Методи Шора з розтягом простору

(лекція для аспірантів)

Стецюк П.І.

stetsyukp@gmail.com

Інститут кібернетики імені В.М. Глушкова
НАН України, Київ

Ужгородський національний університет
17 червня 2021

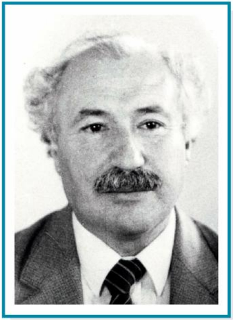
План лекції

1. Коротко про Шора Н.З. (лист проф. С. Бойда)
2. Метод з розтягом простору (1969)
3. Про зовнішню схожість Шора та Ейнштейна
4. Про ідейну схожість Шора та Ейнштейна
5. Оператор розтягу та Оптимальні еліпсоїди
6. Методи з перетворенням простору

План лекції

1. Коротко про Шора Н.З. (лист проф. С. Бойда)
2. Метод з розтягом простору (1969)
3. Про зовнішню схожість Шора та Ейнштейна
4. Про ідейну схожість Шора та Ейнштейна
5. Оператор розтягу та Оптимальні еліпсоїди
6. Методи з перетворенням простору

Шор Наум Зуселевич (1937–2006)



Засновник наукової школи методів негладкої оптимізації (Інститут кібернетики НАН України)

У 1962 році Н.З. Шор розробив перший субградієнтний метод

У 1969 році вперше використав оператор розтягу простору для прискорення градієнтних методів

Методи Н.З. Шора високо оцінені фахівцями,

мають велике теоретичне та прикладне значення, є "ключем"до розв'язання задач великих розмірів

Основні монографії

1. Шор Н.З. *Методы минимизации недифференцируемых функций и их приложения.* – Киев: Наукова думка, 1979.

English translation: Shor N.Z. *Minimization Methods for Non-Differentiable Functions.* – Berlin: Springer-Verlag, 1985.

2. Shor N.Z. *Nondifferentiable optimization and polynomial problems.* – Boston; Dordrecht; London: Kluwer Academic Publishers, 1998.

3. Шор Н.З., Стеценко С.И. *Квадратичные экстремальные задачи и недифференцируемая оптимизация.* – Киев: Наукова думка, 1989.

Лист професора С. Бойда stanford.edu/~boyd/

April 15, 2005

Dear Professor Shor,

We have never met, but your work has very much influenced me for many years now. I started with your small 1985 Springer book on subgradient methods, which I read as a PhD student. I recently read your newer book on nondifferentiable optimization (1998), which I enjoyed very much.

I'm enclosing copies of the three books I've written. The first concerns the design of linear controllers via convex optimization; the second is on linear matrix inequalities; and the third one is a basic textbook on convex optimization. [...] I hope you can see your strong influence in all of these books.

With the best regards,
Stephen P. Boyd

План лекції

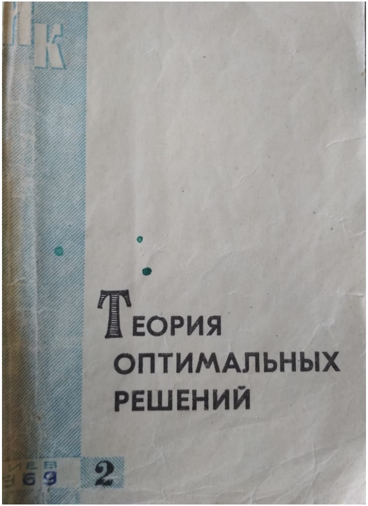
1. Коротко про Шора Н.З. (лист проф. С. Бойда)
2. Метод з розтягом простору (1969)
3. Про зовнішню схожість Шора та Ейнштейна
4. Про ідейну схожість Шора та Ейнштейна
5. Оператор розтягу та Оптимальні еліпсоїди
6. Методи з перетворенням простору

