Scalar Optimization with Linear and Nonlinear Constraints Using Evolution Strategies

To Thanh Binh & Ulrich Korn

Institute of Automation, University of Magdeburg, Germany

Abstract. This paper introduces a new EVOlution Strategy for scalar optimization with LInear and NOnlinear Constraints (EVOSLINOC) which is robust to obtain a good approximation of a feasible global minimum. EVOSLINOC is based on the new concept of $\mathcal{C}$-, $\mathcal{F}$- and $\mathcal{N}$-fitness allowing systematically to handle constraints and (in)feasible individuals. In addition a number of ideas for using (in)feasible niche individuals which enable to explore new feasible areas and to make the population quickly to evolve towards a feasible global minimum is proposed. The performance of the EVOSLINOC can be successfully evaluated on many benchmark optimization problems [11, 10].

1 Introduction

1.1 Optimization problem

The general numerical scalar optimization problems with linear and nonlinear constraints can be formally stated as:

$$\min_{\mathbf{x}} \; f(\mathbf{x})$$

where $\mathbf{x} = (x_1, \ldots, x_n)^T \in \mathcal{F} \subseteq \mathcal{S}$.
The *search space* $\mathcal{S}$ is usually defined as a rectangle of the n-dimensional space $\mathcal{R}^n$ (domains of variables defined by their lower and upper bounds):

$$x_i^{(lower)} \leq x_i \leq x_i^{(upper)}, \forall i = \overline{1, n},$$

whereas the *feasible region* $\mathcal{F} \subseteq \mathcal{S}$ is defined by a set of m additional linear and/or nonlinear constraints ($m \geq 0$):

$$g_j(\mathbf{x}) \leq 0, \forall j = \overline{1, m}.$$

Constraints in terms of equations, e.g. $h(\mathbf{x}) = 0$, are not considered here because they can be replaced by a pair of inequalities:

$$\begin{aligned} -h(\mathbf{x}) - \epsilon &\leq 0, \\ h(\mathbf{x}) - \epsilon &\leq 0, \end{aligned}$$

where an additional parameter ϵ is used to define the precision of the system. Therefore they seem to be special cases of constraints in terms of inequalities.

1.2 Motivation

Most of the current research on applications of evolutionary algorithms (EAs) to solve this optimization problem is based on the concept of penalty functions, i. e, the constrained optimization is converted into the unconstrained optimization of an auxiliary scalar function which is created from the given objective function and constraints [5, 6, 9, 10]. The basic problem of all penalty approaches is *how to design an auxiliary scalar function*. Therefore, they can be very well suitable for some optimization problems but they are disappointing for other ones.

Based on eliminating infeasible individuals from the population the main idea of traditional evolution strategies (ESs) seems to be unsuitable for solving this problem. ESs cannot start until a feasible starting population is generated. The finding feasible individuals is itself a difficult problem especially in cases the ratio between the feasible and search region is too small.

To overcome these drawbacks a new ES for handling linear and nonlinear constraints should satisfy the following requirements:

- It is not necessary to create any auxiliary function and to give a feasible starting point before optimization.
- The handling of the objective function and constraints is performed separately.
- The generating infeasible individuals is allowed.

This paper is organized as follows. The next section introduces a new representation of individuals that is suitable for the given optimization task. The handling feasible and infeasible individuals is discussed in section 3. The first experimental results of the EVOSLINOC on some test cases are presented in section 5.

2 Representation of an individual

For handling linear and nonlinear constraints it is necessary to introduce a suitable representation of an individual. As in traditional ESs, each individual consists of a vector of objective variables $\mathbf{x} = (x_1, x_2, \cdots, x_n)^T$ (a point in the search space), a strategy parameter vector $\mathbf{s} = (s_1, s_2, \cdots, s_n)^T$ (a vector of standard deviations). To evaluate the fitness of an individual the two following measures have to be taken into account:

- an objective function value $f(\mathbf{x})$ (so-called $\mathcal{F}$-fitness in the objective function space),
- a degree of violation of constraints or degree of (in)feasibility (it is called $\mathcal{C}$-fitness in the constraint space).

A problem how to evaluate the $\mathcal{C}$-fitness of an individual will be discussed here.

