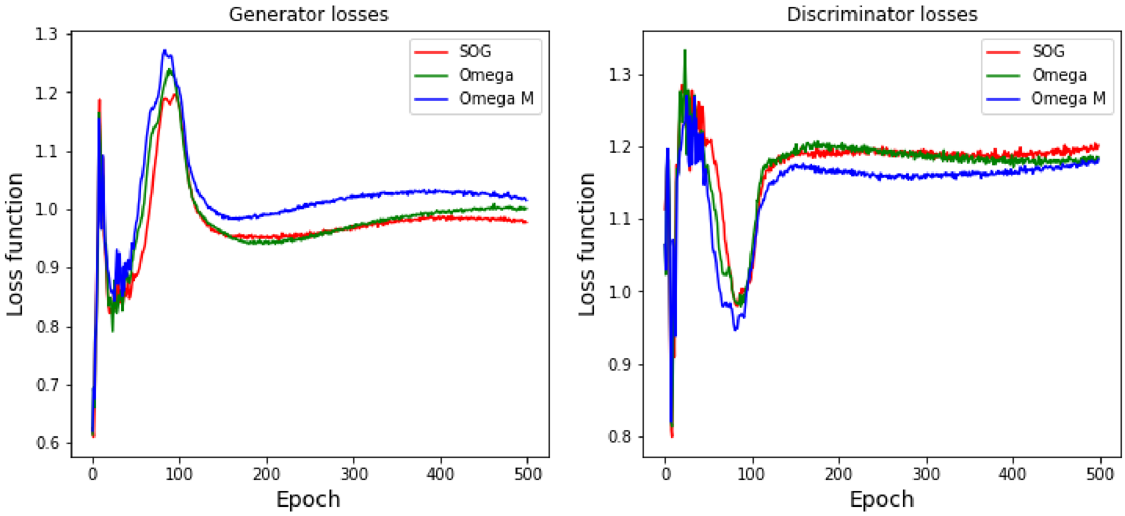


Рис. 5.26. Зміна цільових функцій генератора та дискримінатора за 500 епох з $\alpha = 0.1$, $\beta = 0.5$.

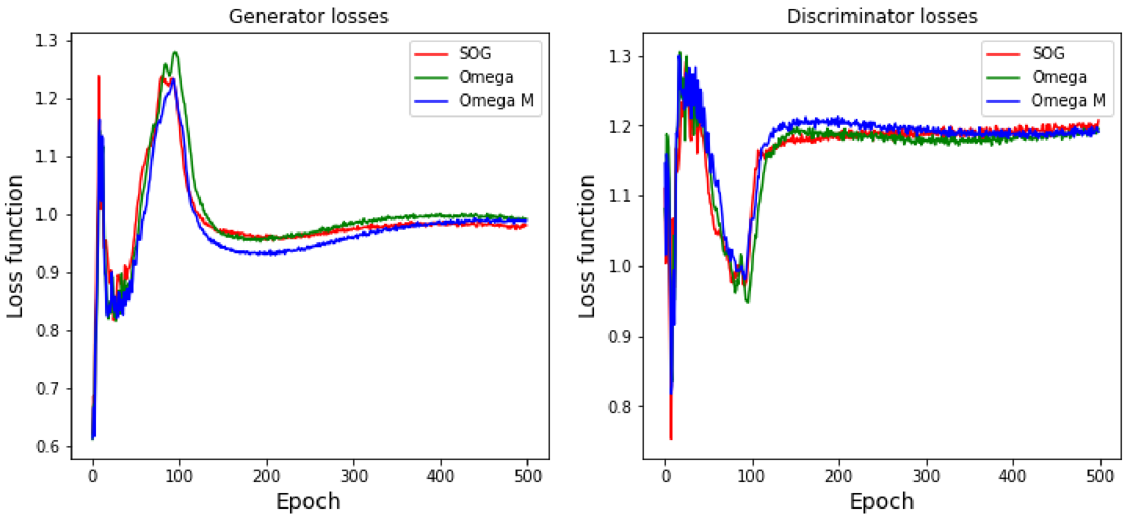


Рис. 5.27. Зміна цільових функцій генератора та дискримінатора за 500 епох з $\alpha = 0.5$, $\beta = 0.1$.

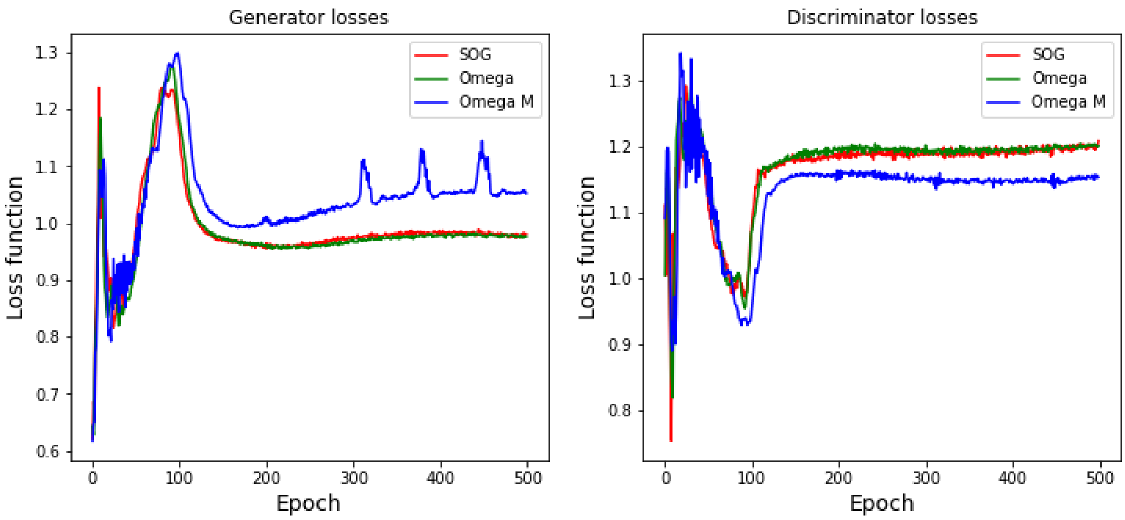


Рис. 5.28. Зміна цільових функцій генератора та дискримінатора за 500 епох з $\alpha = 0.5$, $\beta = 0.9$.

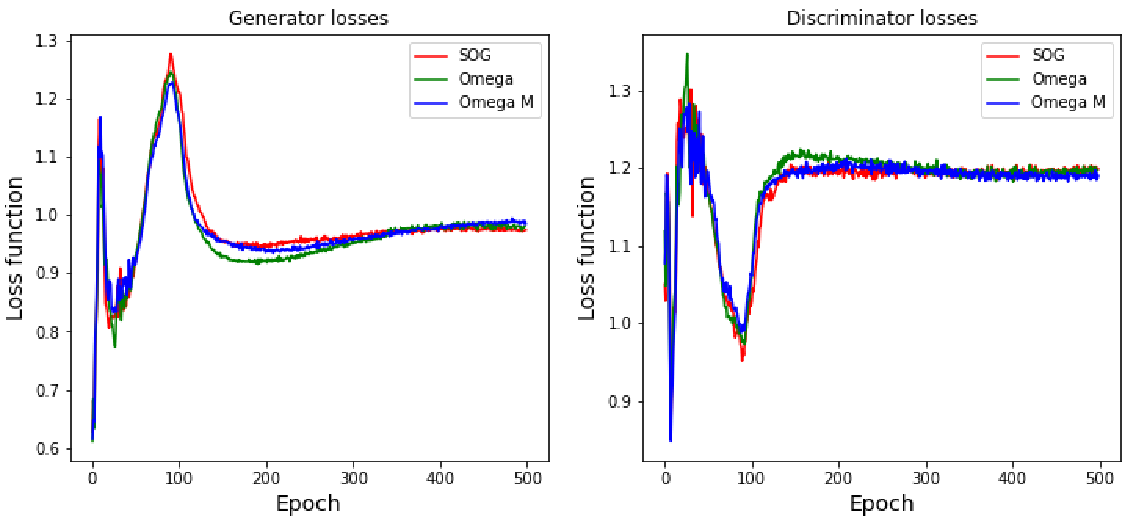


Рис. 5.29. Зміна цільових функцій генератора та дискримінатора за 500 епох з $\alpha = 1.0$, $\beta = 0.1$.