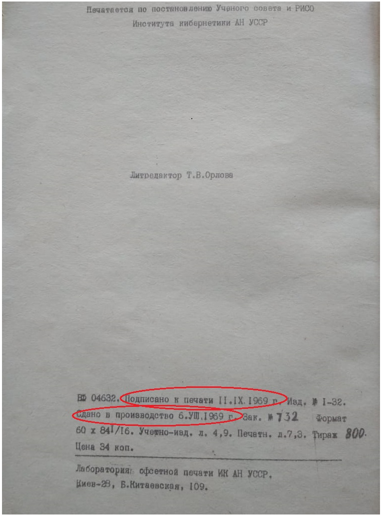
Стаття Шор, Білецький (1969)

Шор Н.З., Билецкий В.И. *Метод растяжения пространства для ускорения сходимости в задачах овражного типа* // Тр. семинара Науч. совета АН УССР по кибернетике "Теория оптимальных решений". – Киев. – 1969. – 2. – С. 3–18.

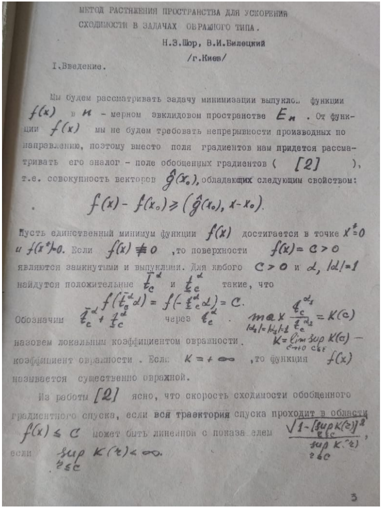
1. Доповідь на семінарі 11.IV.1969,
2. Здано у виробництво 6.VIII.1969,
3. Підписано до друку 11.IX.1969.

Збірник 1969 (титульна та заключна сторінки)





Стаття 1969 (перша та остання сторінки)



МЕТОД РАСТЯЖЕНИЯ ПРОСТРАНСТВА ДЛЯ УСКОРЕНИЯ СХОДИМОСТИ В ЗАДАЧАХ ОВРАЖНОГО ТИПА.

Н.З.Шор, В.И.Билецкий
/г.Киев/

I.Введение.

Мы будем рассматривать задачу минимизации выпукло.. функции $f(x)$ в n - мерном эвклидовом пространстве E_n. От функции $f(x)$ мы не будем требовать непрерывности производных по направлению, поэтому вместо поля градиентов нам придется рассматривать его аналог - поле обобщенных градиентов ([2]), т.е. совокупность векторов $\hat{g}(x_0)$, обладающих следующим свойством:

$$f(x)-f(x_0) \geq (\hat{g}(x_0), x-x_0).$$

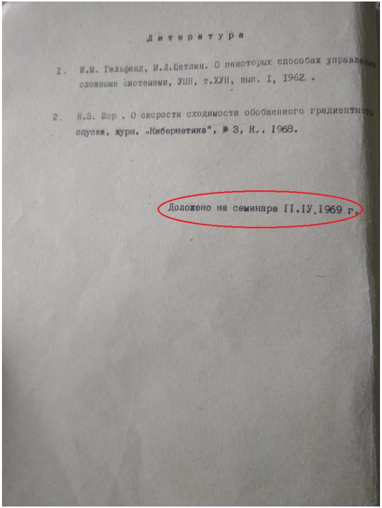
Пусть единственный минимум функции $f(x)$ достигается в точке $x^*=0$ и $f(x^*)=0$. Если $f(x) \not\equiv 0$, то поверхности $f(x)=c>0$ являются замкнутыми и выпуклыми. Для любого $c>0$ и α, $|\alpha|=1$ найдутся положительные $\bar{t}_c^\alpha$ и $\underline{t}_c^\alpha$ такие, что

$$f(\bar{t}_c^\alpha \alpha) = f(-\underline{t}_c^\alpha \alpha) = c.$$

Обозначим $\bar{t}_c^\alpha + \underline{t}_c^\alpha$ через t_c^α. $\max\limits_{|\alpha_1|=|\alpha_2|=1} \frac{t_c^{\alpha_1}}{t_c^{\alpha_2}} = K(c)$ назовем локальным коэффициентом овражности. $K = \lim\limits_{\varepsilon \to 0} \sup\limits_{c \leq \varepsilon} K(c)$ — коэффициент овражности. Если $K=+\infty$, то функция $f(x)$ называется существенно овражной.

