Baltic J. Modern Computing, Vol. 2 (2014), No. 4, pp. 248–259

Quadratic Programming with Complementarity Constraints for Multidimensional Scaling with City-Block Distances

Roger FLETCHER[1], Nerijus GALIAUSKAS[2], Julius ŽILINSKAS[2]

[1] University of Dundee, Department of Mathematics, Dundee, DD1 4HN, Scotland, UK
[2] Vilnius University, Institute of Mathematics and Informatics, Akademijos 4 LT-08663 Vilnius, Lithuania

fletcher@maths.dundee.ac.uk,
{nerijus.galiauskas, julius.zilinskas}@mii.vu.lt

Abstract. In this paper, we consider an optimization problem arising in multidimensional scaling with city-block distances. The objective function of this problem has many local minimum points and may be even non-differentiable at a minimum point. We reformulate the problem into a problem with convex quadratic objective function, linear and complementarity constraints. In addition, we propose an algorithm to find a local solution of this reformulated problem.

Keywords: multidimensional scaling, city-block distances, quadratic programming, active-set method

1 Introduction

In this section, we present a problem arising in multidimensional scaling with city-block distances. Multidimensional scaling (MDS) is a data analysis method (Borg and Groenen, 2005, Živadinović, 2011, Dzemyda et al., 2013). It was first introduced by Torgerson (1952). In applying MDS, data are transformed into points in the one-, two- or three-dimensional Cartesian coordinate system. These points are thought of as an image of the given data. The image is then used to perform an analysis of the data. Let us describe the method in detail.

First of all, let us perceive the data as a set of n ($n > 2$) particular objects. In applying MDS, the first step is to measure relationships between objects. The relationship between objects i and j is usually called dissimilarity and is denoted by δ_{ij}. Here we restrict our attention to the case where $\delta_{ij} \in \mathbb{R}$ and $\delta_{ij} = \delta_{ji} > 0$, $\delta_{ii} = 0$, $1 \leq i < j \leq n$. Let $\Delta = (\delta_{ij})^{n\times n}$ denotes a dissimilarity matrix. If the objects are vectors in a real vector space, dissimilarities between them might easily be measured by applying any distance function, defined on that vector space.

The second step, in applying MDS, is to select a distance function, defined on vector space $\mathbb{R}^m$, where $m \in \{1, 2, 3\}$. On this space, MDS will generate an image of the data. The Minkowski distance of order p ($p \geq 1$) is usually selected:

$$d_p(x_i, x_j) = \left(\sum_{k=1}^{m} |x_{ki} - x_{kj}|^p \right)^{1/p},$$

where $x_i = (x_{1i}, \dots, x_{mi})^T$, $x_j = (x_{1j}, \dots, x_{mj})^T \in \mathbb{R}^m$.

The last and the most difficult step is to choose a loss function, defined on space $\mathbb{R}^{mn}$, and to minimize it on that space. Let us define a loss function as follows:

$$f_p(x) = \sum_{i<j}^{n} \left(d_p(x_i, x_j) - \delta_{ij} \right)^2,$$

where $x = (x_1^T, \dots, x_n^T) \in \mathbb{R}^{mn}$. Function f_p is usually called a least-squares stress function with L^p norm. If $x^* = ({x_1^*}^T, \dots, {x_n^*}^T) \in \mathbb{R}^{mn}$ is a minimizer of function f_p on $\mathbb{R}^{mn}$, a set of points $\{x_1^*, \dots, x_n^*\} \subset \mathbb{R}^m$ is perceived as an image of the given data. Note that function f_p is invariant under a translation and mirroring. Therefore, we usually introduce a set of constraints to center an image at the origin of m-dimensional Cartesian coordinate system:

$$\sum_{i=1}^{n} x_{ki} = 0, \ 1 \leq k \leq m.$$

Let us consider the minimization of function f_1 (a least-squares stress function with city-block distances):

$$\begin{aligned} &\underset{x \in \mathbb{R}^{mn}}{\text{minimize}} \sum_{i<j}^{n} \left(\sum_{k=1}^{m} |x_{ki} - x_{kj}| - \delta_{ij} \right)^2 \\ &\text{subject to} \sum_{i=1}^{n} x_{ki} = 0, \ 1 \leq k \leq m. \end{aligned} \tag{1}$$

Problem (1) is a difficult optimization problem. Objective function f_1 may have many local minimum points, it may be even non-differentiable at a minimum point (Žilinskas and Žilinskas, 2007). There are some methods devoted to solve this problem: Hubert et al., 1992, Groenen et al., 1999, Brusco, 2001, Leung and Lau, 2004, Žilinskas and Žilinskas, 2009. The methods vary from heuristic to deterministic methods. In this paper, we present a reformulation of problem (1) into an optimization problem with convex quadratic objective function, linear and complementarity constraints. An algorithm to find a local solution of the reformulated problem is proposed too. The algorithm is based on the active-set method for convex quadratic programming (Fletcher, 2006, Nocedal and Wright, 2006).

2 Quadratic programming with complementarity constraints

Let us reformulate problem (1) into a certain optimization problem with convex quadratic objective function, linear and complementarity constraints.

