Circumscribed 2d-ellipsoid and Shor's r-algorithm

Stetsyuk Petro
stetsyukp@gmail.com

V.M.Glushkov Institute of Cybernetics, Kiev, Ukraine

OPTIMA 2014, Petrovac, Montenegro
September 28 – October 4

Outline

1. On the ellipsoid method
2. r-algorithms and ellipsoid method
3. 2d-ellipsoid and r-algorithms
4. On using 2d-ellipsoid

Outline

Ellipsoid method was proposed

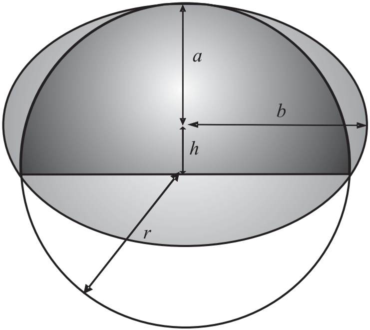
1976 by **Yudin and Nemirovskii** as a method of successive cutting-plane [1].

1977 by **Shor** as a variant of the method with space dilation in the direction of the subgradient [2].

1. Yudin D.B. and Nemirovskii A.S. *Informational complexity and effective methods for the solution of convex extremal problems* // Ekonom. Mat. Metody, 12, No. 2 (1976).

2. Shor N.Z. *Cut-off method with space extension in convex programming problems* // Cybernetics, 13, No. 1 (1977).

The idea of ellipsoid method (1d-ellipsoid $\mathcal{E}_n$)



1d-ellipsoid $\mathcal{E}_n$, containing half of ball S_n in E^n, has the minimal volume if

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

To transform $\mathcal{E}_n$ into a ball we have to stretch the space by a factor of $\alpha = \frac{b}{a} = \sqrt{\frac{n+1}{n-1}}$.

The ratio of 1d-ellipsoid volume to ball volume is equal to

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

Operator of space dilation

is introduced by N.Z. Shor (1969) and 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 $\xi \in E^n$, $\|\xi\|=1$; I_n is the identity $n \times n$-matrix.

Shor's algorithms use the inverse operator

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

which means "compression" of space subgradients.

Outline

Shor's Dream is actual for today ...

„The theory of the whole class of algorithms with space dilation is far from perfect. It seems quite realistic goal – building of such an algorithm, which, by its practical efficacy, was not inferior to r-algorithm and would be just as well founded as the ellipsoid method" [*].

*. **V.I. Gershovich, N.Z. Shor** *Method of ellipsoids, its generalizations and applications* // Cybernetics, 18, No 5 (1982).

Мечта Шора есть актуальной и сегодня ...

„Теория всего класса алгоритмов с растяжением пространства далека от совершенства. Нам кажется достаточно реалистичной целью – построение такого алгоритма, который по своей практической эффективности не уступал бы r-алгоритму и был столь же хорошо обоснован как метод эллипсоидов" [*].

*. **Гершович В.И., Шор Н.З.** *Метод эллипсоидов, его обобщения и приложения.* // Кибернетика, 1982, № 5.

r-Algorithms and ellipsoids (Stetsyuk,1996)

An attempt to explain r-algorithms was made in the paper

P.I. Stetsyuk. "r-Algorithms and ellipsoids," Cybernetics and System Analysis, **32**, No. 1, 93–110 (1996).

Here, for the transformation of a special ellipsoid into a ball anti-zigzag method like one in r-algorithms is used.

Space dilation is realized in the direction of the difference of the two normalized subgradients. This direction is close to the direction of the difference of two subgradients if the norms of the subgradients are approximately equal to each other.

