Mathematical Modelling and Analysis
Volume 14 Number 2, 2009, pages 259–270
Doi:10.3846/1392-6292.2009.14.259-270

ISSN 1392-6292 print
ISSN 1648-3510 online
www.vgtu.lt/mma/

Multidimensional Scaling with City-Block Distances Based on Combinatorial Optimization and Systems of Linear Equations

J. Žilinskas

Institute of Mathematics and Informatics
Akademijos 4, LT-08663 Vilnius, Lithuania
E-mail: `zilinskasjulius@gmail.com`

Received September 30, 2008; revised January 14, 2009; published online April 30, 2009

Abstract. Multidimensional scaling is a technique for exploratory analysis of multidimensional data. The essential part of the technique is minimization of a multimodal function with unfavorable properties like invariants and non-differentiability. In this paper a two-level optimization based on combinatorial optimization and systems of linear equations is proposed exploiting piecewise quadratic structure of the objective function with city-block distances. The approach is tested experimentally and improvement directions are identified.

Key words: Multidimensional scaling; city-block distances; multilevel optimization; combinatorial optimization.

1 Introduction

Multidimensional scaling (MDS) is a technique for exploratory analysis of multidimensional data widely usable in different applications [4, 6]. Pairwise dissimilarities among n objects are given by the matrix (δ_{ij}), $i, j = 1, \ldots, n$. A set of points in an embedding metric space is considered as an image of the set of objects. Normally, an m-dimensional vector space is used, and $\mathbf{x}_i \in \mathbb{R}^m$, $i = 1, \ldots, n$, should be found whose inter-point distances fit the given dissimilarities. Images of the considered objects can be found minimizing a fit criterion [22], e.g. the most frequently used least squares *STRESS* function:

$$S(\mathbf{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left(d\left(\mathbf{x}_i, \mathbf{x}_j\right) - \delta_{ij} \right)^2,$$

where $\mathbf{x} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$, $\mathbf{x}_i = (x_{i1}, x_{i2}, \ldots, x_{im})$. It is supposed that the weights are positive: $w_{ij} > 0, i, j = 1, \ldots, n$; $d(\mathbf{x}_i, \mathbf{x}_j)$ denotes the distance between the points $\mathbf{x}_i$ and $\mathbf{x}_j$. Usually Minkowski distances are used:

$$d_r(\mathbf{x}_i, \mathbf{x}_j) = \Big(\sum_{k=1}^{m} |x_{ik} - x_{jk}|^r \Big)^{1/r}.$$

The equation defines Euclidean distances when $r = 2$, and city-block distances when $r = 1$. The most frequently used distances are Euclidean, but multidimensional scaling with other Minkowski distances in the embedding space can be even more informative [1]. In the present paper the problem with the *STRESS* criterion and city-block distances in the embedding space are considered.

STRESS normally has many local minima. It is invariant with respect to the translation and mirroring. It can be non-differentiable even at a minimum point [18]; the case of city-block metric is different from the other cases of Minkowski metric where positiveness of distances at a local minimum point imply differentiability of *STRESS* [7, 9]. However *STRESS* with city-block distances is piecewise quadratic, and such a structure can be exploited for tailoring of ad hoc global optimization algorithms.

Global optimization algorithms for multidimensional scaling are reviewed with particular emphasis on parallel computing in [27]. A heuristic algorithm based on simulated annealing for two-dimensional city-block scaling was proposed in [5]. Each coordinate axis is partitioned by uniformly spaced points, and a simulated annealing algorithm is used to search the lattice defined by these points aiming to minimize one of two types of *STRESS* either the sum of squares or the sum of corresponding absolute values. The solution found is locally improved by quadratic programming. A two stages approach for city-block MDS was proposed in [11]. The least square regression is used to obtain a local minimum of Stress function in the first stage. Simulated annealing is used in the second stage of the method. A multivariate randomly alternating simulated annealing procedure with permutation and translation phases has been applied to develop an algorithm for multidimensional scaling in any Minkowski metric in [14].

