DOI: 10.15514/ISPRAS-2022-34(4)-14

Theoretical Foundations of an Algorithm of Visualization of a Set of Points of a Multidimensional Space for Use in Anthropotechnical Decision Support Systems

I.A. Alexandrov, ORCID: 0000-0003-1818-5763 <alexandrov@ikti.ru>
V.Zh. Kuklin, ORCID: 0000-0003-2332-895X <vzh.kuklin@gmail.com>
A.N. Muranov, ORCID: 0000-0003-1436-2160 <muranov@ikti.ru>
A.A. Tatarkanov, ORCID: 0000-0001-7334-6318 <tatarkanov@ikti.ru>

Institute of Design-Technology Information Science RAS,
127055 build. 1a., 18, Vadkovsky av., Moscow, Russia

Abstract. Active usage of data collections by experts and decision makers tasked with preparing decision alternatives is an essential characteristic of effectiveness of an anthropotechnic system. In many cases such data analysis may require a standalone visual analysis that implies projection of a multidimensional array of data onto a lower-dimensional space. The article below presents the results of developing the theoretical foundation of such algorithm that is oriented towards an interactive analysis procedure.

Keywords: anthropotechnic systems; decision support system; multidimensional data visualization; methods of interactive information analysis

For citation: Alexandrov I.A., Kuklin V.Zh., Muranov A.N., Tatarkanov A.A. Theoretical Foundations of an Algorithm of Visualization of a Set of Points of a Multidimensional Space for Use in Anthropotechnical Decision Support Systems. Trudy ISP RAN/Proc. ISP RAS, vol. 34, issue 4, 2022. pp. 201-210 (in Russian). DOI: 10.15514/ISPRAS-2022-34(4)-14

Acknowledgements. The results presented are a part of a larger work being completed in accordance with Subsidy Agreement No. 075-11-2022-029 as of April 8, 2022 on the subject: "Creation of a digital cluster for data storage and processing with consequent formation of dynamic stream using AI technologies." with Ministry of Science and Higher Education of the Russian Federation.

Теоретические основы алгоритма визуализации множества точек многомерного пространства для использования в антропотехнических системах поддержки принятия решений

И.А. Александров, ORCID: 0000-0003-1818-5763 <alexandrov@ikti.ru>
В.Ж. Куклин, ORCID: 0000-0003-2332-895X <vzh.kuklin@gmail.com>
А.Н. Муранов, ORCID: 0000-0003-1436-2160 <muranov@ikti.ru>
А.А. Татарканов, ORCID: 0000-0001-7334-6318 <tatarkanov@ikti.ru>

Институт конструкторско-технологической информатики РАН,
127055, Россия, г. Москва, пер. Вадковский, д. 18, стр. 1а

Аннотация. Существенным фактором эффективного функционирования антропотехнической системы является активное использование при подготовке управленческих решений информационных массивов с необходимыми для подготовки вариантов решений лицами, принимающими решения, и привлеченными экспертами. Во многих случаях для анализа подобных данных, наряду с

использованием автоматизированных методов, оказывается необходимым проводить визуальный анализ, используя методы проектирования многомерных данных в пространство визуализации малой размерности. В работе представлены результаты разработки теоретического обоснования подобного алгоритма, ориентированного на интерактивную процедуру анализа.

Ключевые слова: антропотехнические системы; система поддержки принятия решений; многомерная визуализация данных; методы интерактивного анализа информации.

Для цитирования: Александров И.А., Куклин В.Ж., Муранов А.Н., Татарканов А.А. Теоретические основы алгоритма визуализации множества точек многомерного пространства для использования в антропотехнических системах поддержки принятия решений. Труды ИСП РАН, том 34, вып. 4, 2022 г., стр. 201-210. DOI: 10.15514/ISPRAS-2022-34(4)-14

Благодарности: Представленные результаты являются частью работы, выполняемой в соответствии с Соглашением о субсидии № 075-11-2022-029 от 08.04.2022 по теме: «Создание цифрового кластера хранения и обработки данных с последующим формированием динамического потока с использованием технологий искусственного интеллекта». с Министерством науки и высшего образования Российской Федерации.

