

A Construction of the Minimum Volume Ellipsoid Containing a Set of Points using BRKGA Metaheuristic

Antonio A. M. Raposo¹

Federal Institute of Education, Science and Technology of São Paulo (IFSP), Tupã, SP, Brazil

São Paulo State University (UNESP), School of Sciences and Engineering, Tupã, SP, Brazil

Valeska Martins de Souza²

Federal University of Maranhão (UFMA), Mathematics Department, São Luís, MA, Brazil

Luís Roberto A. G. Filho³

São Paulo State University (UNESP), School of Sciences and Engineering, Tupã, SP, Brazil

Abstract. The problem of obtaining a Minimum Volume Enclosing Ellipsoid (MVEE) of a given point set $C = \{x_1, \dots, x_n\} \subseteq \mathbb{R}^n$ is found in several practical applications. This paper proposes an alternative methodology to build these ellipsoids of minimum volume through a metaheuristic based on the Biased Random-Keys Genetic Algorithm (BRKGA) in order to reduce the computational cost in solving n -dimensional MVEE problems. The formulation was implemented in Python and compared with the CVX package implemented in MATLAB for 10 two-dimensional instances. The results showed that BRKGA generated solutions with a high level of accuracy and low computational cost.

Keywords. Minimum Volume Ellipsoid, BRKGA, Löwner-John Ellipsoid, Optimization, Python.

1 Introduction

The problem of obtaining a MVEE a certain number of points C has multiple applications in statistics and optimal design, integer programming, computational geometry, randomized algorithms for computing polytope volume, placing relay antennae, and packing [1]. In the area of robust statistics and data mining, efficiently finding outliers is a challenge that has attracted research interest [9].

The MVEE problem can be computed by several algorithms. Algorithms that use the interior point approach were introduced by Zhang et al., and Khachiyan and Todd. The one provided by Barnes relies on quadratic programming. There are other algorithms of the stochastic type, introduced by Gärtner and Schönherr. Other studies of the problem are by Yildirim et al. [1]. An algorithm based on Genetic Algorithm (GA) is proposed in [1]. Although the authors have obtained satisfactory results for some instances, the strategy used to obtain future generations from crossover and mutation operators, as well as the construction of the fitness of the MVEE problem use random numbers entered arbitrarily by the authors [1].

From the bibliographic review, it is noted that the works developed on the theme do not usually address bio-inspired algorithms in the search for optimal solutions, mainly due to the difficulty of adapting evolutionary operators, as well as the fitness function. Hence, the objective of this paper

¹antonio.raposo@ifsp.edu.br & antonio.raposo@unesp.br.

²valeska.martins@ufma.br.

³gabriel.filho@unesp.br.

is to propose an alternative, computationally competitive, that provides high quality solutions for the MVEE problem using an evolutionary approach based on BRKGA metaheuristic [2]. It requires a suitable representation of the solutions, a modified fitness function as well as suitable crossover and mutation operators, and stopping criteria. The paper is organized as follows. Section 2 is the formulation MVEE problem. Section 3 is the proposed implementation of BRKGA for solving the MVEE. The results are shown in the Section 4. The Section 5 is the conclusion.

2 Formulation of the MVEE Problem

Consider a nonsingular affine map of the form $\Phi(x) := Ax + b$, where $\Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is linear and $b \in \mathbb{R}^n$ is a constant vector. An ellipsoid in \mathbb{R}^n is the image of the unit ball $\mathbb{B}^n \subseteq \mathbb{R}^n$ under $\Phi(x)$. The ellipsoid can be defined by

$$\mathbb{E}_{A,b} = \{x \in \mathbb{R}^n \mid (x - b)^T A^{-1}(x - b) \leq 1\}, \quad (1)$$

where b is the center of the ellipsoid, $A \in \mathbb{R}_{n \times n}$, and $A = A^T \succ 0$.