A two-level minimization method for the two-dimensional embedding space was proposed in [18] where a problem of combinatorial optimization is tackled by evolutionary search at the upper level, and a problem of quadratic programming is tackled at the lower level. The parallel version of the algorithm is proposed and investigated in [16]. Quantitatively the precision of the algorithm was estimated using global minima found by means of the developed parallel version of explicit enumeration algorithm in [17]. Efficiency of the parallel version of the algorithm on computational grids is investigated in [13]. The generalized method for arbitrary dimensionality of the embedding space is developed and experimentally compared with other approaches in [19]. A branch and bound algorithm for the upper level combinatorial problem is proposed in [21].

Dependence of visualization error on the dimensionality of embedding space is investigated in [26]. Visualization of geometrical multidimensional data in three-dimensional space is investigated in [15] and of empirical multidimensional data in [20]. The images in three-dimensional embedding space normally show the structural properties of sets of considered objects with acceptable accuracy, and widening of applications of stereo screens makes three-dimensional visualization very attractive.

In this paper a new two-level approach for multidimensional scaling is presented based on combinatorial optimization and systems of linear equations. In the next section multidimensional scaling with city-block distances based on quadratic programming is presented. In Sect. 3 the approach based on systems of linear equations is proposed. Experimental investigation is described in Sect. 4. Although the number of feasible solutions of the upper level combinatorial problem is larger than in the case of two-level optimization based on quadratic programming, the lower level problems are simpler. The conclusions are drawn in the last section.

2 Multidimensional Scaling with City-Block Distances Based on Quadratic Programming

STRESS with city-block distances $d_1(\mathbf{x}_i, \mathbf{x}_j)$ can be redefined as

$$S(\mathbf{x}) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\Big(\sum_{k=1}^{m} |x_{ik} - x_{jk}| - \delta_{ij}\Big)^2.$$

Let $A(\mathbf{P})$ denotes a set such that

$$A(\mathbf{P}) = \Big\{\mathbf{x} | \ \ x_{ik} \le x_{jk} \text{ for } p_{ki} < p_{kj}, \ i,j = 1,\dots,n, \ k = 1,\dots,m\Big\},$$

where $\mathbf{P} = (\mathbf{p}_1, \dots, \mathbf{p}_m)$, $\mathbf{p}_k = (p_{k1}, p_{k2}, \dots, p_{kn})$ is a permutation of $1,\dots,n$; For $\mathbf{x} \in A(\mathbf{P})$,

$$S(\mathbf{x}) = \sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\Big(\sum_{k=1}^{m} (x_{ik} - x_{jk})\, z_{kij} - \delta_{ij}\Big)^2,$$

where

$$z_{kij} = \begin{cases} 1, & p_{ki} > p_{kj}, \\ -1, & p_{ki} < p_{kj}. \end{cases}$$

Since function $S(\mathbf{x})$ is quadratic over polyhedron $\mathbf{x} \in A(\mathbf{P})$ the minimization problem

$$\min_{\mathbf{x} \in A(\mathbf{P})} S(\mathbf{x})$$

is a quadratic programming problem. It is equivalent to (see [26])

$$\min \Big[-\sum_{k=1}^{m}\sum_{i=1}^{n} x_{ik} \sum_{j=1}^{n} w_{ij}\delta_{ij} z_{kij} + \frac{1}{2}\sum_{k=1}^{m}\sum_{l=1}^{m}\sum_{i=1}^{n}\Big(x_{ik}x_{il} \sum_{t=1, t\ne i}^{n} w_{it} z_{kit} z_{lit} - \sum_{j=1, j\ne i}^{n} x_{ik}x_{jl} w_{ij} z_{kij} z_{lij}\Big)\Big],$$

$$\text{s.t.} \sum_{i=1}^{n} x_{ik} = 0, \ k = 1, \ldots, m,$$

$$x_{\{j|p_{kj}=i+1\},k} - x_{\{j|p_{kj}=i\},k} \geq 0, \ k = 1, \ldots, m, \ i = 1, \ldots, n-1.$$

Taking into account the structure of the minimization problem a two-level minimization algorithm can be applied [18]:

1. To solve a combinatorial problem at the upper level;

2. To solve a quadratic programming problem at the lower level:

$$\min_{\mathbf{P}} S(\mathbf{P}),$$

$$\text{s.t. } S(\mathbf{P}) = \min_{\mathbf{x} \in A(\mathbf{P})} S(\mathbf{x}) \sim \min \left(-\mathbf{c_P}^T \mathbf{x} + \frac{1}{2} \mathbf{x}^T \mathbf{Q_P} \mathbf{x} \right), \ \text{s.t.} \ \begin{array}{c} \mathbf{Ex} = \mathbf{0}, \\ \mathbf{A_P x} \geq \mathbf{0}. \end{array}$$

For the lower level problem a standard quadratic programming method can be applied. The upper level problem is solved with guarantee using explicit enumeration of all feasible solutions or using the branch and bound method [21]. Genetic algorithm is applied for larger problems [18, 19].