1. Introduction

A significant factor of efficient functioning of an antropothecnical system is active usage of bulk data arrays that are utilised for preparation of candidate solutions by the involved decision making persons and experts. In the majority of cases of data analysis, in addition to automated methods, it is necessary to conduct visual data analysis by visualising the multivariable dataset within a space of lesser dimension (e.g. 3 or 4). The paper presents theoretical foundations required to develop such an algorithm for interactive data analysis.

The majority of the data visualisation procedures known nowadays have a disadvantage of being a finite image which corresponds to the minimum of error, while the exact map is unknown. This leads to certain difficulties when a need to update the source data set arises which in turn leads to the need to repeat the procedure from scratch.

A careful analysis shows that if we sacrifice some consistency of the transformation that would give us certain advantages, i.e. a visualisation mode which would allow a user to intervene in an interactive way as well as an explicit optimal projection definition (formula).

The paper considers a task of mapping an initial data set from an n-dimensional space into a (lesser) m-dimensional space (m = 3 or 4) – the visualisation (display) space with locally minimal distortion of mutual location of points in the original (data source) space.

2. Subjects, methods and results of the research paper

2.1 Source data and problem definition

Let's introduce important definitions and denotations.

Let R^k − vector space of dimension *k*, R^n – initial dataspace, R^m − display space ($m < n$), *P* – ($n \times n$) − projection matrix, $P: R^n \rightarrow R^m$, $R^{\perp}$ – (n—m) dimensional subspace R^n, that is an orthogonal complement of R^m, $X = \{x^i\}_{i=1}^N$ − initial ensemble of points in R^n, $\Sigma(R^k)$ – a group of orthogonal ($k \times k$) − matrices in R^k.

Let's define a configuration set *X* as a set of vectors *Z*,

$$Z = \{z_j \,|\, z_j = x^i - x^k, i = \overline{1, N-1}, k = i+1, N\}. \quad (1)$$

Let's consider that we met the following condition $i \neq j \Leftrightarrow x^i \neq x^j$ and lets denote a number of elements *Z* through *J*, $card\ Z = J = N(N—1)/2$.

Let's define a distortion criterion of configuration set *X* like

$$E(f)[\mathrm{X}] = \sum_{i=1}^{N-1}\sum_{k=i+1}^{N} \omega_{ik} \left| \left\| x^i - x^k \right\|^2 - \left\| Pf(x^i) - Pf(x^k) \right\|^2 \right|, \quad (2)$$

and class Φ of non-linear functions vector functions, f: $R^n \rightarrow R^n$,

$$\Phi = \{ f \,/\, f = f(x) = Sx + \sum_{\mu=1}^{m} e_\mu \lambda_\mu \left\| P_\perp x \right\| \},$$

where $S \in \Sigma(R^n)$, $\{e_\mu\}_{\mu=1}^{n}$ – orthonormal basis in R^n,

$P_\perp$ – $(n \times n)$ – projection matrix form R^n in $R^\perp$,

$\omega_{ik} = \omega(x^i, x^k)$ – scalar nonnegative function, defined at R^n.

Let us recall, that for matrix P the following conditions are met

A task of visualization is to define the map $f \in \Phi$, for which

$$E(f)[\mathrm{X}] = \min_{f \in \Phi} E(f)[\mathrm{X}]. \quad (3)$$

Let us note that outlined criterion is generalization of a known continuity criterion of Shepherd ([2–4]).

2.2 Visualization problem in the linear case

For the linear visualisation the task is defined as follow: find a matrix $S \in \Sigma(R^n)$ that satisfies:

$$E(S)[X] = \min_{S \in \Sigma(R^*)} E(S)[X].$$

Under the assumption that the following theorem (theorem 1) is correct.

Theorem 1. Let $X = \{x^k\}_{k=1}^{N} \subseteq R^n$,

$$E(S)[X] = \sum_{i=1}^{N-1}\sum_{k=i+1}^{N} \omega_{ik} \left| \left\| x^i - x^k \right\|^2 - \left\| PS(x^i - x^k) \right\|^2 \right|, \quad (4)$$

where $S \in \Sigma(R^n)$, $\omega^{jk} = \omega(x^i, x^k)$ meets a condition

$\forall S \in \Sigma(R^n)\ \omega_{ik}\ (S_{x^i}, S_{x^k}) = \omega(x^i, x^k)$.