Из работы [2] ясно, что скорость сходимости обобщенного градиентного спуска, если вся траектория спуска проходит в области $f(x) \leq C$ может быть линейной с показателем $\frac{\sqrt{1-[\sup_{\tau \leq c} K(\tau)]^2}}{\sup_{\tau \leq c} K(\tau)}$, если $\sup\limits_{\tau \leq c} K(\tau) < \infty$.

3



Литература

1. И.М. Гельфанд, М.Л.Цетлин. О некоторых способах управления сложными системами, УМН, т.XVII, вып. I, 1962.

2. Н.З. Шор. О скорости сходимости обобщенного градиентного спуска, журн. „Кибернетика", № 3, К., 1968.

Доложено на семинаре II.IV.1969 г.

Про три ідеї Шора (січень, 2012)

75-річчю з дня народження Н.З. Шора присвячена стаття

Сергиенко И.В., Стецюк П.И.

О трех научных идеях Н.З. Шора // Кибернетика и системный анализ. – 2012, № 1. – С. 4–22.

У статті описано три центральні ідеї Н.З. Шора:

узагальнений градієнтний спуск (1962),
використання лінійних неортогональних перетворень простору для поліпшення обумовленості яружних функцій (1969),
двоїстий підхід до отримання та уточнення оцінок цільової функції в неопуклих квадратичних моделях (1985).

План лекції

Стаття в газеті “Сегодня”, 28 грудня 2002 року

ПОЕХАЛ В ГРЕЦИЮ, а чемодан остался в Венгрии...

Академик НАН Украины профессор Наум Шор, родившийся 1 января 1937 года, имеет железную логику и ясный ум. Он 44 года без перерыва работает в Институте кибернетики им. В. М. Глушкова. Сегодня занимает там должность заведующего отделом методов решения сложных задач оптимизации.

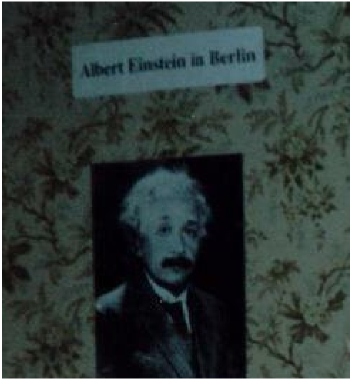
Коллеги из Швейцарии пригласили ученого в музей Эйнштейна в Берне, чтобы сделать это фото и подчеркнуть их сходство

— Что обо мне писать? — засмущался Наум Зуселевич. — Я живу: с работы домой и обратно. Все мои 66 лет забрала математика. Мне повезло, дипломной работой по диффе-

комился 40 лет назад, 1 января 1963 года. Звезды расположились так, что и она Козерог, родилась 9 января. А вот дети — сын и дочь — оба Раки. Гороскоп кошки Алисы не проверяли, достаточно того, что

Фотографія (Ж.Ф. Емменеггер, Берн, 1997)

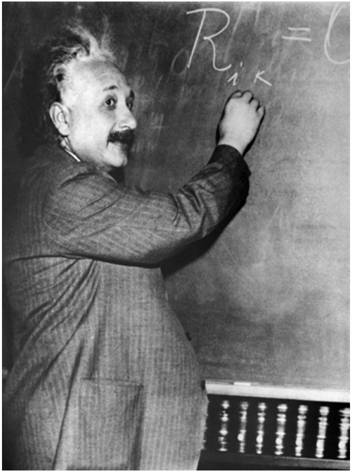
Ще раз про зовнішню схожість, але раніше ...

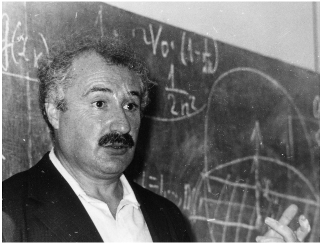


Альберт Ейнштейн в Берліні.

