A Class of Test Functions for Global Optimization

RUDOLF MATHAR
Institute of Statistics, Aachen University of Technology, 52056 Aachen, Germany

and

ANTANAS ŽILINSKAS*
Institute of Mathematics and Informatics & VPU, Studentu 39, Vilnius, Lithuania

(Received: 16 March 1993; accepted: 22 September 1993)

Abstract. We suggest weighted least squares scaling, a basic method in multidimensional scaling, as a class of test functions for global optimization. The functions are easy to code, cheap to calculate, and have important applications in data analysis. For certain data these functions have many local minima. Some characteristic features of the test functions are investigated.

Key words: Global optimization, test functions, multidimensional scaling

1. Introduction

Experimentally testing global optimization algorithms is an important counterpart to the theoretical research in the field. However, at present there does neither exist a generally accepted methodology of testing nor a collection of test functions. In contrary, for local optimization there are well grounded recommendations for testing as well as widely used collections of test functions (cf.[1],[5],[6]). Some of these recommendations are useful also for global optimization, but the set of test functions for global optimization should clearly be different. A collection of test functions with special structure, e.g. indefinite quadratic problems, may be found in [2]. Some recently published general test functions are collected and a discussion of testing is presented in [8], but these functions are rather artificial, and often too simple in certain aspects.

The test functions should possess at least the following properties:

- They should be cheap to evaluate and simple to code.
- Certain features of the functions should be easy to control, in order to obtain simple and difficult examples from the same setup.
- Test functions should be relevant for practical problems.

*This paper was written while the second author was a visiting Professor at Aachen University of Technology, funded by the Deutsche Forschungsgemeinschaft.

2. Multidimensional Scaling

All these requirements are valid for the following functions

$$f(x) = \sum_{i<j}^{n} w_{ij}(\delta_{ij} - d_{ij}(X))^2 \tag{1}$$

where $X = (x_{11}, \ldots, x_{n1}, x_{12}, \ldots, x_{n2}) \in \mathbb{R}^{2n}$, and $d_{ij}(X)$ denotes the Euclidian distance between the points (x_{i1}, x_{i2}) and (x_{j1}, x_{j2}). The matrix (δ_{ij}) represents given pairwise dissimilarities between n objects, and the matrix (w_{ij}) contains nonnegative weights. We call the corresponding matrix $\mathbf{X} = (x_{ij})_{1\leq i\leq n,\, 1\leq j\leq 2} \in \mathbb{R}^{n\times 2}$ a *configuration* of n points in $\mathbb{R}^2$. In multidimensional scaling the objective function (1) is called STRESS, see, e.g. [3],[4]. It formalizes the problem to find n points in a metric space $(\mathcal{X}, \alpha)$ such that the interpoint distances fit the given dissimilarities. This type of problem also occurs when visualizing global optimization algorithms (cf. [8]).

In order to search for a local minimum we may accept the representation (1). But in case of global search this representation has a serious drawback. The mutual distances between the two dimensional points (x_{i1}, x_{i2}), $i = 1, \ldots, n$, are invariant w.r.t. translations and rotations. This means that the set of global minimum points of $f(X)$, and also sets of local minimum and maximum points, are the union of certain orbits, i.e. a continuum of points. It is well known that such subsets with constant function values (especially if they are not very different from the optimum) cause difficulties for almost all global optimization algorithms.

The problem may be reformulated in the following way. Let us introduce the restrictions

$$\sum_{i=1}^{n} x_{i1} = 0, \quad \sum_{i=1}^{n} x_{i2} = 0, \tag{2}$$

which obviously exclude invariance with respect to translations. To achieve rotational invariance we represent the i-th two dimensional point in polar co-ordinates as $(\rho_i \cos\psi_i,\ \rho_i \sin\psi_i)$. The set of two dimensional points is rotated by an angle θ such that the columns of the configuration are orthogonal

$$\sum_{i=1}^{n} x_{i1}x_{i2} = 0. \tag{3}$$

This gives

$$\sum_{i=1}^{n} x_{i1}x_{i2} = \sum_{i=1}^{n} \rho_i^2 \cos(\psi_i + \theta)\sin(\psi_i + \theta) =$$
$$\frac{1}{4}\sum_{i=1}^{n} \rho_i^2(\sin(2\theta)\cos(2\psi_i) + \cos(2\theta)\sin(2\psi_i)) = 0,$$