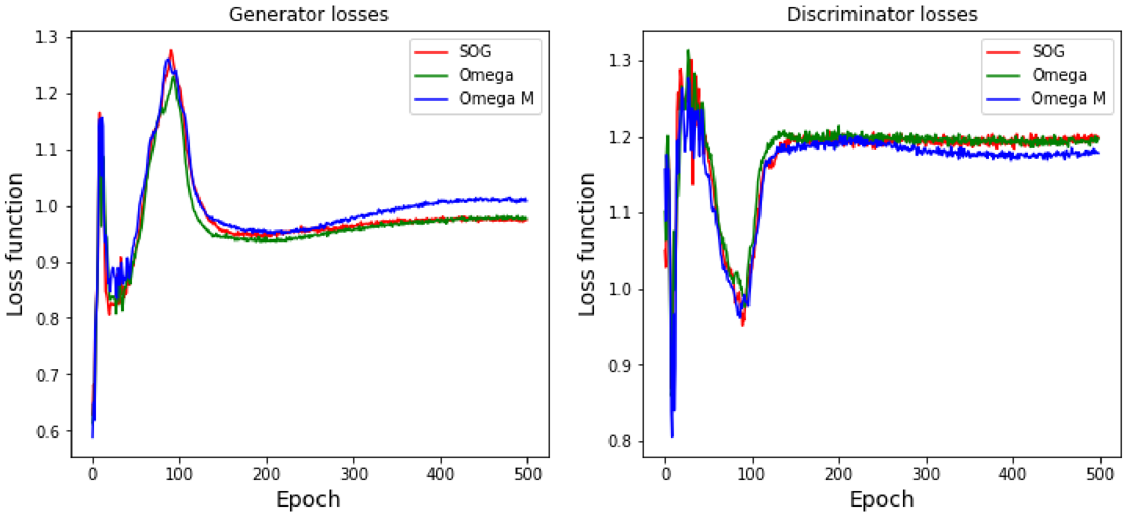


Рис. 5.30. Зміна цільових функцій генератора та дискримінатора за 500 епох з $\alpha = 1.0$, $\beta = 0.5$.

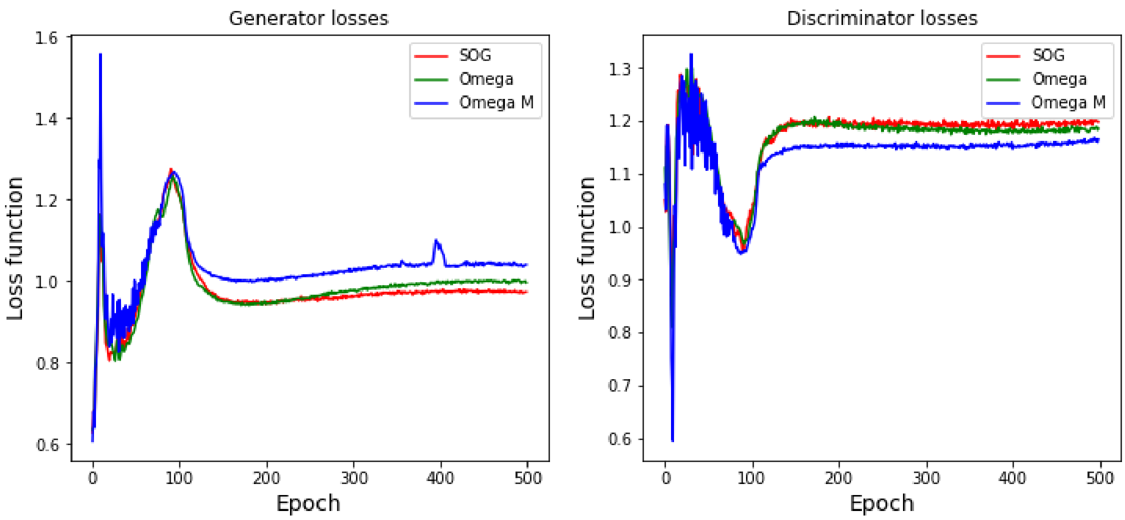


Рис. 5.31. Зміна цільових функцій генератора та дискримінатора за 500 епох з $\alpha = 1.0$, $\beta = 0.9$.

5.4. Висновки до розділу 5

У даному розділі описано експерименти, проведені для перевірки теорем 2.2, 2.4, 3.2 та 3.4, що містять оцінки швидкості збіжності розглянутих алгоритмів. Розглядались білінійні та квадратичні сідлові задачі. Результати проведених експериментів підтверджують теоретичні оцінки.

Також проведено експерименти з навчанням простих GANs за допомогою двох рандомізованих варіантів методу екстраполяції з минулого з експоненційним усередненням. Результати мають попередній характер. Зауважимо, що задача про теоретичні гарантії збіжності цих варіантів є відкритою.

ВИСНОВКИ

В даній роботі розроблено та теоретично обгрунтовано нові алгоритми для розв'язання варіаційних нерівностей. Зокрема, отримано такі нові результати:

— доведено теореми сильної збіжності алгоритму екстраполяції з минулого та алгоритму операторної екстраполяції для варіаційних нерівностей в гільбертовому просторі, що задовольняють умову рівномірної монотонності;

— доведено теореми слабкої збіжності та сублінійні оцінки швидкості збіжності алгоритмів;

— розроблено та теоретично обгрунтовано адаптивні та регуляризовані варіанти алгоритму екстраполяції з минулого та алгоритму операторної екстраполяції для варіаційних нерівностей в гільбертовому просторі;

— отримано лінійні оцінки швидкості збіжності алгоритму екстраполяції з минулого та алгоритму операторної екстраполяції для варіаційних нерівностей в гільбертовому просторі, що задовольняють умову узагальненої сильної монотонності;

— розроблено варіанти алгоритму екстраполяції з минулого та алгоритму операторної екстраполяції з дивергенцією Брегмана для розв'язання варіаційних нерівностей з монотонними та ліпшицевими операторами, що діють в скінченновимірному дійсному лінійному просторі;

— розроблено та теоретично обгрунтовано варіант алгоритму операторної екстраполяції для варіаційних нерівностей в 2-рівномірно опуклому та рівномірно гладкому банаховому просторі.

Окремі результати роботи були впроваджені у навчальний процес кафедри обчислювальної математики факультету комп'ютерних наук та кібернетики Київського національного університету імені Тараса Шевченка. У подальшому одержані результати можуть бути використані при розробці алгоритмів для сідлових задач, варіаційних нерівностей, ігрових задач, що пов'язані з моделями математичної економіки та машинним навчанням.