First of all, let $\mathbb{X}$ denotes a feasible set of problem (1). For every $x \in \mathbb{X}$ let us define a vector-valued function $\tilde{d}^{\pm} : \mathbb{X} \to \mathbb{R}_+^{mn(n-1)}$ such that $\tilde{d}^{\pm}(x) = (\tilde{d}_{112}^{+}(x), \tilde{d}_{112}^{-}(x), \ldots, \tilde{d}_{m12}^{+}(x), \tilde{d}_{m12}^{-}(x), \ldots, \tilde{d}_{m(n-1)n}^{+}(x), \tilde{d}_{m(n-1)n}^{-}(x))^T$ and

$$\begin{aligned} \tilde{d}_{kij}^{+}(x) &= \begin{cases} x_{ki} - x_{kj}, & x_{ki} > x_{kj}, \\ 0, & \text{otherwise}, \end{cases} \\ \tilde{d}_{kij}^{-}(x) &= \begin{cases} -(x_{ki} - x_{kj}), & x_{ki} < x_{kj}, \\ 0, & \text{otherwise}. \end{cases} \end{aligned} \tag{2}$$

Let us build a set $\tilde{\mathbb{Y}} = \left\{ \begin{pmatrix} x \\ \tilde{d}^{\pm}(x) \end{pmatrix} \in \mathbb{R}^{mn^2} : x \in \mathbb{X} \right\}$. Note that elements of every vector in $\tilde{\mathbb{Y}}$ satisfy the following conditions:

$$\begin{gathered} \textstyle\sum_{l=1}^{n} x_{kl} = 0,\ \tilde{d}_{kij}^{+}(x) - \tilde{d}_{kij}^{-}(x) = x_{ki} - x_{kj} \text{ and} \\ \tilde{d}_{kij}^{+}(x)\tilde{d}_{kij}^{-}(x) = 0,\ \tilde{d}_{kij}^{+}(x), \tilde{d}_{kij}^{-}(x) \geq 0, \end{gathered} \tag{3}$$

for all $1 \leq i < j \leq n$, $1 \leq k \leq m$.

Let $d^{\pm} = (d_{112}^{+}, d_{112}^{-}, \ldots, d_{m12}^{+}, d_{m12}^{-}, \ldots, d_{m(n-1)n}^{+}, d_{m(n-1)n}^{-})^T$ denotes a non-negative vector from the space $\mathbb{R}_+^{mn(n-1)}$. Let us build another set

$$\mathbb{Y} = \left\{ \begin{pmatrix} x \\ d^{\pm} \end{pmatrix} \in \mathbb{R}^{mn^2} : \begin{array}{l} x \in \mathbb{X}, \\ d_{kij}^{+} - d_{kij}^{-} = x_{ki} - x_{kj}, \\ d_{kij}^{+} d_{kij}^{-} = 0, d_{kij}^{+}, d_{kij}^{-} \geq 0, \\ \text{for all } 1 \leq i < j \leq n,\ 1 \leq k \leq m \end{array} \right\} \tag{4}$$

and let us prove that $\tilde{\mathbb{Y}} = \mathbb{Y}$. If $\tilde{y} = \begin{pmatrix} x \\ \tilde{d}^{\pm}(x) \end{pmatrix} \in \tilde{\mathbb{Y}}$ with some $x \in \mathbb{X}$, then it follows from (3) that $\tilde{y} \in \mathbb{Y}$. Thus $\tilde{\mathbb{Y}} \subset \mathbb{Y}$. On the other hand, if $y = \begin{pmatrix} x \\ d^{\pm} \end{pmatrix} \in \mathbb{Y}$, then $d_{kij}^{+} = x_{ki} - x_{kj} \geq 0$, if $d_{kij}^{-} = 0$, and $d_{kij}^{-} = -(x_{ki} - x_{kj}) \geq 0$, if $d_{kij}^{+} = 0$. Thus, vector $d^{\pm}$ satisfies equalities, defined by (2), and, consequently, $d^{\pm} = \tilde{d}^{\pm}(x)$ with some $x \in \mathbb{X}$. Note that, if $\tilde{d}^{\pm}(x_i) \neq \tilde{d}^{\pm}(x_j)$, then $x_i \neq x_j$ for all $x_i, x_j \in \mathbb{X}$. It means that set $\mathbb{Y}$ is not larger that set $\tilde{\mathbb{Y}}$. Therefore, $y \in \tilde{\mathbb{Y}}$ and $\mathbb{Y} \subset \tilde{\mathbb{Y}}$. Hence, it follows that $\tilde{\mathbb{Y}} = \mathbb{Y}$.

Next, for every element in $\mathbb{Y}$ let us define a function $g : \mathbb{Y} \to \mathbb{R}$ such that

$$g \begin{pmatrix} x \\ d^{\pm} \end{pmatrix} = \sum_{i<j}^{n} \left(\sum_{k=1}^{m} (d_{kij}^{+} + d_{kij}^{-}) - \delta_{ij} \right)^2. \tag{5}$$

It is clear that $g \begin{pmatrix} x \\ d^{\pm} \end{pmatrix} = f_1(x)$ whenever $\begin{pmatrix} x \\ d^{\pm} \end{pmatrix} \in \mathbb{Y}$, because $d_{kij}^{+} + d_{kij}^{-} = |x_{ki} - x_{kj}|$.

Finally, we can state that problem (1) is equivalent to the following optimization problem with quadratic objective function, linear and complementarity constraints:

$$
\begin{array}{ll}
\underset{\substack{x \in \mathbb{R}^{mn},\\ d^{\pm} \in \mathbb{R}^{mn(n-1)}}}{\text{minimize}} & \sum\limits_{i<j}^{n} \left(\sum\limits_{k=1}^{m} (d_{kij}^{+} + d_{kij}^{-}) - \delta_{ij} \right)^2 \\
\text{subject to} & \sum\limits_{i=1}^{n} x_{ki} = 0, \ 1 \leq k \leq m, \\
& x_{ki} - x_{kj} = d_{kij}^{+} - d_{kij}^{-}, \ 1 \leq k \leq m, \ 1 \leq i < j \leq n, \\
& d_{kij}^{+} d_{kij}^{-} = 0, \ 1 \leq k \leq m, \ 1 \leq i < j \leq n, \\
& d_{kij}^{+}, d_{kij}^{-} \geq 0, \ 1 \leq k \leq m, \ 1 \leq i < j \leq n.
\end{array}
\tag{6}
$$