A minimum point of a quadratic programming problem is not necessary a local minimizer of the initial problem of minimization of *STRESS*, if it is on the boundary of polyhedron. Local search may be applied [18]:

- go to the neighbour polyhedron on the opposite side of the active inequality constrains;

- perform quadratic programming;

- repeat while better values are found and some inequality constrains are active.

Although such a heuristic local search improves performance of hybrid optimization algorithm, it does not guarantee that the found minimum point is a local minimizer of minimization of *STRESS*. Therefore it cannot be used to count the number of local minimizers.

3 Multidimensional Scaling with City-Block Distances Based on Systems of Linear Equations

Let us define a different decomposition of optimization problem which is more convenient to derive a two-level minimization problem with combinatorial problem at the upper level and a system of linear equations at the lower level. Let us change the variables to

$$y_{lk} = x_{\{j|p_{kj}=l+1\},k} - x_{\{j|p_{kj}=l\},k}, \ \ k = 1, \ldots, m, \ \ l = 1, \ldots, n_k,$$

where $n_k < n$ is the number of different values of kth coordinate minus one. Here $p_{ki} = p_{kj}$ may be allowed even when $i \neq j$. In the case $\mathbf{p}_k$ is a permutation of $1, \ldots, n$, this is not allowed and $n_k = n - 1$.

Polyhedron $\mathbf{x} \in A(\mathbf{P})$ can be defined by $y_{lk} \geq 0$, $k = 1, \dots, n$, $l = 1, \dots, n_k$. Interior of the polyhedron can be defined by $y_{lk} > 0$. For $\mathbf{x} \in A(\mathbf{P})$, *STRESS* with city-block distances can be rewritten in the following form:

$$S(\mathbf{x}) = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \left(\sum_{k=1}^{m} \sum_{l=1}^{n_k} y_{lk} z_{lkij} - \delta_{ij} \right)^2,$$

where

$$z_{lkij} = \begin{cases} 1, & \min(p_{ki}, p_{kj}) \leq l < \max(p_{ki}, p_{kj}), \\ 0, & \text{otherwise}. \end{cases}$$

The quadratic function $S(\mathbf{x})$ can be written in the following form:

$$\begin{aligned} S(\mathbf{x}) = & \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \Big(\sum_{k=1}^{m} \sum_{l=1}^{n_k} y_{lk} z_{lkij} - \delta_{ij} \Big)^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \delta_{ij}^2 \\ & -2 \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \delta_{ij} \sum_{k=1}^{m} \sum_{l=1}^{n_k} y_{lk} z_{lkij} + \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \Big(\sum_{k=1}^{m} \sum_{l=1}^{n_k} y_{lk} z_{lkij} \Big)^2. \end{aligned}$$

The first summand is a constant, and need not be taken into account in minimization. The second summand is a linear function which can be rewritten as follows

$$-2 \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \delta_{ij} \sum_{k=1}^{m} \sum_{l=1}^{n_k} y_{lk} z_{lkij} = -2 \sum_{k=1}^{m} \sum_{l=1}^{n_k} y_{lk} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \delta_{ij} z_{lkij}.$$

Similarly the third summand can be written as a quadratic function

$$\begin{aligned} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \Big(\sum_{k=1}^{m} \sum_{l=1}^{n_k} y_{lk} z_{lkij} \Big)^2 &= \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \sum_{k=1}^{m} \sum_{l=1}^{n_k} \sum_{u=1}^{m} \sum_{v=1}^{n_u} y_{lk} y_{vu} z_{lkij} z_{vuij} \\ &= \sum_{k=1}^{m} \sum_{l=1}^{n_k} \sum_{u=1}^{m} \sum_{v=1}^{n_u} y_{lk} y_{vu} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} z_{lkij} z_{vuij}. \end{aligned}$$