The MVEE or Löwner-John ellipsoid [4] of a set \mathbb{C} is the minimum volume ellipsoid that contains set \mathbb{C} . Computing the MVEE of a set \mathbb{C} is equivalent to computing the MVEE of a polytope that is defined by the convex hull of those points. Thus, the MVEE problem can be defined as

$$\begin{aligned} & \text{minimize} && f(A, b) = \log(\det(A^{-1})) \\ & \text{subject to} && \|Ax_i + b\|_2 \leq 1, \quad i = 1, \dots, k, \\ & && A = A^T \succ 0. \end{aligned} \quad (2)$$

The volume of $\mathbb{E}_{A,b}$ is given by the formula [9]:

$$V_{\mathbb{E}_{A,b}} = \frac{\pi^{n/2} \sqrt{\det(A^{-1})}}{\Gamma(n/2 + 1)}, \quad (3)$$

where $\Gamma(\cdot)$ is the standard gamma function of calculus.

As mentioned earlier, there are several methods and software that do not use bio-inspired algorithms to obtain solutions to the problem (2). For example, the CVX package implemented in MATLAB [3], uses the Path-Follow Infeasible Interior-Point algorithm to solve (2). Instances of the model described in (2) can be submitted to this package directly. However, n -dimensional problems have high computational costs. Figure 1 illustrates the solutions obtained by the CVX package and through the proposed technique based on BRKGA. The data shown in the Figure 1 is the set of points $x_i \in \mathbb{R}^2, i = 1, \dots, 30$ randomly generated in the domain $[-10, 10] \times [-10, 10]$.

From the Figure 1, it can be seen that both CVX and BRKGA obtained satisfactory results for the two-dimensional problem. However, as the size of the problem increases, the computational cost of iterative algorithms increases as well. In this way, a way to reduce the computational cost, but maintaining an acceptable level of precision of the solutions is to explore the metaheuristics, as well as to implement them in programming languages of high numerical performance, e. g., C/C++, Julia or Python [5].

3 Formulation of the MVEE Problem via BRKGA

The proposed technique to minimize (2) is based on BRKGA [2]. BRKGA is an evolutionary metaheuristic for combinatorial optimization problems, whose solutions are represented by a vector of m random keys, where a random key is a real number, generated randomly, in the continuous

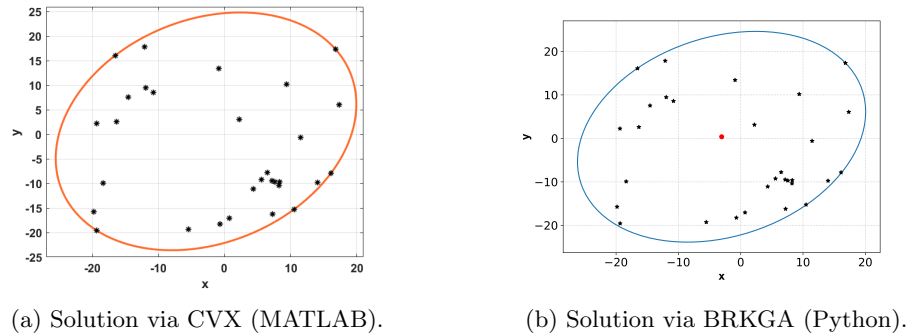


Figure 1: MVEE problem implemented in CVX *vs.* BRKGA.

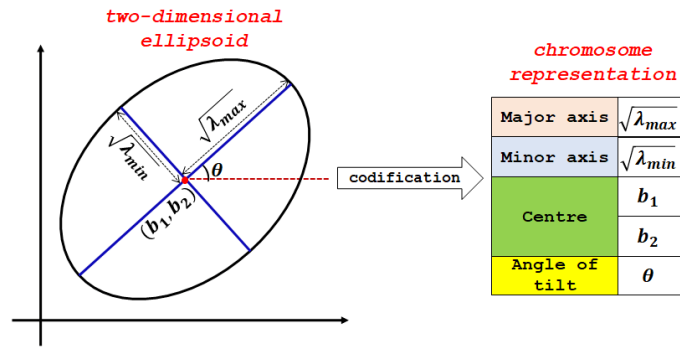
interval [0.1). A decoder maps this vector to a solution to the optimization problem and calculates the cost of this solution. The BRKGA population is made up of \mathcal{P} vectors of random keys, with $|\mathcal{P}| = m$. In each generation, the vectors are partitioned into a small set with the best individuals, called the elite set $\mathcal{S}_{\mathcal{E}}$, and the rest are allocated in a set called non-elite $\mathcal{S}_{\mathcal{NE}}$ ($|\mathcal{S}_{\mathcal{NE}}| = |\mathcal{P}| - |\mathcal{S}_{\mathcal{E}}|$). All elements of $\mathcal{S}_{\mathcal{E}}$ are copied without change for the next generation population. A small number of random key vectors are mutated and are also added to the next generation population. The rest of the next generation population consists of solutions generated by the uniform combination of solution pairs (crossover), in which one solution is elite and the other non-elite [2].