Then a matrix $\hat{S} \in \Sigma(R^n)$, meets a condition $E(\hat{S})[\mathrm{X}] = \min_{S \in \Sigma(R^n)} E(S)[\mathrm{X}]$, is defined in the following way:

$$\hat{S} = \lim_{k \to \infty} S_k, \quad S_0 = I, \quad \forall k \geq 0 \quad S_{k+1} = R_k S_k, \quad (5)$$

where $R_k = R_k(i, j, \varphi)$ is Jacobi rotation method, $i \leq m, j \leq m+1$,

$$\varphi = \frac{1}{2} arctg \frac{2a_{ij}^{(k)}}{a_{ii}^{(k)} - a_{jj}^{(k)}}, |\varphi| \leq \frac{\pi}{4},$$

$$a_{\mu\nu}^{(k)} = \Sigma_{p<}^{N-1} \Sigma_q^N \omega_{pq} (x_\mu^{p(k)} - x_\mu^{q(k)})(x_\nu^{p(k)} - x_\nu^{q(k)}),$$

$$x^{p(k)} = (x_1^{p(k)}, x_2^{p(k)}, \ldots, x_n^{p(k)}) = S_k x^p, \ p = 1, N$$

at that the matrix $\hat{S}$ is defined with accuracy up to arbitrary left multiplier represented in the form of:

$$\hat{S} = \begin{pmatrix} S_m & O \\ O^{\mathrm{T}} & S_\perp \end{pmatrix}, \quad (6)$$

where O – $(m \times (n—m))$ – zero matrix, $S_m \in \Sigma(R^m)$, $S_\perp \in \Sigma(R^\perp)$.

Theorem 1 is a foundation of a multidimensional data set visualisation computation procedure with subsequent minimisation of distortion on a set of orthogonal transformations of the source data space.

Practical application of the procedure consists of iterative application of the following steps:

- calculation of a matrix $R(i, j, \varphi)$ subject to the requirement of minimisation of the distortion

(i.e. residual), transformation of matrix A and formulation of a resulting transform - matrix $\hat{S}$;

- transformation of the source data set in R^n and its visualisation (if required) by projecting it onto the visualisation (display) space, output ot the displaying device and making a decision whether a subsequent iteration is required.

The result of the procedure application is a matrix of an optimising transformation. Additionally, we know the value of the distortion criterion. The block representation of matrix A allows us to assess the distance between the achieved value of the distortion criterion and its global minimum. Block diagonalisation of the matrix A is a collateral result of the procedure. Let's emphasise that any method of block diagonalisation of matrix A gives us a solution to the visualisation problem. Hence, in order to solve the linear visualisation problem any diagonalisation procedure conducted according to the aforementioned description will suffice, including, but not limited to, the principal component analysis. Nevertheless, the procedure above possesses the following peculiarities:

- While the optimal visualisation is being constructed it is possible to output it while maintaining the monotonically decreasing of distortion criteria (it is known that existing PCA implementations do not have such a property) [4]. It is due to the fact that the most efficient realizations are based on matrix diagonalization, which are optimal either in terms of memory minimization or in terms of solution time [5]. At that intermediary results as a rule are not available for a researcher);
- Solution procedure does not require obligatory matrix A diagonalisation and is limited only by the condition of block diagonalization, that shortens the time of solution;
- Class of task to which the procedure above can be applied is not limited to the case considered but also allows customisation of the dot product as well as of the norm function $\omega(x^i, x^k)$ depending on the task specifics.

However, in the general case ($n >> m$) [6] a small value of distortion criterion cannot be achieved with the use of linear transformation. This means that the input of traits that are orthogonal to the visualisation space is significant for the distortion criterion. Consequently, we need to resort to non-linear methods in order to reduce the distortion criterion.

