

The generalized ellipsoid method and its implementation

Petro Stetsyuk¹, Andreas Fischer² and Olga Khomyak¹
stetsyukp@gmail.com

¹V.M. Glushkov Institute of Cybernetics, Kiev, Ukraine

²Technische Universität Dresden, Dresden, Germany

The X International Conference "Optimization and Applications"
OPTIMA 2019, Petrovac, Montenegro
September 30 – October 4

Outline

- 1 GEM (basic idea)
- 2 The GEM description
- 3 Special cases of GEM
- 4 Algorithm emshor
- 5 Acceleration using 2d-ellipsoid

Outline

- 1 GEM (basic idea)
- 2 The GEM description
- 3 Special cases of GEM
- 4 Algorithm emshor
- 5 Acceleration using 2d-ellipsoid

EM (Yudin-Nemirovski-Shor)

The Ellipsoid method (EM) 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).

Operator of space dilation

is introduced by N.Z. Shor (1969) and has the following form

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

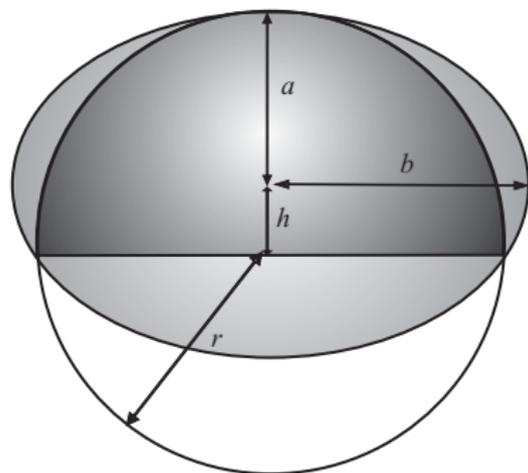
Here: α is the coefficient of space dilation in the normed direction $\xi \in \mathbb{R}^n$, $\|\xi\|=1$; I_n is the identity $n \times n$ -matrix.

Shor's algorithms use the „inverse“ operator

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

which means "compression" of subgradients space.

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



The 1d-ellipsoid \mathcal{E}_n , containing half of ball S_n in \mathbb{R}^n , has parameters

$$b = \left(\alpha + \frac{1}{\alpha} \right) \frac{r}{2}, \quad h = \left(1 - \frac{1}{\alpha^2} \right) \frac{r}{2},$$

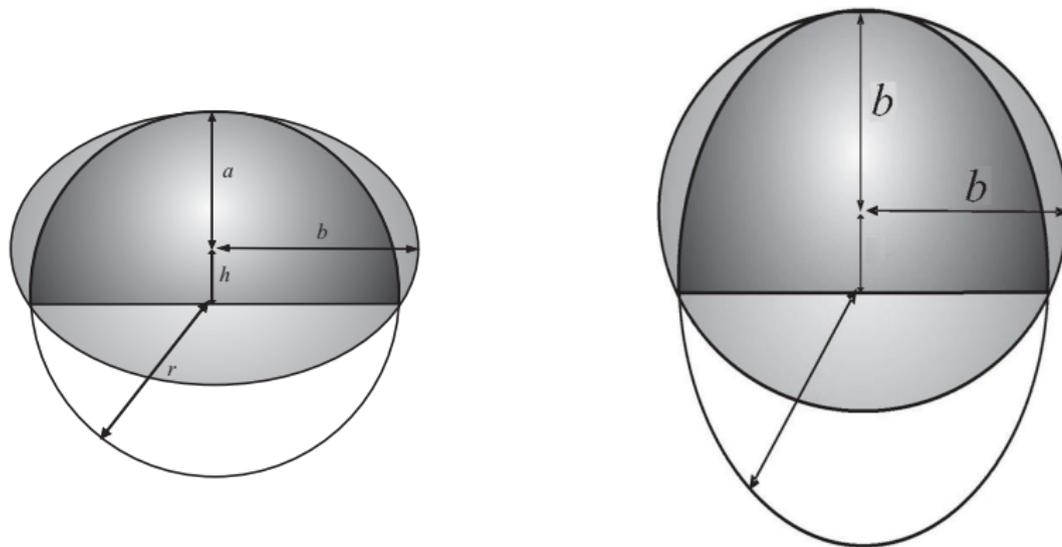
where $\alpha = \frac{b}{a}$ and r – radius of ball.