Let us write problem (6) into a matrix form. First of all, suppose that $y = (x_{11}, \ldots, x_{m1}, \ldots, x_{mn}, d_{112}^{+}, d_{112}^{-}, \ldots, d_{m12}^{+}, d_{m12}^{-}, \ldots, d_{m(n-1)n}^{+}, d_{m(n-1)n}^{-})^T \in \mathbb{R}^{mn^2}$ and $step = mn(n-1)/2$. Let $I(L)^l = \begin{pmatrix} L & & \\ & \ddots & \\ & & L \end{pmatrix} \in \mathbb{R}^{l_1 l \times l_2 l}$ denotes a rectangular block diagonal matrix, where $L \in \mathbb{R}^{l_1 \times l_2}$. For the sake of simplicity, let $I^l = I(1)^l \in \{0,1\}^{l \times l}$ denotes an identity matrix and $O^{p \times q} \in \{0\}^{p \times q}$ denotes a zero matrix. Then problem (6) may be rewritten as:

$$
\begin{array}{ll}
\underset{y \in \mathbb{R}^{mn^2}}{\text{minimize}} & \frac{1}{2} y^T A y - b^T y \\
\text{subject to} & c_i^T y = 0, \ i \in \mathbb{E} = \{1, \ldots, m + step\}, \\
& c_i^T y \geq 0, \ i \in \mathbb{I} = \{m + step + 1, \ldots, m + 3step\}, \\
& y_{(mn+i)} y_{(mn+1+i)} = 0, \ i = 1, 3, 5, \ldots, 2step - 1,
\end{array}
\tag{7}
$$

where:

- $A = \begin{pmatrix} O^{mn \times mn} & O^{mn \times 2step} \\ O^{2step \times mn} & I(E)^{n(n-1)/2} \end{pmatrix} \in \{0,1\}^{mn^2 \times mn^2}$ and $E \in \{1\}^{2m \times 2m}$.
- $b = \begin{pmatrix} O^{mn \times 1} \\ \Delta^{12} \\ \vdots \\ \Delta^{(n-1)n} \end{pmatrix} \in \mathbb{R}_{+}^{mn^2}$ and $\Delta^{ij} = \begin{pmatrix} \delta_{ij} \\ \vdots \\ \delta_{ij} \end{pmatrix} \in \mathbb{R}_{++}^{2m}, 1 \leq i < j \leq n.$
- $C = (c_1 \ \ldots \ c_{m+3step}) = \begin{pmatrix} C^1 & C^2 & O^{mn \times 2step} \\ O^{2step \times m} & I \begin{pmatrix} -1 \\ 1 \end{pmatrix}^{step} & I^{2step} \end{pmatrix} \in$ $\{-1, 0, 1\}^{mn^2 \times (m+3step)}$, where:
  - $C^1 = \begin{pmatrix} I^m \\ \vdots \\ I^m \end{pmatrix} \in \{0,1\}^{mn \times m}.$

$$\circ\ C^2 = \begin{pmatrix} I^m & I^m & \dots & I^m & & & & & \\ -I^m & & & & I^m & I^m & \dots & I^m & \\ & -I^m & & & -I^m & & & & \\ & & & & & -I^m & & & \\ & & \ddots & & & & \ddots & & \dots \\ & & & & & & & & I^m \\ & & & -I^m & & & & -I^m & -I^m \end{pmatrix} \in$$
$\{-1,0,1\}^{mn \times step}$.

In order to better understand the structure of matrices A, C and vector b, let us show an example. Suppose that $n = 3$, $m = 2$ and $\Delta = \left(\begin{smallmatrix} 0 & 7 & 12 \\ 7 & 0 & 3 \\ 12 & 3 & 0 \end{smallmatrix}\right)$. In this case:

- $A = \left(\begin{array}{cccccc|cccccccccccc}
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
\hline
0&0&0&0&0&0&\mathbf{1}&\mathbf{1}&\mathbf{1}&\mathbf{1}&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&\mathbf{1}&\mathbf{1}&\mathbf{1}&\mathbf{1}&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&\mathbf{1}&\mathbf{1}&\mathbf{1}&\mathbf{1}&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&\mathbf{1}&\mathbf{1}&\mathbf{1}&\mathbf{1}&0&0&0&0&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&\mathbf{1}&\mathbf{1}&\mathbf{1}&\mathbf{1}&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&\mathbf{1}&\mathbf{1}&\mathbf{1}&\mathbf{1}&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&\mathbf{1}&\mathbf{1}&\mathbf{1}&\mathbf{1}&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&\mathbf{1}&\mathbf{1}&\mathbf{1}&\mathbf{1}&0&0&0&0\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&\mathbf{1}&\mathbf{1}&\mathbf{1}&\mathbf{1}\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&\mathbf{1}&\mathbf{1}&\mathbf{1}&\mathbf{1}\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&\mathbf{1}&\mathbf{1}&\mathbf{1}&\mathbf{1}\\
0&0&0&0&0&0&0&0&0&0&0&0&0&0&\mathbf{1}&\mathbf{1}&\mathbf{1}&\mathbf{1}
\end{array}\right) \in \{0,1\}^{18\times 18}$.
- $b = \left(\begin{array}{cccccc|cccccccccccc} 0&0&0&0&0&0&\mathbf{7}&\mathbf{7}&\mathbf{7}&\mathbf{7}&\mathbf{12}&\mathbf{12}&\mathbf{12}&\mathbf{12}&\mathbf{3}&\mathbf{3}&\mathbf{3}&\mathbf{3} \end{array}\right)^T \in \mathbb{R}^{18}$.
- $C = \left(\begin{array}{cc|cccccc|cccccccccccc}
\mathbf{1}&0&\mathbf{1}&0&\mathbf{1}&0&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&\mathbf{1}&0&\mathbf{1}&0&\mathbf{1}&0&0&0&0&0&0&0&0&0&0&0&0&0&0\\
\mathbf{1}&0&-\mathbf{1}&0&0&0&\mathbf{1}&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&\mathbf{1}&0&-\mathbf{1}&0&0&0&\mathbf{1}&0&0&0&0&0&0&0&0&0&0&0&0\\
\mathbf{1}&0&0&0&-\mathbf{1}&0&-\mathbf{1}&0&0&0&0&0&0&0&0&0&0&0&0&0\\
0&\mathbf{1}&0&0&0&-\mathbf{1}&0&-\mathbf{1}&0&0&0&0&0&0&0&0&0&0&0&0\\
\hline
0&0&-\mathbf{1}&0&0&0&0&0&\mathbf{1}&0&0&0&0&0&0&0&0&0&0&0\\
0&0&\mathbf{1}&0&0&0&0&0&0&\mathbf{1}&0&0&0&0&0&0&0&0&0&0\\
0&0&0&-\mathbf{1}&0&0&0&0&0&0&\mathbf{1}&0&0&0&0&0&0&0&0&0\\
0&0&0&\mathbf{1}&0&0&0&0&0&0&0&\mathbf{1}&0&0&0&0&0&0&0&0\\
0&0&0&0&-\mathbf{1}&0&0&0&0&0&0&0&\mathbf{1}&0&0&0&0&0&0&0\\
0&0&0&0&\mathbf{1}&0&0&0&0&0&0&0&0&\mathbf{1}&0&0&0&0&0&0\\
0&0&0&0&0&-\mathbf{1}&0&0&0&0&0&0&0&0&\mathbf{1}&0&0&0&0&0\\
0&0&0&0&0&\mathbf{1}&0&0&0&0&0&0&0&0&0&\mathbf{1}&0&0&0&0\\
0&0&0&0&0&0&-\mathbf{1}&0&0&0&0&0&0&0&0&0&\mathbf{1}&0&0&0\\
0&0&0&0&0&0&\mathbf{1}&0&0&0&0&0&0&0&0&0&0&\mathbf{1}&0&0\\
0&0&0&0&0&0&0&-\mathbf{1}&0&0&0&0&0&0&0&0&0&0&\mathbf{1}&0\\
0&0&0&0&0&0&0&\mathbf{1}&0&0&0&0&0&0&0&0&0&0&0&\mathbf{1}
\end{array}\right) \in \{-1,0,1\}^{18\times 20}$.

Note that matrix A is singular and $A = ((2m)^{-1/2}A)^T((2m)^{-1/2}A)$. Hence, matrix A is positive semidefinite and the objective function is not-strictly convex quadratic function.

3 Modified Active-Set (MAS)

Suppose that $m, n \in \mathbb{N}$ ($n > 2, m < 4$) and $\Delta = (\delta_{ij}) \in \mathbb{R}^{n\times n}$ ($\delta_{ij} = \delta_{ji} > 0, \delta_{ii} = 0$, $1 \le i < j \le n$) are given quantities. In this section, we present an algorithm to find a local minimizer of g on $\mathbb{Y}$. Objective function g is defined by (5) and feasible set $\mathbb{Y}$ –

by (4). The algorithm is based on the active-set method for linearly constrained convex quadratic programming (Fletcher, 2006). In this method, a finite sequence of feasible points is constructed such that objective values at these points are decreasing. Note that in our case, feasible set $\mathbb{Y}$ is defined by both linear and complementarity constraints. Therefore, we modified the active-set method so that points of the sequence would satisfy both linear and complementarity constraints. Unfortunately, the modification gives a local minimizer of quadratic function g on $\mathbb{Y}$ instead of a global. We called the algorithm as "Modified Active-Set" (MAS). Pseudocode of MAS is presented below (see Algorithm 1). The pseudocode is written by using the matrix form of problem (6). Let us describe the main steps of MAS in detail. The number after word "line" or "lines" indicates a line number in the pseudocode.

Initialization (lines 1–10). An initial feasible point $y^0 \in \mathbb{Y}$ is selected (line 1) and set $\mathbb{W}^0 \subset \{i \in \mathbb{E} \cup \mathbb{I} : c_i^T y^0 = 0\}$ is constructed (lines 2–9). A feasible point $y^0 \in \mathbb{Y}$ may be selected by using various techniques. One of the techniques is presented in the next section. Set $\mathbb{W}^0$ contains indices of all equality constraints (line 2) and indices of some inequality constraints (lines 3–9). Because of $y^0_{(mn+i)} y^0_{(mn+1+i)} = 0$ for all $i \in \{1, 3, 5, \ldots, 2step - 1\}$, it follows that

$$c^T_{(m+step+i)} y^0 = 0 \text{ and/or } c^T_{(m+step+1+i)} y^0 = 0 \tag{8}$$

for all $i \in \{1, 3, 5, \ldots, 2step - 1\}$. If only one of two equalities (8) holds, either index $(m + step + i)$ or index $(m + step + 1 + i)$ is added to set $\mathbb{W}^0$ (lines 4–6, 9). If both of two equalities (8) hold, randomly selected index is added (lines 7–9). Note that $\mathbb{W}^0$, constructed in this way, defines a set of linearly independent constraints.

Calculations (lines 11–37). A finite sequence of feasible points $\{y^1, y^2, \ldots, y^K = y^*\} \subset \mathbb{Y}$ is constructed such that y^* is a local minimizer of g on $\mathbb{Y}$. Every element of the sequence is defined by the following formula: $y^{k+1} = y^k + \alpha^k p^k$, $k = 0, 1, \ldots, K-1$. Vector $p^k \in \mathbb{R}^{mn^2}$ is called a step-direction and number $\alpha^k \in [0, 1]$ is called a step-length.