2.3 Visualization problem in the nonlinear case

Considering nonlinear mapping of a general form

$$f(x) = f_0 + f_1(x) + f_2(x) + \cdots,$$

where f_0 is a constant vector, which can be equal to zero without losing generality

f_1 – linear,

f_2 – quadric forms defined on the components of the vector x, let's note that using summands, corresponding to forms of higher order – a third one and further results into a significant increment of computational complexity. Let's note that the existing nowadays procedures of visualization also use the methods of minimization not higher that the second order, that is described by the example $f(x) = Ax + f_2(x)$.

For the majority of known procedures of visualization, the drawback is that the result is only a finite image (that is the final result of the procedure), where a criterion minimum was achieved. While the transforming image itself remains unknown, that results into the complexity while trying to refill the initial data set by additional data and get the result of designing without repeated procedure in general.

The analysis shows that giving up the consistency of the transformation to some extent, it is possible to get some advantages in return, that is interactive mode of visual data processing with the intervention mode of a user for him to run the analysis and optimum scale designing image in explicit form.

Let's consider that a linear optimal map for X is the identity mapping. Moreover let's define a class of allowed mappings Φ as follows:

$$\Phi = \{ f \,/\, f = f(x) = x + \sum_{\mu=1}^{n} \lambda_\mu e_\mu \left\| P_\perp x \right\| \}, \qquad (7)$$

where $P_\perp$ is a projection from R^n in $R^\perp$,

λ_μ – unknown coefficients of the mapping, $\mu = \overline{1,n}$,

$\{e_\mu\}_{\mu=1}^{n}$ – orthonormal basis in R^n.

Then a task of visualization will be as following.

Let the set $\mathrm{X} = \{x^i\}_{i=1}^{N} \subseteq R^n$, be defined, then the distortion criterion of the configuration of set X is

$$E(f)[\mathrm{X}] = \Sigma_{i<}^{N-1} \Sigma_k^N \omega_{ik} \left| \left\| x^i - x^k \right\|^2 - \left\| Pf(x^i) - Pf(x^k) \right\|^2 \right|, \qquad (8)$$

and moreover, it is assumed that

$$E(I)[\mathrm{X}] = \min_{S \in \Sigma(R^n)} E(S)[\mathrm{X}],$$

where I – an identity matrix.

To define mapping $\hat{f} \in \Phi$, for which

$$E(\hat{f})[\mathrm{X}] = \min_{f \in \Phi} E(f)[\mathrm{X}]. \qquad (9)$$

Let's transform a formula (8) to a form more convenient for analysis. Omitting the intermediary results, we get

$$E(f)[\mathrm{X}] = \Sigma_{i=1}^{N-1} \Sigma_{k=i+1}^{N} \zeta_{ik} \left| t^2 + 2t(\eta_{ik}, e)^2 - \zeta_{ik} \right|,$$

where $t = (\Sigma_{\mu=1}^{n} \lambda_\mu^2)^{1/2}$

$$e = \frac{1}{t}(\lambda_1, \lambda_2, \ldots, \lambda_m, 0, \ldots 0) \in R^m, \left\| e \right\| = 1$$

$$\xi_{ik} = \omega_{ik} a_{ik}^2 \geq 0,$$

$$\eta_{ik} = P(x^i - x^k) a_{ik}^{-1} \in R^m,$$

$$\zeta_{ik} = \left\| P_\perp (x^i - x^k) \right\| a_{ik}^{-1},$$

$$a_{ik} = \left\| P_\perp x^i \right\| - \left\| P_\perp x^k \right\|.$$

Let's note that ζ_{ik}, η_{ik}, ξ_{ik}, do not depend on t and e, and that is why the problem of criterion $E(f)[\mathrm{X}]$ minimization can be considered as a problem of finding a scalar $\hat{t} \geq 0$ and a unitary vector, for which

$$h(t,e) = \min_{t \geq 0, \|e\|=1} h(t,e), \quad h(t,e) = \Sigma_{j=1}^{J} \alpha_j^2 \left| t^2 + 2t(r_j, e) - \beta_j^2 \right|$$