To transform \mathcal{E}_n into a „new“ ball we have to dilate the space with coefficient $\alpha = \frac{b}{a}$.

The ratio of \mathcal{E}_n volume to S_n volume equals

$$q_n(\alpha) = \frac{\text{vol}(\mathcal{E}_n)}{\text{vol}(S_n)} = \frac{1}{\alpha} \left(\frac{b}{r} \right)^n = \frac{1}{\alpha} \left(\frac{1}{2} \left(\alpha + \frac{1}{\alpha} \right) \right)^n.$$

1d-ellipsoid before and after space dilation



The 1d-ellipsoid \mathcal{E}_n is a ball in the dilated space.
 The radius of this ball $r := b$ and we can repeat this procedure (one iteration of ellipsoid method).

What is the GEM?

The generalized ellipsoid method (GEM)

is an algorithm with dilation of the n -dimensional space, where the space dilation coefficient satisfies the inequality

$$\alpha + \frac{1}{\alpha} < 2\sqrt[n]{\alpha}, \quad (1)$$

for which the volume reduction factor

$$q_n(\alpha) = \frac{1}{\alpha} \left(\frac{1}{2} \left(\alpha + \frac{1}{\alpha} \right) \right)^n < 1.$$

Outline

- 1 GEM (basic idea)
- 2 The GEM description**
- 3 Special cases of GEM
- 4 Algorithm emshor
- 5 Acceleration using 2d-ellipsoid

The GEM application

The GEM is used to determine x^* for the problem:

Let a mapping $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be given. We assume that there is $x^* \in \mathbb{R}^n$ so that $g(x)^\top (x - x^*) \geq 0$ for all $x \in \mathbb{R}^n$ and $g(x) \neq 0$ for all $x \neq x^*$.

The GEM can be use for the following problems:

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

The B -form of GEM

Step 0. Choose $x_0 \in \mathbb{R}^n$, a matrix $B_0 \in \mathbb{R}^{n \times n}$, and r_0 so that

$$\|B_0^{-1}(x_0 - x^*)\| \leq r_0.$$

Moreover, choose α according to (1) and set $k := 0$.

Step 1. Set $g_k := g(x_k)$. If $g_k = 0$, then set $x^* := x_k$ and STOP.

Step 2. Calculate

$$x_{k+1} := x_k - h_k B_k \xi_k, \quad \text{where} \quad \xi_k := \frac{B_k^\top g_k}{\|B_k^\top g_k\|}, \quad h_k := \frac{1}{2} \left(1 - \frac{1}{\alpha^2}\right) r_k.$$

Step 3. Update

$$B_{k+1} := B_k + \left(\frac{1}{\alpha} - 1\right) (B_k \xi_k) \xi_k^\top \quad \text{and} \quad r_{k+1} := \frac{1}{2} \left(\alpha + \frac{1}{\alpha}\right) r_k.$$

Step 4. Set $k := k + 1$ and go to Step 1.

The convergence of the B -form of GEM

Theorem 1

Let x_k and x_{k+1} be generated by the B -form of GEM. Then,

$$x^* \in \mathcal{E}_k \quad (2)$$

is satisfied. Moreover, the ratio of volumes of the ellipsoids \mathcal{E}_{k+1} and \mathcal{E}_k does not depend on k and is equal to

$$q_n(\alpha) := \frac{\text{vol}(\mathcal{E}_{k+1})}{\text{vol}(\mathcal{E}_k)} = \frac{1}{\alpha} \left(\frac{1}{2} \left(\alpha + \frac{1}{\alpha} \right) \right)^n < 1. \quad (3)$$

Here the ellipsoid \mathcal{E}_k is defined as

$$\mathcal{E}_k := \{x \mid \|B_k^{-1}(x_k - x)\| \leq r_k\}.$$

The H -form of GEM

Step 0. Choose $x_0 \in \mathbb{R}^n$, $H_0 = B_0 B_0^\top \in \mathbb{R}^{n \times n}$, and r_0 so that

$$(x_0 - x^*)^\top H_0^{-1} (x_0 - x^*) \leq r_0^2.$$

Moreover, choose α according to (1) and set $k := 0$.

Step 1. Set $g_k := g(x_k)$. If $g_k = 0$, then set $x^* := x_k$ and STOP.

