Проективні методи для монотонних варіаційних нерівностей та задачі оптимізації

Маліцький Юрій Валерійович

Київський національний університет імені Тараса Шевченка
факультет кібернетики
кафедра обчислювальної математики

16 грудня 2014 р.

Дисертація на здобуття наукового ступеня кандидата фізико-математичних наук за спеціальністю 01.05.02 — математичне моделювання та обчислювальні методи

Науковий керівник: д. ф.-м.н., проф. Семенов В.В.

Структура

Вступ

$H, \langle \cdot, \cdot \rangle$ — гільбертовий простір $\quad C \subseteq H$ — замкнена опукла множина

$F \colon H \to H$ називається

1. монотонним, якщо $\langle F(x) - F(y), x - y \rangle \geq 0$
2. L –ліпшицевим, якщо $\|F(x) - F(y)\| \leq L \|x - y\|$

Завдання: розв'язати варіаційну нерівність (VIP)

знайти $x^* \in C$ такий, що $\quad \langle F(x^*), x - x^* \rangle \geq 0 \quad \forall x \in C$

Важливість

До VIP відноситься

- мінімізація $\min_{x \in C} f(x)$

$$\langle \nabla f(x^*), x - x^* \rangle \geq 0 \quad \forall x \in C$$

- пошук сідлових точок

$$\min_x \max_y f(x, y)$$

- задачі на рівновагу
- задачі матфізики, теорії керування

Класичні проективні алгоритми

Екстраґрадієнтний метод (Корпелевич)

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

Метод Ценга

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

Метод Попова

$$\begin{cases} x_{n+1} = P_C(x_n - \lambda F(y_n)) \\ y_{n+1} = P_C(x_{n+1} - \lambda F(y_n)), \end{cases}$$

2.2. Метод Попова

$$\begin{cases} x_{n+1} = P_C(x_n - \lambda F(y_n)) \\ y_{n+1} = P_C(x_{n+1} - \lambda F(y_n)), \end{cases}$$

- доведена збіжність у нескінченновимірному просторі
- покращення довжини кроку: замість $\lambda \in (0, \frac{1}{3L}]$ стало $\lambda \in (0, \frac{\sqrt{2}-1}{L})$

2.3. Субґрадієнтний метод Попова

Маліцький, Семенов (2014)

$$
\begin{cases}
T_n = \{w \in H \mid \langle x_n - \lambda F(y_{n-1}) - y_n, w - y_n \rangle \leq 0\} \\
x_{n+1} = P_{T_n}(x_n - \lambda F(y_n)) \\
y_{n+1} = P_C(x_{n+1} - \lambda F(y_n)),
\end{cases}
$$

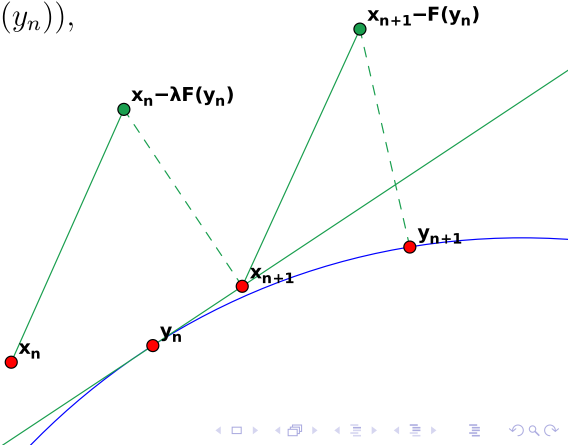
де $\lambda \in (0, \frac{\sqrt{2}-1}{L})$

2.3. Субґрадієнтний метод Попова

Маліцький, Семенов (2014)

$$\begin{cases} T_n = \{w \in H \mid \langle x_n - \lambda F(y_{n-1}) - y_n, w - y_n \rangle \le 0\} \\ x_{n+1} = P_{T_n}(x_n - \lambda F(y_n)) \\ y_{n+1} = P_C(x_{n+1} - \lambda F(y_n)), \end{cases}$$

