ELLIPSOID METHODS WITH SPACE SCALING*

P.I. STETSYUK[1], A. FISCHER[2], O.M. KHOMIAK[1]

[1] V.M. Glushkov Institute of Cybernetics,Academy of Sciences of Ukraine, Kyiv, Ukraine
[2]Institute of Numerical Mathematics, Technische Universität Dresden, Germany
e-mail: stetsyukp@gmail.com, andreas.fischer@tu-dresden.de, khomiak.olha@gmail.com

1. Introduction

The ellipsoid method for minimizing a convex function was proposed independently by Yudin and Nemirovski [8] in 1976 and by Shor in 1977 [6]. In the latter paper, the ellipsoid method is presented as a special case of methods with space dilation in the direction of a subgradient. Space dilation methods were proposed by Shor at the end of the sixties [5] with the aim of speeding up the convergence of gradient methods. Based on the framework of methods with space dilation, we suggest an ellipsoid method with space scaling by a parameter $\lambda > 0$. For certain values of λ, existing variants of the ellipsoid method are obtained, namely those by Shor [6], Nemirovski and Yudin [3, page 76]), and Khachiyan [2, Lemma 4]. For details see Table 3.

The ellipsoid method with space scaling is presented in Section 2. It is constructed to find an ε-approximation to the minimum point of the convex function $f : \mathbb{R}^n \to \mathbb{R}$. Moreover, Section 2 also includes theorems on the convergence of the proposed ellipsoid method with space scaling. Finally, results of computational experiments for the three indicated variants of this method applied to a piecewise linear function f are shown in Section 3.

Section 2 presents the computational scheme of the algorithm and theorems on its convergence. Then, in Section 3, we show results of computational experiments for finding the ε-approximation to the minimum point of a convex piece-wise linear function for small ε using the above three variants of the ellipsoid method.

2. Algorithm EM22B for minimizing convex function

Again, let $f : \mathbb{R}^n \to \mathbb{R}$ be a convex function. Its minimum value is denoted by $f^* = f(x^*)$, where x^* is a minimum point. For any $x \in \mathbb{R}^n$, let $g(x)$ be a subgradient of f at x, i.e.,

$$(x - x^*)^\top g(x) \geq f(x) - f^* \quad \text{for all } x \in \mathbb{R}^n \tag{1}$$

is satisfied.

The EM22B algorithm is designed to find a ε-approximation to x^*, i.e., a point x^*_ε for which $f(x^*_\varepsilon) - f^* \leq \varepsilon$, where $\varepsilon > 0$ is given. Its name refers to **e**llipsoid **m**ethod, the year **2022**, and the use of the ***B***-form (see subsequent paragraph). The algorithm depends on the chosen scaling parameter λ, the starting point x_0, a radius r_0, and the desired approximation precision ε.

For the space dilation, we are here using the B-form of such techniques, see [4]. To this end, let us consider two spaces $X = Y = \mathbb{R}^n$, a nonsingular matrix $A \in \mathbb{R}^{n\times n}$, and the space transformation from $x \in X$ to $y \in Y$ defined by $y := Ax$. Then, with $B := A^{-1}$, we get $x = By$

*This work was partially supported by Volkswagen Foundation (grant No.97 775) and CRDF Global (grant G-202102-68020).

for the inverse transformation. The following algorithm updates the B-matrix in each step.

Algorithm EM22B(λ, x_0, r_0, ε)

Step 0. Choose $\lambda > 0$, $x_0 \in \mathbb{R}^n$, $r_0 > 0$, $\varepsilon > 0$ such that $\|x_0 - x^*\| \le r_0$. Set $B_0 := I_n \in \mathbb{R}^{n\times n}$ (denoting the identity matrix) and $k := 0$.

Step 1. If $\left\|B_k^\top g(x_k)\right\| r_k \le \varepsilon$, then STOP: $k^* := k$, $x_\varepsilon^* := x_k$.

Step 2. Compute $x_{k+1} := x_k - \frac{r_k}{n+1} B_k \xi_k$, where $\xi_k := \frac{B_k^\top g(x_k)}{\left\|B_k^\top g(x_k)\right\|}$.

Step 3. Update $B_{k+1} := \lambda \left(B_k + \left(\sqrt{\frac{n-1}{n+1}} - 1\right)(B_k \xi_k)\, \xi_k^\top\right)$ and $r_{k+1} := \frac{1}{\lambda} \frac{n}{\sqrt{n^2-1}} r_k$.