subject to $\beta_j^2 \geq 1$, $r_j \in R^m$, $e \in R^m$. Let's introduce notations

$$J_+(t) = \{ j \,/\, t^2 + 2t(r_j, e) \geq \beta_j^2 \}, \; J_-(t) = \{ j \,/\, t^2 + 2t(r_j, e) \geq \beta_j^2 \},$$

$$t_j = \sqrt{\beta_j^2 + (r_j, e)^2} - (r_j, e) \geq 0 \qquad (10)$$

and dispose of the module in the formulae for $h(t, e)$,

$$h(t,e) = G_0 t^2 + G_1 t - G_2$$

$$G_0 = \sum_{J_+(t)} \alpha_j^2 - \sum_{J_-(t)} \alpha_j^2$$

$$G_1 = \sum_{J_+(t)} \alpha_j^2 (r_j, e) - \sum_{J_-(t)} \alpha_j^2 (r_j, e)$$

$$G_2 = \sum_{J_+(t)} \alpha_j^2 \beta_j^2 - \sum_{J_-(t)} \alpha_j^2 \beta_j^2$$

Without limitation of generality it is possible to consider that t_j are ordered,

$$0 = t_0 < t_1 \leq t_2 \leq \ldots \leq t_j < t_{j+1} < \ldots$$

Let's consider a function $h''(t_k + 0, e)$. It is obvious that

$$h''(+0, e) < 0, \quad h''(t_j + 0, e) > 0$$

that is why exists k_0, of the kind that

$$\forall t \leq t_{k_0} \;\; h''(t - 0, e) < 0, \;\; \forall t \geq t_{k_0} \;\; h''(t + 0, e) \geq 0. \qquad (11)$$

And as the function $h''(t,e)$ is piecewise constant, then

$$\forall k \leq k_0 \;\; \min_{t \in [t_{k-1}, t_k]} h(t,e) = \min\{ h(t_{k-1}, e), h(t_k, e) \}$$

Then it is easy to get the following

$$h'(t_{k+1} + 0, e) = h'(t_k + 0, e) + (2t_{k+1} - t_k) h''(t_k + 0, e) + \; + 4\alpha_{k+1}^2 \sqrt{\beta_{k+1}^2 + (r_{k+1}, e)^2}$$

from which, considering (10), we get

$$h'(t_{k1} + 0, e) \geq 0, \; h''(t_{k1} + 0, e) \geq 0 \Rightarrow \forall k \geq k_1, \; h'(t_k + 0, e) \geq 0. \qquad (12)$$

Following the previous considerations, it is clear that it is correct that

Lemma 1. Let's

$$h(t,e) = \Sigma_{j=1}^{J} \alpha_j^2 \left| t^2 + 2t(r_j, e) - \beta_j^2 \right|,$$

where $t \geq 0$, $\forall j \;\; r_j \in R^m$, $e \in R^m$, $\left\| e \right\| = 1$, values k_0 and k_1 depend on the conditions

$$k_0 = \min\{ k \,/\, h''(t_{k-1} + 0, e) \leq 0, \quad h''(t_k + 0, e) \geq 0 \},$$

$$k_1 = \min\{ k \,/ \sum_{j=1}^{k-1} \alpha_j^2 \sqrt{\beta_j^2 + (r_j, e)^2} - \sum_{j=k+1}^{J} \alpha_j^2 \sqrt{\beta_j^2 + (r_j, e)^2} \geq 0 \}$$

then

$$\exists \hat{t} = t(e) : \forall t \geq 0 \;\; h(\hat{t}, e) \leq h(t, e), \qquad (13)$$

Such that t ∈ $\{0, t_1, \ldots, t_{k0}\} \cup [t_{k0}, t_{k1}]$, where interval is $[t_{k0}, t_{k1}]$ empty, if $k_0 \geq k_1$.

Lemma 1 gives the definition of a function $h(t, e)$ absolute minimum, where e is fixed. At the same time the procedure of finding the minimum is not iterative. Let's consider a case of arbitrate (varying e).