СПИСОК ВИКОРИСТАНИХ ДЖЕРЕЛ

1. Semenov V. V., Denisov S. V., Sandrakov G. V., Kharkov O. S. Convergence of the Operator Extrapolation Method for Variational Inequalities in Banach Spaces. Cybernetics and Systems Analysis. 2022. Vol. 58. Issue 5. P. 740–753.
2. Харьков О. С. Оцінки ефективності для методів з дивергенцією Брегмана. Журнал обчисл. та прикл. матем. 2023. № 2. С. 83–93.
3. Семенов В. В., Харьков О. С. Алгоритм екстраполяції з минулого для варіаційних нерівностей в гільбертовому просторі. Журнал обчисл. та прикл. матем. 2023. № 2. С. 52–82.
4. Семенов В. В., Харьков О. С. Регуляризований алгоритм операторної екстраполяції. Журнал обчисл. та прикл. матем. 2023. № 1. С. 15–27.
5. Семенов В., Харьков О. Метод операторної екстраполяції для варіаційних нерівностей в банахових просторах. Фізико-математичне моделювання та інформаційні технології. 2023. Вип. 37. С. 118–122.
6. Денисов С. В., Семенов В. В., Харьков О. С. Слабка збіжність методу операторної екстраполяції для варіаційних нерівностей в рівномірно опуклих банахових просторах. Журнал обчисл. та прикл. матем. 2022. № 2. С. 42–49.
7. Семенов В. В., Сірик Д. С., Харьков О. С. Збіжність методу операторної екстраполяції. Доповіді НАН України. 2021. № 4. С. 28–35.
8. Семенов В. В., Сирык Д. С., Харьков О. С. Адаптивный метод операторной экстраполяции. Фізико-математичне моделювання та інформаційні технології. 2021. Вип. 33. С. 143–147.
9. Семенов В. В., Денисов С. В., Сирык Д. С., Харьков О. С. Сходимость метода экстраполяции из прошлого и метода операторной экстраполяции-

ции. Міжнародний науково-технічний журнал «Проблеми керування та інформатики». 2021. Том 66. № 3. С. 58–72.

10. Харьков О. С., Ведель Я. І., Семенов В. В. Методи для задач векторного узагальненого оптимального керування системами з розподіленими параметрами. Журнал обчисл. та прикл. матем. 2020. № 2 (134). С. 71–98.
11. Kharkov O., Semenov V. The regularized operator extrapolation method. Information Technology and Implementation (Satellite): Conference Proceedings, November 21, 2023, Kyiv, Ukraine. P. 269–270.
12. Харьков О., Семенов В. Швидкість збіжності нових алгоритмів для варіаційних нерівностей. Modeling, control and information technologies: Proceedings of VI International scientific and practical conference. P. 159–160. https://doi.org/10.31713/MCIT.2023.048.
13. Семенов В. В., Харьков О. С. Швидкість збіжності методу операторної екстраполяції. Матеріали Всеукраїнської наукової конференції «Сучасні проблеми прикладної математики та комп'ютерних наук», присвяченої 50-річчю кафедри теорії оптимальних процесів, 7–9 листопада 2023, Львів. С. 203–205.
14. Semenov V., Kharkov O. The regularized operator extrapolation algorithm for monotone variational inequalities. Intelligent Solutions–S: Proceedings of the International Symposium, September 28, 2023, Kyiv–Uzhorod, Ukraine. P. 88–90.
15. Семенов В., Харьков О. Лінійна швидкість збіжності алгоритмів екстраполяції з минулого та операторної екстраполяції для варіаційних нерівностей. Математика та інформаційні технології. Матеріали міжнародної наукової конференції, присвяченої 55-річчю факультету математики та інформатики, 28–30 вересня 2023 р. — Чернівці: Чернівецький нац. ун-т, 2023. С. 301–302.
16. Kharkov O. S., Semenov V. V. The regularized operator extrapolation

method for variational inequalities and saddle point problems. XXXVIII International Conference «Problems of decision making under uncertainties (PDMU-2023)». Abstracts. September 11–15, 2023, Polyana, Ukraine. P. 55.

17. Semenov V., Denysov S., Kharkov O. About weak convergence of the operator extrapolation method. XXXVII International Conference «Problems of decision making under uncertainties (PDMU-2022)». Abstracts. November 23–25, 2022, Sheki–Lankaran, Republic of Azerbaijan. P. 98.
18. Semenov V., Siryk D., Kharkov O. About convergence of the operator extrapolation method. XXXVI International Conference «Problems of decision making under uncertainties (PDMU-2021)». Abstracts. May 11–14, 2021, Skhidnytsia, Ukraine. P. 91.
19. Denisov S. V., Kharkov O., Semenov V., Vedel Ya. About regularized adaptive extra-proximal algorithm for equilibrium problems in Hadamard spaces. XXXV International Conference «Problems of decision making under uncertainties (PDMU-2020)». Abstracts. May 11–15, 2020, Baku–Sheki, Republic of Azerbaijan. P. 32.
20. Маліцький Ю. В. Ефективні проективні методи для варіаційних нерівностей та задач структурної оптимізації. Автореферат дис. канд. фіз.-мат. наук: 01.05.02. Київ, 2015. 20 с.
21. Чабак Л. М. Проективні алгоритми для варіаційних нерівностей та задач рівноважного програмування. Автореферат дис. канд. фіз.-мат. наук: 01.05.02. Київ, 2018. 20 с.
22. Bauschke H. H., Combettes P. L. Convex Analysis and Monotone Operator Theory in Hilbert Spaces. Berlin-Heidelberg-New York: Springer, 2011. 408 p.
23. Ляшко С. І., Номіровський Д. А., Семенов В. В. та ін. Математичні моделі та обчислювальні процеси. К.: ВПЦ «Київський університет», 2019. 209 с.

24. Семенов В. В. Варіаційні нерівності: теорія та алгоритми. Київ: ВПЦ «Київський університет», 2021. 167 с.
25. Lions J. L., Stampacchia G. Variational inequalities. Commun. Pure Appl. Math. 1967. Vol. XX. P. 493–519.
26. Stampacchia G. Le probleme de Dirichlet pour les equations elliptiques du second ordre a coefcients discontinus. Ann. Inst. Fourier. 1965. Vol. 15, no. 1. P. 189–257.
27. Fichera G. Problemi Elastostatici con Vincoli Unilaterali, il Problema di Signorini con Ambigue Condizioni al Contorno. Mem. Accad. Naz. Lincei. 1964. Vol. VIII, no. 7. P. 91–140.
28. Hartman P., Stampacchia G. On some nonlinear elliptic di erential functional equations. Acta Math. 1966. Vol. 115. P. 153–188.
29. Лионс Ж.-Л. Некоторые методы решения нелинейных краевых задач. Москва: Мир, 1972. 587 с.
30. Лионс Ж.-Л. Оптимальное управление системами, описываемыми уравнениями с частными производными. Москва: Мир, 1972. 414 с.
31. Киндерлерер Д., Стампаккья Г. Введение в вариационные неравенства и их приложения. Москва: Мир, 1983. 256 с.
32. Байокки К., Капело А. Вариационные и квазивариационные неравенства. Москва: Наука, 1988. 448 с.
33. Вайнберг М.М. Вариационный метод и метод монотонных операторов в теории нелинейных уравнений. Москва: Наука, 1972. 416 с.
34. Иваненко В.И., Мельник В.С. Вариационные методы в задачах управления для систем с распределенными параметрами. Киев: Наук. думка, 1988. 285 с.
35. Moreau J. J. Fonctionnelles convexes. Sem. sur les Equations aux Derivees Partielles, College de France. 1967.
36. Brezis H. Problemes unilateraux. J. Math. Pures Appl. 1972. Vol. 51. P. 1–168.