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

Obviously, updating the B-matrix (see Step 3) requires $O(n^2)$ operations. This is due to the the use of space scaling parameter λ and the space dilation operator $R_\alpha(\xi) : \mathbb{R}^n \to \mathbb{R}^n$ defined as

$$R_\alpha(\xi) := I_n + (\alpha - 1)\xi\xi^\top, \tag{2}$$

where $\alpha > 0$ and $\xi \in \mathbb{R}^n$ with $\|\xi\| = 1$ is the direction of dilation. For each k, the above algorithm uses ξ_k as new direction of dilation, see Step 2. Denoting the inverse dilation operator as $R_\alpha^{-1}(\xi)$ and setting $\beta := 1/\alpha$, we have

$$R_\alpha^{-1}(\xi) = R_\beta(\xi), \quad B_{k+1} = \lambda B_k R_\beta(\xi_k), \quad \text{and} \quad \beta = \sqrt{\frac{n-1}{n+1}}.$$

This also shows the roles of the parameter λ and the space dilation operator (2) for updating B-matrices.

Theorem 1. For any $(\lambda, x_0, r_0, \varepsilon) \in (0,\infty) \times \mathbb{R}^n \times (0,\infty) \times (0,\infty)$, algorithm EM22B is well-defined and generates a sequence $\{x_k\}_{k=0}^{k^*}$. With $A_k := B_k^{-1}$, it holds that

$$\|A_k(x_k - x^*)\| \le r_k \quad \text{for } k = 0, 1, 2, \ldots, k^*. \tag{3}$$

The proof of Theorem 1 can be carried out similar to the one of Theorem 1 [1].

For any $k \in \{0, 1, \ldots, k^*\}$, the set

$$E_k := \{x \in \mathbb{R}^n \mid \|A_k(x_k - x)\| \le r_k\}$$

is an ellipsoid that, due to (3), contains x^*. For its volume, we have $vol(E_k) = v_0 r_k^n / \det A_k$, where v_0 is the volume of the Euclidean n-dimensional unit ball and $\det A_k$ denotes the determinant of A_k. The rate of convergence of the EM22B algorithm is determined by the ratio of the volumes of two consecutively generated ellipsoids.

Theorem 2. There is $k^* \in \mathbb{N}$ so that algorithm $EM22B$ stops at Step 1 for $k = k^*$. For each k with $1 \le k \le k^*$, the ratio of the volumes of the ellipsoids E_k and E_{k-1} is a constant q_n with

$$q_n = \frac{vol(E_k)}{vol(E_{k-1})} = \sqrt{\frac{n-1}{n+1}} \left(\frac{n}{\sqrt{n^2-1}}\right)^n < \exp\left\{-\frac{1}{2n}\right\} < 1. \tag{4}$$

Moreover, $f(x_{k^*}) - f^* \le \varepsilon$ is satisfied.

The proof of Theorem 2 can be done similar to that of [1, Theorem 2]. An important part is as follows.

Using Theorem 2, the Cauchy-Schwarz inequality, and $\|\xi_k\| = 1$, we get for any $k \in \{0, \ldots, k^*\}$

$$r_k \ge \|A_k(x_k - x^*)\| = \|B_k^{-1}(x_k - x^*)\|\|\xi_k\| \ge (B_k^{-1}(x_k - x^*))^\top \xi_k = \frac{(x_k - x^*)^\top g(x^k)}{\|B_k^\top g(x_k)\|}. \tag{5}$$

If the algorithm stops in Step 1, then $\left\|B_k^\top g(x_{k^*})\right\| r_{k^*} \le \varepsilon$ is fulfilled. This, (5), and (1) imply

$$\varepsilon \ge r_{k^*} \|B_{k^*}^\top g(x_{k^*})\| \ge (x_{k^*} - x^*)^\top g(x_{k^*}) \ge f(x_{k^*}) - f^*.$$

3. Computational experiment

The computational experiments were carried out on a computer with an AMD Ryzen 5 4500U 2.38 GHz processor and 16 GB memory on a Windows 10 system using GNU Octave version 6.2.0. The values of λ we considered correspond to three versions of the ellipsoid method by Shor 1977 [6], Khachiyan 1980 [2], Nemirovski and Yudin 1979 [3], see Table 1 below.

Table 1. Characteristics of three variants of the ellipsoid method.