Let $c_i(\mathbf{x}) = \max\{g_i(\mathbf{x}), 0\}$, $\forall i = \overline{1, m}$, the $\mathcal{C}$-fitness of an individual can be charaterized by a vector:

$$\mathcal{C}(\mathbf{x}) = (c_1(\mathbf{x}), c_2(\mathbf{x}), \cdots, c_m(\mathbf{x})), \tag{1}$$

(a point in the constraint space $\mathcal{R}_c^m$). Clearly, $\mathcal{C} \equiv \mathbf{0}$ for feasible individuals and $\mathcal{C} > \mathbf{0}$ for infeasible ones. In this sense the original point of the constraint space corresponds to the feasible region of the search space. An advantage to using this $\mathcal{C}$-fitness is that it enables precisely to represent an individual in the constraint space. However it has the following disadvantages:

- When the number of constraints is big, a large amount of memory is necessary to save $\mathcal{C}(\mathbf{x})$ for every individual.
- It takes much time to make a decision which infeasible individuals are better than other ones or should be selected for the next generation due to the fact that the comparision between infeasible individuals essential is the comparision between vectors $\mathcal{C}(\mathbf{x})$.

To avoid these problems the following measure of violation based on the *distance* between a point $(c_1(\mathbf{x}), c_2(\mathbf{x}), \cdots, c_m(\mathbf{x}))$ and the original point in the constraint space, is used [12]:

$$\mathcal{C}(\mathbf{x}) = \Big(\sum_{i=1}^{m} [c_i(\mathbf{x})]^p\Big)^{\frac{1}{p}}, \quad (p > 0). \tag{2}$$

The second measure of violation has experimentally shown to be as good as the first one and acceptable. Therefore it should be used to design the EVOSLINOC. By this way, the population can be divided into classes; individuals of the same class have the same value of $\mathcal{C}$ (i. e. the same distance to the feasible region) and they are said to be individuals of the $\mathcal{C}$-class. Using this concept, the 0-class includes all feasible individuals; individuals of the higher classes are *farther* from the feasible region than ones of the lower classes.

Thus, an individual can be represented as follows:

$$\mathcal{I}nd \stackrel{\text{def}}{=} (\mathbf{x}, \mathbf{s}, f(\mathbf{x}), \mathcal{C}(\mathbf{x})).$$

3 Handling feasible and infeasible individuals

Different to traditional ESs, both feasible and infeasible individuals can live in the population simultaneously. The EVOSLINOC allows mutation and reproduction operators to generate both feasible and infeasible offspring. Therefore it is necessary to check whether offspring are better than their parent (by mutation and reproduction) or to select better individuals for the next generation (by selection).

Up to now, there is no general method for comparision between infeasible and feasible individuals or between two infeasible individuals together [10]. For the EVOSLINOC the following selection criteria are recommended:

Criterion 1. An individual of the $\mathcal{C}_1$-class ($\mathcal{C}_1 > 0$) is said to be better than the other of the $\mathcal{C}_2$-class iff $\mathcal{C}_1 < \mathcal{C}_2$.

In other words, infeasible individuals of higher classes are worse than the other of lower classes. Using this criterion the new ES will drive the population towards the feasible region before trying to search for a feasible global minimum.

Criterion 2. Among individuals of the same class, better individuals have lower $\mathcal{F}$-fitness values.

During the search process for the feasible region the population at some stage of the evolution process may contain some feasible and infeasible individuals. If the criterion 1 is extended for use with 0-class (that means feasible individuals are always better than infeasible ones), the population would be soon feasible because all feasible individuals will be kept in the population unless a feasible offspring with a lower $\mathcal{F}$-fitness is generated. In many optimization problems the feasible region is non-convex or the ratio between the feasible and search region (denoted by ρ) is too small (for example, a feasible region defined by constraints in terms of functional equations) so that feasible offspring cannot be generated even from feasible individuals after many generations. In these cases the feasible population very slowly converges to a feasible global minimum. The following criteria help the evolution strategy avoid such situation and is based upon the intuitive conviction as below:

For many optimization problems, feasible individuals are *not always* better than all infeasible ones. For example, in *Fig. 1* the infeasible individual marked by 'b' lies nearer to the feasible global minimum than the feasible one marked by 'a'. For the same strategy parameter vector $\mathbf{s}$ it is hopeful that the infeasible individual 'b' generates feasible offspring which are better than offspring of the feasible individual 'a'. Here the infeasible individual 'b' is *not better* than other feasible individuals, but it is at least *acceptable* for the next generation. Such infeasible individuals should be kept in the population and can be included into a special class of infeasible individuals in a neigbourhood of the feasible region.