37. Patriksson M. Nonlinear programming and variational inequality problems: A unified approach. Dordrecht: Kluwer Academic Publishers, 1999. xiv + 334 p.
38. Гловински Р., Лионс Ж.-Л., Тремольер Р. Численное исследование вариационных неравенств. Москва: Мир, 1979. 576 с.
39. Экланд И., Темам Р. Выпуклый анализ и вариационные проблемы. Москва: Мир, 1979. 399 с.
40. Обен Ж.-П., Экланд И. Прикладной нелинейный анализ. Москва: Мир, 1988. 510 с.
41. Bensoussan A. Points de Nash dans le cas de fontionnelles quadratiques et jeux differentiels lineaires a N personnes. SIAM Journal on Control. 1974. N. 12. P. 460–499.
42. Dafermos S. Traffic equilibria and variational inequalities. Transportation Science. 1982. Vol. 16. P. 231–240.
43. Harker P. T., Pang J.-S. Finite-dimensional variational inequality and nonlinear complementarity problems: a survey of theory, algorithms and applications. Math. Programming. 1990. N. 48. P. 161–220.
44. Polyak R. Finding Nonlinear Production–Consumption Equilibrium. arXiv preprint arXiv:2204.04496. 2022.
45. Nagurney A. Network economics: A variational inequality approach. Dordrecht: Kluwer Academic Publishers, 1999. 325 p.
46. Sandrakov G. V. Homogenization of variational inequalities for non-linear diffusion problems in perforated domains. Izvestiya Mathematics. 2005. Vol. 69. Issue 5. P. 1035–1059.
47. Дюво Г., Лионс Ж.-Л. Неревенства в механике и физике. Москва: Наука, 1980. 382 с.
48. Brezis H., Sibony M. Equivalence de deux inequations variationnelles et applications. Archive Rat. Mech. Anal. 1971. Vol. 41. P. 254–265.

49. Lewy H., Stampacchia G. On the regularity of the solution of variational ineqaulity. Pure Appl. Math. 1969. Vol. 22. P. 153–188.
50. Мосолов П. П., Мясников В. П. Вариационные методы в теории вязкопластических сред. ПММ. 1965. Т. 29. № 3. С. 468–492.
51. Nemirovski A. 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 J. on Optim. 2004. Vol. 15. P. 229–251.
52. Wang J.-K., Abernethy J., Levy K. Y. No-Regret Dynamics in the Fenchel Game: A Unified Framework for Algorithmic Convex Optimization. arXiv preprint arXiv:2111.11309. 2021.
53. Goodfellow I., Pouget-Abadie J., Mirza M., Xu B., Warde-Farley D., Ozair Sh., Courville A., Bengio Y. Generative Adversarial Networks. Advances in Neural Information Processing Systems. 2014. P. 2672–2680.
54. Arjovsky M., Chintala S., Bottou L. Wasserstein GAN. arXiv preprint arXiv:1701.07875. 2017.
55. Gidel G., Berard H., Vincent P., Lacoste-Julien S. A Variational Inequality Perspective on Generative Adversarial Networks. arXiv preprint arXiv:1802.10551. 2018.
56. Daskalakis C., Ilyas A., Syrgkanis V., Zeng H. Training GANs with optimism. arXiv preprint arXiv:1711.00141. 2018.
57. Chavdarova T., Gidel G., Fleuret F., Lacoste-Julien S. Reducing Noise in GAN Training with Variance Reduced Extragradient. arXiv preprint arXiv:1904.08598. 2019.
58. Mishchenko K., Kovalev D., Shulgin E., Richtarik P., Malitsky Y. Revisiting Stochastic Extragradient. arXiv preprint arXiv:1905.11373. 2019.
59. Liu M., Zhang W., Mroueh Y., Cui X., Ross J., Yang T., Das P. A decentralized parallel algorithm for training generative adversarial nets. arXiv preprint arXiv:1910.12999. 2019.

60. Browder F. E. Fixed-point theorems for noncompact mappings in Hilbert space. Proc. Natl. Acad. Sci. U.S.A. 1965. 53. P. 1272–1276.

61. Browder F. E. Nonexpansive nonlinear operators in a Banach space. Proc. Natl. Acad. Sci. 1965. 54. P. 1041–1044.

62. Kirk W. A. A fixed point theorem for mappings which do not increase distances. Amer. Math. Monthly. 1965. 72. P. 1004–1006.

63. Göhde D. Zum Prinzip der kontraktiven Abbildung. Math. Nachr. 1965. 30. P. 251–258.

64. Mann W. R. Mean value methods in iteration. Proc. Amer. Math. Soc. 1953. P. 506–510.

65. Красносельский М. А. Два замечания о методе последовательных приближений. УМН. 1955. Том 10. Выпуск 1 (63). С. 123–127.

66. Opial Z. Weak convergence of the sequence of successive approximations for nonexpansive mappings. Bull. Amer. Math. Soc. 1967. Vol. 73. P. 591–597.

67. Halpern B. Fixed points of nonexpanding maps. Bull. Amer. Math. Soc. 1967. Vol. 73. P. 957–961.

68. Lions P.-L. Approximation de points fixes de contractions. C. R. Acad. Sci. Paris, Ser. A-B. 1977. 284. P. A1357–A1359.

69. Wittmann R. Approximation of fixed points of nonexpansive mappings. Arch. Math. 1992. 58. P. 486–491.

70. Xu H. K. Another control condition in an iterative method for nonexpansive mappings. Bull. Austral. Math. Soc. 2002. 65 P. 109–113.

71. Xu H. K. Iterative algorithms for nonlinear operators. J. London Math. Soc. 2002. 2. P. 240-256.

72. Suzuki T. A sufficient and necessary condition for Halpern-type strong convergence to fixed points of nonexpansive mappings. Proc. of the AMS. 2007. V. 135. N. 1. P. 99–106.

73. Семенов В. В. Два методи апроксимації нерухомої точки фейєрівського оператора. Журнал обчисл. та прикл. матем. 2013. № 1 (111). С. 46–56.
74. Nakajo K., Takahashi W. Strong convergence theorems for nonexpansive mappings and nonexpansive semigroups. J. Math. Anal. Appl. 2003. 279. P. 372–379.
75. Takahashi W., Takeuchi Y., Kubota R. Strong convergence theorems by hybrid methods for families of nonexpansive mappings in Hilbert spaces. J. Math. Anal. Appl. 2008. 341. P. 276—286.
76. Konnov I. V. Combined relaxation methods for variational inequalities. Berlin, Heidelberg, New York: Springer–Verlag, 2001. 181 p.
77. Урясьев С. П. Адаптивные алгоритмы стохастической оптимизации и теории игр. Москва: Наука, 1990. 184 с.
78. Goldstein A. A. Convex programming in Hilbert space. Bull. Amer. Math. Soc. 1964. Vol. 70. P. 709–710.
79. Левитин Е. С., Поляк Б. Т. Методы минимизации при наличии ограничений. Ж. вычисл. матем. и матем. физ. 1966. Т. 6. № 5. С. 787–823.
80. Корпелевич Г. М. Экстраградиентный метод для отыскания седловых точек и других задач. Экономика и математические методы. 1976. № 4. С. 747–756.
81. Антипин А. С. О методе выпуклого программирования, использующем симметрическую модификацию функции Лагранжа. Экономика и математические методы. 1976. № 6. С. 1164–1173.
82. Nadezhkina N., Takahashi W. Weak convergence theorem by an extragradient method for nonexpansive mappings and monotone mappings. Journal of Optimization Theory and Applications. 2006. Vol. 128. P. 191–201.
83. Зыкина А. В., Меленьчук Н. В. Двухшаговый экстраградиентный метод для вариационных неравенств. Изв. вузов. Матем. 2010. № 9. С.

82–85.