Step 2. Calculate

$$x_{k+1} := x_k - h_k \frac{H_k g_k}{\sqrt{g_k^\top H_k g_k}}, \quad \text{where } h_k := \frac{1}{2} \left(1 - \frac{1}{\alpha^2} \right) r_k.$$

Step 3. Update

$$H_{k+1} := H_k + \left(\frac{1}{\alpha^2} - 1 \right) \frac{H_k g_k g_k^\top H_k}{g_k^\top H_k g_k} \quad \text{and} \quad r_{k+1} := \frac{1}{2} \left(\alpha + \frac{1}{\alpha} \right) r_k.$$

Step 4. Set $k := k + 1$ and go to Step 1.

The H -form derivation

It is easy to show that

$$x_{k+1} = x_k - h_k B_k \xi_k = x_k - h_k \frac{H_k g_k}{\sqrt{g_k^\top H_k g_k}},$$

$$B_{k+1} B_{k+1}^\top = H_k + \left(\frac{1}{\alpha^2} - 1 \right) \frac{H_k g_k g_k^\top H_k}{g_k^\top H_k g_k}$$

holds for all iterates with $k = 0, 1, 2, \dots$ for the B -form and defines

$$H_{k+1} := B_{k+1} B_{k+1}^\top$$

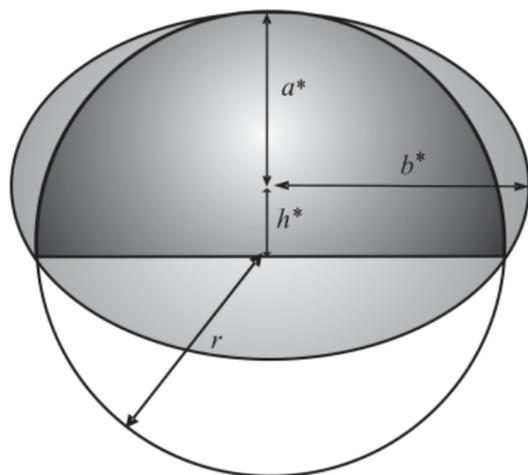
for these k . Here \mathcal{E}_k has an equivalent representation

$$\mathcal{E}_k = \{x \mid \|B_k^{-1}(x_k - x)\| \leq r_k\} = \{x \mid (x_k - x) H_k^{-1} (x_k - x) \leq r_k^2\}.$$

Outline

- 1 GEM (basic idea)
- 2 The GEM description
- 3 Special cases of GEM**
- 4 Algorithm emshor
- 5 Acceleration using 2d-ellipsoid

EM (Yudin-Nemirovski-Shor, 1976–1977)



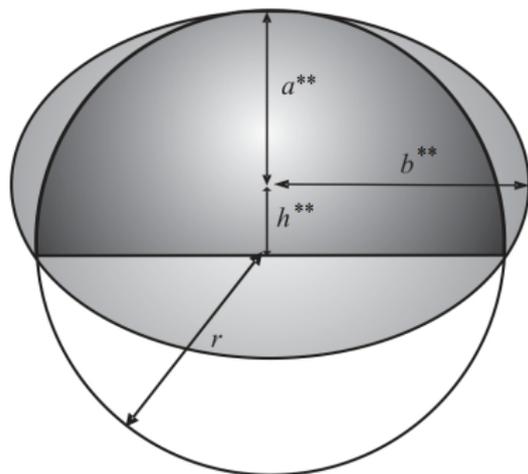
$$a^* = r \frac{n}{n+1},$$

$$b^* = r \frac{n}{\sqrt{n^2 - 1}},$$

$$h^* = r \frac{1}{n+1}.$$

$$\alpha^* = \frac{b^*}{a^*} = \sqrt{\frac{n+1}{n-1}}, \quad q_n^* = \frac{n}{n+1} \left(\frac{n}{\sqrt{n^2 - 1}} \right)^{n-1}.$$