Let $\lambda = te = (\lambda_1, \lambda_2, \ldots, \lambda_m)$, then

$$h(\lambda) = h(t,e) - \sum_{j=1}^{J} \alpha_j^2 \left| (\lambda, \lambda) + 2(\lambda, r_j) - \beta_j^2 \right|.$$

Let's denote by e_1 the current value of e, and let's $t^{(1)}$ meets (12) at $e = e_1$. Let's consider a problem of λ defining, for which $h(\lambda) < h(\lambda^{(1)})$, where $\lambda^{(1)} = t^{(1)} e_1$. For this we will consider the following two cases – $t^{(1)} \neq t_k$ and $t^{(1)} = t_k$, where k – has a value that meets (9). Let's start with (any) $\forall k \; t^{(1)} \neq t_k$. Then at some neighbourhoods of a point $\lambda^{(1)}$ the function $h(\lambda)$ is twice continuously differentiated, consequently, antigradient v

$$v = -grad\ h(\lambda)\big|_{\lambda=\lambda^{(1)}} = -\left(\frac{\partial h(\lambda^{(1)})}{\partial \lambda_1}, \frac{\partial h(\lambda^{(1)})}{\partial \lambda_2}, \ldots, \frac{\partial h(\lambda^{(1)})}{\partial \lambda_m}\right) \quad (14)$$

Is either equal to zero, then $\lambda^{(1)}$ there is a point of a function local minimum $h(\lambda)$ or defines the directions of maximum decrease and then

$$\exists \theta > 0 : h(\lambda^{(1)} + \theta v) \le h(\lambda^{(1)}). \quad (15)$$

Let then $t^{(1)} = t_k$ for some k. Let's represent $h(\lambda)$ in the form of

$$h(\lambda) = h_0(\lambda) + h_k(\lambda)$$

where

$$h_k(\lambda) = \alpha_k^2 \left|(\lambda, \lambda) + 2(\lambda, r_k) - \beta_k^2\right|,$$

And as $h_0(\lambda)$ in some neighbourhoods of a point $\lambda^{(1)}$ is twice continuously differentiated, let's define (it should be grad in the formulae)

$$v = -grad\ h(\lambda)\big|_{\lambda=\lambda^{(1)}} \quad (16)$$

If $v \ne 0$, let's assume that $v_0 = v\|v\|^{-1}$, then at $\theta > 0$, $\theta \ll 1, \|w\| = 1$, we will get

$$(v_0, w) > 0 \Rightarrow h_0(\lambda^{(1)} + \theta w) - h_0(\lambda^{(1)}) = -\theta(v_0, w) + o(\theta)$$

and as

$$h_k(\lambda^{(1)} + \theta w) - h_k(\lambda^{(1)}) = \alpha_k^2 \mid \theta w, 2(\lambda^{(1)} + r_k) + \theta w) \mid$$

consequently,

$$h(\lambda^{(1)} + \theta w) - h(\lambda^{(1)}) = -\theta[(v_0, w) - \alpha_k^2 \mid w, 2(\lambda^{(1)} + r_k)) \mid] + o(\theta),$$

and if exists w such that $\|w\| = 1$ and

$$(v_0, w) - \alpha_k^2 \mid (w, 2(\lambda^{(1)} + r_k)) > 0 \mid, \quad (17)$$

then

$$\exists \theta > 0 : h(\lambda^{(1)}) > h(\lambda^{(1)} + \theta w)$$

Immediate calculation shows that given:

$$v_0 \ne (\lambda^{(1)} + r_k)\left\|\lambda^{(1)} + r_k\right\|^{-1} \quad (18)$$

Vector w, defined by the formulae

$$w = [\rho^2 v_0 - \sigma(\lambda^{(1)} + r_k)][\rho\sqrt{\rho^2 - \sigma^2}\,]^{-1}, \quad (19)$$

where

$$\rho = \left\|\lambda^{(1)} + r_k\right\|, \sigma = (v_0, \lambda^{(1)} + r_k)$$

meets (17). If (18) is not fulfilled but the condition is correct

$$\alpha_k^2 \left\|\lambda^{(1)} + r_k\right\| < \frac{1}{2},$$

then (17) is correct at $w = v_0$.