λ	Update of the B-matrix	Multiplications	Updated radius	Reference
1	$B_1 = B + \left(\sqrt{\frac{n-1}{n+1}} - 1\right)(B\xi)\xi^\top$	$3n^2 + n$	$r_1 = \frac{n}{\sqrt{n^2-1}} r$	[6]
$\frac{n}{\sqrt{n^2-1}}$	$B_2 = \frac{n}{\sqrt{n^2-1}} B_1$	$4n^2 + n$	$r_2 = r$	[2]
$\sqrt[n]{\frac{n+1}{n-1}}$	$B_3 = \left(\frac{n+1}{n-1}\right)^{\frac{1}{2n}} B_1$	$4n^2 + n$	$r_3 = \left(\frac{n-1}{n+1}\right)^{\frac{1}{2n}} r_1$	[3]

Algorithm EM22B is applied to $f(x) = \sum_{i=1}^{10} 2^{i-1}|x_i - 1|$, a convex piecewise linear function. For small ε, Table 2 shows results for computing x^*_ε with $f(x^*_\varepsilon) \leq f^* + \varepsilon$. Though the three versions of EM22B are equivalent, we observe slight differences in the number of iterations for $\varepsilon \in \{10^{-7}, 10^{-8}\}$ due to accumulation of numerical errors. A study of such effects for different f, n, and ε is intended. Algorithm EMB22 can be accelerated by tighter ellipsoidal approximations [7] and applied to convex programs or saddle point problems for convex-concave functions.

Table 2. Results for applying EM22B with $x_0 = (0, \ldots, 0)^\top$, $r_0 = 10$.

		Shor [6]			Khachiyan [2]			Nemirovski and Yudin [3]		
ε	f^*_ε	k^*	$\Vert B^*_k\Vert$	r^*_k	k^*	$\Vert B^*_k\Vert$	r^*_k	k^*	$\Vert B^*_k\Vert$	r^*_k
1.0e-04	2.2e-06	3124	6.4e-13	6.6e+07	3124	4.2e-06	10	3124	2.6e+01	1.6e-06
1.0e-06	2.0e-09	4024	8.1e-17	6.1e+09	4024	4.9e-08	10	4024	2.8e+01	1.8e-08
1.0e-07	6.9e-09	4474	9.0e-19	5.8e+10	4474	5.2e-09	10	**4490**	2.8e+01	1.7e-09
1.0e-08	6.5e-10	**4827**	2.6e-20	3.4e+11	**4934**	5.0e-10	10	**4953**	2.8e+01	1.7e-10

Keywords: Ellipsoid Method, Convex Function, Scaling, Operator Of Space Dilation.

AMS Subject Classification: 90C25, 90C30, 90C90.

References

[1] Izmailov A.F., Stetsyuk P.I., Fischer A., Emshor algoritm and its octave implementation, *Computer Mathematics*, No.1, 2019, pp.132-142.
[2] Khachian L.G., Polynomial algorithms in linear programming, *USSR Computational Mathematics and Mathematical Physics*, Vol.20, No.1, 1980, pp.53-72.
[3] Nemirovski A., Yudin D., *Problem Complexity and Method Efficiency in Optimization*, John Wiley, New York, 1983, 388 p., translated from book published by Nauka, Moscow, 1979.
[4] Sergienko I.V., Stetsyuk P.I., On N.Z. Shor's three scientific ideas, *Cybernetics and Systems Analysis*, Vol.48, No.1, 2012, pp.2-16.
[5] Shor N.Z., Utilization of the operation of space dilation in the minimization of convex functions, *Cybernetics and Systems Analysis*, Vol.6, No.1, 1972, pp.7-15, translated from *Kibernetika*, Vol.6, No.1, 1970, pp.6-12.
[6] Shor N.Z., Cut-off method with space extension in convex programming problems, *Cybernetics*, Vol.13, No.1, 1977, pp.94-96.
[7] Stetsyuk P., Fischer A., Khomyak O., The generalized ellipsoid method and its implementation, *Optimization and Applications. OPTIMA 2019. Communications in Computer and Information Science*, (M. Jaćimović, M. Khachay, V. Malkova, M. Posypkin, editors), Vol.1145, 2020, pp.355-370.
[8] Yudin D.B., Nemirovski A.S., Informational complexity and effective methods of solution of convex extremal problems, *Èkonomika i Matematicheskie Metody*, Vol.12, No.2, 1976, pp.357-369 (in Russian).