TABLE I. The matrix of distances

***	1.27	1.69	2.04	3.09	3.20	2.86	3.17	3.21	2.38
0.49	***	1.43	2.35	3.18	3.22	2.56	3.18	3.18	2.31
1.74	1.47	***	2.43	3.26	3.27	2.58	3.18	3.18	2.42
2.26	2.08	2.67	***	2.85	2.88	2.59	3.12	3.17	1.94
2.14	1.84	0.84	2.67	***	1.55	3.12	1.31	1.70	2.85
1.85	1.39	1.32	1.92	1.39	***	3.06	1.64	1.36	2.81
1.62	1.43	1.70	1.74	1.39	1.48	***	3.00	2.95	2.56
0.79	0.82	1.59	1.96	2.04	1.78	1.66	***	1.32	2.91
1.87	1.69	1.66	1.76	2.01	1.47	1.84	1.68	***	2.97
1.85	1.50	1.67	2.28	1.38	1.11	1.06	2.11	2.04	***

and a solution is achieved if

$$\tan(2\theta) = -\frac{\sum \rho_i^2 \sin(2\psi_i)}{\sum \rho_i^2 \cos(2\psi_i)}. \tag{4}$$

The restrictions (2) and (3) exclude the drawback mentioned above not completely. The angle θ can be uniquely defined by (4) only up to additive constants $-\pi/2$, $\pi/2$, π. Moreover, because we may choose left or right orientation of the coordinate system arbitrarily the degrees of freedom are doubled. In summary this yields 8 possible coordinate representations of each configuration **X**. Therefore, the problem

$$\min f(X), \quad \text{such that} \quad \sum_{i=1}^{n} x_{i1} = \sum_{i=1}^{n} x_{i2} = \sum_{i=1}^{n} x_{i1}x_{i2} = 0,$$

will have at least 8 global minimum points, and 8 symmetric points at each local minimum. The introduction of more complicated restrictions could possibly reduce the number of symmetric points, but also leads to a more complicated feasible region. On the other hand, the symmetry of points can be easily taken into account, e.g. in case of computing the lower bounds in a Lipshitzian approach or introducing additional terms of a tunneling function.

Most global optimization algorithms may be applied only for problems with a bounded feasible region. Because of (2) and the inherent symmetry, the corresponding interval should be symmetric. An appropriate choice is the $2n$-dimensional cube $[-a, a]^{2n}$, where for example $a = \alpha \max_{i,j} \delta_{ij}$ for some sufficiently large $\alpha > 0$ ensures that the global minimum is attained in $[-a, a]^{2n}$.

3. Test Results

Let us consider two $2n$ dimensional test functions for $n = 10$, $w_{ij} = 1$, and δ_{ij} defined by the upper and lower triangular matrix in Table I, respectively.

To save space we have combined the data for both cases into a single array in Table I. The lower triangular matrix corresponds to the distances between ten randomly generated points in the five dimensional cube $(-1, 1)^5$. The upper triangular matrix corresponds to the experimental data on Cola testing from [3],[4].

To reveal structural characteristics of the optimization problem thousand local searches were performed by means of the algorithm [7] with starting points generated randomly according to a uniform distribution over the cube $[-1.2, 1.2]^{20}$. In general the starting points were not feasible with respect to (2) and (3). The interval constraints were chosen wide enough to impose no extra restrictions.

At the points X with $x_{i1} = x_{j1}$ and $x_{i2} = x_{j2}$ the objective function is nondifferentiable. However, during the experiments nondifferentiable points never occured, indicating that for (1) descend trajectories under NLPQL (cf.[7] with differentiable starting points do not contain nondifferentiable points.

For the first test function (lower triangular matrix) we detected 133 different local minimum points. The best function value 4.85 was found 156 times at 8 different points (25,25,24,21,19,17,14,11 times each). For example, two points where the best minimum was attained are:

$$X_1^* =$$
$$(-0.064\ \ -0.071\ \ -1.11\ \ 1.60\ \ -1.28\ \ -0.30\ \ 0.21\ \ 0.40\ \ 1.10\ \ -0.49$$
$$-1.160\ \ -0.710\ \ -0.30\ \ 0.74\ \ 0.41\ \ 0.37\ \ 0.83\ \ 1.04\ \ -0.22\ \ 1.07)$$

$$X_2^* =$$
$$(-1.16\ \ -0.71\ \ -0.29\ \ 0.74\ \ 0.41\ \ 0.37\ \ 0.83\ \ 1.04\ \ -0.21\ \ 1.06$$
$$0.064\ \ 0.072\ \ 1.10\ \ -1.60\ \ 1.28\ \ 0.30\ \ -0.21\ \ -0.40\ \ -1.10\ \ 0.49)$$