Therefore $\min_{\mathbf{x} \in A(\mathbf{P})} S(\mathbf{x})$ is equivalent to

$$\begin{aligned} \min \Big(& -2 \sum_{k=1}^{m} \sum_{l=1}^{n_k} y_{lk} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \delta_{ij} z_{lkij} \\ & + \sum_{k=1}^{m} \sum_{l=1}^{n_k} \sum_{u=1}^{m} \sum_{v=1}^{n_k} y_{lk} y_{vu} \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} z_{lkij} z_{vuij} \Big), \\ & \text{s.t. } y_{lk} \geq 0, \quad k = 1, \dots, m, \quad l = 1, \dots, n_k. \end{aligned}$$

The coordinate values of image points can be found from the corresponding minimum point of the constrained quadratic problem:

$$x_{ik}^* = \sum_{l=1}^{p_{ki}-1} y_{lk}^*.$$

Therefore similarly as in the previous section, a two-level minimization problem with combinatorial problem at the upper level and a quadratic programming problem at the lower level can be defined:

$$\min_{\mathbf{P}} S(\mathbf{P}),$$

$$\text{s.t. } S(\mathbf{P}) = \min_{\mathbf{x}\in A(\mathbf{P})} S(\mathbf{x}) \sim \min\Big(-\mathbf{c_P}^T\mathbf{y} + \frac{1}{2}\mathbf{y}^T\mathbf{Q_P}\mathbf{y}\Big), \text{ s.t. } \mathbf{y} \geq \mathbf{0},$$

where

$$c_{lk} = 2\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\delta_{ij}z_{lkij}, \quad Q_{lk\ vu} = 2\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}z_{lkij}z_{vuij}.$$

It can be seen that the linear equality and inequality constraints have been avoided comparing with the decomposition presented in the previous section and only the bound constraints $\mathbf{y} \geq \mathbf{0}$ have been left. Such constraints are checked and managed easier. Moreover the number of variables is at least by m smaller.

Since $\mathbf{Q_P}$ is positive definite, objective function of quadratic problem is convex and has a unique minimizer. Therefore a standard quadratic programming method can be applied for the lower level problem.

The upper level function is defined over the set of m-tuple of permutations of $1, \ldots, n$ representing sequences of coordinate values of image points. The number of feasible solutions of the upper level combinatorial problem is $(n!)^m$. A solution of MDS with city-block distances is invariant with respect to mirroring when changing direction of coordinate axes or exchanging of coordinates [21]. The feasible set can be reduced taking into account these symmetries. The number of feasible solutions can be reduced to $(n!/2)^m$ refusing mirrored solutions changing direction of each coordinate axis. It can be further reduced to approximately $(n!/2)^m/m!$ refusing mirrored solutions with exchanged coordinates.

In the case of $m = 1$ and $n = 3$ all possible permutations are "123", "132", "231", "213", "312" and "321". Here every numeral represents a value of p_{1i}. To refuse mirrored solutions with changed direction of coordinate axes, permutations with $p_{k1} > p_{k2}$ can be forbidden as for example "123" and "321" are equivalent. In this case allowed permutations are "123", "132" and "231". In the case of $m = 2$ and $n = 3$ allowed permutations would be "123/123", "123/132", "123/231", "132/123", "132/132", "132/231", "231/123", "231/132" and "231/231". Here every numeral represents p_{ki} and "/" separates coordinates. To refuse mirrored solutions with exchanged coordinates some restrictions on permutations may be set. Let us define the order for permutations of $1, 2, 3$ as "123" $\prec$ "132" $\prec$ "231". A permutation $\mathbf{p}_k$ cannot precede $\mathbf{p}_l$ for $k > l$ ($l < k \Rightarrow \mathbf{p}_l \preceq \mathbf{p}_k$). Therefore tuples of permutations "132/123", "231/123" and "231/132" are not allowed, as they represent symmetric solutions to "123/132", "123/231" and "132/231" respectively. Therefore in this case allowed permutations are "123/123", "123/132", "123/231", "132/132", "132/231" and "231/231".