Наум Шор в Києві.

Ейнштейн пише послання Шору ...





"... щоб символом $R_\alpha(\xi)$ він позначив оператор розтягу простору"

План лекції

Про схожість ідей … або що спільного?

між

операторором Шора для розтягу простору

$$R_\alpha(\xi) = I + (\alpha - 1)\xi\xi^T, \qquad \text{де} \qquad \alpha > 1, \qquad (1)$$

(I – одинична $n\times n$-матриця, вектор $\xi \in E^n$ такий, що $\|\xi\|=1$)

та

формулою Ейнштейна для енергії в релятивістській динаміці

$$E = mc^2, \qquad \text{де} \qquad m = \frac{m_0}{\sqrt{1 - v^2/c^2}} \qquad (2)$$

Куля S_n та еліпсоїд $\mathcal{E}_n$ в просторі E^n

Образом кулі $S_n = \{x \in E^n : \|x - x_0\| \leq r\}$

у перетвореному просторі $y = R_\alpha(\xi)x$

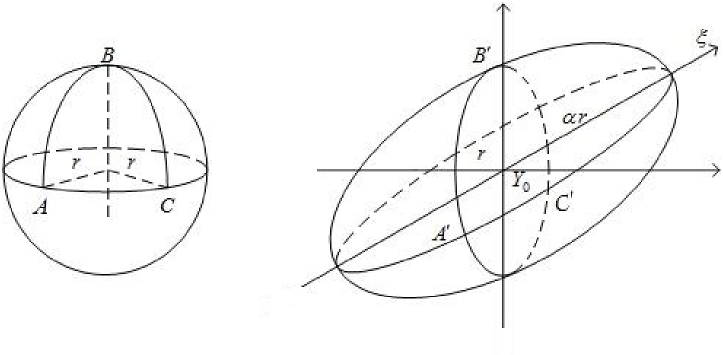
є еліпсоїд $\mathcal{E}_n = \{y \in E^n : \|R_\beta(\xi)(y - y_0)\| \leq r\}$.

Об'єми кулі S_n та еліпсоїда $\mathcal{E}_n$ дорівнюють

$$\text{vol}(S_n) = v_0 r^n \qquad \text{та} \qquad \text{vol}(\mathcal{E}_n) = \frac{v_0 r^n}{\det R_\beta(\xi)},$$

де v_0 – об'єм одиничної n-вимірної кулі.

Приклад: куля S_3 та еліпсоїд $\mathcal{E}_3$



Після розтягу простору в напрямі $\xi \in E^3$ куля S_3 (зліва) є еліпсоїдом $\mathcal{E}_3$ (справа)

Відношення об'ємів кулі S_n та еліпсоїда $\mathcal{E}_n$

Відношення об'ємів $\mathcal{E}_n$ і S_n обчислюється за формулою

$$\frac{\mathrm{vol}(\mathcal{E}_n)}{\mathrm{vol}(S_n)} = \frac{\frac{v_0 r^n}{\det R_\beta(\xi)}}{v_0 r^n} = \frac{1}{\det R_\beta(\xi)} = \frac{1}{\beta} = \alpha > 1,$$

де v_0 – об'єм одиничної n-вимірної кулі.

Його можна переписати так

$$\mathrm{vol}(\mathcal{E}_n) = \frac{\mathrm{vol}(S_n)}{\beta} = \frac{\mathrm{vol}(S_n)}{\sqrt{1-(1-\beta^2)}} \tag{3}$$

Зв'язок об'ємів з формулою Ейнштейна

Із (3) випливає

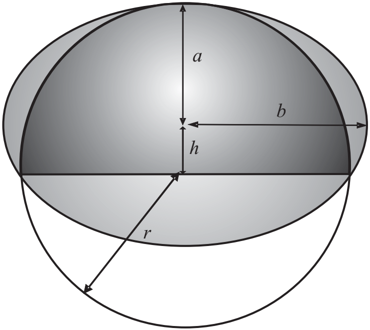
$$\mathrm{vol}(\mathcal{E}_n) = \frac{\mathrm{vol}(S_n)}{\sqrt{1-(1-\beta^2)}} \quad .=. \quad \frac{\mathrm{vol}(S_n)}{\sqrt{1-v^2/c^2}}. \qquad \text{(Shor)}$$

