An evolutionary annealing-simplex algorithm for global optimisation of water resource systems

Andreas Efstratiadis and Demetris Koutsoyiannis

Department of Water Resources, Faculty of Civil Engineering, National Technical University of Athens, Heroon Polytechniou 5, 15780 Zographou, Greece

Abstract The evolutionary annealing-simplex algorithm is a probabilistic heuristic global optimisation technique that joins ideas from different methodological approaches, enhancing them with some original elements. The main concept is based on a controlled random search scheme, where a generalised downhill simplex methodology is coupled with a simulated annealing procedure. The algorithm combines the robustness of simulated annealing in rugged problems, with the efficiency of hill-climbing methods in simple search spaces. The following-up procedure is based on a simplex-searching scheme. The simplex is reformulated at each generation going either downhill or uphill, according to a probabilistic criterion. In the first case, it moves towards the direction of a candidate local minimum via a generalised Nelder-Mead strategy. In the second case, it expands itself along the uphill direction, in order to escape from the current local minimum. In all possible movements, a combination of deterministic as well as stochastic transition rules is applied. The evolutionary annealing-simplex algorithm was first examined in a variety of typical benchmark functions and then it was applied in two global optimisation problems taken from water resources engineering, the calibration of a hydrological model and the optimisation of a multiple reservoir systems' operation. The algorithm has been proved very reliable in locating the global optimum, requiring reasonable computational effort.

Keywords global optimisation, evolutionary algorithms, simulated annealing, downhill simplex, controlled random search, effectiveness, efficiency.

Introduction

A wide number of hydroinformatics applications are formulated as nonlinear unconstrained programming problems, commonly referred as global optimisation problems. Given that most of such problems are intrinsically nonconvex and multimodal, they do not exist deterministic optimisation methods that can guarantee the determination of the globally optimum solution. During the last decades, a variety of probabilistic schemes have been developed for solving such problems. These methods involve the evaluation of the objective function at a usually random sample of points, followed by subsequent evolutions of the sample using a combination of random and deterministic rules.

In many of the applications that are met in practice, a highly accurate solution is neither possible nor feasible. Particularly, it may be impossible because of uncertainties and inaccuracies in the underlying model or data, or it may be infeasible due to the unacceptable high computational effort required to attain it, especially when the evaluation of the objective function is time consuming. Hence, the performance of a global optimisation solver depends upon two characteristics, the effectiveness and the efficiency (*Duan et al.*, 1992). The first one indicates the capability of locating the global optimum starting from any initial point (or population of points), whereas the second one indicates the convergence speed.

Hydroinformatics 2002: Proceedings of the Fifth International Conference on Hydroinformatics, Cardiff, UK

© IWA Publishing and the authors. ISBN 1 84339 021 3 (set)

The scope of this paper is to review the most representative global optimisation techniques (section 2) and to propose a heuristic global optimisation scheme, the evolutionary annealing-simplex algorithm, which ensures both effectiveness and efficiency by combining mechanisms from different methodologies in an original manner (section 3). The algorithm is tested in some theoretical and real-world problems, by comparing its performance against that of a recent but already widespread method, the shuffled complex evolution algorithm (section 4). Last section summarises the main issues and presents the final conclusions.

Overview of nonlinear optimisation algorithms

Deterministic methods

We assume a nonlinear unconstrained minimisation problem of n control variables. Historically, the first attempt to solve such problems was via deterministic descend methods (also called local optimisation methods), gradient-based as well as direct. Gradient methods require the partial derivatives of the function to be known, whereas direct search algorithms are derivative-free. Their concept is to employ subsequent line minimisations, starting from an initial feasible point. Although local search methods are very effective for simple, unimodal spaces, in real-world problems they easily get trapped in non-optimal regions due to the existence of many local optima, flat surfaces or ridges.

Among a variety of local search algorithms, one should distinguish the downhill simplex method of *Nelder and Mead* (1965). The core of the method is an evolving pattern of $n + 1$ points (the vertices of a simplex) that span the n-dimensional space. The simplex explores the feasible space either by reflecting, contracting or expanding away from the actually worst vertex, or by shrinking toward the best vertex. An appropriate sequence of such movements converges to the nearest local minimum.

The shortcoming of deterministic local optimisation algorithms to cope with nonconvex, multimodal and even ill-posed functions leaded researchers to develop probabilistic schemes that use stochastic transition rules and do not impose any special requirement about the nature of the objective function. These methods are also called global, given that they guarantee asymptotic convergence to the global optimum of the function.