3.1 Solution representation

The BRKGA was designed for combinatorial optimization problems, having a wide application in engineering problems [6], [7], [8], but it also has applications in the field of continuous optimization. Consider $f(A, b)$ defined in (2). In \mathbb{R}^2 , A is a 2×2 matrix corresponding to four variables and b is a 2×1 column vector corresponding to two variables. In this way, in \mathbb{R}^n , a solution for (2) can be represented as a one-dimensional array with $n(n + 1)/2$ inputs of matrix A ($A = A^T$) and inputs of vector b , that is $(n^2 + 3n)/2$, or 5 in \mathbb{R}^2 . In other words, chromosomes of length 5 will have to be generated to form a population of BRKGA individuals.

A symmetric matrix A is positive definite if, and only if, all eigenvalues of A are positive. In this way, the matrix A is totally characterized by its eigenvalues and by the angle of tilt that the largest eigenvector supports A , that is, the largest axis of the ellipsoid, forms with the x axis. Based on this information and the chromosome representation proposed in [1], a partitioned chromosome is defined as coding for a candidate solution in BRKGA: (i) the first n positions correspond to the total number of axes of the ellipsoid ordered from the largest to the shorter length; (ii) the next n positions correspond to the centre of the ellipsoid, that is, the components of the vector $b \in \mathbb{R}^n$, and; (iii) the last position of the chromosome corresponds to the angle of tilt θ of the largest axis of the ellipsoid forms with the x axis.

The representation for a i -th d -dimensional chromosome P_i is $P_i = [p_{i1}, \dots, p_{id}]$ and illustrated in Figure 2, where i is an individual from the population of ellipsoids. Note in the Figure 2b that the chromosome's length can be generalized to $d = 2n + 1$, where n is a dimension of the ellipsoid. Moreover, $\sqrt{\lambda_{max}}$ and $\sqrt{\lambda_{min}}$ represent the largest and smallest axis of the ellipsoid, with λ_{max} and λ_{min} being the largest and smallest eigenvalue of A . The proposed chromosomal representation in this paper prevents matrix A from being represented by $n(n + 1)/2$ genes on the chromosome ($A = A^T, A \in \mathbb{R}_{n \times n}$), reducing by $n(n - 1)/2 - 1$ the number of necessary genes, that is, of memory space.



(a) Two-dimensional chromosome.

Axes			Centre			Angle of tilt
$\sqrt{\lambda_{max}}$...	$\sqrt{\lambda_{min}}$	b_1	...	b_n	θ
$\underbrace{\hspace{10em}}_n$			$\underbrace{\hspace{10em}}_n$			

(b) d -dimensional chromosome.

Figure 2: Chromosome representation of an ellipsoid.

3.2 Genetic operators for MVEE problem

From the initial ellipsoid solutions, randomly generated, BRKGA starts the process of generating new individuals by partitioning the current generation into an elite (\mathcal{S}_E) and non-elite (\mathcal{S}_{NE}) population. The next step in the process of generating new individuals is to apply genetic operators of mutation and crossover. This implementation is performed in a small number of random key vectors by alternating one or more genes. In the BRKGA proposed in this paper, if a gene mutates, the alteration process consists of randomly generating a value for that gene (axis of the ellipsoid, centre or angle of tilt). The next step in the process of population evolution consists of obtaining, from the crossover operator, new individuals according to a uniform combination of solution pairs, one solution belonging to the \mathcal{S}_E and one belonging to the \mathcal{S}_{NE} . The crossover produces a new child chromosome with potentially different axes, centres and angles of tilt. The values of the gene of this chromosome are selected randomly among the values of the gene of the parents according to (4). If the children generated have axes that are not in decreasing order, an ordering subroutine is performed in the first n positions of the child chromosome to correct this imperfection.