Step-direction p^k is a global minimizer of function $\frac{1}{2}(y^k+p)^T A(y^k+p) - b^T(y^k+p)$ on set $\{p \in \mathbb{R}^{mn^2} : c_i^T(y^k+p) = 0,\ i \in \mathbb{W}^k\}$. Note that $\mathbb{W}^k \subset \{i \in \mathbb{E} \cup \mathbb{I} : c_i^T y^k = 0\}$ contains indices of all equality and some inequality constraints which are active at point $y^k \in \mathbb{Y}$. Therefore, step-direction p^k is found by creating and solving the following equality constrained quadratic program:

$$\begin{array}{ll} \underset{p \in \mathbb{R}^{mn^2}}{\text{minimize}} & \frac{1}{2} p^T A p + (Ay^k - b)^T p \\ \text{subject to} & c_i^T p = 0,\ i \in \mathbb{W}^k. \end{array} \tag{9}$$

Matrix A is positive semidefinite and vector $(Az - b)$ belongs to the range of matrix A (column space of A) for all $z \in \mathbb{R}^{mn^2}$. Therefore, the objective function is bounded from bellow on its domain $\mathbb{R}^{mn^2}$. Hence, problem (9) will always have a solution. According to the first-order necessary optimality conditions (KKT conditions), if p^k is a solution of problem (9), then there is a Lagrange multiplier vector $\lambda^k \in \mathbb{R}^{|\mathbb{W}^k|}$ such that the following system (KKT system) is satisfied:

$$\begin{pmatrix} -A & c_{\mathbb{W}^k} \\ c^T_{\mathbb{W}^k} & 0 \end{pmatrix} \begin{pmatrix} p^k \\ \lambda^k \end{pmatrix} = \begin{pmatrix} Ay^k - b \\ 0 \end{pmatrix}. \tag{10}$$

Algorithm 1 Modified Active-Set (MAS)

Input: $m, n \in \mathbb{N}$ ($n > 2, m < 4$), $\Delta = (\delta_{ij}) \in \mathbb{R}^{n \times n}$ ($\delta_{ii} = 0, \delta_{ij} = \delta_{ji} > 0, 1 \le i, j \le n$)
Output: $y^* = \mathrm{argmin}_{\mathrm{local}}\{\frac{1}{2}y^T Ay - b^T y : c_{\mathbb{E}}^T y = 0, c_{\mathbb{I}}^T y \ge 0$ and $y_{(mn+i)}y_{(mn+1+i)} = 0,$ $i = 1, 3, 5, \ldots, 2step - 1\}$

1: $y^0 \leftarrow$ any feasible point
2: $\mathbb{W}^0 \leftarrow \mathbb{E}$
3: **for all** $i \in \{1, 3, \ldots, 2step - 1\}$ **do**
4: $\quad j \leftarrow 0$
5: $\quad$ **if** $y^0_{(mn+1+i)} = 0$ **then**
6: $\quad\quad j \leftarrow 1$
7: $\quad\quad$ **if** $y^0_{(mn+i)} = 0$ **then**
8: $\quad\quad\quad j \leftarrow$ random number from set $\{0, 1\}$
9: $\quad \mathbb{W}^0 \leftarrow \mathbb{W}^0 \cup \{m + step + i + j\}$
10: $stop \leftarrow 0$; $k \leftarrow 0$
11: **while** $stop \ne 1$ **do**
12: $\quad \begin{pmatrix} p^k \\ \lambda^k \end{pmatrix} \in \left\{ \begin{pmatrix} p \\ \lambda \end{pmatrix} \in \mathbb{R}^{mn^2+|\mathbb{W}^k|} : \begin{pmatrix} -A & c_{\mathbb{W}^k} \\ c_{\mathbb{W}^k}^T & 0 \end{pmatrix} \begin{pmatrix} p \\ \lambda \end{pmatrix} = \begin{pmatrix} Ay^k - b \\ 0 \end{pmatrix} \right\}$
13: $\quad$ **if** $p^k \ne 0$ **then**
14: $\quad\quad \tilde{\mathbb{I}} \leftarrow \{i \in \mathbb{I} : i \notin \mathbb{W}^k$ and $c_i^T p^k < 0\}$
15: $\quad\quad \alpha^k \leftarrow \min\{1, \min\{-(c_i^T y^k)/(c_i^T p^k) : i \in \tilde{\mathbb{I}}\}\}$
16: $\quad\quad$ **if** $\alpha^k = -(c_{i^*}^T y^k)/(c_{i^*}^T p^k) \le 1$ with some $i^* \in \tilde{\mathbb{I}}$ **then**
17: $\quad\quad\quad \mathbb{W}^{k+1} \leftarrow \mathbb{W}^k \cup \{i^*\}$
18: $\quad\quad y^{k+1} \leftarrow y^k + \alpha^k p^k$
19: $\quad$ **else**
20: $\quad\quad \lambda \leftarrow 0 \in \mathbb{R}^{m+3step}$; $j \leftarrow 0$
21: $\quad\quad$ **for all** $i \in \mathbb{W}^k$ **do**
22: $\quad\quad\quad j \leftarrow j + 1$; $\lambda_i \leftarrow \lambda_j^k$
23: $\quad\quad \tilde{\mathbb{I}} \leftarrow \{i \in \mathbb{W}^k \cap \mathbb{I} : \lambda_i < 0\}$
24: $\quad\quad next \leftarrow 1$
25: $\quad\quad$ **while** $next = 1$ **do**
26: $\quad\quad\quad$ **if** $\tilde{\mathbb{I}} \ne \emptyset$ **then**
27: $\quad\quad\quad\quad i^* \leftarrow \mathrm{argmin}\{\lambda_i : i \in \tilde{\mathbb{I}}\}$
28: $\quad\quad\quad\quad j \leftarrow 1$
29: $\quad\quad\quad\quad$ **if** $(i^* - (m + step))\%2 = 0$ **then**
30: $\quad\quad\quad\quad\quad j \leftarrow -1$
31: $\quad\quad\quad\quad$ **if** $(i^* + j) \in \mathbb{W}^k$ **then**
32: $\quad\quad\quad\quad\quad \mathbb{W}^{k+1} \leftarrow \mathbb{W}^k \setminus \{i^*\}$; $next \leftarrow 0$
33: $\quad\quad\quad\quad$ **else**
34: $\quad\quad\quad\quad\quad \tilde{\mathbb{I}} \leftarrow \tilde{\mathbb{I}} \setminus \{i^*\}$
35: $\quad\quad\quad$ **else**
36: $\quad\quad\quad\quad y^* \leftarrow y^k$; $next \leftarrow 0$; $stop \leftarrow 1$
37: $\quad k \leftarrow k + 1$

Because of KKT system (10) has at least one solution, we choose any of them (line 12).