If the minimum point of *STRESS* is not on the boundary of polyhedron $\mathbf{x} \in A(\mathbf{P})$ then it can be found solving a system of linear equations searching where the gradient of the quadratic function is zero. If it is on the boundary of polyhedron then the gradient is zero at the point which is not in the polyhedron. However it is possible to define $A(\mathbf{P'})$ (corresponding to either polyhedron $A(\mathbf{P})$ or its faces and edges):

$$A(\mathbf{P'}) = \left\{ \mathbf{x} \,\middle|\, \begin{array}{l} x_{ik} < x_{jk} \text{ for } p'_{ki} < p'_{kj}, \\ x_{ik} = x_{jk} \text{ for } p'_{ki} = p'_{kj}, \end{array} \quad i,j = 1,\dots,n,\ k = 1,\dots,m \right\}.$$

Here p'_{ki} may be equal to p'_{kj} even if $i \neq j$, and therefore they can define polyhedrons $A(\mathbf{P'})$ which are faces and edges of polyhedron $A(\mathbf{P})$. If $p'_{ki} = p'_{kj}$ for $i \neq j$, then $x_{ik} = x_{jk}$ and one of these variables is eliminated from the function. It is possible to find the minimum points of *STRESS* solving systems of linear equations searching where the gradient of the reduced quadratic function is zero in polyhedrons $A(\mathbf{P'})$.

Two-level problem with quadratic programming at the lower level can be redefined as a two-level minimization problem with combinatorial problem at the upper level and a system of linear equations at the lower level:

$$\min_{\mathbf{P'}} S(\mathbf{P'}),$$

$$\text{s.t. } S(\mathbf{P'}) = \min_{\mathbf{x} \in A(\mathbf{P'})} S(\mathbf{x}) \sim \mathbf{y}^T \mathbf{Q}_{\mathbf{P'}} = \mathbf{c}_{\mathbf{P'}}{}^T, \text{ s.t. } \mathbf{y} > \mathbf{0}.$$

If the solution of system of linear equations is not in polyhedron $A(\mathbf{P'})$, then the lower level problem is not feasible. This can be easily checked testing if $\mathbf{y}^* > \mathbf{0}$, which is computationally cheaper than to check linear inequality constraints if decomposition presented in the previous section was used.

In the case of $m = 1$ and $n = 3$ possible sequences for $\mathbf{P'}$ are “123”, “122”, “132”, “121”, “231”, “112”, “111”, “221”, “213”, “212”, “312”, “211”, “321”. To refuse mirrored solutions with changed direction of coordinate axes, similarly to the case of permutations $\mathbf{P}$, sequences with $p_{k1} > p_{k2}$ and $p_{k1} = p_{k2} \& p_{k1} > p_{k3}$ can be forbidden as for example “112” and “221” are equivalent as well as “123” and “321”. In this case allowed permutations are “123”, “122”, “132”, “121”, “231”, “112” and “111”.

The upper level problem may be solved with guarantee using explicit enumeration of all feasible solutions or a branch and bound algorithm similar to [21, 25]. A branch and bound template [2, 3] may be used for implementation. Metaheuristic algorithms may be applied when problems are too large to be solved with guarantee.

4 Experimental Investigation

The primal goal of the experimental investigation is to check if the proposed approach is correct. Explicit enumeration for the upper level combinatorial problems is used. However it is also interesting to compare performance of the proposed approach with the approach based on quadratic programming. The number of lower level problems solved is used for comparison.

At this stage of the research computational time is not used for comparison. This is because a system of linear equations is solved using LU decomposition and then the bound constraints are checked. Such an approach is enough for testing the correctness, but requires approximately the same computational time as solution of quadratic programming problem by a standard algorithm. It can be expected that iterative algorithms for solution of bound constrained system of linear equations may be faster as they can be stopped immediately when it is found that the solution cannot be feasible. In the case of quadratic programming problem the smallest value of quadratic function at the boundary of the feasible region should still be found. Investigation of algorithms for solution of bound constrained system of linear equations as well as quadratic programming problem is one of the directions of future research.