The region of attraction of the best points may be estimated as 15.6%. The largest numbers of repeatedly found other points were: 30 times with function value 5.53, 24 times with 5.53, 23 times with 5.53, 21 times with 5.29. 17 points were found twice, 30 points were obtained once. The mean value of the detected local minima was 5.60 with standard deviation 0.88.

For the second test function (the upper triangular matrix in Table I) the number of attained different local minimum points was 457. The best function value 11.75 was attained 48 times at 15 different points (9,8,7,6,4,4,2 times and 8 points once). For example, three points where the best minimum was attained are:

$$X_1^* =$$
$$(-1.74\ \ -1.59\ \ -1.53\ \ -1.10\ \ 1.49\ \ 1.82\ \ -0.27\ \ 1.58\ \ 1.62\ \ -0.30$$
$$-0.31\ \ 0.32\ \ 0.97\ \ -1.38\ \ -0.90\ \ -0.32\ \ 1.93\ \ 0.32\ \ 0.88\ \ -1.50)$$

$$X_2^* =$$
$$(0.31\ \ -0.32\ \ -0.97\ \ 1.38\ \ 0.90\ \ 0.32\ \ -1.93\ \ -0.32\ \ -0.88\ \ 1.50$$
$$-1.74\ \ -1.59\ \ -1.53\ \ -1.10\ \ -1.49\ \ 1.82\ \ -0.27\ \ 1.58\ \ 1.62\ \ -0.30)$$

$$X_3^* =$$
$$\begin{pmatrix} -0.32 & 0.31 & 0.97 & -1.37 & -0.90 & 0.30 & 1.93 & -0.27 & 0.86 & -1.51 \\ 1.74 & 1.59 & 1.53 & 1.10 & -1.53 & -1.48 & 0.27 & -1.85 & -1.66 & 0.29 \end{pmatrix}$$

The region of attraction of the best points may be estimated as 4.8%. The largest numbers of repeatedly found other points were: 13 times with function value 11.77, 12 times with 12.14, 9 times with function values 11.77, 13.06, 11.81, 12.14. The mean value of the detected local minima was 13.03 with standard deviation 1.05.

In summary the problem seems to be more difficult if the dissimilarities δ_{ij} are far from being Euclindean. The collection of test functions may be generated randomly, e.g. by means of distances between random points in an m-dimensional Euclidean space. Data on practical problems may be found in references given in [3]. More difficult test problems may be obtained using in (1) Minkovski distances instead of Euclidean ones and a higher dimensional embedding space.

The test results with these functions will hopefully lead to the development of improved algorithms for multidimensional scaling and visualisation of multidimensional data.

References

1. Crowder, H. Dembo, R.S., and Mulvey, J.M., (1978), Reporting computational experiments in mathematical programming, *Mathematical Programming* **15**, 316–329.
2. Floudas, C. and Pardalos, P. (1990), *A Collection of Test Problems for Constrained Global Optimization Algorithms*, Lecture Notes in Computer Science **455**, Springer, Berlin.
3. Green, P.E., Carmone, F.J. jr., and Smith, S.M. (1989), *Multidimensional Scaling, Concepts and Applications*, Allyn and Bacon, Boston.
4. Groenen, P., Mathar, R, and Heiser W (1993), The Majorization Approach to Multidimensional Scaling for Minkowski Distances, *Journal of Classification*, (to appear 1994).
5. Hock, W. and Schittkowski, K. (1981), *Test Examples for Nonlinear Programming* Codes, Lecture Notes in Economics and Mathematical Systems **187**, Springer, Berlin.
6. Jackson, R. and Mulvey, J. (1978), A critical review of comparsion of mathematical programming algorithms and software, *Journal of Research of the National Bureau of Standards* **83**, 563–584.
7. Schittkovski, K (1985), NLPQL: A Fortran subroutine solving constrained nonlinear programming problems, *Annals of Operations Research* **5**, 485–500.
8. Törn, A. and Žilinskas, A. (1989), *Global Optimization*, Lecture Notes in Computer Science, **350**, Springer, Berlin.