Suppose that step-direction p^k is not equal to zero (line 13). In this case, a new feasible point $y^{k+1} = y^k + \alpha^k p^k$ is constructed along this direction. Objective value at the new point is less than or equal to the value at point y^k (Nocedal and Wright, 2006). Step-length α^k is defined by the following formula:

$$\alpha^k = \min\{1, \min\{-(c_i^T y^k)/(c_i^T p^k) : i \notin \mathbb{W}^k \text{ and } c_i^T p^k < 0\}\}.$$

The definition of the step-length ensures that point y^{k+1} belongs to feasible set $\mathbb{Y}$. If the new point touches an inequality constraint that is not active at point y^k (index of that constraint does not belong to set $\mathbb{W}^k$), index of that constraint is added to set $\mathbb{W}^k$ (lines 16–17).

Next, suppose that step-direction p^k is equal to zero (line 19). In this case, we verify if point y^k is a solution of problem (6) or not. It is done by checking the KKT dual feasibility condition. In other words, on set $\mathbb{W}^k$ we select indices of those inequality constraints which have negative Lagrange multipliers (lines 20–23). Let $\tilde{\mathbb{I}}$ denotes a set of these selected indices. If set $\tilde{\mathbb{I}}$ is empty, then Lagrange multipliers, corresponding to inequality constraints, active at y^k, are equal to zero or positive. Hence, point y^k is a solution of problem (6) and calculations are stopped. If $\tilde{\mathbb{I}}$ is not empty, set $\mathbb{W}^k$ contains at least one index of inequality constraint with negative Lagrange multiplier. In this case, a regular active-set method from set $\mathbb{W}^k$ removes an index of inequality constraint with the most negative Lagrange multiplier (lines 27, 32). Note that we are considering an optimization problem with a number of complementarity constraints $y_{(mn+i)}y_{(mn+1+i)} = 0$, $i = 1, 3, \ldots, 2step - 1$. Each complementarity constraint i corresponds to one (or both) of these equalities:

$$c_{(m+step+i)}^T y = 0,\ c_{(m+step+1+i)}^T y = 0.$$

Thus, one (or both) of indices $(m + step + i)$ or $(m + step + 1 + i)$ belongs to set $\mathbb{W}^k$, because $y^k \in \mathbb{Y}$. If both of these indices would be removed from set $\mathbb{W}^k$, at the next iteration it might be the following situation (but not necessarily): $p_{(mn+i)}^{k+1} \neq 0$, $p_{(mn+1+i)}^{k+1} \neq 0$ and $\alpha^{k+1} > 0$. It is clear that then $y_{(mn+i)}^{k+1} y_{(mn+1+i)}^{k+1} \neq 0$ and $y^{k+1} \notin \mathbb{Y}$. If after some number of iterations one of the above indices will appear on set $\mathbb{W}^l$ with some $l > k$, there is no guarantee that complementarity constraint i will be satisfied again. Hence, in order to preserve all complementarity constraints at every iteration, we forbid the removal of an index that does not have a pair on set $\mathbb{W}^k$ (lines 28–34). If the removal of index $(m + step + i)$ or $(m + step + 1 + i)$ is forbidden, the index is eliminated from set $\tilde{\mathbb{I}}$ (line 34). Then the analysis of $\tilde{\mathbb{I}}$ is continued until $\tilde{\mathbb{I}}$ becomes empty (a local solution is found) or the removal is allowed (lines 31–32).

4 Numerical investigation of MAS

We implemented algorithm "Modified Active-Set" (MAS) and performed a numerical investigation of the algorithm. In this section, we give a few details on the implementation of MAS. Namely, we describe how 1) the selection of an initial feasible point $y^0 \in \mathbb{Y}$ and 2) the solution of KKT systems (10) were implemented. The main purpose of the numerical investigation was to evaluate relative errors in different solutions of problem (1). If $x^* \in \mathbb{X}$ is a solution of problem (10), then the relative error in x^* we denote by symbol E^* and define it by the following formula:

$$E^* = E(x^*) = \sqrt{\frac{f_1(x^*)}{\sum_{i<j}^{n} \delta_{ij}^2}}.$$

The relative error evaluates the quality of a solution. The closer E^* is to zero, the better solution (image) is found. Results of the investigation are compared with corresponding results, obtained by using algorithm SMOOTH (Groenen et al., 1999).