The accuracy of fit evaluated via minimum of $S(\mathbf{x})$ depends on n and $\delta_{ij}, i, j = 1, \ldots, n$. To reduce this undesirable impact, a relative error is used in the presentation of the results:

$$f(\mathbf{x}) = \sqrt{S(\mathbf{x}) \Big/ \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} \delta_{ij}^2}\,.$$

Several sets of multidimensional points corresponding to well understood geometric objects have been used for experimental investigation: the sets of vertices of multidimensional simplices and cubes [24]. Below we use shorthand 'simplex' and 'cube' for the sets of their vertices.

A frequently used test problem for MDS algorithms is based on experimental testing of several soft drinks [8]. 38 students have tested ten different brands of soft drinks. Each pair was judged on its dissimilarity on a 9 point scale (1 – very similar, 9 – completely different). The accumulated dissimilarities have been used as a practical data set in our experiments. This problem is referred as 'cola' problem in the results below, $n = 10$ in this problem.

Problems of analysis of pharmacological binding affinity data [23] have been used as other practical data sets. Binding affinity data is represented through a matrix, one dimension formed by different ligands tested in a series of experiments while the other dimension represents different proteins. Dissimilarities of proteins are computed as city-block distances between vectors of the $\log_{10}$-transformed binding affinities representing properties of the proteins. Dissimilarities of ligands are computed as city-block distances between vectors of the $\log_{10}$-transformed binding affinities representing ligands. 'ruusk' represents binding affinity data of [12] analyzed as properties of three human and five zebrafish α_2-adrenoceptor proteins, $n = 8$; 'hwa' represents binding affinity data of [10] analyzed as properties of ligands, $n = 9$.

The results of experimental investigation are shown in Table 1. All values of relative errors f^* corresponding to minima of *STRESS* coincide for both approaches – the same solutions of the problems have been found. The numbers of systems of linear equations (NSLE) are quite larger than the numbers of quadratic programming problems (NQPP) what is not surprising. The ratio of the numbers increases exponentially depending on n and m. This means that such an approach based on systems of linear equations may be faster than one based on quadratic programming problems only if solution of bound constrained system of linear equations is by at least the number of variables

Table 1. Results of explicit enumeration.

m	n	f^* unit simplices	standard simplices	cubes	empirical data sets	NQPP	NSLE
1	3	0.0000	0.3333			3	7
1	4	0.3651	0.4082	0.4082		12	38
1	5	0.4140	0.4472			60	271
1	6	0.4554	0.4714			360	2342
1	7	0.4745	0.4879			2520	23647
1	8	0.4917	0.5000	0.4787	ruusk 0.2975	20160	272918
1	9	0.5018	0.5092		hwa 0.0107	181440	3543631
1	10	0.5113	0.5164		cola 0.3642	1814400	51123782
1	11	0.5176	0.5222			19958400	811316287
2	3	0.0000	0.0000			6	28
2	4	0.0000	0.0000	0.0000		78	741
2	5	0.0000	0.1907			1830	36856
2	6	0.1869	0.2309			64980	2743653
2	7	0.2247	0.2621			3176460	279602128

times faster than solution of quadratic programming problem.

If the lower level problem is feasible, the solution of system of linear equations is in polyhedron $A(\mathbf{P'})$, therefore the minimum point of *STRESS* is in polyhedron and there is no need to find the minimum point on the faces and edges of this polyhedron. Therefore if the system defined by "123/123" has solution with $\mathbf{y} > \mathbf{0}$, it is not necessary to solve systems defined by "123/122", "123/112", "122/122", "122/112", "112/112", "123/111", "111/111". The numbers of feasible systems of linear equations (NFSLE) are shown in Table 2. All lower level problems are feasible for the problems of standard simplices when $m = 1$. This means that in this case all possible polyhedrons $A(\mathbf{P})$ contain minima points and therefore *STRESS* function has $n!$ minima points. The numbers of feasible lower level problems are smaller for other data sets and for the case of $m \neq 1$, but the numbers are still quite large. Moreover, even if only $y_{12} > 0$ solving the system of linear equations defined by "123/123", it is not necessary to evaluate the system defined by "123/112". This encourages development of an algorithm which takes into account the results of the solution of lower level problems. Not only the number of required to solve systems would be reduced if this will be taken into account, but the number of feasible lower level problems might be used to count the number of minima points of *STRESS* function.