$$P_k = w^{\mathcal{S}_E} + u(t^{\mathcal{S}_{NE}} - w^{\mathcal{S}_E}), \tag{4}$$

where $w^{\mathcal{S}_E}$ is a chromosome randomly drawn from the \mathcal{S}_E ; $t^{\mathcal{S}_{NE}}$ is a chromosome randomly drawn from the \mathcal{S}_{NE} ; u is a uniformly distributed number in the interval $[0, 1]$, and; P_k is the k -th chromosome resulting from the crossover between the chromosomes $w^{\mathcal{S}_E}$ and $t^{\mathcal{S}_{NE}}$.

3.3 Fitness function

As shown in [1], applying metaheuristics to solve (2) can produce inaccurate individuals or solutions with the search space, leading the algorithm to stagnation or inaccurate results. The fitness function is based on the technique proposed in [1], but unlike this technique, the fitness function proposed in this paper is independent on the number of points that the optimal ellipsoid must contain, i.e., no empirical constant is created as a function of the number of points one wishes

to surround with a minimum volume ellipsoid. Let $F_{\mathbb{E}}$ the fitness of ellipsoid \mathbb{E} represented by the i -th chromosome P_i , \mathbb{C} the set of points to attach and \mathbb{C}_E the set of points that ellipsoid \mathbb{E} contains, then $F_{\mathbb{E}}$ can be defined as

$$F_{\mathbb{E}}(P_i) = \begin{cases} [e^{(|\mathbb{C}|-|\mathbb{C}_E|)/10^{\beta-1}} + V_{\mathbb{E}_{A,b}} - f(A,b)]^2 & , \text{ if } |\mathbb{C}_E| < |\mathbb{C}| \\ [V_{\mathbb{E}_{A,b}} - f(A,b)]^2 & , \text{ otherwise.} \end{cases} \quad (5)$$

The β constant in (5) represents the largest exponent in which the cardinality of \mathbb{C} can be expressed on a decimal basis. This constant has been included in (5) to avoid overflowing. The reset process can also be included in the BRKGA framework [2]. At this stage, the BRKGA is reinitialized every 50 generations. The BRKGA flowchart used to solve (5) is shown in Figure 3.

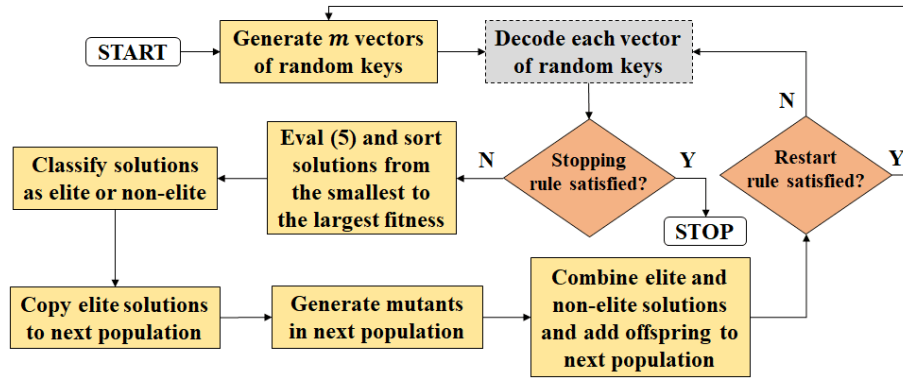


Figure 3: BRKGA flowchart for solving the MVEE problem.

4 Tests and Results

The proposed methodology was implemented using the Python programming language [5] with the following computational platform: desktop with Intel Core i7-7500U processor, 2.70 GHz CPU and 8 GB of RAM. The tests were performed for two-dimensional problems considering 10 instances. The CVX package, implemented in MATLAB [3], was used to compare the solutions obtained by BRKGA. The BRKGA parameters [6], [7]: mutation rate: 0.5; crossover rate: 0.7; population size (N_{pop}): 40; elite population size: $0.2N_{pop}$; population size for crossover: $0.6N_{pop}$; population size for mutation: $0.2N_{pop}$, and; maximum number of generations (g_{max}): 10000. The g_{max} for BRKGA is used as a stop rule, but if the standard deviation of the fitness function is less than or equal to 0.01 for every 2000 consecutive generations, then the BRKGA must be stopped.