Чи не подібно на формулу (2) ? ...

$$m = \frac{m_0}{\sqrt{1-v^2/c^2}}. \qquad \text{(Einstein)}$$

План лекції

Оптимальний 1d-еліпсоїд (Шор, 1977)



Еліпсоїд $\mathcal{E}_n$, що містить півкулю в E^n, має мінімальний об'єм, якщо

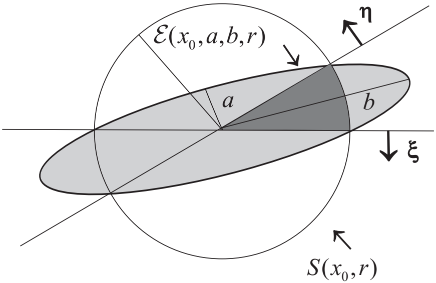
$$a = \frac{n}{n+1} r, \quad b = \frac{n}{\sqrt{n^2-1}} r, \quad h = \frac{1}{n+1} r.$$

Щоб перетворити $\mathcal{E}_n$ в кулю потрібно розтянути простір з коефіцієнтом $\alpha = \frac{b}{a} = \sqrt{\frac{n+1}{n-1}}$.

На кожній ітерації МЕ об'єм еліпсоїда зменшується в

$$q(n) = \frac{\mathrm{vol}(\mathcal{E}_n)}{\mathrm{vol}(S_n)} = \frac{a}{r}\left(\frac{b}{r}\right)^{n-1} = \sqrt{\frac{n-1}{n+1}}\left(\frac{n}{\sqrt{n^2-1}}\right)^n \le 1 - \frac{1}{2n},$$

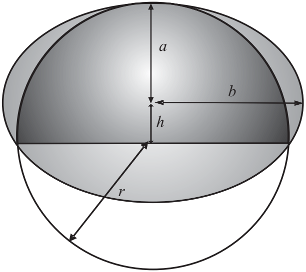
Оптимальний 2d-еліпсоїд (Стецюк, 1996)

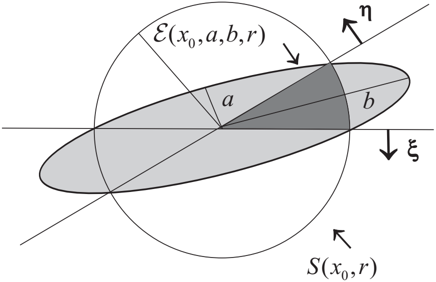


Перетворення 2d-еліпсоїда в кулю вимагає розтягу простору в напрямі $\frac{\xi-\eta}{\|\xi-\eta\|}$ з коеф. $\alpha_1=\frac{1}{\sqrt{1+(\xi,\eta)}} > 1$ і подальшого зжимання простору в напрямі $\frac{\xi+\eta}{\|\xi+\eta\|}$ з коеф. $\alpha_2=\frac{1}{\sqrt{1-(\xi,\eta)}} < 1.$

$$q = \frac{\mathrm{vol}(\mathcal{E}(x_0,a,b,r))}{\mathrm{vol}(S(x_0,r))} = \left(\frac{a}{r}\right)\left(\frac{b}{r}\right) = \sqrt{1-(\xi,\eta)^2}.$$

Де можна дізнатися про 1d та 2d-еліпсоїди?





1. Стецюк П.И. *Методы эллипсоидов и r-алгоритмы.* – Кишинэу, Эврика, 2014. – 488 с.
2. Стецюк П.І. та ін. *Субградієнтні алгоритми та задачі на комбінаторних конфігураціях.* – Київ: ПУЛЬСАРИ, 2019. – 235 с.