де $\lambda \in (0, \frac{\sqrt{2}-1}{L})$



2.4. Узагальнений метод Попова

$$\begin{cases} x_{n+1} = \Pi_C\big((\nabla f)^{-1}(\nabla f(x_n) - \lambda F(y_n))\big), \\ y_{n+1} = \Pi_C\big((\nabla f)^{-1}(\nabla f(x_{n+1}) - \lambda F(y_n))\big), \end{cases}$$

де брегманівська відстань:

$$D(x, y) := f(x) - f(y) - \langle \nabla f(y), x - y \rangle$$

брегманівська проекція точки x на C:

$$\bar{x} = \Pi_C(x) := \arg\min_{y \in C} D(y, x)$$

Переваги:
Для деяких множин зручніше рахувати Π_C, ніж P_C

2.5. Гібридний алгоритм

Сильна збіжність?
Маліцький, Семенов (2014)

$$\begin{cases} z_{n+1} = P_C(x_n - \lambda F(z_n)) \\ \left. \begin{array}{l} C_n = \dots \\ Q_n = \dots \end{array} \right\} \text{явно задані напівпростори} \\ x_{n+1} = P_{C_n \cap Q_n} x_0 \end{cases}$$

Переваги: одне значення F та одна проекція P_C

2.6. Циклічний алгоритм

Оператор $T\colon H \to H$ фейєрівський, якщо

$$\operatorname{Fix} T := \{x \in H : Tx = x\} \neq \varnothing$$
$$\|Tx - z\| \leqslant \|x - z\| \quad \forall z \in \operatorname{Fix} T$$

Завдання:

$$\langle F(x^*), x - x^* \rangle \geqslant 0 \quad \forall x \in C,$$

де $C = \cap_i \operatorname{Fix} T_i$ і F — сильно монотонний

Алгоритм:

$$\begin{cases} y_n = \lambda x_n + (1-\lambda) T_{[n]} x_n \\ x_{n+1} = y_n - \alpha_n F(y_n), \end{cases}$$

2.7. Проективні відбиваючі алгоритми

$$x_{n+1} = P_C(x_n - \lambda F(2x_n - x_{n-1})),$$

де $\lambda \in (0, \frac{\sqrt{2}-1}{L})$

2.7. Проективні відбиваючі алгоритми

$$x_{n+1} = P_C(x_n - \lambda F(2x_n - x_{n-1})),$$

де $\lambda \in (0, \frac{\sqrt{2}-1}{L})$

Переваги:

- простота
- одна проекція P_C
- одне значення F
- можливість ефективно визначати довжину кроку λ

F — сильно монотонний $\Longrightarrow$ швидкість збіжності (x_n) — принаймні R –лінійна

Нестаціонарний проективний відбиваючий алгоритм

1. Обираємо $\alpha \in (0, \sqrt{2} - 1)$, x_0, x_1.
2. Для даних x_n, x_{n-1} y_{n-1}:

   2.1 обчислюємо y_n і λ_n:

$$y_n = 2x_n - x_{n-1}$$

$$\lambda_n = \frac{\alpha \|y_n - y_{n-1}\|}{\|F(y_n) - F(y_{n-1})\|}$$

   2.2 обчислюємо x_{n+1}:

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

Структурна оптимізація

Завдання:

$$\min_{x \in \mathbb{R}^d} f(x) + g(x)$$

- f — опукла диференційована
- g — опукла

Структурна оптимізація

Завдання:

$$\min_{x \in \mathbb{R}^d} f(x) + g(x)$$

- f — опукла диференційована
- g — опукла

Включає:

- умовну оптимізацію
- регуляризацію
- представлення множини за допомогою бар'єрної функції

Проксимальні методи

Проксимальний оператор:

$$\operatorname{prox}_g x = \operatorname{argmin}_y \left\{ g(y) + \frac{1}{2} \|x - y\|^2 \right\}$$

Проксимальні методи

Проксимальний оператор:

$$\operatorname{prox}_g x = \operatorname{argmin}_y \left\{ g(y) + \frac{1}{2} \|x - y\|^2 \right\}$$