The Figures 4a and 4b show the final solutions obtained by the CVX package and through BRKGA for 500 points, respectively. The figures 4c and 4d show the 40 initial solutions generated by BRKGA for 100 and 500 points, respectively. The Tables 1 and 2 illustrates the results obtained for 10 instances through BRKGA and CVX considering only one execution of each algorithm.

From Table 1, BRKGA's low cost stands out for problems involving up to 100 points. However, for problems with hundreds or thousands of points to be involved by a minimum volume ellipsoid, there is an increase in the computational cost of BRKGA in relation to the CVX package. Another result to be highlighted are the good approximations obtained by BRKGA for $V_{\mathbb{E}_{A,b}}$ using a relative number of generations as can be seen in Tables 1 and 2. In addition, Table 1 shows that the relative errors between $V_{\mathbb{E}_{A,b}}$ obtained through BRKGA and CVX are below 5 % for all case studies, being below 1 % for more than half of the analyzed instances.

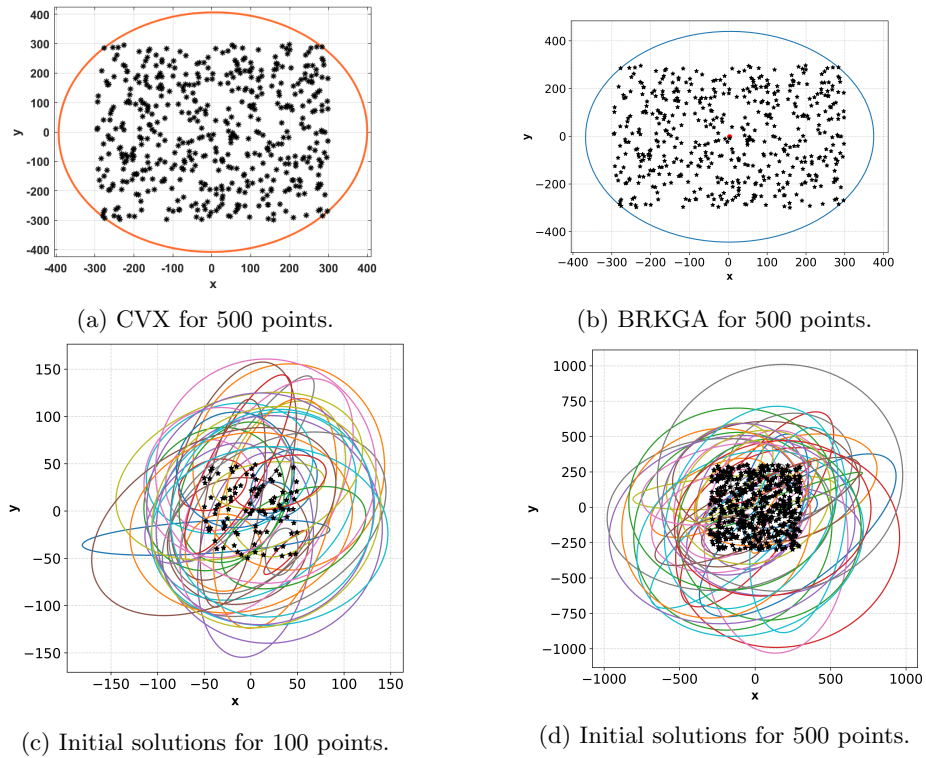


Figure 4: CVX and BRKGA solutions for MVEE problem.

Table 1: BRKGA and CVX results for 10 instances.