Criterion 3 (Extra class). Infeasible individuals up to the $\mathcal{C}_{extra}$-class (i. e. individuals of $\mathcal{C}$-classes so that $0 < \mathcal{C} \leq \mathcal{C}_{extra}$) are said to be in the same class (called the *extra class*).

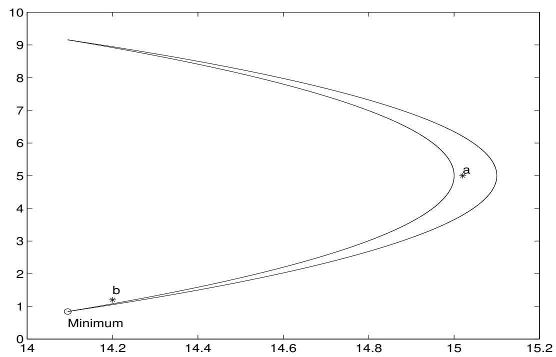

Fig. 1. An infeasible individual (marked 'b') of an extra class

The problem of how to choose the best value of $\mathcal{C}_{extra}$ is far from trivial. With a lower value of $\mathcal{C}_{extra}$ it is more difficult to locate offspring in the extra class. Otherwise the choosing a higher value of $\mathcal{C}_{extra}$ leads to generate more infeasible individuals (in higher classes). Our first experiments showed that the value $\mathcal{C}_{extra}$=0.1 was preferable.

The following criterion allows feasible individuals to generate infeasible offspring:

Criterion 4. An infeasible offspring of a feasible individual is said to be viable iff it belongs to the so-called extension $\mathcal{C}_{extension}$-class defined by:

$$\mathcal{C}_{extension} = \max\{\mathcal{C}_{extra}, \mathcal{C}_{pop}\},$$

where $\mathcal{C}_{pop}$ is the highest class in the population (corresponding to the infeasible individual with the biggest distance to the original point of the constraint space).

4 Niche individuals

Traditional ESs [2, 7] have shown that so-called *niche* (feasible) individuals play an important role to overcome local minima by multi modal optimization problems. Therefore it is also significant to use niche (in)feasible individuals for the new evolution strategy. A niche feasible individual can be characterized by the two following properties [2, 7]:

- Their $\mathcal{F}$-fitness is as small as possible.
- They are as far as possible from the current best feasible individual.

Mathematically, a niche feasible individual should have a small value of the so-called niche fitness ($\mathcal{N}_f$-fitness) [2, 7]:

$$\mathcal{N}_f(\mathbf{x}) = \frac{f(\mathbf{x}) - f(\mathbf{x}_{fbest})}{||\mathbf{x} - \mathbf{x}_{fbest}||^\beta}, \quad (\mathbf{x} \neq \mathbf{x}_{fbest}) \tag{3}$$

where:

- $\mathbf{x}_{fbest}$ and $f(\mathbf{x}_{fbest})$ are an objective variable vector and the $\mathcal{F}$-fitness of the best feasible individual, respectively.
- $||.||$ denotes a norm in the n-dimensional parameter space.
- β is a scalar value ($\beta = 1, 2, \cdots$).

The left problem is to choose niche infeasible individuals from the current population. For optimization problems for which feasible individuals are hard to find, the population of some first generations contains infeasible individuals only. In this situation the $\mathcal{C}$-fitness is used to help the evolution strategy in shifting the population towards feasibility. To avoid problems for which the infeasible population converges to (possible) local minima of the $\mathcal{C}$-fitness, the following criterion is recommended to choose niche infeasible individuals:

Criterion 5. If there is no feasible individual in a population, niche infeasible individuals should have an as small value of $\mathcal{C}$-fitness as possible and be as far as possible from the best current infeasible individual.