Проксимальний градієнтний метод:

$$x_{n+1} = \operatorname{prox}_{\lambda g}(x_n - \lambda \nabla f(x_n)),$$

де $\lambda \in (0, \frac{1}{L})$,

L — консанта Ліршиця ∇f

зазвичай невідома

Пошук підходящого кроку

Для даних x_n, λ_{n-1}

`repeat`

$z = \mathrm{prox}_{\lambda g}(x_n - \lambda \nabla f(x_n))$

`break if`

$$f(z) \le f(x_n) + \langle \nabla f(x_n), z - x_n \rangle + \frac{1}{2\lambda} \|x_n - z\|^2$$

`update` $\lambda := 0.5\lambda$

`return` $\lambda_n := \lambda$, $x_{n+1} := z$

Пошук підходящого кроку

Для даних x_n, λ_{n-1}

```
repeat
  z = prox_{λg}(x_n − λ∇f(x_n))
  break if
          f(z) ≤ f(x_n) + ⟨∇f(x_n), z − x_n⟩ + 1/(2λ) ‖x_n − z‖²
  update λ := 0.5λ
return λ_n := λ, x_{n+1} := z
```

repeat

$z = \mathrm{prox}_{\lambda g}(x_n - \lambda \nabla f(x_n))$

break if

$$f(z) \le f(x_n) + \langle \nabla f(x_n), z - x_n \rangle + \frac{1}{2\lambda} \|x_n - z\|^2$$

update $\lambda := 0.5\lambda$

return $\lambda_n := \lambda$, $x_{n+1} := z$

Недоліки:

- вимагає $f(z)$ і $\mathrm{prox}_{\lambda g}$
- (λ_n) — незростаюча

3.1. Відбиваючий проксимальний алгоритм

Обираємо $\alpha \in (0, \sqrt{2}-1)$, $\theta \in (0, \frac{1}{2})$ і велике $\bar{\lambda} > 0$.

1: Для даних x_n, x_{n-1}, y_{n-1} визначаємо

$$\lambda(y) := \min\left\{ \frac{\alpha \left\| y - y_{n-1} \right\|}{\left\| \nabla f(y) - \nabla f(y_{n-1}) \right\|}, (2-\theta)\lambda_{n-1}, \bar{\lambda} \right\}$$

Знаходимо будь-яке $\tau_n \in \left(0, \frac{\theta}{2}\right]$ таке, що

$$\begin{aligned} y_n &= x_n + \tau_n (x_n - x_{n-1}) \\ \lambda_n &= \lambda(y_n) \geq \tau_n \lambda_{n-1} \end{aligned}$$

2: Обчислюємо x_{n+1}

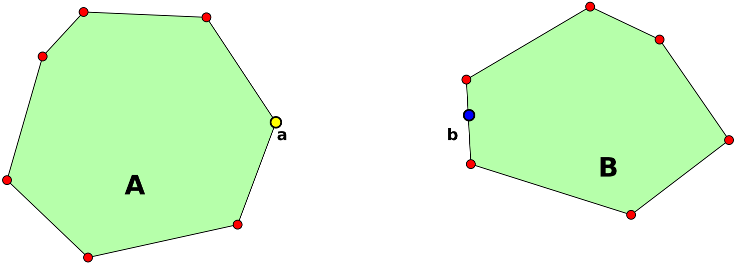
$$x_{n+1} = \operatorname{prox}_{\lambda_n g}(x_n - \lambda_n \nabla f(y_n))$$

3.2. Пошук найближчої пари

$A, B \subset \mathcal{H}$ — непорожні опуклі замкнені множини

Задача: знайти $a \in A$ і $b \in B$ такі, що:

$$||a - b|| = d(A, B) = \inf_{x \in A, y \in B} ||x - y||.$$



Нове формулювання

Нехай $A = \bigcap_{i=1}^{m} A_i \neq \varnothing, \quad B = \bigcap_{i=1}^{m} B_i \neq \varnothing$
Задача еквівалентна до

$$0 \in F(z) + \sum_{i=1}^{m} N_{A_i \times B_i}(z),$$

де

$$z = \begin{pmatrix} x \\ y \end{pmatrix}, \quad F(z) = \begin{pmatrix} x - y \\ y - x \end{pmatrix}$$

Отже, можна застосувати алгоритм Дугласа-Речфорда!