AEM (Stetsyuk, 2003)



$$a^{**} = r \left(1 - \frac{1}{n} \left(\sqrt{1 + \frac{1}{n^2}} - \frac{1}{n} \right) \right),$$

$$b^{**} = r \sqrt{1 + \frac{1}{n^2}},$$

$$h^{**} = \frac{r}{n} \left(\sqrt{1 + \frac{1}{n^2}} - \frac{1}{n} \right).$$

$$\alpha^{**} = \frac{b^{**}}{a^{**}} = \sqrt{1 + \frac{1}{n^2}} + \frac{1}{n}, \quad Q_n^* = \left(1 + \frac{1}{n^2} \right)^{n/2} \left(\sqrt{1 + \frac{1}{n^2}} - \frac{1}{n} \right).$$

Comparison of EM and AEM iterations

n	q_n^*	Q_n^*	Q_n^*/q_n^*	k_q	k_Q
2	0.7698004	0.7725425	1.0035621	177	179
3	0.8437500	0.8441633	1.0004898	407	408
4	0.8813189	0.8814234	1.0001186	730	730
5	0.9042245	0.9042600	1.0000392	1144	1144
6	0.9196855	0.9197001	1.0000159	1651	1651
7	0.9308347	0.9308416	1.0000074	2249	2250
8	0.9392592	0.9392628	1.0000038	2940	2940
9	0.9458508	0.9458528	1.0000021	3723	3723
10	0.9511498	0.9511510	1.0000012	4598	4598

k_q, k_Q are smallest integers with $(q_n^*)^{k_q} \leq 10^{-10n}$, $(Q_n^*)^{k_Q} \leq 10^{-10n}$

Outline

- 1 GEM (basic idea)
- 2 The GEM description
- 3 Special cases of GEM
- 4 Algorithm emshor**
- 5 Acceleration using 2d-ellipsoid

Algorithm emshor can be used

for a convex function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ to find the minimizer x_ε^* , such that $f(x_\varepsilon^*) \leq f^* + \varepsilon$ for some predefined $\varepsilon > 0$.

The **emshor** algorithm can be derived from the B -form of GEM to $g := g_f$ with $\alpha := \sqrt{\frac{n+1}{n-1}}$, $B_0 := I_n$ and stopping criterion:

$$\begin{aligned}
 f(x_k) - f^* &\leq (x_k - x^*)^\top g_f(x_k) \\
 &= (B_k^{-1}(x_k - x^*))^\top B_k^\top g_f(x_k) \\
 &\leq \|B_k^{-1}(x_k - x^*)\| \|B_k^\top g_f(x_k)\| \\
 &\leq r_k \|B_k^\top g_f(x_k)\|.
 \end{aligned} \tag{4}$$

The subgradient $g_f = g_f(x)$ satisfies

$$(x - x^*)^\top g_f(x) \geq f(x) - f(x^*) = f(x) - f^* \geq 0 \quad \text{for } x \in \mathbb{R}^n. \tag{5}$$

Algorithm emshor

Step 0. Choose $x_0 \in \mathbb{R}^n$ and r_0 so that $\|x_0 - x^*\| \leq r_0$.

Moreover, choose $\varepsilon > 0$, set $B_0 := I_n$ and $k := 0$.

Step 1. If $\|B_k^\top g_f(x_k)\| r_k \leq \varepsilon$, then set $k^* := k$, $x_\varepsilon^* := x_k$ and STOP.

Step 2. Calculate

$$x_{k+1} := x_k - h_k B_k \xi_k, \quad \text{where} \quad \xi_k := \frac{B_k^\top g_f(x_k)}{\|B_k^\top g_f(x_k)\|}, \quad h_k := \frac{1}{n+1} r_k.$$

Step 3. Update

$$B_{k+1} := B_k + \left(\sqrt{\frac{n-1}{n+1}} - 1 \right) (B_k \xi_k) \xi_k^\top \quad \text{and} \quad r_{k+1} := \frac{n}{\sqrt{n^2 - 1}} r_k.$$

Step 4. Set $k := k + 1$ and go to Step 1.

The convergence of the emshor

Theorem 2

Let x_k and x_{k+1} be generated by the **emshor** algorithm. Then,

$$x^* \in \mathcal{E}_k$$

is satisfied and the ratio of the volumes of the ellipsoids \mathcal{E}_{k+1} and \mathcal{E}_k does not depend on k and is equal to

$$q_n := \frac{\text{vol}(\mathcal{E}_{k+1})}{\text{vol}(\mathcal{E}_k)} = \frac{n}{n+1} \left(\frac{n}{\sqrt{n^2-1}} \right)^{n-1} < \exp \left\{ -\frac{1}{2n} \right\} < 1.$$

Moreover, if algorithm **emshor** stops, then $f(x_\varepsilon^*) \leq f^* + \varepsilon$ holds.

