Naum Shor – pioneer of the subgradient method

Petro Stetsyuk[1], Dumitru Solomon[2], Serhii Zhabin[3]

[1]V.M. Glushkov Institute of Cybernetics, Kyiv, Ukraine
[2]Institute of Mathematics and Informatics, Chisinau, Moldova
[3]G.M. Dobrov Institute for Scientific and Technological Potential and Science History Studies, Kyiv, Ukraine

The Seminar “The Theory of Optimal Solutions“
Kyiv, September 13, 2022

Dedicated to the memory of academician
Naum Zuselevich Shor
(to the 85th anniversary of his birth)

Newspaper „Segodnya“, December 28, 2002

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

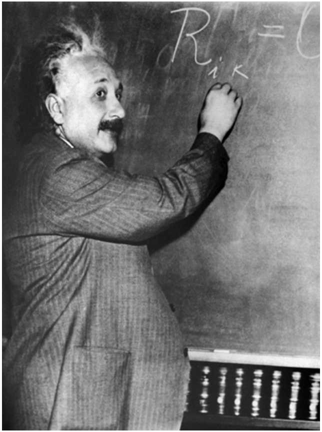
Академик НАН Украины профессор Наум Шор, родившийся 1 января 1937 года, имеет железную логику и ясный ум. Он 44 года без перерыва работает в Институте кибернетики им. В. М. Глушкова. Сегодня занимает там должность заведующего отделом методов решения сложных задач оптимизации.

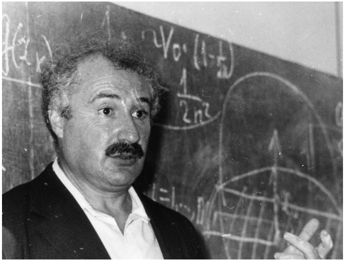
Коллеги из Швейцарии пригласили ученого в музей Эйнштейна в Берне, чтобы сделать это фото и подчеркнуть их сходство

Что обо мне писать? — засмущался Наум Зуселевич. — Я живу с ра-

комился 40 лет назад, 1 января 1963 года. Звезды расположились так, что и она Ко-

Einstein writes a message to Shor



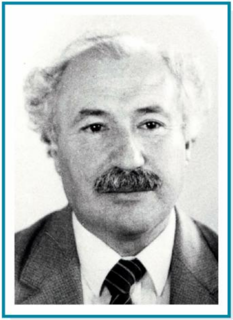


to denote space dilation operator as $R_\alpha(\xi)$

Outline

Content

Naum Z. Shor (1937–2006)



Shor founded scientific school of non-smooth optimization methods (Institute of Cybernetics, Kyiv, Ukraine)

In 1962 he developed the first subgradient method.

In 1969 he used the space dilation operator for acceleration of gradient methods convergences.

N.Z. Shor's methods

are still of a great theoretical and applied importance, and "a key" for solution of large-scale problems.

Shor's basic monographies

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.

Shor‘s Three key Ideas

Sergienko I.V., Stetsyuk P.I.

On N.Z. Shor’s three scientific ideas. Cybernetics and Systems Analysis 48, 2–16 (2012).

The paper is devoted to the 75th anniversary of N.Z.Shor.

This paper described Shor‘s three key ideas:
generalized gradient descent (1962),
the use of linear nonorthogonal space transformations to improve the conditionality of ravine functions (1969),
dual approach for finding bounds of the objective function in nonconvex quadratic models (1985).
Examples of the application of these ideas in methods and algorithms developed at the V.M. Glushkov Institute of Cybernetics of the NAS of Ukraine are given.

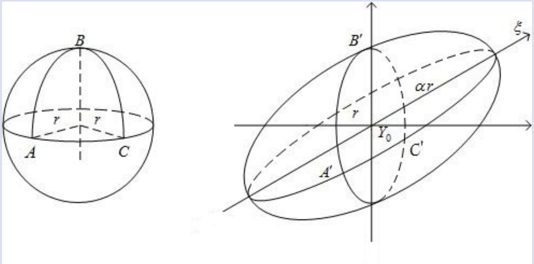
Shor's space dilation operator

Space dilation operator has the following form

$$R_\alpha(\xi) = I_n + (\alpha - 1)\xi\xi^T, \quad \text{where} \quad \alpha > 1.$$

Here: α is the coefficient of space dilation in the normed direction $\xi \in \mathbb{R}^n$, $\|\xi\| = 1$; I_n is the identity $n \times n$-matrix.

Example in $\mathbb{R}^3$: Ball (left) dilate to ellipsoid (right)



Content

Subgradient methods with space transformation

purposed to solve the following problem

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

where $f(x)$ – a convex function (smooth, non-smooth)

We have: x_0 – starting point, B_0 – $n \times n$-matrix,

Iterations $k = 1, 2, \ldots$ have the following form

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

where h_k – step-size, $g_f(x_k)$ – a subgradient of function $f(x)$ at the point x_k, T_k – $n \times n$-matrix.

The most known methods

1. **r-algorithms** (robust to accumulation of errors);
2. **ellipsoid methods**(convergence - geometric progression);
3. **subgradient methods with space transformation** (use Fejer-type steps, Polyak's steps and others).