4. Чисельні експерименти.
4.1. Варіаційні нерівності

Приклад 1. Нехай $C = \mathbb{R}^m$ і $F(x) = Ax$, де $A \in \mathbb{R}^{m\times m}$

$$A = \begin{pmatrix} 0 & 0 & \dots & 0 & -1 \\ 0 & 0 & \dots & -1 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 1 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \end{pmatrix}$$

$$\langle Ax^*, x - x^* \rangle \geq 0 \quad \forall x \in C$$

Розв'язок: $x^* = (0, 0, \dots, 0)$

4.1. Варіаційні нерівності

$$\lambda = 0.4, \; e = 10^{-3}$$

m	EGM		SubEGM		SubPopov		RPGM	
	iter.	time	iter	time	iter	time	iter.	time
500	128	0.4	128	0.3	90	0.1	91	0.1
1000	132	1.4	132	1.3	93	0.5	94	0.5
2000	137	5.9	137	5.8	96	2.1	97	2.2
4000	142	32.0	142	32.0	99	11.7	100	11.9
8000	147	350.3	147	345.8	102	118.9	103	118.6

4.1. Варіаційні нерівності

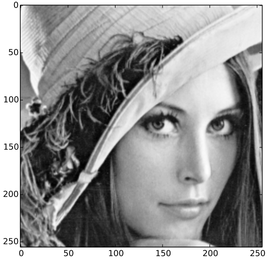
$\lambda = 0.4,\ e = 10^{-3}$

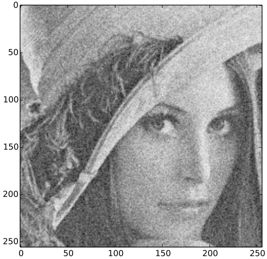
m	EGM		SubEGM		SubPopov		RPGM	
	iter.	time	iter	time	iter	time	iter.	time
500	128	0.4	128	0.3	90	0.1	91	0.1
1000	132	1.4	132	1.3	93	0.5	94	0.5
2000	137	5.9	137	5.8	96	2.1	97	2.2
4000	142	32.0	142	32.0	99	11.7	100	11.9
8000	147	350.3	147	345.8	102	118.9	103	118.6

$\lambda = 0.7,\ e = 10^{-3}$

m	EGM		SubEGM	
	iter.	time	iter.	time
500	69	0.3	69	0.2
1000	71	0.8	71	0.7
2000	73	3.1	73	3.0
4000	76	16.7	76	16.8
8000	78	142.0	78	138.2

4.1. Варіаційні нерівності. Обробка зображень.





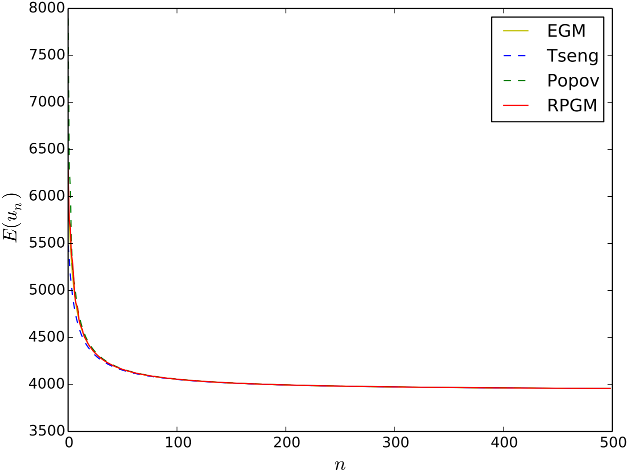
ROF модель:

$$\min_{u \in X} E(u) := \|Au\|_{2,1} + \frac{\gamma}{2} \|u - f\|^2$$

$\Leftrightarrow$

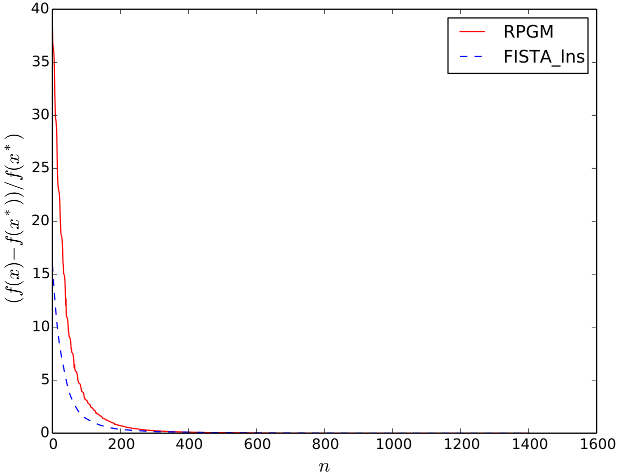
$$\min_{u \in X} \max_{p \in Y} \langle Au, p \rangle_Y + \frac{\gamma}{2} \|u - f\|^2 - \delta_K(p)$$

Обробка зображень



4.2. Оптимізація. Мінімізація

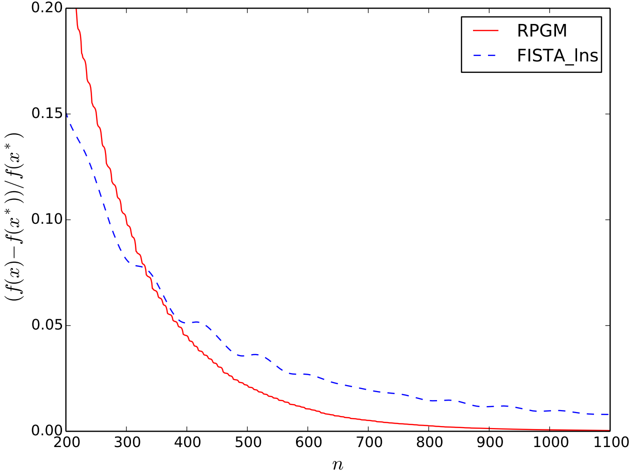
$$f(x) = \frac{1}{2}\|x\|^2 + \sum_{i=1}^{m} q_i(e^{x_i} - x_i - 1) \to \min$$



4.2. Оптимізація

Мінімізація

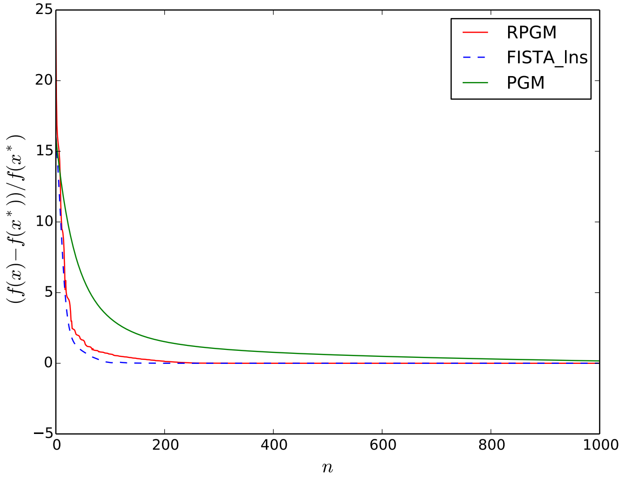
$$f(x) = \frac{1}{2}\|x\|^2 + \sum_{i=1}^{m} q_i(e^{x_i} - x_i - 1) \to \min$$



4.2. Оптимізація

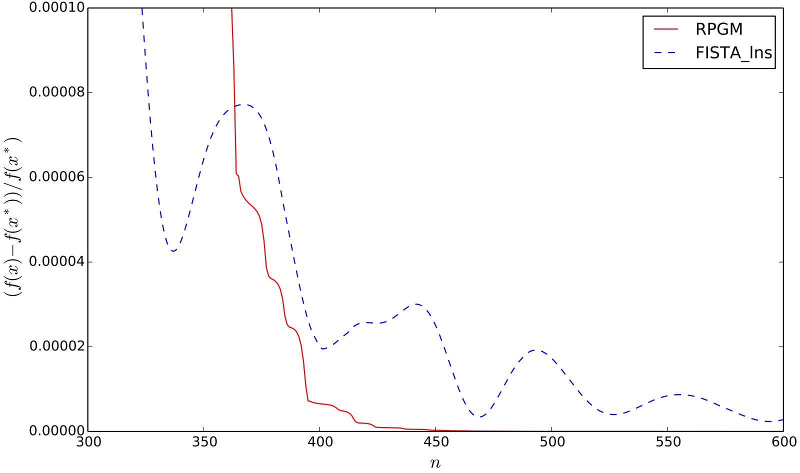
«пошук базису» (basis pursuit)