Theorem 1 (Stetsyuk, OPTIMA 2014)

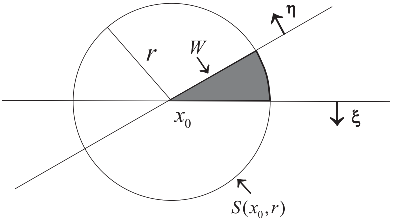
Let B_k *be* $n \times n$*-matrix, such that* $\|B_k^{-1}(x_k - x^*)\| \le r$*;* g_1 *and* g_2 *are* n*-dimensional vectors, subject to* $(x_k - x^*, g_1) \ge 0$ *and* $(x_k - x^*, g_2) \ge 0$*. If* $-\|B_k^T g_1\|\|B_k^T g_2\| < (B_k^T g_1, B_k^T g_2) < 0$ *and matrix* B_{k+1} *is recalculated according to the rule*

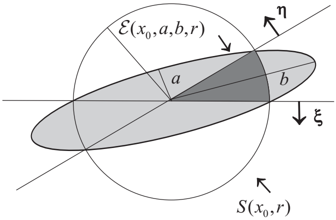
$$B_{k+1} = B_k R_{\beta_1}\left(\frac{\xi - \eta}{\|\xi - \eta\|}\right) R_{\beta_2}\left(\frac{\xi + \eta}{\|\xi + \eta\|}\right), \ \xi = \frac{B_k^T g_1}{\|B_k^T g_1\|}, \eta = \frac{B_k^T g_2}{\|B_k^T g_2\|},$$

where $\beta_1 = \sqrt{1 + (\xi, \eta)}$ *and* $\beta_2 = \sqrt{1 - (\xi, \eta)}$*, then matrix* B_{k+1} *has the following properties: (i)* $\|B_{k+1}^{-1}(x_k - x^*)\| \le r$*;*
(ii) $\det(B_{k+1}) = \det B_k \sqrt{1 - (\xi, \eta)^2}$*; (iii)* $(B_{k+1}^T g_1, B_{k+1}^T g_2) = 0$*.*

Outline

The convex set W and 2d-ellipsoid $\mathcal{E}(x_0, a, b, r)$



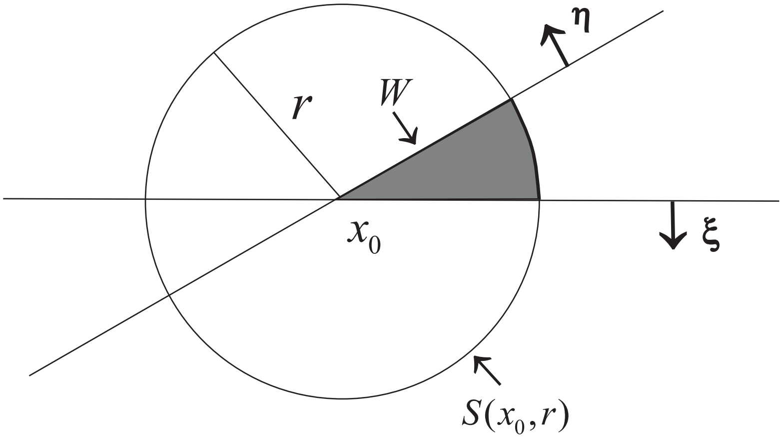


the set W is the intersection of a ball $S(x_0, r)$ with half-spaces $P(x_0, \xi)$ and $P(x_0, \eta)$

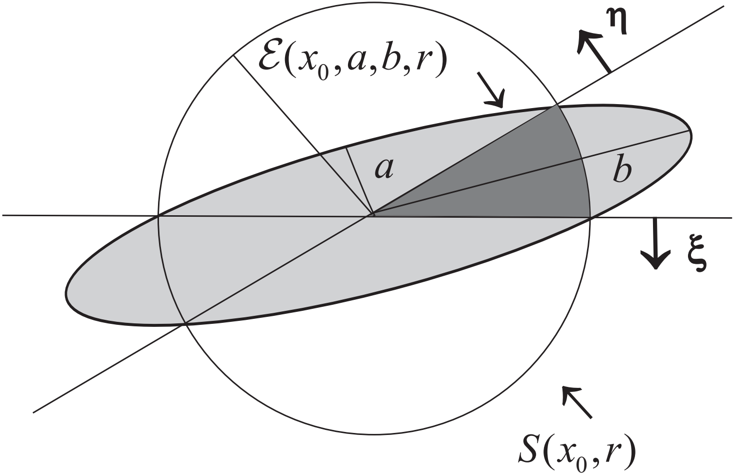
2d-ellipsoid contains the set W and has the minimum volume

STETSYUK (1996) r-Algorithms and ellipsoids, Cybernetics and System Analysis, **32**, No. 1.