Monte Carlo methods

A primitive global optimisation method is the pure random sampling. In order to implement the generation of evaluation points in a more systematic way, adaptive random sampling (ARS) techniques have been developed. The general strategy of ARS consists of generating the next point as a perturbation around the actual one and accept it only if it improves the objective function (*Rubinstein*, 1986). An advanced ARS strategy is the controlled random search technique, proposed by *Price* (1965) in several versions. Price introduced the concept of an evolving population of feasible points that became the basis of most modern global optimisation methods. At each step, a simplex is formed from a sample of the population, which is reformed by reflecting one of its vertices through its centroid.

A specific category of Monte Carlo methods is the multistart strategy, which consists of running several independent trials of a local search algorithm. In an ideal case, these methods aim at starting the local search once in every region of attraction of local optima that may be identified via clustering analysis (*Solomatine*, 1999).

Evolutionary algorithms

The family of evolutionary algorithms (EAs), inspired from the mechanism of natural evolution, introduced some important modifications to random search. In EAs, the searching procedure is implemented at stages called generations. At each one, a population of randomly generated points evolves by applying the selection, recombination and mutation operators. The most popular interpretation of EAs is genetic algorithms (*Holland*, 1975; *Goldberg*, 1989; *Michalewich*, 1992), where variables are represented on a chromosome-like (usually binary string) structure. This specific structure enables the documentation of the global convergence property of genetic algorithms via the schema theory (*Goldberg*, 1989).

Simulated annealing

Simulated annealing (SA) is based on an analogy with the homonymous thermodynamical process. For slowly cooled thermodynamical systems (e.g., metals), nature is able to find the minimum state of energy, while the system may end in an amorphous state of higher energy if it is cooled quickly. This principle is expressed by the Boltzmann probability distribution:

$$p(E) \sim \exp(-E / k\,T) \tag{1}$$

The energy of a system in thermal equilibrium at a given temperature T is probabilistically distributed among all different states E. The system may switch to a new energy state with probability p, irrespective of whether it is higher or lower. Therefore, nature's minimisation strategy is to allow the system sometimes to go uphill as well as downhill, so that it has a chance to escape from a local energy minimum in favour of finding a better, more global minimum. However, the lower the temperature, the less likely is a significant uphill step.

Kirkpatrick et al. (1983) transferred the principles of annealing in optimisation, by introducing a control parameter, analogue of temperature, and an annealing cooling schedule that describes its gradual reduction. Assuming a large enough initial temperature and a proper schedule, SA slowly converges to the globally optimal solution.

The shuffled complex evolution method

The shuffled complex evolution (SCE) method (*Duan et al.*, 1992) is a heuristic global optimisation scheme that became quickly one of the most popular among water resources engineers. According to the algorithm, a random set of points is sampled and partitioned into complexes. Each of them is allowed to evolve in the direction of global improvement, using competitive evolution techniques based on the downhill simplex method. At periodic stages, the entire set of points is shuffled and reassigned to new complexes, to enable information sharing. The combination of competitive evolution and shuffling ensures that the information gained by each of the individual complexes is shared through the entire population. This results in a robust searching scheme that conducts an efficient search on the feasible space.

The evolutionary annealing-simplex algorithm

Coupling simulated annealing with the simplex method

An interesting category of global optimisation techniques aims to combine the robustness of SA in rugged problems, with the efficiency of direct search methods in simple search spaces. The strengths and weaknesses of the two approaches are complementary. SA avoids local optima by jumping away from them, but it sacrifices efficiency (i.e., running time) by doing

so; on the other hand, direct search methods converge quickly to the nearest local optimum, but they have no way of getting out of it.

A well-known local optimiser is the downhill simplex algorithm. However, there are only few references in literature on how to incorporate the SA strategy to this method. For example, *Press et al.* (1992) introduced into the objective function a logarithmically distributed random variable associated with every vertex, proportional to the temperature. The simplex behaves between a random walk and an ordered downhill motion, depending on the ratio $\Delta f / T$, where Δf is the difference in values of the function at two vertices and T is the temperature. *Pan and Wu* (1998) improved this method by incorporating some follow-up strategies to escape from local optima. Finally, *Kvaniscka and Pospichal* (1997) proposed a simplex-based algorithm where the generation of reflection points is randomised and their returning to the population is implemented according to a probabilistic criterion.

Description of the evolutionary annealing-simplex algorithm

The evolutionary annealing-simplex (EAS) algorithm presented herein is a heuristic method, developed on the basis of the above schemes. Its concept is based on a control random search technique, where a generalised Nelder-Mead method is coupled with an annealing strategy. The core of the evolution is a simplex implementing the typical Nelder-Mead movements as well as some original ones, according to a combination of deterministic and stochastic rules. The main characteristic is that none of the movements is totally deterministic, enabling thus the exploration of rough search spaces. A typical iteration cycle consists of the following:

(1) A simplex $S = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_{n+1}\}$ is formulated by randomly selecting its vertices from the current population P, consisting of $m \geq n + 1$ points, where $\mathbf{x}_1$ corresponds to the best (lowest) and $\mathbf{x}_{n+1}$ to the worst (highest) value of the objective function, f.