$$\frac{1}{2}\left\|Ax-b\right\|^2+\gamma\left\|x\right\|_1\to\min$$



4.2. Оптимізація «пошук базису» (basis pursuit)

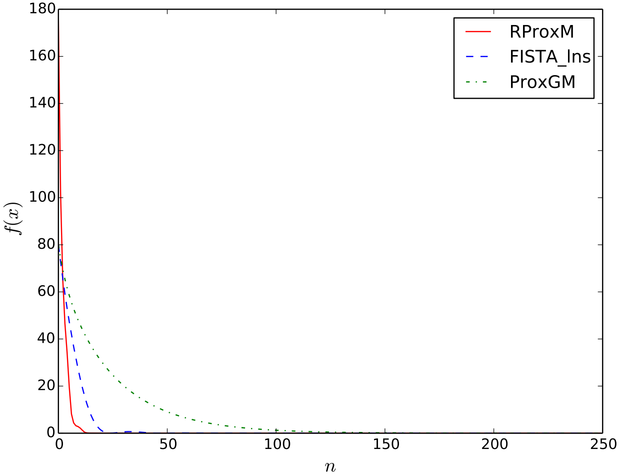
$$\frac{1}{2}\left\|Ax-b\right\|^2+\gamma\left\|x\right\|_1\to\min$$



4.2. Оптимізація

Мінімізація – 2

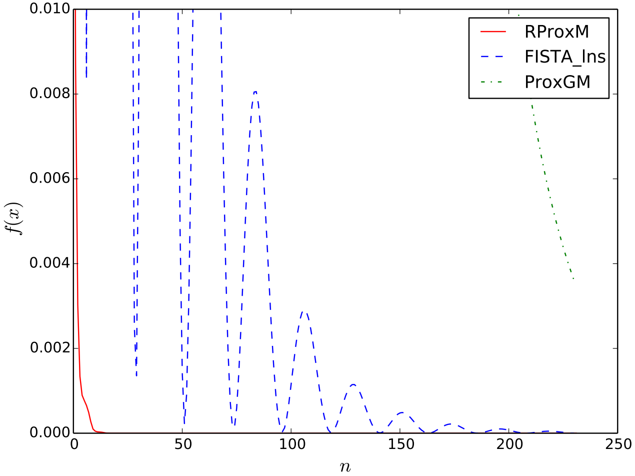
$$f(x) = \frac{L-\mu}{8}\left(x_1^2 + \sum_{i=1}^{m-1}(x_i - x_{i+1})^2\right) + \frac{\mu}{2}\left\|x\right\|^2 \to \min$$



4.2. Оптимізація

Мінімізація – 2

$$f(x) = \frac{L-\mu}{8}\left(x_1^2 + \sum_{i=1}^{m-1}(x_i - x_{i+1})^2\right) + \frac{\mu}{2}\|x\|^2 \to \min$$



4.3. Задача пошуку найближчої пари

Многогранники:

d	N	Цикл.	Парал.	Д.-Р.	Гібр.
2	10	0.2	0.1	>50	0.1
	100	0.4	0.7	>50	1.6
	1000	0.7	11	>50	13
20	10	0.2	0.7	9	0.2
	100	1.0	3.4	>50	2.4
	1000	3.6	17.8	>50	26
200	10	6.5	4.1	>50	0.4
	100	>50	>50	>50	12

Кулі:

d	N	Цикл.	Парал.	Д.-Р.	Гібр.
2	10	6	14	>50	0.1
2	10	25	0.4	>50	0.2
2	100	>50	4	>50	3
10	50	>50	25	>50	0.6

Висновки

- побудовано і доведено збіжність алгоритмів для VIP:
  - узагальнений алгоритм Попова;
  - субґрадієнтний алгоритм Попова;
  - відбиваючий метод;
  - нестаціонарний відбиваючий метод;
  - гібридний метод;
  - циклічний метод.
- одержано нові проксимальні методи для задачі структурної оптимізації;
- одержано варіант алгоритму Дугласа-Речфорда для задачі пошуку найближчої пари;
- проілюстровано чисельними експериментами.