Points (C)	Domain	$V_{E_{A,b}}$ BRKGA (I)	Time (s) BRKGA	$V_{E_{A,b}}$ CVX (II)	Time (s) CVX	Relative error (%) (I) and (II)
5	$[20, 30] \times [20, 30]$	3.0714E01	16.1245	3.0416E01	2.6094	0.9794
10	$[20, 30] \times [20, 30]$	6.3528E01	14.3237	6.3087E01	2.3906	0.6986
20	$[-10, 10] \times [-10, 10]$	3.6565E02	33.9806	3.5165E02	2.6563	3.9802
30	$[-20, 20] \times [-20, 20]$	1.7163E03	25.4369	1.6661E03	2.9219	3.0113
40	$[-50, 0] \times [-50, 0]$	2.5977E03	25.1283	2.5827E03	3.3281	0.5800
50	$[-50, 50] \times [-50, 50]$	1.0964E04	47.6350	1.0932E04	3.2813	0.2917
100	$[-50, 50] \times [-50, 50]$	1.2769E04	47.8286	1.2439E04	3.1094	2.6540
200	$[-300, 300] \times [-300, 300]$	4.6288E05	92.8977	4.5785E05	3.5625	1.0977
500	$[-300, 300] \times [-300, 300]$	5.2932E05	253.2720	5.2637E05	4.8594	0.5601
1000	$[-500, 700] \times [-300, 800]$	2.0622E06	547.3884	2.0488E06	5.4375	0.6533

5 Conclusions

This paper proposed an alternative methodology for a construction of the MVEE a set of points using BRKGA metaheuristic. The BRKGA was implemented in Python and had its performance compared to the CVX package implemented in MATLAB. The results for two-dimensional problems

Table 2: Generations of the BRKGA for 10 instances.

Points ($ \mathcal{C} $)	5	10	20	30	40	50	100	200	500	1000
Generations	5582	3525	6858	3881	3032	4574	2863	3115	6055	3886

showed that BRKGA generated solutions with a high level of accuracy and low computational cost. The authors intend to extend the methodology to n -dimensions in future works, as well as to evaluate the performance of other metaheuristics in solving the MVEE problem.

Acknowledgments

Thanks to São Paulo State University (UNESP) for the post-doctoral internship granting to the first author (Process 3134/2020); and to National Council for Scientific and Technological Development (CNPq) for the research productivity scholarship awarded to the last author (Process 315228/2020-2).

References

- [1] Alsabeh, R. A.- and Salhi, A. An Evolutionary Approach to Constructing the Minimum Volume Ellipsoid Containing a Set of Points and the Maximum Volume Ellipsoid Embedded in a Set of Points, *Journal of Physics: Conf. Series*, volume 1530, pages 1-15, 2020. DOI: [10.1088/1742-6596/1530/1/012087](https://doi.org/10.1088/1742-6596/1530/1/012087).
- [2] Gonçalves, J. F. and Resende, M. G. C. Biased random-key genetic algorithms for combinatorial optimization, *Journal of Heuristics*, volume 17, pages 487-525, 2011. DOI: [10.1007/s10732-010-9143-1](https://doi.org/10.1007/s10732-010-9143-1).
- [3] Grant, M. and Boyd, S. The CVX Users' Guide Release 2.2, *CVX Research*, 2020.
- [4] John, F. Extremum problems with inequalities as subsidiary condition, *Studies and Essays Presented to R. Courant on his 60th Birthday*, Wiley Interscience, New York, pages 187-204, 1948.
- [5] Python Software Foundation. The Python Tutorial. [Online]. Available: <https://docs.python.org/3/tutorial/index.html>, 2021.
- [6] Raposo, A. A. M., Rodrigues, A. B., and Da Silva, M. G. Alocação Ótima de Medidores para a Estimção de Estado Considerando a Reconfiguração da Rede de Distribuição, *Anais do XXII Congresso Brasileiro de Automática*, João Pessoa, pages 1-8, 2018.
- [7] Raposo, A. A. M., Rodrigues, A. B., and Da Silva, M. G. Robust meter placement for state estimation considering distribution network reconfiguration for annual energy loss reduction, *Electric Power Systems Research*, volume 182, pages 1-9, 2020. DOI: doi.org/10.1016/j.epsr.2020.106233.
- [8] Raposo, A. A. M. Reconfiguração e Alocação de Medidores em Redes de Distribuição. Tese de Doutorado, PPGEE/UFMA, 2020.
- [9] Sun, P. and Freund, R. M. Computation of Minimum-Volume Covering Ellipsoids, *Operations Research*, pages 690-706, 2004. DOI: [10.1287/opre.1040.0115](https://doi.org/10.1287/opre.1040.0115).