84. Nesterov Y. Dual Extrapolation and Its Applications to Solving Variational Inequalities and Related Problems. Math. Program. 2007. V. 109. No. 2-3. P. 319–344.
85. Censor Y., Gibali A., Reich S. The subgradient extragradient method for solving variational inequalities in Hilbert space. Journal of Optitmization Theory and Applications. 2011. 148. P. 318–335.
86. Lyashko S. I., Semenov V. V., Voitova T. A. Low-cost modification of Korpelevich's methods for monotone equilibrium problems. Cybernetics and Systems Analysis. 2011. № 4. P. 631–639.
87. Tseng P. A modified forward-backward splitting method for maximal monotone mappings. SIAM Journal on Optimization. 2000. Vol. 38. P. 431–446.
88. Брэгман Л. М. Релаксационный метод нахождения общей точки выпуклых множеств и его применение для решения задач выпуклого программирования. Ж. вычисл. матем. и матем. физ. 1967. №3. С. 620–631.
89. Хоботов Е. Н. О модификации экстраградиентного метода для решения вариационных неравенств и некоторых задач оптимизации. Ж. вычисл. матем. и матем. физ. 1987, том 27. № 10. С. 1462–1473.
90. Marcotte P. Application of Khobotov's algorithm to varational inequalities and network equilibrium problems. Indorm. Systems Oper. Res. 1991. Vol. 29. P. 258–270.
91. Denisov S. V., Semenov V. V., Chabak L. M. Convergence of the Modified Extragradient Method for Variational Inequalities with Non-Lipschitz Operators. Cybernetics and Systems Analysis. 2015. Vol. 51. Issue 5. P. 757–765.
92. Bach F., Levy K. Y. A Universal Algorithm for Variational Inequalities Adaptive to Smoothness and Noise. arXiv preprint arXiv:1902.01637.

2019.

93. Stonyakin F. S., Vorontsova E. A., Alkousa M.S.: New Version of Mirror Prox for Variational Inequalities with Adaptation to Inexactness. In: Jaćimović M., Khachay M., Malkova V., Posypkin M. (eds.) Optimization and Applications. OPTIMA 2019. Communications in Computer and Information Science, vol 1145. Springer, Cham. 2020. P. 427–442.
94. Denisov S. V., Semenov V. V., Stetsyuk P. I. Bregman Extragradient Method with Monotone Rule of Step Adjustment. Cybernetics and Systems Analysis. 2019. Vol. 55. Issue 3. P. 377–383.
95. Denisov S. V., Nomirovskii D. A., Rublyov B. V., Semenov V. V. Convergence of Extragradient Algorithm with Monotone Step Size Strategy for Variational Inequalities and Operator Equations. Journal of Automation and Information Sciences. 2019. Vol. 51. Iss. 6. P. 12–24.
96. Бакушинский А. Б., Гончарский А. В. Некорректные задачи. Численные методы и приложения. Москва: Изд-во МГУ, 1989. 200 с.
97. Nadezhkina N., Takahashi W. Strong convergence theorem by a hybrid method for nonexpansive mappings and Lipschitz-continuous monotone mappings. SIAM J. Optim. 2006. Vol. 16. P. 1230–1241.
98. Censor Y., Gibali A., Reich S. Strong convergence of subgradient extragradient methods for the variational inequality problem in Hilbert space. Optimization Methods and Software. 2011. Vol. 26. P. 827–845.
99. Апостол Р. Я., Гриненко А. А., Семенов В. В. Ітераційні алгоритми для монотонних дворівневих варіаційних нерівностей. Журнал обчисл. та прикл. матем. 2012. № 1 (107). С. 3–14.
100. Verlan D. A., Semenov V. V., Chabak L. M. A Strongly Convergent Modified Extragradient Method for Variational Inequalities with Non-Lipschitz Operators. Journal of Automation and Information Sciences. 2015. Vol. 47. Issue 7. P. 31–46.

101. Vedel Y. I., Denisov S. V., Semenov V. V. Regularized Adaptive Extra-Proximal Algorithm for Equilibrium Problem in Hadamard Spaces. Journal of Automation and Information Sciences. 2020. Vol. 52. Issue 9. P. 12–26.
102. Попов Л. Д. Модификация метода Эрроу-Гурвица поиска седловых точек. Математические заметки. 1980. Т. 28, № 5. С. 777–784.
103. Malitsky Yu. V., Semenov V. V. An extragradient algorithm for monotone variational inequalities. Cybernetics and Systems Analysis. 2014. Vol. 50. P. 271–277.
104. Semenov V. V. A Version of the Mirror descent Method to Solve Variational Inequalities. Cybernetics and Systems Analysis. 2017. Vol. 53. P. 234–243.
105. Semenov V. V. A variant of mirror descent method for solving variational inequalities. In: Polyakova, L. N. (ed.) Constructive Nonsmooth Analysis and Related Topics (dedicated to the memory of V. F. Demyanov). IEEE, 2017. P. 281–284.
106. Denisov S. V., Dudar V. V., Semenov V. V., Vedel Y. I. A New Mirror-prox Algorithm For Variational Inequalities. Журнал обчисл. та прикл. матем. 2017. № 1 (124). С. 15–29.
107. Ведель Я. И., Семенов В. В. Двухэтапный метод с расхождением Брэгмана для решения вариационных неравенств. Журнал обчисл. та прикл. матем. 2018. № 3 (129). С. 18–27.
108. Nomirovskii D. A., Rublyov V. V., Semenov V. V. Convergence of Two-Stage Method with Bregman Divergence for Solving Variational Inequalities. Cybernetics and Systems Analysis. 2019. Vol. 55. P. 359–368.
109. Semenov V. V., Denisov S. V., Kravets A. V. Adaptive Two-Stage Bregman Method for Variational Inequalities. Cybernetics and Systems Analysis. 2021. Vol. 57. Issue 6. P. 959–967.

110. Чабак Л.М., Ведель Я. И., Дударь В. В., Семенов В. В. Конечное число итераций в двухэтапных алгоритмах для решения вариационных неравенств. Журнал обчисл. та прикл. матем. 2017. № 2 (125). С. 91–99.
111. Ведель Я. И., Семенов В. В., Чабак Л. М. О двухэтапном проксимальном алгоритме для решения задачи о равновесии. Журнал обчисл. та прикл. матем. 2019. № 2 (131). С. 23–31.
112. Ведель Я. И., Голубева Е. Н., Семёнов В. В. Конечное число итераций в двухэтапных алгоритмах для решения задач о равновесии. Журнал обчисл. та прикл. матем. 2019. № 3 (132). С. 21–32.
113. Malitsky Yu. Projected Reflected Gradient Methods for Monotone Variational Inequalities. SIAM Journal on Optimization. 2015. Vol. 25. P. 502–520.
114. Dung N. V., Vu B. C. Convergence analysis of the stochastic reflected forward-backward splitting algorithm. arXiv preprint arXiv:2102.08906. 2021.
115. Malitsky Y., Tam M. K. A Forward-Backward Splitting Method for Monotone Inclusions Without Cocoercivity. SIAM J. on Optim. 2020. Vol. 30. P. 1451–1472.
116. Csetnek E. R., Malitsky Y., Tam M. K. Shadow Douglas–Rachford Splitting for Monotone Inclusions. arXiv preprint arXiv: 1903.03393. 2019.
117. Mokhtari A., Ozdaglar A., Pattathil S. A unified analysis of extra-gradient and optimistic gradient methods for saddle point problems: proximal point approach. arXiv preprint arXiv:1901.08511. 2019.
118. Mokhtari A., Ozdaglar A., Pattathil S. Convergence rate of $O(1/k)$ for optimistic gradient and extra-gradient methods in smooth convex-concave saddle point problems. arXiv preprint arXiv:1906.01115. 2020.
119. Kotsalis G., Lan G., Li T. Simple and optimal methods for stochastic variational inequalities, I: Operator extrapolation. arXiv preprint arXiv:

2011.02987. 2020.