In applying algorithm MAS, an initial feasible point $y^0 \in \mathbb{Y}$ has to be selected. The selection was implemented by using an algorithm, presented in Algorithm 2. In this algorithm, a mn-vector of uniformly distributed random numbers between u and v ($u, v \in \mathbb{R}$) is picked (lines 1–3 of Algorithm 2). Then, based on this mn-vector, point $y^0 \in \mathbb{R}^{mn^2}$ is constructed such that the following equalities and inequalities are satisfied:

$$\sum_{i=1}^{n} y^0_{(k+(i-1)m)} = 0 \tag{11a}$$

$$y^0_{(k+(i-1)m)} - y^0_{(k+(j-1)m)} = y^0_{l(k,i,j)} - y^0_{(l(k,i,j)+1)} \tag{11b}$$

$$y^0_{l(k,i,j)} y^0_{(l(k,i,j)+1)} = 0 \tag{11c}$$

$$y^0_{l(k,i,j)}, y^0_{(l(k,i,j)+1)} \geq 0 \tag{11d}$$

where $l(k,i,j) = mn+2k-1+2m(j-i+(2n-i)(i-1)/2-1)$ and $1 \leq i < j \leq n$, $1 \leq k \leq m$. Steps of Algorithm 2, defined in lines 4–7, ensures that point y^0 satisfies

Algorithm 2 A point selection on set $\mathbb{Y}$

Input: $m, n \in \mathbb{N}$ ($n > 2, m < 4$)
Output: $y \in \mathbb{Y}$

```
 1: z ← 0 ∈ R^{mn}
 2: for i = 1, mn do
 3:     z_i ← random number uniformly distributed over the interval [u, v] ⊂ R
 4: for k = 1, m do
 5:     s ← (1/n) Σ_{j=1}^{n} z_(k+(j−1)m)
 6:     for i = 1, n do
 7:         y_(k+(i−1)m) ← z_(k+(i−1)m) − s
 8: l ← mn + 1
 9: for i = 1, (n − 1) do
10:     for j = (i + 1), n do
11:         for k = 1, m do
12:             s ← y_(k+(i−1)m) − y_(k+(j−1)m)
13:             if s < 0 then
14:                 y_l ← 0; y_(l+1) ← −s
15:             else
16:                 y_l ← s; y_(l+1) ← 0
17:             l ← l + 2
```

equalities (11a). Steps, defined in lines 8–17, ensures that point y^0 satisfies equalities and inequalities (11b)–(11d). Thus, in this way constructed point y^0 belongs to feasible set $\mathbb{Y}$.

The solution of every KKT system (10) was implemented by using the null-space method (Fletcher and Johnson, 1997). In applying the method, first of all, we have to select a basis matrix of the null-space of $c_{\mathbb{W}^k}$. In our case, the basis matrix was selected by using a QR decomposition of $c_{\mathbb{W}^k}$: $c_{\mathbb{W}^k} = QR$, where $Q \in \mathbb{R}^{mn^2 \times mn^2}$, $R \in \mathbb{R}^{mn^2 \times |\mathbb{W}^k|}$. It is not hard to check that, if $Q = (Q_1\, Q_2)$, where $Q_1 \in \mathbb{R}^{mn^2 \times |\mathbb{W}^k|}$, $Q_2 \in \mathbb{R}^{mn^2 \times (mn^2 - |\mathbb{W}^k|)}$, then $Q_2^T c_{\mathbb{W}^k} = 0$, i.e., matrix Q_2 is a basis matrix of the null-space of $c_{\mathbb{W}^k}$. If set $\mathbb{W}^k$ was updated by adding or removing an index, a basis matrix of the null-space of $c_{\mathbb{W}^{k+1}}$ was obtained from factors Q and R by using a particular updating technique (Hammarling and Lucas, 2008). When the basis matrix is selected, a reduced KKT system (to find a step-direction p^k) and another system of linear equations (to find a vector of Lagrange multipliers λ^k) are constructed. The reduced KKT system was solved by using a LAPACK (version 3.5.0) subroutine for linear least squares problems. Another system was solved by using a LAPACK subroutine for regular systems of linear equations.

Results of the numerical investigation of MAS are presented in Table 1. Numerical experiments were conducted on a computer with Intel(R) Core(TM)2 Duo processor, running at 2.40 GHz. Algorithm MAS was implemented by using Fortran 95 and compiled with gfortran (version 4.8.2) on XUbuntu 14.04 (64-bit, kernel version 3.13.0-24-generic). The investigation was performed as follows:

1. Data sets were selected from Žilinskas (2006, 2007). Let us remember that a dissimilarity matrix is one of the most important input data for any algorithm for multidimensional scaling. We selected a set of dissimilarity matrices, obtained by measuring relationships between certain objects in geometry (cube, regs, simp) and in pharmacology (hwa12, hwa21, ruusk, uhlen). In addition, a well known dissimilarity matrix, obtained by measuring dissimilarities between some soft drinks, was selected too (cola).
2. For every selected data set, a set of 30 values of E^* was generated by using algorithm SMOOTH. Average time, required to find 1 value of E^*, is presented in column $t(s)$ of Table 1. In order to find 1 value of E^*, SMOOTH generated a set of 10 solutions of problem (1) (column #x^*). Relative error E^* was calculated in the solution with the smallest objective value.
3. By using the same data, another set of 30 values of E^* was generated by using algorithm MAS. In this case, to find 1 value of E^*, the average time $t(s)$, received by using SMOOTH, was used, i.e., MAS had been generating solutions of problem (1), until the time limit was reached. Again, relative error E^* was calculated in the solution with the smallest objective value.

The smallest and the biggest values of the sets of E^* are presented in corresponding columns of Table 1. The mean and the standard deviation (std) are presented too. Note that in almost all cases, MAS generated more solutions of problem (1) than SMOOTH during the same time. However, the mean and standard deviation of relative errors are much bigger by using MAS. It means that MAS is very sensitive to the choice of an initial feasible point.

Table 1: Results of numerical investigation of algorithm MAS and corresponding results, obtained by using algorithm SMOOTH.