It is proved in [9] that the distances between image points are positive at a local minimum point of *STRESS*. Therefore it is possible to avoid coincidence of image points. This may be performed avoiding for example systems defined by "112" and "112/112". This would also reduce the number of lower level problems and make the approach more attractive, but further investigation should be performed.

Table 2. The numbers of feasible lower level problems.

m	n	NSLE	NFSLE			
			unit simplices	standard simplices	cubes	empirical data sets
1	3	7	7	7		
1	4	38	38	38	38	
1	5	271	219	271		
1	6	2342	1907	2342		
1	7	23647	17026	23647		
1	8	272918	186599	272918	221975	ruusk 210195
1	9	3543631	2378067	3543631		hwa 1350181
1	10	51123782	34412855	51123782		cola 49208660
1	11	811316287	519221622	811316287		
2	3	28	21	25		
2	4	741	380	426	421	
2	5	36856	11139	20670		
2	6	2743653	606634	1414761		
2	7	279602128	44663090	118027269		

5 Conclusions

Two-level optimization for multidimensional scaling with city-block distances based on combinatorial optimization and systems of linear equations is proposed. The approach exploits piecewise quadratic structure of the objective function. The lower level problems are bound constrained systems of linear equations.

For investigation of the approach, the upper level combinatorial problem is solved with guarantee using explicit enumeration of all feasible solutions. Development of a branch and bound as well as metaheuristic algorithms for the upper level combinatorial problem is a direction for future research.

The approach has been tested solving geometrical and empirical data sets. The numbers of lower level problems are larger than in the case of similar approach based on combinatorial optimization and quadratic programming. However, lower level problems of systems of linear equations are simpler than quadratic programming problems.

The numbers of lower level problems can be reduced significantly taking into account results of their solutions. This is the primal direction for future research. Investigation of algorithms for solution of bound constrained system of linear equations as well as quadratic programming problem is another direction for future research.

6 Acknowledgements

The research is supported by Lithuanian State Science and Studies Foundation within the project B-03/2007 "Global optimization of complex systems using high performance computing and GRID technologies".

References

[1] P. Arabie. Was Euclid an unnecessarily sophisticated psychologist? *Psychometrika*, **56**(4):567–587, 1991. (doi:10.1007/BF02294491)

[2] M. Baravykaitė and R. Čiegis. An implementation of a parallel generalized branch and bound template. *Math. Model. Anal.*, **12**(3):277–289, 2007. (doi:10.3846/1392-6292.2007.12.277-289)

[3] M. Baravykaitė, R. Čiegis and J. Žilinskas. Template realization of generalized branch and bound algorithm. *Math. Model. Anal.*, **10**(3):217–236, 2005.

[4] I. Borg and P. J. F. Groenen. *Modern Multidimensional Scaling: Theory and Applications*. Springer, New York, 2nd edition, 2005.

[5] M. J. Brusco. A simulated annealing heuristic for unidimensional and multidimensional (city-block) scaling of symmetric proximity matrices. *Journal of Classification*, **18**(1):3–33, 2001. (doi:10.1007/s00357-001-0003-4)

[6] T. F. Cox and M. A. A. Cox. *Multidimensional Scaling*. Chapman & Hall/CRC, Boca Raton, 2nd edition, 2001.

[7] J. de Leeuw. Differentiability of Kruskal's stress at a local minimum. *Psychometrika*, **49**(1):111–113, 1984. (doi:10.1007/BF02294209)

[8] P. Green, F. Carmone and S. Smith. *Multidimensional Scaling: Concepts and Applications*. Allyn and Bacon, Boston, 1989.

[9] P. J. F. Groenen, R. Mathar and W. J. Heiser. The majorization approach to multidimensional scaling for Minkowski distances. *Journal of Classification*, **12**(1):3–19, 1995. (doi:10.1007/BF01202265)

[10] J. Hwa, R. M. Graham and D. M. Perez. Identification of critical determinants of α_1-adrenergic receptor subtype selective agonist binding. *Journal of Biological Chemistry*, **270**(39):23189–23195, 1995.