(2) From the subset $\{\mathbf{x}_2, \ldots, \mathbf{x}_{n+1}\}$ a vertex $\mathbf{w}$ is selected that is candidate to be replaced. The candidate point is the vertex that maximises the modified function:

$$g(\mathbf{x}_i) = f(\mathbf{x}_i) + u\,T \tag{2}$$

where u denotes a uniform random number from the interval [0, 1] and T is the actual temperature of the system.

(3) A new point $\mathbf{r}$ is generated by reflecting the simplex away from $\mathbf{w}$ according to:

$$\mathbf{r} = \mathbf{g} + (0.5 + u)\,(\mathbf{g} - \mathbf{w}) \tag{3}$$

where $\mathbf{g}$ is the centroid of the subset $S - \{\mathbf{w}\}$ and u is a uniform random number.

(4) If $f(\mathbf{r}) < f(\mathbf{w})$, the new point $\mathbf{r}$ replaces the vertex $\mathbf{w}$. Moreover, if $f(\mathbf{r}) < f(\mathbf{x}_1)$, i.e. the new point is better than the current best vertex, a sequence of expansion (line minimisation) steps are implemented according to the equation:

$$\mathbf{x}_{\text{new}} = \mathbf{g} + \varphi^{[s]}\,(\mathbf{r} - \mathbf{g}) \tag{4}$$

where $\varphi^{[s]} = \varphi^{[s-1]} + u$, with $\varphi^{[0]} = 1$. The expansion continues as long as the function value improves, accelerating thus the local searching procedure. On the other hand, if $f(\mathbf{r}) > f(\mathbf{x}_1)$, the simplex is outside contracted as follows:

$$\mathbf{x}_{\text{new}} = \mathbf{g} + (0.25 + 0.5u)\,(\mathbf{r} - \mathbf{g}) \tag{5}$$

If either the expansion or the outside contraction successes, $\mathbf{x}_{\text{new}}$ replaces $\mathbf{r}$.

(5) If $g(\mathbf{r}) > g(\mathbf{w})$, the reflection point $\mathbf{r}$ is not accepted, the actual temperature is reduced by a factor λ and the simplex is inside contracted according to the equation:

$$\mathbf{x}_{\text{new}} = \mathbf{g} - (0.25 + 0.5u)\,(\mathbf{g} - \mathbf{w}) \tag{6}$$

If $f(\mathbf{x}_{\text{new}}) > f(\mathbf{x}_{n+1})$, i.e. the new point is worse than the current worst vertex, the simplex shrinks toward the best vertex $\mathbf{x}_1$, such as $\mathbf{x}'_i = 0.5(\mathbf{x}_1 + \mathbf{x}_i)$.

(6) If $g(\mathbf{r}) < g(\mathbf{w})$, the reflection point $\mathbf{r}$ is accepted even if it deteriorates the function value. Next, a given number of uphill movements are implemented according to (4). The difference of subsequent function values is an approximation of the gradient. Whenever the gradient becomes negative, the simplex escapes from the region of attraction of the current local minimum and the new point replaces $\mathbf{r}$. Otherwise, if any trial uphill movement success, a random point is generated on the boundaries of the population P, and replaces $\mathbf{r}$ according to a mutation probability p_m. The new point is generated as follows:

$$\mathbf{x}_{new} = \mathbf{c} + d\,\mathbf{y} / ||\mathbf{y}|| \quad (7)$$

where $\mathbf{c}$ is the centroid of P, d is the maximum Euclidean distance of the members of P from the centroid and $\mathbf{y}$ is a random direction in the n-dimensional space.

The algorithm stops if the relative distance between the current best f_{min}, and worst, f_{max}, function value in P becomes smaller than a given tolerance, ε. The initial temperature is set equal to $f_{max} - f_{min}$, while at the beginning of each cycle it is re-evaluated so that it never exceeds $\xi\,(f_{max} - f_{min})$, where $\xi \geq 1$ is a control parameter of the annealing schedule.

Experimental results

The EAS algorithm was first tested in 8 typical benchmark functions (Figure 1), most of them adapted from *De Jong* (1975). For each function, 100 independent runs were made. The success percentage was used as indicator of the algorithm's effectiveness, whereas the average number of function evaluations was used as indicator of its efficiency. Input arguments were $m = 4n + 1$, $\lambda = 0.95$, $\xi = 5$, $p_m = 10\%$ and $\varepsilon = 1\%$. For the sake of comparison, the results from the SCE method, having the same population length, are also included. The analysis showed that both methods faced with success almost all benchmark problems. The EAS algorithm was proved slightly more effective, while the SCE method was proved slightly more efficient.