План лекції

1. Коротко про Шора Н.З. (лист проф. С. Бойда)
2. Метод з розтягом простору (1969)
3. Про зовнішню схожість Шора та Ейнштейна
4. Про ідейну схожість Шора та Ейнштейна
5. Оператор розтягу та Оптимальні еліпсоїди
6. Методи з перетворенням простору

Субградієнтні методи з перетворенням простору

призначені для розв'язання задачі

$$f^* = f(x^*) = \min_{x \in R^n} f(x),$$

где $f(x)$ – опукла функція (гладка, негладка)

Задані: x_0 – стартова точка, B_0 – $n \times n$-матриця,

Ітерації $k=1, 2, \ldots,$ мають вид

$$x_{k+1} = x_k - h_k B_k \frac{B_k^T g_f(x_k)}{\|B_k^T g_f(x_k)\|}, \quad B_{k+1} = B_k T_k, \qquad \text{(Shor69)}$$

де h_k – крок, $g_f(x_k)$ – довільний субградієнт функції $f(x)$ в точці x_k, T_k – $n \times n$-матриця.

Найбільш відомі методи

1. r-**алгоритми** (стійкі до накопичення помилок обчислень);
2. **методи еліпсоїдів** (збіжність – геометрична прогресія);
3. **субградієнтні методи з перетворенням простору** (використовують феєровські кроки, кроки Поляка).

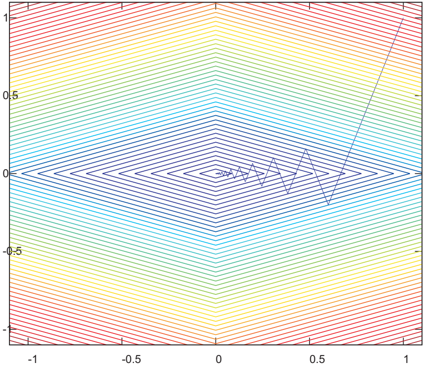
Методи мають прискорену збіжність

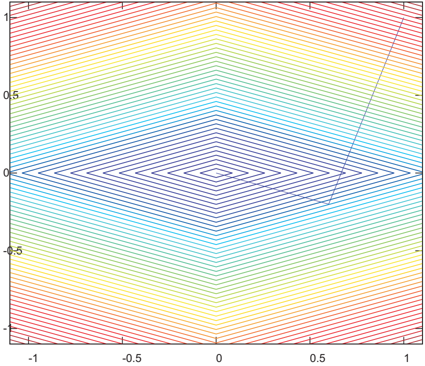
для яружних опуклих функцій (гладких та негладких)

та використовуються

як оптимізаційні ядра для прикладних задач.

Прискорена збіжність: кусочно лінійна функція





1. Траєкторія методу **ams**

2. ...прискореного методу **amsg2**

для яружної функції $f_1(x_1, x_2) = |x_1| + 10|x_2|,\ x_0 = (1, 1).$

Програмні реалізації методів

як оптимізаційні ядра використовують:

1. **Кисельова О.М.** (Дніпро, задачі розміщення);
2. **Стоян Ю.Г.** (Харків, задачі упаковки);
3. **Соломон Д.І.** (Кишинеу, транспортні задачі);
4. **Шарий С.П.** (Новосибірськ, інтервальний аналіз);
5. **Міца О.В.** (Ужгород, багатошарові оптичні покриття).

Серед них

octave-функції – **ralgb5**, **amsg2p**, **emshor** та інші.

Про серію книг«НДО и ее приложения»

1. Соломон Д.И. *Дробное программирование и недифференцируемая оптимизация.* – Кишинэу, Эврика, 2010. – 556 с.

2. Стецюк П.И. *Методы эллипсоидов и r-алгоритмы.* – Кишинэу, Эврика, 2014. – 488с.

3. Киселева Е.М., Коряшкина Л.С. *Непрерывные задачи оптимального разбиения множеств и r-алгоритмы.* – Киев, Наукова думка, 2015. – 400 с.

4. Стецюк П.И. *Двойственные оценки в квадратичных экстремальных задачах.* – Кишинэу, Эврика, 2018. – 504с.

Запитання?

ДЯКУЮ ЗА УВАГУ!