Computing experience with emshor

The emshor algorithm:

1. finds very precise approximations to the minimizer of the ravine convex function;
2. the number of iterations grows slightly faster than n^2 (to reduce the deviation of the function value from f^* by a factor of 10, the EM requires $\approx 4.6n^2$ iterations).

We show this for a ravine piecewise linear convex function

$$f(x) = \sum_{i=1}^n 2^{i-1} |x_i - 1|, \quad x^* = (1, 1, \dots, 1)^T \in \mathbb{R}^n, \quad f^* = 0. \quad (6)$$

The results for minimization of function (6)

$$x_0 := 0 \in \mathbb{R}^n, r_0 = 5$$

n	$\varepsilon = 10^{-3}$		$\varepsilon = 10^{-6}$		$\varepsilon = 10^{-9}$	
	itn	$f(x_{itn})$	itn	$f(x_{itn})$	itn	$f(x_{itn})$
5	519	6.1e-06	873	1.1e-07	1201	1.2e-10
10	2484	8.7e-05	3829	7.2e-08	5246	7.9e-11
15	6561	6.5e-06	9667	6.0e-08	12786	1.5e-11
20	13101	4.8e-05	18714	3.5e-09	23416	2.0e-11

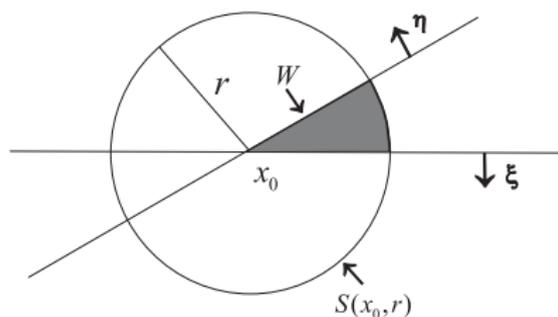
$$x_0 := 0 \in \mathbb{R}^n, r_0 = 500$$

n	itn	$f(x_{itn})$	itn	$f(x_{itn})$	itn	$f(x_{itn})$
5	747	1.1e-05	1080	1.6e-07	1392	1.7e-10
10	3429	9.0e-05	4810	9.3e-08	6185	6.3e-11
15	8615	5.6e-05	11704	6.5e-08	14805	2.4e-11
20	16729	1.8e-06	22404	4.4e-08	27161	1.4e-11

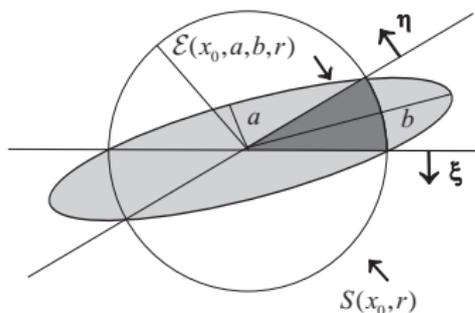
Outline

- 1 GEM (basic idea)
- 2 The GEM description
- 3 Special cases of GEM
- 4 Algorithm emshor
- 5 Acceleration using 2d-ellipsoid**

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



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



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



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

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 equals

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

If the angle between vectors ξ and η becomes closer to π (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 dilate space in the direction of $\frac{\xi - \eta}{\|\xi - \eta\|}$ with

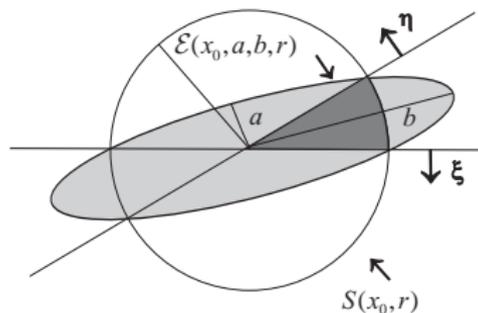
$$\alpha_1 = \frac{r}{a} = \frac{1}{\sqrt{1 + (\xi, \eta)}} > 1,$$

then we „dilate“ space in the direction of $\frac{\xi + \eta}{\|\xi + \eta\|}$ with