Table 1: Effectiveness and efficiency (in parenthesis) indices for the EAS and SCE methods.

Test function	n	Number of optima	EAS	SCE-UA
Sphere	2	1	100 (4128)	100 (5159)
Hozaki	2	2	100 (324)	100 (296)
Goldestein-Price	2	4	100 (552)	99 (449)
Rozenbrock	2	1	100 (619)	100 (1191)
Rozenbrock	10	1	26 (10847)	99 (11105)
Griewank	10	> 1000	91 (2768)	100 (5574)
Michalewicz	2	> 100	51 (1409)	44 (438)
Integer step	10	1	100 (3324)	1 (2350)
Mean effectiveness			83.5	80.0

Next, two real-world applications were examined, taken from the domain of water resources engineering. The first was the calibration of a parsimonious hydrological model of 4 parameters, specifically the imperviousness of the basin, the storage capacity of the soil moisture reservoir and the recession coefficient of the soil moisture and groundwater. Input data was the precipitation and the potential evapotranspiration for a simulation period of 93 months. Two case studies were examined; the first one was based on real output data (i.e., historical runoff series) whereas the second one was based on synthetic output data, where runoff was calculated assuming arbitrary parameter values. The reason of using synthetic

data was the elimination of errors due to the data as well as the model structure. For both cases, the EAS algorithm managed to achieve the same effectiveness rate (84%), which was better than the rate achieved by the SCE method (80% and 72%, respectively).

The second application was the maximisation of the mean annual energy profit of a hypothetical hydrosystem, comprising of two hydroelectric reservoirs in parallel. A high-dimensional methodology was implemented, assuming as control variables the step-by-step reservoir target releases. Hence, for a simulation period of 16 years, the total number of control variables was 384. The results were 44.7 monetary unions for the SCE method and 46.5 unions for the EAS algorithm. The characteristic of this problem was the flat-type response surface of the objective function, due to the use of desirable and not real magnitudes as control variables, which makes extremely difficult the location of the gradient of the function, particularly when the number of parameters is large.

Summary and conclusions

The current trend in optimisation research is the combination of techniques obtained from diverse methodological approaches, in order to develop more robust search schemes. The EAS algorithm follows the above concept, by trying to couple the robustness of SA in rough problems, with the efficiency of the downhill simplex method in simple search spaces. By enhancing the typical Nelder-Mead procedure with new movements such as climbing and mutation, and by introducing to the original movements a stochastic component, it not only makes possible to easily escape from local optima but also to accelerate the searching procedure, especially in high-dimensional applications. After extended analysis, the algorithm was proved at least as effective and efficient as the SCE method, which is now widely used in the region of water resources systems optimisation.

References

De Jong, K. (1975). *An analysis of the behaviour of a class of genetic adaptive systems*. PhD Thesis, University of Michigan. Ann Arbor.

Duan, Q., S. Sorooshian, and V. Gupta, (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. *Water Resources Research*, 28(4), 1015-1031.

Goldberg, D. E., (1989). *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley Publishing Company.

Holland, J. H., (1975). *Adaptation in Natural and Artificial Systems*. University of Michigan Press.

Kirkpatrick, S., C. D. Gelatt, and M. P. Vecchi, (1983). Optimization by simulated annealing. *Science*, 220, 671-680.

Kvasnicka, V., and J. Pospichal, (1997). A hybrid of simplex method and simulated annealing. *Chemometrics and Intelligent Laboratory Systems*, 39, 161-173.

Michalewicz, Z., (1996). *Genetic Algorithms + Data Structures = Evolution Programs*. Springer-Verlag, New York.

Nelder, J. A., and R. Mead, (1965). A simplex method for function minimization. *Computer Journal*, 7(4), 308-313.

Pan, L., and L. Wu, (1998). A hybrid global optimization method for inverse estimation of hydraulic parameters: annealing-simplex method. *Water Resources Research*, 34(9), 2261-2269.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, (1992). *Numerical Recipes in C*. 2nd edition, Cambridge University Press, Cambridge, U. K.

Price, W. L., (1965). A controlled random search procedure for global optimization. *Computer Journal*, 7, 347-370.

Rubinstein, R. Y., (1986). *Month Carlo Optimization, Simulation and Sensitivity of Queuing Networks*. John Wiley.

Solomatine, D. P., (1999). Two strategies of adaptive cluster covering with descent and their comparison to other algorithms. *Journal of Global Optimization*, 14(1), 55-78.