The convex set W



2d-ellipsoid $\mathcal{E}(x_0, a, b, r)$



Properties of 2d-ellipsoid $\mathcal{E}(x_0, a, b, r)$

2d-ellipsoid has the following parameters:

$$a = r\sqrt{1 + (\xi, \eta)} < r; \quad b = r\sqrt{1 - (\xi, \eta)} > r.$$

(i) If $(\xi, \eta) < 0$ then 2d-ellipsoid contains the convex set W.

(ii) The ratio of 2d-ellipsoid volume to ball volume is equal to

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

If the angle between vectors ξ and η becomes closer to 180 degrees then the ratio q becomes smaller.

Transformation of 2d-ellipsoid into a ball

corresponds to the updating of matrix

$$B_{k+1}^{-1} = R_{\alpha_2}\left(\frac{\xi+\eta}{\|\xi+\eta\|}\right) R_{\alpha_1}\left(\frac{\xi-\eta}{\|\xi-\eta\|}\right) B_k^{-1},$$

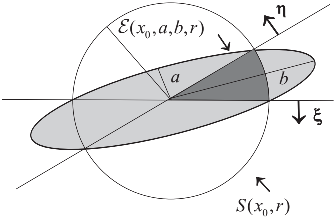
i.e. we stretch space in the direction of $\frac{\xi-\eta}{\|\xi-\eta\|}$ with coefficient

$$\alpha_1 = \frac{r}{a} = \frac{1}{\sqrt{1+(\xi,\eta)}} > 1$$

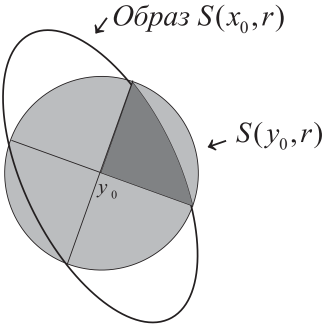
and we compress space in the direction of $\frac{\xi+\eta}{\|\xi+\eta\|}$ with coefficient

$$\alpha_2 = \frac{r}{b} = \frac{1}{\sqrt{1-(\xi,\eta)}} < 1.$$

2d-ellipsoid before and after transformation



2d-ellipsoid



is a ball in the transformed space

Closeness to Shor's r-algorithm

(iii) Images of vectors ξ and η in the transformed space are orthogonal

This feature allows to „extend" cone of feasiable directions of the function decrease for the subgradient process in the transformed space of variables, similar to Shor's r-algorithm.

Coefficient of space dilation in the direction

of the difference of the two normalized subgradients is

$$\alpha = \frac{\alpha_1}{\alpha_2} = \frac{\sqrt{1-(\xi,\eta)}}{\sqrt{1+(\xi,\eta)}} = \sqrt{1-\frac{2(\xi,\eta)}{1+(\xi,\eta)}}$$

if $(\xi,\eta) = -1/2$, then $\alpha = \sqrt{3}$,
if $(\xi,\eta) = -\sqrt{2}/2$, then $\alpha = \sqrt{6}$,
if $(\xi,\eta) = -9/10$, then $\alpha = \sqrt{19}$,
if $(\xi,\eta) = -99/100$, then $\alpha = \sqrt{199}$, ...

Outline

Where can we use 2d-ellipsoid?

**for construction of accelerated ellipsoid methods
for the following problem categories:**

1) convex programming problems,
2) finding saddle points of concave-convex functions,
3) special cases of variational inequalities,
4) linear and non-linear complimentarity problems.