[11] P. L. Leung and K. Lau. Estimating the city-block two-dimensional scaling model with simulated annealing. *European Journal of Operational Research*, **158**(2):518–524, 2004. (doi:10.1016/S0377-2217(03)00357-6)

[12] J. O. Ruuskanen, J. Laurila, H. Xhaard, V.-V. Rantanen, K. Vuoriluoto, S. Wurster, A. Marjamäki, M. Vainio, M. S. Johnson and M. Scheinin. Conserved structural, pharmacological and functional properties among the three human and five zebrafish α_2-adrenoceptors. *British Journal of Pharmacology*, **144**(2):165–177, 2005. (doi:10.1038/sj.bjp.0706057)

[13] A. Varoneckas, A. Žilinskas and J. Žilinskas. Parallel multidimensional scaling using grid computing: assessment of performance. *Information Technology and Control*, **37**(1):52–56, 2008.

[14] J. F. Vera, W. J. Heiser and A. Murillo. Global optimization in any Minkowski metric: a permutation-translation simulated annealing algorithm for multidimensional scaling. *Journal of Classification*, **24**(2):277–301, 2007. (doi:10.1007/s00357-007-0020-1)

[15] A. Žilinskas and J. Žilinskas. On visualization of multidimensional data using three-dimensional embedding space. *Technological and Economic Development of Economy*, **12**(4):353–359, 2006.

[16] A. Žilinskas and J. Žilinskas. Parallel hybrid algorithm for global optimization of problems occurring in MDS based visualization. *Computers & Mathematics with Applications*, **52**(1-2):211–224, 2006. (doi:10.1016/j.camwa.2006.08.016)

[17] A. Žilinskas and J. Žilinskas. Parallel genetic algorithm: assessment of performance in multidimensional scaling. In *GECCO '07: Proceedings of the 9th annual conference on Genetic and evolutionary computation*, pp. 1492–1501. ACM, 2007. (doi:10.1145/1276958.1277229)

[18] A. Žilinskas and J. Žilinskas. Two level minimization in multidimensional scaling. *Journal of Global Optimization*, **38**(4):581–596, 2007. (doi:10.1007/s10898-006-9097-x)

[19] A. Žilinskas and J. Žilinskas. A hybrid method for multidimensional scaling using city-block distances. *Mathematical Methods of Operations Research*, **68**(3):429–443, 2008. (doi:10.1007/s00186-008-0238-5)

[20] A. Žilinskas and J. Žilinskas. Three-dimensional visualization by means of multidimensional scaling. In L. Sakalauskas, G. W. Weber and E. K. Zavadskas(Eds.), *The 20th International Conference EURO Mini Conference Continuous Optimization and Knowledge-Based Technologies (EurOPT-2008), May 20-23, 2008, Neringa, Lithuania*, pp. 71–76, 2008.

[21] A. Žilinskas and J. Žilinskas. Branch and bound algorithm for multidimensional scaling with city-block metric. *Journal of Global Optimization*, **43**(2-3):357–372, 2009. (doi:10.1007/s10898-008-9306-x)

[22] A. Žilinskas and J. Žilinskas. Optimization-based visualization. In C. A. Floudas and P. M. Pardalos(Eds.), *Encyclopedia of Optimization*, pp. 2785–2791. Springer, 2nd edition, 2009. (doi:10.1007/978-0-387-74759-0_478)

[23] J. Žilinskas. Multidimensional scaling in protein and pharmacological sciences. In I. D. L. Bogle and J. Žilinskas(Eds.), *Computer Aided Methods in Optimal Design and Operations*, pp. 139–148. World Scientific, 2006.

[24] J. Žilinskas. Reducing of search space of multidimensional scaling problems with data exposing symmetries. *Information Technology and Control*, **36**(4):377–382, 2007.

[25] J. Žilinskas. Branch and bound with simplicial partitions for global optimization. *Math. Model. Anal.*, **13**(1):145–159, 2008. (doi:10.3846/1392-6292.2008.13.145-159)

[26] J. Žilinskas. On dimensionality of embedding space in multidimensional scaling. *Informatica*, **19**(3):447–460, 2008.

[27] J. Žilinskas. Parallel global optimization in multidimensional scaling. In R. Čiegis, D. Henty, B. Kågström and J. Žilinskas(Eds.), *Parallel Scientific Computing and Optimization*, volume 27 of *Springer Optimization and Its Applications*, pp. 69–82. Springer, 2009. (doi:10.1007/978-0-387-09707-7_6)