120. Semenov V., Denysov S. Convergence of Adaptive Operator Extrapolation Method for Operator Inclusions In Banach Spaces. Proceedings of The Fifth International Workshop on Computer Modeling and Intelligent Systems (CMIS-2022). Zaporizhzhia, Ukraine, May 12, 2022. CEUR Workshop Proceeedings, vol. 3137. 2022. P. 186–199.
121. Popov L. D. On schemes for the formation of a master sequence in a regularized extragradient method for solving variational inequalities. Russian Mathematics. 2004. Vol. 48. Issue 1. P. 67–76.
122. Alber Y., Ryazantseva I. Nonlinear Ill Posed Problems of Monotone Type. Dordrecht: Springer, 2006. 410 p.
123. Alber Y. I. Metric and generalized projection operators in Banach spaces: properties and applications. In: Theory and Applications of Nonlinear Operators of Accretive and Monotone Type, vol. 178. New York: Dekker, 1996. P. 15–50.
124. Iiduka H., Takahashi W. Weak convergence of a projection algorithm for variational inequalities in a Banach space. Journal of Mathematical Analysis and Applications. 2008. Vol. 339, N 1. P. 668–679.
125. Cholamjiak P., Shehu Y. Inertial forward-backward splitting method in Banach spaces with application to compressed sensing. Appl. Math. 2019. Vol. 64. P. 409–435.
126. Shehu Y. Single projection algorithm for variational inequalities in Banach spaces with application to contact problem. Acta Math. Sci. 2020. Vol. 40. P. 1045–1063.
127. Yang J., Cholamjiak P., Sunthrayuth P. Modified Tseng's splitting algorithms for the sum of two monotone operators in Banach spaces. AIMS Mathematics. 2021. Vol. 6, Iss. 5. P. 4873–4900.
128. Denisov S., Semenov V. Convergence of Adaptive Forward-Reflected-Backward Algorithm for Solving Variational Inequalities. Selected

Papers of the II International Scientific Symposium «Intelligent Solutions» (IntSol-2021). Workshop Proceedings, Kyiv–Uzhhorod, Ukraine, September 28-30, 2021. CEUR Workshop Proceedings, vol. 3106. 2021. P. 116–127.

129. Denysov S., Semenov V. Theoretical Bound of the Complexity of Some Extragradient-Type Algorithms for Variational Inequalities in Banach Spaces. Selected Papers of the VIII International Scientific Conference «Information Technology and Implementation» (IT&I-2021). Workshop Proceedings, Kyiv, Ukraine, December 1-3, 2021. CEUR Workshop Proceedings, vol. 3179. 2022. P. 144–155.
130. Semenov V. V., Denisov S. V. Convergence of the Method of Extrapolation from the Past for Variational Inequalities in Uniformly Convex Banach Spaces. Cybernetics and Systems Analysis. 2022. Vol. 58. Iss. 4. P. 564–575.
131. Vedel Y., Semenov V., Denisov S. A Novel Algorithm with Self-adaptive Technique for Solving Variational Inequalities in Banach Spaces. In: Olenev N. N., Evtushenko Y. G., Jaćimović M., Khachay M., Malkova V. (eds.) Advances in Optimization and Applications. OPTIMA 2021. Communications in Computer and Information Science, vol 1514. Springer, Cham, 2021. P. 50–64.
132. Kassay G., Radulescu V. D. Equilibrium Problems and Applications. London: Academic Press, 2019. xx + 419 p.
133. Nikaido H., Isoda K. Note on noncooperative convex games. Pacific Journal of Mathematics. 1955. Vol. 5. P. 807–815.
134. Blum E., Oettli W. From optimization and variational inequalities to equilibrium problems. Math. Stud. 1994. 63. P. 123–145.
135. Muu L. D. and Oettli W. Convergence of an adaptive penalty scheme for finding constrained equilibria. Nonlinear Anal. TMA. 1992. 18. P. 1159–1166.

136. Антипин А. С. Равновесное программирование: проксимальные методы. Журн. вычисл. матем. и матем. физ. 1997. 37, № 11. С. 1327—1339.
137. Antipin A. S. Equilibrium programming: gradient methods. Part 2. Autom. Remote Control. 1997. 58 (8). P. 1337–1347.
138. Antipin A. Equilibrium programming problems: prox-regularization and prox-methods. In: Recent Advances in Optimization. Lecture Notes in Economics and Mathematical Systems, vol. 452, Springer, Heidelberg, 1997. P. 1–18.
139. Antipin A. S., Flam S. D. Equilibrium programming using proximal-like algorithms. Math. Program. 1997. 78. P. 29–41.
140. Mastroeni G. On auxiliary principle for equilibrium problems. In: Daniele, P. et al. (eds.) Equilibrium Problems and Variational Models. Kluwer Academic Publishers, Dordrecht, 2003. P. 289-298.
141. Quoc T. D., Muu L. D., Hien N. V. Extragradient algorithms extended to equilibrium problems. Optimization. 2008. Vol. 57. P. 749–776.
142. Van N. T. T., Strodiot J. J., Nguyen V.H. A bundle method for solving equilibrium problems. Math. Program. 2009. 116 (1–2), Ser. B. P. 529–552.
143. Войтова Т. А., Денисов С. В., Семенов В. В. Сильно збіжний модифікований варіант методу Корпелевич для задач рівноважного програмування. Журнал обчисл. та прикл. матем. 2011. № 1 (104). С. 10–23.
144. Anh P. N. Strong convergence theorems for nonexpansive mappings and Ky Fan inequalities. J. Optim. Theory Appl. 2012. 154. P. 303–320.
145. Vuong P. T., Strodiot J. J, Nguyen V. H. Extragradient methods and linesearch algorithms for solving Ky Fan inequalities and fixed point problems. J. Optim. Theory Appl. 2012. 155. P. 605–627.
146. Quoc T. D., Anh P. N., Muu L. D. Dual extragradient algorithms to equilibrium problems. J. Glob. Optim. 2012. 53. P. 139–159.
147. Vuong P. T., Strodiot J. J., Nguyen V. H.: On extragradient-viscosity methods for solving equilibrium and fixed point problems in a Hilbert

space. Optimization. 2015. 64 (2). P. 429–451.

148. Nguyen T. P. D., Strodiot J. J., Nguyen V. H., Nguyen T. T. V. A family of extragradient methods for solving equilibrium problems. J. Ind. Manag. Optim. 2015. 11. P. 619–630.
149. Anh P. N., Hai T. N., Tuan P. M. On Ergodic Algorithms for Equilibrium Problems. J. Glob. Optim. 2016. 64 (1). P. 179–195.
150. Ведель Я. И., Семенов В. В. Новый двухэтапный проксимальный алгоритм для решения задачи о равновесии. Журнал обчисл. та прикл. матем. 2015. № 1 (118). С. 15–23.
151. Lyashko S. I., Semenov V. V. A New Two-Step Proximal Algorithm of Solving the Problem of Equilibrium Programming. In: B. Goldengorin (ed.) Optimization and Its Applications in Control and Data Sciences. Springer Optimization and Its Applications, vol. 115. Springer, Cham, 2016. P. 315–325.
152. Ведель Я. І., Семенов В. В. Адаптивні алгоритми для задач про рівновагу в просторах Адамара. Доповіді НАН України. 2020. № 8. С. 26–34.
153. Vedel Y. I., Sandrakov G. V., Semenov V. V. An Adaptive Two-Stage Proximal Algorithm for Equilibrium Problems in Hadamard Spaces. Cybernetics and Systems Analysis. 2020. Vol. 56. Issue 6. P. 978–989.
154. Semenov V., Vedel Y. Convergence of adaptive methods for equilibrium problems in Hadamard spaces. Proceedings of the 7th International Conference «Information Technology and Interactions» (IT&I-2020). Workshops Proceedings Kyiv, Ukraine, December 02-03, 2020. CEUR Workshop Proceedings, vol. 2845. 2021. P. 321–335.
155. Вартузова М. В., Семенов В. В., Чабак Л. М. Новый алгоритм с расстоянием Брэгмана для решения задачи о равновесии. Журнал обчисл. та прикл. матем. 2016. № 3 (123) С. 9–21.