Similar to *Eq. 3* the niche fitness of infeasible individuals ($\mathcal{N}_i$-fitness) can be evaluated by:

$$\mathcal{N}_i(\mathbf{x}) = \frac{\mathcal{C}(\mathbf{x}) - \mathcal{C}(\mathbf{x}_{ibest})}{||\mathbf{x} - \mathbf{x}_{ibest}||^\beta}, \quad (\mathbf{x} \neq \mathbf{x}_{ibest}) \tag{4}$$

where $\mathbf{x}_{ibest}$ and $\mathcal{C}(\mathbf{x}_{ibest})$ are an objective variable vector and the $\mathcal{C}$-fitness of the best infeasible individual, respectively. Notice that traditional ESs [2] have also shown that the use of niche individuals may help the ES in finding many global minima with the same objective function value located at different points of the parameter space. For constrained optimization problems for which the feasible region is disconnected the $\mathcal{C}$-fitness (see *Eq. 2*) has the (global) smallest value 0 at different subsets of the feasible region. Therefore it is meaningful to use niche infeasible individuals to explore other subsets of the feasible region.

When there exist some feasible individuals in a population, the ES tries to generate infeasible individuals in a neighbourhood of the feasible region (in a $\mathcal{C}_{extension}$-class). The best infeasible individuals are not only infeasible ones with lower $\mathcal{C}$-fitness, but they are other ones lying nearer to global feasible minima (niche infeasible individuals). Because global feasible minima are unknown, it can be expected that they are still far from the best current feasible individual. Therefore:

Criterion 6. If at least one feasible individual exists in the population, niche infeasible individuals should have an as small value of $\mathcal{C}$-fitness as possible and be as far as possible from the best current feasible individual.

The niche fitness of infeasible individuals ($\mathcal{N}_{if}$-fitness) can be evaluated by:

$$\mathcal{N}_{if}(\mathbf{x}) = \frac{\mathcal{C}(\mathbf{x})}{||\mathbf{x} - \mathbf{x}_{fbest}||^{\beta}}. \tag{5}$$

Our initial experiments showed that this criterion was useful for many optimization problems.

5 Some test cases

In this section, some test cases have been carefully selected to illustrate the efficiency of the EVOSLINOC. Most of them are described as benchmark optimization problems and used to compare or demonstrate the current research methods for the optimization with linear and nonlinear constraints [8, 10].
For all test cases the following important parameters of the ES were used:

- Population size = 100
- Number of niche feasible individuals = 5
- Number of niche infeasible individuals = 5
- Number of parents for mutation and reproduction = 10
- Number of offspring per mutation = 5.

5.1 Test Case 1

This optimization problem (taken from [6], Test Case 6 of [10]) is

$$minimize\ \ f(\mathbf{x}) = (x_1 - 10)^3 + (x_2 - 20)^3$$

subject to nonlinear constraints:

$$(x_1 - 5)^2 + (x_2 - 5)^2 \geq 100$$
$$(x_1 - 6)^2 + (x_2 - 5)^2 \leq 82.81$$

and bounds: $13 \leq x_1 \leq 100$ and $0 \leq x_2 \leq 100$. In this case, the feasible region is nonconvex and $\rho = 0.0066$[1]. The known global solution is $\mathbf{x}^* = (14.095, 0.84296)^T$ and $f(\mathbf{x}^*) = -6961.81381$. From an infeasible starting point $\mathbf{x}_0 = (20.1, 5.84)^T$ lying far from the feasible region the evolution of the population is shown in *Fig. 2*. For all runs the search for the feasible region ends

[1] From now, ρ was determined experimentally by generating one milion random points in the search space $\mathcal{S}$ and checking if they belong to $\mathcal{F}$