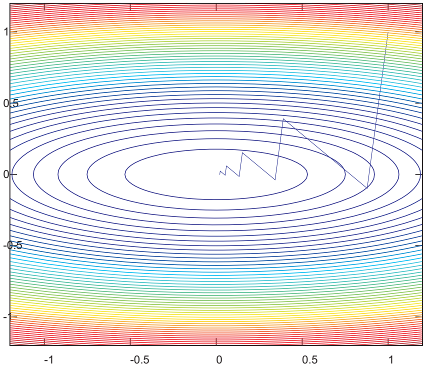
What can we expect from such methods?

It is possible to guarantee the rate of convergence close to the rate of r-algorithms.

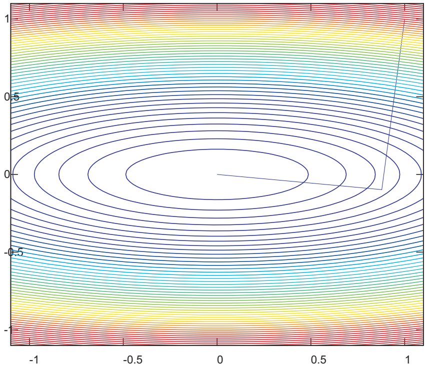
This fact is proved by subgradient methods with space transformation [*], which speed up the well-known Polyak's subgradient method. They are effective for ill-conditioned functions.

*. **P.I. Stetsyuk** *Ellipsoid methods and r-algorithms*, Evrika, Chisinau (2014)

Example 1. (quadratic function)

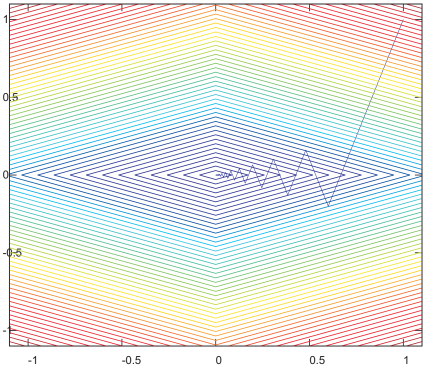


1.1 Trajectory of Polyak's method

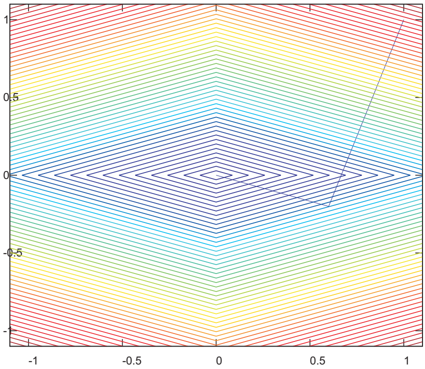


1.2 Trajectory of amsg2 method

Example 2. (piecewise linear function)

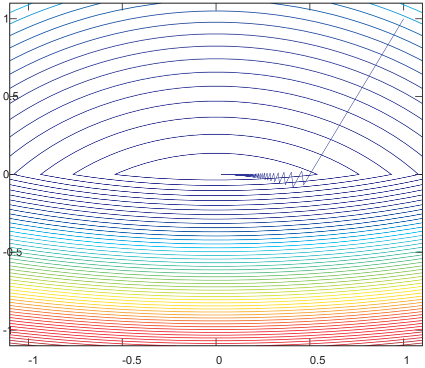


2.1 Trajectory of Polyak's method

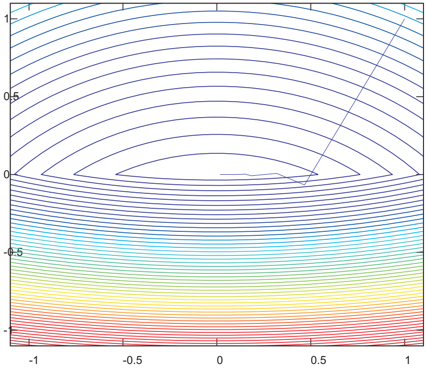


2.2 Trajectory of amsg2 method

Example 3 (piecewise quadratic function)



3.1 Trajectory of Polyak's method



3.2 Trajectory of amsg2 method

Questions?

THANK YOU
FOR YOUR ATTENTION!