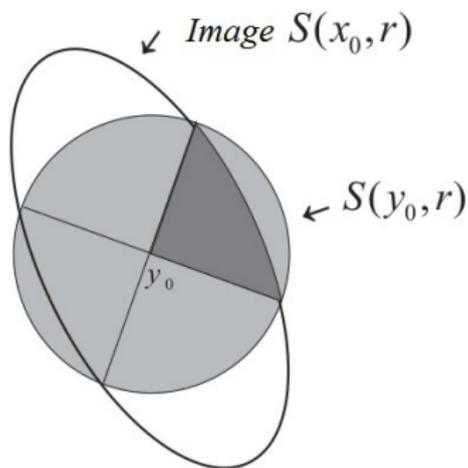
$$\alpha_2 = \frac{r}{b} = \frac{1}{\sqrt{1 - (\xi, \eta)}} < 1.$$

2d-ellipsoid before and after transformation

In the transformed space $Y = R_{\alpha_2} \left(\frac{\xi + \eta}{\|\xi + \eta\|} \right) R_{\alpha_1} \left(\frac{\xi - \eta}{\|\xi - \eta\|} \right) X \equiv \mathbb{R}^n$



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



becomes a ball $S(y_0, r)$

Conclusion

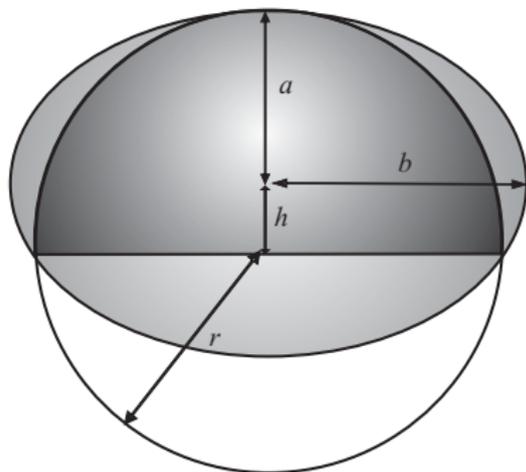
The accelerated variants of ellipsoid methods on the basis of the 1d-ellipsoid and the 2d-ellipsoid can be used for solving a variety of problems: convex programming problems, problems of finding saddle points of convex-concave functions, and special cases of variational inequalities.

Thanks

Volkswagen Foundation
grant No 90 306

THANK YOU
FOR YOUR ATTENTION!

BACKUP SLIDES: 1d-ellipsoid (Shor, 1977)



Minimal volume 1d-ellipsoid \mathcal{E}_n , containing half-ball in \mathbb{R}^n , has parameters

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

Dilating the space by a factor of

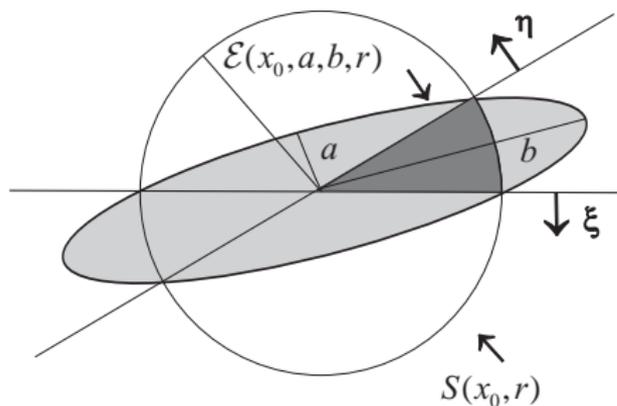
$$\alpha = \frac{b}{a} = \sqrt{\frac{n+1}{n-1}}$$

we transform \mathcal{E}_n into ball.

The ratio of \mathcal{E}_n volume to S_n volume equals

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

BACKUP SLIDES: 2d-ellipsoid (Stetsyuk, 1996)



The space transformation is two space dilations:

in the direction $\frac{\xi-\eta}{\|\xi-\eta\|}$ with $\alpha_1 = \frac{r}{a} = \frac{1}{\sqrt{1+(\xi,\eta)}} > 1$,

in the direction $\frac{\xi+\eta}{\|\xi+\eta\|}$ with $\alpha_2 = \frac{r}{b} = \frac{1}{\sqrt{1-(\xi,\eta)}} < 1$.

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