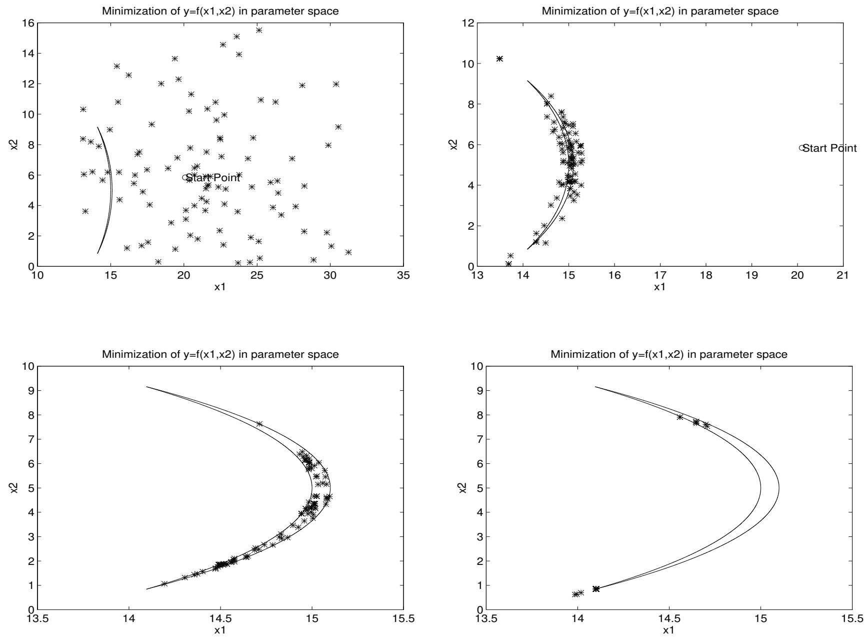

Fig. 2. Evolution of the population for Test Case 1

after less than 10 generations; the population moves then into the feasible global minimum and concentrates on it in less than 400 generations. It is interesting to note that the ES is more robust and stable by finding the global feasible minimum than other tools [11, 2, 7].

5.2 Test Case 2

The optimization problem (taken from [6, 11]) is

$$minimize \;\; f(\mathbf{x}) = -x_1 - x_2,$$

subject to nonlinear constraints:

$$\begin{aligned} 2x_1^4 - 8x_1^3 + 8x_1^2 + 2 - x_2 &\geq 0 \\ 4x_1^4 - 32x_1^3 + 88x_1^2 - 96x_1 + 36 - x_2 &\geq 0 \end{aligned}$$

and bounds: $0 \leq x_1 \leq 3$ and $0 \leq x_2 \leq 4$. Different to the test case 1, the feasible region is almost disconnected. The EVOSLINOC enables to explore a new subset of the feasible region and to bring the population towards the global feasible solution at $\mathbf{x}^* = (2.3295, 3.1783)^T$ with $f(\mathbf{x}^*) = -5.5079$. The global solution can be found after 10 generations (see *Fig. 3*).

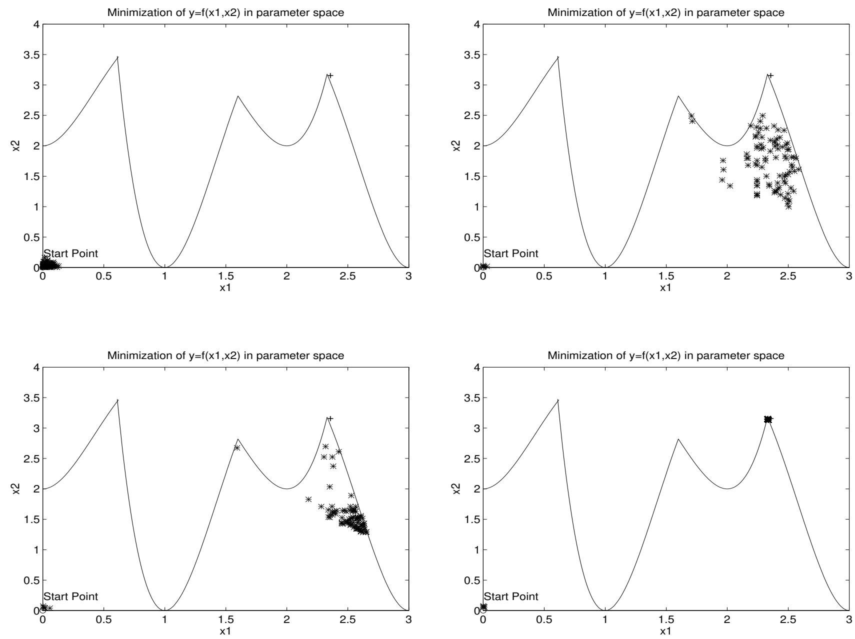

Fig. 3. Evolution of the population for Test Case 2

5.3 Test Case 3

This optimization problem has been used to illustrate *the death penalty method* for an evolution strategy [1] and to compare some evolutionary algorithms ([11], Test Case 7 of [10]). Minimize a function:

$$f(\mathbf{x}) = x_1^2 + x_2^2 + x_1 x_2 - 14x_1 - 16x_2 + (x_3 - 10)^2 + 4(x_4 - 5)^2 + (x_5 - 3)^2 + \\ 2(x_6 - 1)^2 + 5x_7^2 + 7(x_8 - 11)^2 + 2(x_9 - 10)^2 + (x_{10} - 7)^2 + 45;$$

subject to the constraints:

$$\begin{aligned}
105 - 4x_1 - 5x_2 + 3x_7 - 9x_8 &\geq 0 \\
-10x_1 + 8x_2 + 17x_7 - 2x_8 &\geq 0 \\
8x_1 - 2x_2 - 5x_9 + 2x_{10} + 12 &\geq 0 \\
3x_1 - 6x_2 - 12(x_9 - 8)^2 + 7x_{10} &\geq 0 \\
-3(x_1 - 2)^2 - 4(x_2 - 3)^2 - 2x_3^2 + 7x_4 + 120 &\geq 0 \\
-x_1^2 - 2(x_2 - 2)^2 + 2x_1x_2 - 14x_5 + 6x_6 &\geq 0 \\
-5x_1^2 - 8x_2 - (x_3 - 6)^2 + 2x_4 + 40 &\geq 0 \\
-.5(x_1 - 8)^2 - 2(x_2 - 4)^2 - 3x_5^2 + x_6 + 30 &\geq 0
\end{aligned}$$

and bounds: $-10 \leq \mathbf{x} \leq 10$. This function is quadratic and has the global minimum at:

$$\mathbf{x}^* = (2.172, 2.364, 8.774, 5.096, .9907, 1.4306, 1.322, 9.829, 8.280, 8.376)^T,$$

where $f(\mathbf{x}^*) = 24.3062$.

Up to now, the best solution which had the value of 25.653 was found in [1], but it is necessary to initialize a population by feasible individuals [11]. The EVOSLINOC reached the value less than 25.00 for all test runs and the best solution among them had the value of 24.61.

5.4 Test Case 4

The Test Case is selected to compare the result by using the EVOSLINOC with results of other methods ([5], [4], Test Case 4 of [10]). Minimize a function:

$$f(\mathbf{x}) = 5.3578547x_3^2 + 0.8356891x_1x_5 + 37.293239x_1 - 40792.141;$$

subject to three double nonlinear inequalities:

$$\begin{aligned}
0 \leq\ & 85.334407 + 0.0056858x_2x_5 + 0.00026x_1x_4 - 0.0022053x_3x_5 \leq 92 \\
90 \leq\ & 80.51249 + 0.0071317x_2x_5 + 0.0029955x_1x_2 + 0.0021813x_3^2 \leq 110 \\
20 \leq\ & 9.300961 + 0.0047026x_3x_5 + 0.0012547x_1x_3 + 0.0019085x_3x_4 \leq 25
\end{aligned}$$

and bounds: $78 \leq x_1 \leq 102$, $33 \leq x_2 \leq 45$, $27 \leq x_i \leq 45$, $\forall i = 3, 4, 5$. EVOSLINOC provided the best value: $f_{min} = -31025$ at

$$\mathbf{x}^* = (78.0001, 33.0072, 27.0760, 44.9872, 44.9591)^T$$

that is better than in [5] ($f_{min} = -30005.7$) and in [4] ($f_{min} = -30665.5$).

5.5 Test Case 5

It is the Test Case 9 in [10] to illustrate *the method of superiority of feasible points* and can be described as below:

$$\min_{\mathbf{x}} f(\mathbf{x}) = (x_1-10)^2+5(x_2-12)^2+x_3^4+3(x_4-11)^2+ \\ 10x_5^6+7x_6^2+x_7^4-4x_6x_7-10x_6-8x_7;$$

subject to the constraints:

$$\begin{aligned} 127-2x_1^2-3x_2^4-x_3-4x_4^2-5x_5 &\geq 0 \\ 282-7x_1-3x_2-10x_3^2-x_4+x_5 &\geq 0 \\ 196-23x_1-x_2^2-6x_6^2+8x_7 &\geq 0 \\ -4x_1^2-x_2^2+3x_1x_2-2x_3^2-5x_6+11x_7 &\geq 0 \end{aligned}$$

and bounds: $-10 \leq x_i \leq 10,\ \forall i=\overline{1,7}$. The function has its global minimum at

$$\mathbf{x}^* = (2.330499, 1.951372 - 0.47754, 4.3657, -0.62448, 1.0381, 1.5942)^T,$$

where $f(\mathbf{x}^*)$=680.6300573. The EVOSLINOC approached very closely the global optimum after less than 400 generations for all runs.

6 Conclusion

In this paper the EVOSLINOC allowing to handle linear and nonlinear constraints was proposed. Experiments on many benchmark optimization problems indicated that the EVOSLINOC was robust and gave good performances. Interesting is that the paper provides some new ideas for effectively handling (in)feasible individuals: niche infeasible individuals can be *better* than some feasible ones. The use of niche infeasible individuals enables the ES to avoid an unnecessary concentration of the population at local minima and quickly to shift the population towards the feasible global minima. The EVOSLINOC was implemented in a MATLAB–based[2] environment [3]. It is still under development, and there are some possible modification and extensions in the future version.

References

1. T. Bäck, F. Hoffmeister, and H.-P. Schwefel. A survey of evolution strategies. In R.K. Belew and L.B. Booker (Eds.), Proceedings of the 4^{th} International Conference on Genetic Algorithms, Morgan Kaufmann, pages 2–9, 1991.

[2] MATLAB is the Trademark of the MathWorks, Inc.

2. T. Binh. Eine Entwurfsstrategie für Mehrgrößensysteme zur Polgebietsvorgabe. PhD thesis, Institute of Automation, University of Magdeburg, Germany, 1994.
3. T. Binh, U. Korn, and J. Kliche. *Evolution Strategy Toolbox for use with MATLAB*. Technical report, Institute of Automation, University of Magdeburg, Germany, Mar. 1996.
4. D. Himmelblau. Applied Nonlinear Programming. McGraw-Hill, 1992.
5. A. Homaifar, S. Lai, and X. Qi. Constrained optimization via genetic algorithms. Simulation 62, 4:242–254, 1994.
6. J. Joines and C. Houck. On the use of non-stationary penalty functions to solve nonlinear constrained optimization problems with gas. In Z. Michalewicz, J. D. Schaffer, H.-P. Schwefel, D. B. Fogel, and H. Kitano (Eds.), Proceedings of the First IEEE International Conference on Evolutionary Computation, pages 579–584, 1994.
7. J. Kahlert. Vektorielle Optimierung mit Evolutionsstrategien und Anwendung in der Regelungstechnik. Forschungsbericht VDI Reihe 8 Nr.234, 1991.
8. Z. Michalewicz. Genetic algorithms, numerical optimization and constraints. In L. J. Eshelman (Ed.), Proceedings of the 6^{th} International Conference on Genetic Algorithms,Morgan Kaufmann, pages 151–158, 1995.
9. Z. Michalewicz. Heuristic methods for evolutionary computation techniques. Journal of Heuristics 1(2), pages 177–206, 1995.
10. Z. Michalewicz, D. Dasgupta, R. Riche, and M. Schoenauer. Evolutionary algorithms for constrained engineering problems. In Y. Davidor, H.-P. Schwefel, and R. Manner (Eds.), Proceedings of the 6^{rd} Conference on Parallel Problems Solving from Nature, Springer Verlag, 1996.
11. Z. Michalewicz and C. Nazhiyath. Genocop III: A co–evolutionary algorithm for numerical optimization problems with nonlinear constraints. In D. B. Fogel (Ed.), Proceedings of the 2^{th} IEEE International Conference on Evolutionary Computation, pages 647–651, 1995.
12. J. Richardson, M. Palmer, G. Liepins, and M. Hilliard. Some guidelines for genetic algorithms with penalty functions. Proceedings of the 3^{rd} International Conference on Genetic Algorithms, Los Altos, CA, Morgan Kaufmann Publishers, pages 191–197, 1989.

This article was processed using the LaTeX macro package with LLNCS style