156. Chabak L., Semenov V., Vedel Y. A New Non-Euclidean Proximal Method for Equilibrium Problems. In: Chertov O., Mylovanov T., Kondratenko Y., Kacprzyk J., Kreinovich V., Stefanuk V. (eds) Recent Developments in Data Science and Intelligent Analysis of Information. ICDSIAI 2018. Advances in Intelligent Systems and Computing, vol. 836. Springer, Cham, 2019. P. 50–58.
157. Семенов В. В. О сходимости методов решения двухуровневых вариационных неравенств с монотонными операторами. Журнал обчисл. та прикл. матем. 2010. № 1 (100). С. 121–129.
158. Войтова Т. А., Семенов В. В. Метод решения двухэтапных операторных включений. Журнал обчисл. та прикл. матем. 2010. № 3 (102). С. 34–39.
159. Денисов С. В., Семенов В. В. Проксимальний алгоритм для дворівневих варіаційних нерівностей: сильна збіжність. Журнал обчисл. та прикл. матем. 2011. № 3 (106). С. 27–32.
160. Denisov S., Semenov V., Vedel Y. Adaptive Algorithm for Variational Inequality Problem Over the Set of Solutions the Equilibrium Problem. 2020 IEEE 2nd International Conference on Advanced Trends in Information Theory (ATIT), Kyiv, Ukraine, 2020. P. 325–329.
161. Ведель Я. И., Денисов С. В., Семёнов В. В. Алгоритм для вариационного неравенства на множестве решений задачи о равновесии. Журнал обчисл. та прикл. матем. 2020. № 1 (133). С. 18–30.
162. Vedel Y. I., Denisov S. V., Semenov V. V. An Adaptive Algorithm for the Variational Inequality Over the Set of Solutions of the Equilibrium Problem. Cybernetics and Systems Analysis. 2021. Vol. 57. Issue 1. P. 91–100.
163. Colao V., Lopez G., Marino G., Martin-Marquez V. Equilibrium problems in Hadamard manifolds. Journal of Mathematical Analysis and Applications. 2012. Vol. 388. P. 61–77.

164. Khatibzadeh H., Mohebbi V. Monotone and pseudo-monotone equilibrium problems in Hadamard spaces. Journal of the Australian Mathematical Society. 2019. P. 1–23. https://doi.org/10.1017/S1446788719000041.
165. Khatibzadeh H., Mohebbi V. Approximating solutions of equilibrium problems in Hadamard spaces. Miskolc Mathematical Notes. 2019. Vol. 20. No. 1. P. 281–297. https://doi.org/10.18514/MMN.2019.2361.
166. Ведель Я. И., Семёнов В. В., Чабак Л. М. Двухэтапный проксимальный алгоритм для задачи о равновесии в пространстве Адамара. Доповіді НАН України. 2020. № 2. С. 7–14.
167. Vedel Y. I., Sandrakov G. V., Semenov V. V., Chabak L. M. Convergence of a Two-Stage Proximal Algorithm for the Equilibrium Problem in Hadamard Spaces. Cybernetics and Systems Analysis. 2020. Vol. 56. Issue 5. P. 784–792.
168. Vedel Y., Semenov V. Adaptive Extraproximal Algorithm for the Equilibrium Problem in Hadamard Spaces. In: Olenev N., Evtushenko Y., Khachay M., Malkova V. (eds.) Optimization and Applications. OPTIMA 2020. Lecture Notes in Computer Science, vol 12422. Springer, Cham, 2020. P. 287–300.
169. Beck A. First-Order Methods in Optimization. Philadelphia: Society for Industrial and Applied Mathematics, 2017. 479 p.
170. Ryu E. K., Yin W. Large-Scale Convex Optimization. Algorithms and Analyses via Monotone Operators. Cambridge: Cambridge University Press, 2023. 303 p.
171. Rudin L. I., Osher S., Fatemi E. Nonlinear total variation based noise removal algorithms. Physica D: Nonlinear Phenomena. 1992. Vol. 60. N. 1–4. P. 259–268.
172. Chambolle A., Lions P.-L. Image recovery via total variation minimization and related problems. Numerische Mathematik. 1997. Vol. 76. N. 2. P. 167–188.

173. Ben-Tal A., El Ghaoui L., Nemirovski A. Robust optimization. Princeton University Press, 2009. 576 p.
174. Madry A., Makelov A., Schmidt L., Tsipras D., Vladu A. Towards deep learning models resistant to adversarial attacks. arXiv preprint arXiv:1706.06083. 2017.
175. Mainge P.-E. Strong convergence of projected subgradient methods for nonsmooth and nonstrictly convex minimization. Set-Valued Analysis. 2008. Vol. 16. P. 899–912.
176. Khanh P. D., Vuong P. T. Modified projection method for strongly pseudomonotone variational inequalities. J. Glob. Optim. 2014. Vol. 58. P. 341–350.
177. Browder F. Existence and approximation of solutions of nonlinear variational inequalities. Proc. Nat. Acad. Sci. USA. 1966. Vol. 56. No. 4. P. 1080–1086.
178. Browder F. E. Convergence of approximants of fixed points of nonexpansive non-linear mappings in Banach spaces. Arch. Rational Mech. Anal. 1967. Vol. 24. P. 82–90.
179. Combettes P. L., Hirstoaga S. A. Equilibrium Programming in Hilbert Spaces. J. Nonlinear Convex Anal. 2005. Vol. 6. P. 117-–136.
180. Gorbunov E., Loizou N., Gidel G. Extragradient Method: $O\left(1/K\right)$ Last-Iterate Convergence for Monotone Variational Inequalities and Connections With Cocoercivity. arXiv preprint arXiv: 2110.04261. 2021.
181. Yoon T., Ryu E. K. Accelerated algorithms for smooth convex-concave minimax problems with $O\left(1/k^2\right)$ rate on squared gradient norm. Proceedings of the 38th International Conference on Machine Learning. Proceedings of Machine Learning Research. 2021. Vol. 139. P. 12098–12109.
182. Vishnoi N. K. Algorithms for Convex Optimization. Cambridge: Cambridge University Press, 2021. 340 p.

183. Juditsky A., Nemirovski A., Tauvel C. Solving variational inequalities with Stochastic Mirror-Prox algorithm. Stochastic Systems. 2011. Vol. 1. No. 1. P. 17–58.

184. Allen-Zhu Z., Orecchia L. Linear coupling: An ultimate unification of gradient and mirror descent. arXiv preprint arXiv:1407.1537. 2014.

185. Cohen M. B., Sidford A., Tian K. Relative Lipschitzness in Extragradient Methods and a Direct Recipe for Acceleration. arXiv preprint arXiv: 2011.06572. 2021.

186. Beauzamy B. Introduction to Banach Spaces and Their Geometry. Amsterdam: North-Holland, 1985. 307 p.

187. Aoyama K., Kohsaka F. Strongly relatively nonexpansive sequences generated by firmly nonexpansive-like mappings. Fixed Point Theory Appl. 2014. 95. https://doi.org/10.1186/1687-1812-2014-95.

188. Xu H. K. Inequalities in Banach spaces with applications. Nonlinear Anal. 1991. Vol. 16. Iss. 12. P. 1127–1138.

189. Duchi J., Shalev-Shwartz S., Singer Y., Chandra T. Efficient projections onto the ℓ_1-ball for learning in high dimensions. In A. McCallum and S. Roweis, editors, Proc. of the 25th Int. Conf. Machine Learning (ICML'08). Helsinki, Finland, July 5–9, 2008. P. 272–279.