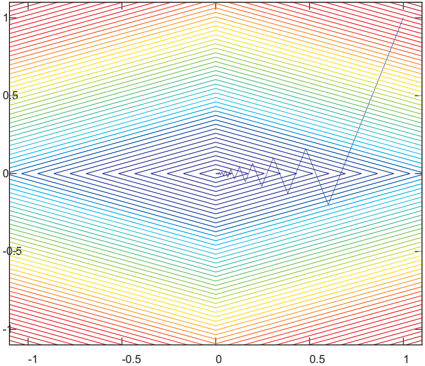
The methods have accelerated convergence

for convex ravine functions (smooth and non-smooth).

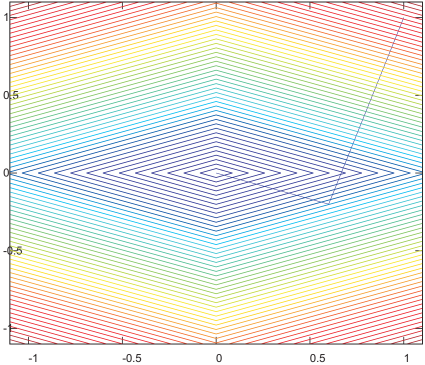
Implementations of the methods in Octave:

1. **ralgb5a** (30 rows), 2. **emshor** (15), 3. **amsg2p** (25).

Accelerated convergence: piecewise linear function



1. Trajectory for **ams** method



2. ...accelerated **amsg2** method

for ravine function $f_1(x_1, x_2) = |x_1| + 10|x_2|,\ x_0 = (1, 1).$

Implementations of r-algorithms

were used for solving:

1. large-scale block optimization problems with various decomposition schemes;
2. minimax and matrix optimization problems;
3. for calculating dual bounds in multiextremal and combinatorial optimization problems;
4. ... and others

They were core for application packages, C and SA

PLANER (1983, **19**, 362–382), DISNEL (1991, **27**, 354–366).

Content

Boris Polyak „Introduction to Optimization“ (1987)

Boris Teodorovich says:

5.3 The Subgradient Method, p. 138

„The fundamental algorithms for minimizing smooth functions, the gradient as well as Newton's algorithms, are based on linear or quadratic approximation of the function given by the first terms of a Taylor series. However this method is unfeasible for nondifferentiable functions, for such a function cannot be well approximated either by a linear or by a quadratic function.“

Boris Polyak „Introduction to Optimization" (1987)

5.3 The Subgradient Method, p. 139

„Methods for minimizing nonsmooth functions cannot be further developed without new, innovative techniques. N.Z. Shor suggests – however surprisingly – a direct analog of the gradient method, with the gradient replaced by an arbitrary subgradient of the function $f(x)$:

$$x_{k+1} = x_k - \gamma_k g_f(x_k). \tag{3}$$

... the values of the function in method (3) cannot decrease monotonically. In this case, however, another function, viz. the distance to the minimum point, decreases monotonically. This is the key idea of the subgradient method (3)."

Stephen Boyd, Stanford University

1. BOYD, S., BARRATT, C. *Linear Controller Design: Limits of Performance* (1991)

2. BOYD, S. AND ALL. *Linear Matrix Inequalities in System and Control Theory* (1994)

3. BOYD, S., VANDENBERGHE, L. *Convex Optimization* (2004)

Boyd's letter to Shor (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

„Optimization Methods and Software" (2008), dedicated to the memory of N. Shor

B. Mordukhovich

M. Solodov

M. Todd

"In 1972, Shor introduced the fundamental generalized differential notion for locally Lipschitzian functions, which he called "the set of almost-gradients"..."

B. Mordukhovich, M. Solodov, M. Todd (2008)

„...It was defined as the collection of limiting points of the usual gradients of the Lipschitz continuous function in question, which is differentiable almost everywhere by the classical Rademacher theorem. This limiting set was later widely used, under the name of B-gradient and B-Jacobian for the case of vector functions, in developing nonsmooth versions of Newton's method. It is worth mentioning that in the same paper of 1972, Shor also introduced and utilized the convex hull of the set of almost-gradients, which he called the 'set of generalized almost gradients'. This latter set was subsequently rediscovered by Clarke and was widely used in nonsmooth optimization under the name of (Clarke's) generalized gradient for Lipschitzian functions."

Content

Shor's basic monographies

- Шор Н.З. *Методы минимизации недифференцируемых функций и их приложения*. Киев: Наукова думка, 1979. 200 с.
- Shor N.Z. *Minimization Methods for Non-Differentiable Functions*. Berlin: Springer-Verlag, 1985. 178 p.
- Шор Н.З. Стеценко С.И. *Квадратичные экстремальные задачи и недифференцируемая оптимизация*. Киев: Наукова думка, 1989. 208с.
- Shor N.Z. *Nondifferentiable optimization and polynomial problems*. Boston; Dordrecht; London: Kluwer Academic Publishers, 1998. 394 p.

Shor's selected papers

- Шор Н.З. *Методы недифференцируемой оптимизации и сложные экстремальные задачи*. Кишинэу, Эврика, 2008. 270 с.
- Шор Н.З. *Методы минимизации негладких функций и матричные задачи оптимизации*. Кишинэу, Эврика, 2009. 240 с.
- Шор Н.З. *Алгоритмы последовательной и негладкой оптимизации*. Кишинэу, Эврика, 2012. 272с.

Questions?

THANK YOU FOR YOUR ATTENTION!

e-mail: stetsyukp@gmail.com