Публікації та тези конференцій

1. Маліцький Ю.В. Нові теореми сильної збіжності проксимального методу для задачі рівноважного програмування/ Маліцький Ю.В., Семенов В.В. // Журнал обчислювальної та прикладної математики. – 2010р. – №3. – С.79-88.
2. Малицкий Ю.В. О проксимальном методе для задачи о равновесии/ Малицкий Ю.В., Семенов В.В. // Конференція «Прийняття рішень в умовах невизначеності». – Ялта, 2010. – С. 96-97.
3. Маліцький Ю.В. Апроксимація нерухомої точки неперервної напівгрупи нерозтягуючих операторів/ Маліцький Ю.В., Семенов В.В. // Конференція «Обчислювальна та прикладна математика». – Київ, 2011. – С. 105.
4. Войтова Т. Методи регуляризації та декомпозиції варіаційних задач/ Войтова Т., Маліцький Ю., Семенов В. // Міжнародна молодіжна математична школа «Питання оптимізації обчислень». – Київ, 2011. – С. 30-31.
5. Маліцький Ю.В. Пошук нерухомої точки ліпшицевої напівгрупи нерозтягуючих операторів/ Маліцький Ю.В. // Журнал обчислювальної та прикладної математики. – 2012р. – №1. – С.35-39.

Публікації та тези конференцій

6. Malitsky Y. The approximation of common fixed point of finite number of Fejer mappings in Hilbert space/ Malitsky Y. // Конференція «Обчислювальна та прикладна математика». – Київ, 2012. – С. 22.
7. Малицкий Ю.В. Схема внешних аппроксимаций для вариационных неравенств на множестве неподвижных точек фейеровских операторов/ Малицкий Ю.В., Семенов В.В. // Доповіді НАН України. – 2013р. – № 7. – С. 47-52.
8. Маліцький Ю.В. Новый экстраградиентный метод для вариационных неравенств/ Маліцький Ю.В., Семенов В.В. // Конференція «Сучасна інформатика, проблеми досягнення та перспективи розвитку». – Київ, 2013. – С.156-157.
9. Malitsky Yu.V. A Variant of Tseng's Splitting Method for Monotone Inclusion Problem/ Malitsky Yu.V. // Конференція «Обчислювальна та прикладна математика». – Київ, 2013. – С. 30.
10. Маліцький Ю.В. Новий екстраґрадієнтний метод для варіаційних нерівностей/ Маліцький Ю.В., Семенов В.В. // Конференція «Прийняття рішень в умовах невизначеності». – Східниця, 2013. – С. 165.

Публікації та тези конференцій

11. Malitsky Yu.V. A hybrid method without extrapolation step for solving variational inequality problems/ Malitsky Yu.V., Semenov V.V. // Journal of Global Optimization. – 2014р. – DOI:10.1007/s10898-014-0150-x – С. 1-10.
12. Малицкий Ю.В. Новый экстраградиентный метод для вариационных неравенств/ Малицкий Ю.В., Семенов В.В. // Кибернетика и системный анализ. – 2014р. – № 2. – С. 125-131.
13. Malitsky Yu.V. A new hybrid method for solving variational inequality problems/ Malitsky Yu. V., Semenov V. V. // Доповіді НАН України. – 2014р. – № 4. – С. 49-55.
14. Malitsky Yu. Douglas-Rachford algorithm for best approximation pair/ Malitsky Yu. // Конференція «Обчислювальна та прикладна математика». – Київ, 2014. – С.128.
15. Malitsky Yu. Projected reflected gradient methods for monotone variational inequality/ Malitsky Yu. // SIAM Journal on Optimization. – 2015 (в підготовці)

Дякую за увагу!