190. Ramirez J., Sukumaran R., Bertrand Q., Gidel G. Omega: Optimistic EMA Gragients. arXiv preprint arXiv:2306.07905. 2023.

191. Kingma D. P., Ba J. Adam: A Method for Stochastic Optimization. arXiv preprint arXiv:1412.6980. 2014.

ДОДАТОК 1. СПИСОК ПУБЛІКАЦІЙ ЗДОБУВАЧА ЗА ТЕМОЮ ДИСЕРТАЦІЇ ТА ВІДОМОСТІ ПРО АПРОБАЦІЮ РЕЗУЛЬТАТІВ ДИСЕРТАЦІЇ

Список публікацій здобувача за темою дисертації

Статті у наукових фахових виданнях:

1. Semenov V. V., Denisov S. V., Sandrakov G. V., Kharkov O. S. Convergence of the Operator Extrapolation Method for Variational Inequalities in Banach Spaces. Cybernetics and Systems Analysis. 2022. Vol. 58. Issue 5. P. 740–753.
2. Харьков О. С. Оцінки ефективності для методів з дивергенцією Брегмана. Журнал обчисл. та прикл. матем. 2023. № 2. С. 83–93.
3. Семенов В. В., Харьков О. С. Алгоритм екстраполяції з минулого для варіаційних нерівностей в гільбертовому просторі. Журнал обчисл. та прикл. матем. 2023. № 2. С. 52–82.
4. Семенов В. В., Харьков О. С. Регуляризований алгоритм операторної екстраполяції. Журнал обчисл. та прикл. матем. 2023. № 1. С. 15–27.
5. Семенов В., Харьков О. Метод операторної екстраполяції для варіаційних нерівностей в банахових просторах. Фізико-математичне моделювання та інформаційні технології. 2023. Вип. 37. С. 118–122.
6. Денисов С. В., Семенов В. В., Харьков О. С. Слабка збіжність методу операторної екстраполяції для варіаційних нерівностей в рівномірно опуклих банахових просторах. Журнал обчисл. та прикл. матем. 2022. № 2. С. 42–49.

7. Семенов В. В., Сірик Д. С., Харьков О. С. Збіжність методу операторної екстраполяції. Доповіді НАН України. 2021. № 4. С. 28–35.
8. Семенов В. В., Сирык Д. С., Харьков О. С. Адаптивный метод операторной экстраполяции. Фізико-математичне моделювання та інформаційні технології. 2021. Вип. 33. С. 143–147.
9. Семенов В. В., Денисов С. В., Сирык Д. С., Харьков О. С. Сходимость метода экстраполяции из прошлого и метода операторной экстраполяции. Міжнародний науково-технічний журнал «Проблеми керування та інформатики». 2021. Том 66. № 3. С. 58–72.
10. Харьков О. С., Ведель Я. І., Семенов В. В. Методи для задач векторного узагальненого оптимального керування системами з розподіленими параметрами. Журнал обчисл. та прикл. матем. 2020. № 2 (134). С. 71–98.

Матеріали конференцій:

11. Kharkov O., Semenov V. The regularized operator extrapolation method. Information Technology and Implementation (Satellite): Conference Proceedings, November 21, 2023, Kyiv, Ukraine. P. 269–270.
12. Харьков О., Семенов В. Швидкість збіжності нових алгоритмів для варіаційних нерівностей. Modeling, control and information technologies: Proceedings of VI International scientific and practical conference. P. 159–160. https://doi.org/10.31713/MCIT.2023.048.
13. Семенов В. В., Харьков О. С. Швидкість збіжності методу операторної екстраполяції. Матеріали Всеукраїнської наукової конференції «Сучасні проблеми прикладної математики та комп'ютерних наук», присвяченої 50-річчю кафедри теорії оптимальних процесів, 7-9 листопада 2023, Львів. С. 203–205.
14. Semenov V., Kharkov O. The regularized operator extrapolation algorithm for monotone variational inequalities. Intelligent Solutions–S: Proceedings of the International Symposium, September 28, 2023, Kyiv–

Uzhorod, Ukraine. P. 88–90.

15. Семенов В., Харьков О. Лінійна швидкість збіжності алгоритмів екстраполяції з минулого та операторної екстраполяції для варіаційних нерівностей. Математика та інформаційні технології. Матеріали міжнародної наукової конференції, присвяченої 55-річчю факультету математики та інформатики, 28-30 вересня 2023 р. — Чернівці: Чернівецький нац. ун-т, 2023. С. 301–302.
16. Kharkov O. S., Semenov V. V. The regularized operator extrapolation method for variational inequalities and saddle point problems. XXXVIII International Conference «Problems of decision making under uncertainties (PDMU-2023)». Abstracts. September 11-15, 2023, Polyana, Ukraine. P. 55.
17. Semenov V., Denysov S., Kharkov O. About weak convergence of the operator extrapolation method. XXXVII International Conference «Problems of decision making under uncertainties (PDMU-2022)». Abstracts. November 23-25, 2022, Sheki–Lankaran, Republic of Azerbaijan. P. 98.
18. Semenov V., Siryk D., Kharkov O. About convergence of the operator extrapolation method. XXXVI International Conference «Problems of decision making under uncertainties (PDMU-2021)». Abstracts. May 11-14, 2021, Skhidnytsia, Ukraine. P. 91.
19. Denisov S. V., Kharkov O., Semenov V., Vedel Ya. About regularized adaptive extra-proximal algorithm for equilibrium problems in Hadamard spaces. XXXV International Conference «Problems of decision making under uncertainties (PDMU-2020)». Abstracts. May 11-15, 2020, Baku–Sheki, Republic of Azerbaijan. P. 32.

Апробація результатів дисертації

Матеріали дисертаційної роботи доповідались та обговорювались на наукових конференціях:

— X International Conference «Information Technology and Implementation – 2023» (Київ, 20-21 листопада 2023);
— 18th International Conference «Mathematical Modeling and Simulation Systems» (MODS2023) (Чернігів, 13-15 листопада 2023);
— VI International Scientific-Practical Conference «Modeling, control and information technology» (Рівне, 9-11 листопада 2023);
— XXVII Всеукраїнська наукова конференція «Сучасні проблеми прикладної математики та комп'ютерних наук» (Львів, 7-9 листопада 2023);
— Міжнародна наукова конференція «Математика та інформаційні технології», присвячена 55-річчю факультету математики та інформатики (Чернівці, 28-30 вересня 2023);
— III International Scientific Symposium «Intelligent Solutions» (Satellite) (Київ–Ужгород, 28 вересня 2023);
— Міжнародний науковий симпозіум «Питання оптимізації обчислень (ПОО–XLVIII)», який присвячений 100-річчю від дня народження академіка В. М. Глушкова (Львів, 19-22 вересня 2023);
— XXXVIII International Conferences «Problems of decision making under uncertainties» (Поляна, 11-15 вересня 2023);
— XXXVII International Conferences «Problems of decision making under uncertainties» (Шекі–Ленкорань, Азербайджан, 23-25 листопада 2022);
— IX Міжнародна конференція «Обчислювальна та прикладна математика», присвячена 100-річчю академіка Івана Івановича Ляшка (Київ, 10-11 жовтня 2022);
— Міжнародна наукова конференція «Питання оптимізації обчислень (ПОО–XLVII)», присвячена 30-річчю незалежності України (Львів, 21-24 вересня 2021);
— XXXVI International Conferences «Problems of decision making under uncertainties» (Східниця, 11-14 травня 2021);

— XXXV International Conferences «Problems of decision making under uncertainties» (Баку–Шекі, Азербайджан, 11-15 травня 2020).

А також доповідались та обговорювались на таких семінарах:

— науковий семінар кафедри обчислювальної математики факультету комп'ютерних наук та кібернетики Київського національного університету імені Тараса Шевченка (Київ, 2023).