Data set	**m**	**n**	**min** E^*	**max** E^*	**mean** E^*	**std** E^*	$t(s)$	**#**x^*	**Algorithm**
cola	1	10	0.3645	0.3645	0.3645	0.0000	0.318	10	SMOOTH
			0.3656	0.3959	0.3762	0.0075	0.319	414	MAS
cola	2	10	0.1679	0.1694	0.1694	0.0003	1.692	10	SMOOTH
			0.1729	0.2089	0.1837	0.0078	1.702	100	MAS
cube	3	4	0.0001	0.0012	0.0009	0.0002	0.519	10	SMOOTH
			0.0000	0.0000	0.0000	0.0000	0.519	975	MAS
cube	3	8	0.0012	0.0013	0.0013	0.0000	1.449	10	SMOOTH
			0.0000	0.0000	0.0000	0.0000	1.463	62	MAS
hwa12	1	9	0.0109	0.0109	0.0109	0.0000	0.074	10	SMOOTH
			0.0107	0.0107	0.0107	0.0000	0.075	55	MAS
hwa12	2	9	0.0108	0.0110	0.0110	0.0001	0.655	10	SMOOTH
			0.0000	0.0027	0.0001	0.0005	0.713	40	MAS
hwa21	1	12	0.1790	0.1790	0.1790	0.0000	0.211	10	SMOOTH
			0.1790	0.1871	0.1821	0.0023	0.214	48	MAS
regs	1	13	0.5311	0.5311	0.5311	0.0000	0.857	10	SMOOTH
			0.5311	0.5311	0.5311	0.0000	0.858	379	MAS
regs	2	9	0.2991	0.2991	0.2991	0.0000	1.382	10	SMOOTH
			0.2991	0.3031	0.2996	0.0011	1.387	184	MAS
regs	3	7	0.0945	0.0945	0.0945	0.0000	1.566	10	SMOOTH
			0.0945	0.0945	0.0945	0.0000	1.571	134	MAS
ruusk	1	8	0.2975	0.2975	0.2975	0.0000	0.201	10	SMOOTH
			0.2975	0.3292	0.3112	0.0087	0.201	480	MAS
ruusk	2	8	0.1096	0.1096	0.1096	0.0000	1.130	10	SMOOTH
			0.1097	0.1306	0.1198	0.0055	1.133	172	MAS
ruusk	2	20	0.0524	0.0555	0.0546	0.0010	4.074	10	SMOOTH
			0.0523	0.1322	0.0850	0.0232	4.962	3	MAS
ruusk	3	8	0.0189	0.0254	0.0214	0.0018	2.386	10	SMOOTH
			0.0188	0.0411	0.0289	0.0066	2.402	86	MAS
simp	1	13	0.5279	0.5281	0.5279	0.0000	1.203	10	SMOOTH
			0.5279	0.5279	0.5279	0.0000	1.205	314	MAS
simp	2	9	0.2759	0.2759	0.2759	0.0000	0.972	10	SMOOTH
			0.2759	0.2808	0.2760	0.0009	0.977	88	MAS
simp	3	7	0.0015	0.0016	0.0016	0.0000	1.673	10	SMOOTH
			0.0000	0.0000	0.0000	0.0000	1.679	150	MAS
uhlen	1	12	0.2112	0.2112	0.2112	0.0000	0.199	10	SMOOTH
			0.2112	0.2251	0.2151	0.0036	0.201	52	MAS
uhlen	2	12	0.0825	0.0909	0.0874	0.0023	2.407	10	SMOOTH
			0.0840	0.1248	0.1033	0.0105	2.440	37	MAS

5 Conclusions

In this paper, we considered the minimization of a least-squares stress function with city-block distances. We reformulated the problem into an optimization problem with convex quadratic objective function, linear and complementarity constraints. Moreover, we presented an algorithm to solve the reformulated problem by using the active-set method with some modifications. This algorithm allowed us to find a local minimizer of the stress function with city-block distances.

Acknowledgements

This research was funded by a grant (No. MIP-063/2012) from the Research Council of Lithuania.

References

Borg I., Groenen P. J. F. (2005). *Modern Multidimensional Scaling: Theory and Applications*. Second Edition. Springer, New York, NY.

Brusco M. J. (2001) A simulated annealing heuristic for unidimensional and multidimensional (city-block) scaling of symmetric proximity matrices. *Journal of Classification*. 18(1), 3-33.

Dzemyda G., Kurasova O., Žilinskas J. (2013). *Multidimensional Data Visualization: Methods and Applications*. Springer, New York, NY.

Fletcher R., Johnson T. (1997). On the stability of null-space methods for KKT systems. *SIAM Journal on Matrix Analysis and Applications*. 18(4), 938-958.

Fletcher, R. (2006). *Practical Methods of Optimization*. Second Edition. Wiley.

Groenen, P. J. F., Heiser W. J., Meulman, J. J. (1999). Global optimization in least-squares multidimensional scaling by distance smoothing. *Journal of Classification*. 16, 225-254.

Hammarling S., Lucas C. (2008) Updating the QR factorization and the least squares problem. The University of Manchester, MIMS EPrint:2008.111.

Hubert L., Arabie P., Hesson-Mcinnis M. (1992) Multidimensional scaling in the city-block metric: a combinatorial approach. *Journal of Classification*. 9(2), 211-236.

Leung P. L., Lau K. (2004) Estimating the city-block two-dimensional scaling model with simulated annealing. *European Journal of Operational Research*. 158(2), 518-524.

Nocedal, J., Wright, S. J. (2006). *Numerical Optimization*. Second Edition. Springer, New York, NY.

Torgerson W. S. (1952) Multidimensional scaling: I. Theory and method. *Psychometrika*. 17(4), 401-419.

Žilinskas J. (2006) Multidimensional scaling in protein and pharmacological sciences. *Computer Aided Methods in Optimal Design and Operations, Series on Computers and Operations Research*. 7, 139-148.

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

Žilinskas A., Žilinskas J. (2007) Two level minimization in multidimensional scaling. *Journal of Global Optimization*. 38, 581-596.

Žilinskas A., Žilinskas J. (2009) Branch and bound algorithm for multidimensional scaling with city-block metric. *Journal of Global Optimization*. 43(2-3), 357-372.

Živadinović N. K. (2011) Multidimensional Scaling: An Introduction. *International Encyclopedia of Statistical Science*. 878-879.

Received September 11, 2014 , revised October 21, 2014, accepted October 22, 2014