Finally we get that $\lambda^{(1)}$ is a point of a function local minimum $h(\lambda)$, if one of the conditions is met:

$$\left.\begin{aligned} \forall k \in \{1,2,\ldots,J\}\ t^{(1)} \ne t_k,\ grad\ h(\lambda^{(1)}) = 0 \\ \exists k : t^{(1)} \ne t_{(k)},\ grad\ h_0(\lambda^{(1)}) = 0 \\ \exists k : t^{(1)} \ne t_{(k)},\ grad\ h_0(\lambda^{(1)}) = \eta(\lambda^{(1)} + r_k) \ne 0, \alpha_k^2\rho \ge \frac{1}{2} \end{aligned}\right\} \quad (20)$$

If none of the conditions are met, then the function $h(\lambda)$ decreases towards one of the vectors defined by one of the formulas − (14), (16) or (19).

Let's denote v as a unit vector of the decrease of function $h(\lambda)$. Let's consider a step choice in this direction. Let's represent $h(\lambda)$ in the form of

$$h(\lambda) = h_0(t^1 + \Theta w) + h_k(\lambda), \quad (21)$$

where

$$k = \max\{j / t^1 \ge t_j, 1 \le j \le J\}. \quad (22)$$

Let's note that the set of admissible values θ, such as

$$\theta > 0, h(\lambda^{(1)} + \theta v) < h(\lambda^{(1)})$$

as nonempty by the definition of v. Now let's define θ_0,

$$\theta_0 = \min\{\theta / \exists j : h_j(\lambda^{(1)} + \theta v) = 0\},$$

Then for $\theta > \theta_0$ representation (21) is nonexistent, therefore, θ_0 is a minimal element of multiple positive roots of equation of the type

$$\theta^2 + 2\theta(\lambda^{(1)} + r_k, v) - a_j = 0, \quad (23)$$

where

$$a_j = (\lambda^{(1)}, \lambda^{(1)}) + 2(\lambda^{(1)}, r_j) - \beta_j^2.$$

As (2) $\forall j \le k\ a_j \ge 0$, $\forall j > k\ a_j < 0$, consequently at $j > k$ equation (23) has a positive root, which is defined by a formulae

$$\theta_j = \sqrt{(v, \lambda^{(1)} + r_j)^2 - a_j} - (v, \lambda^{(1)} + r_j).$$

At $j \le k$ the condition of having a positive root is as following,

$$(v, (\lambda^{(1)} + r_j)) < 0, (v, \lambda^{(1)} + r_j)^2 > a_j,$$

and then

$$\theta = \mid (v, \lambda^{(1)} + r_j) \mid - \sqrt{(\lambda^{(1)} + r_j v)^2 - a_j}.$$

Choosing θ_0 from the condition $\theta_0 = \min_{1\le j\le J} \theta_j$ and supposing that

$$\bar{t} = \left\|\lambda^{(1)} + \theta_0 v\right\|,$$

$$e_2 = \frac{1}{2}(\lambda^{(1)} + \theta_0 v),$$

we will get

$$h(\bar{t}, e_2) < h(t^{(1)}, e_1) = h(\lambda^{(1)}).$$

Applying lemma 1 at $e = e_2$, we will get that

$$\exists t^{(2)} \ge 0 : h(t^{(2)}, e_2) \le h(\bar{t}, e_1) < h(t^{(1)}, e_1).$$

following which it is consequent that

$$h(\lambda^{(2)}) = h(t^{(2)}, e_2) < h(t^{(1)}, e_1) = h(\lambda^{(1)}).$$

From the aforementioned considerations it is consequent that the following theorem is correct

Theorem 2. Let $X = \{x^i\}_{i=1}^N \subseteq R^n$, $f \in \Phi$,

$$E(\lambda) = E(f)[X] = \sum_{i=1}^{N-1} \sum_{k=i+1}^{N} \omega_{ik} \left| \left\| x^i - x^k \right\|^2 - \left\| Pf(x^i) - Pf(x^k) \right\|^2 \right|,$$

and $\lambda^{(0)} = (\lambda_1^0, \lambda_2^0, \ldots, \lambda_m^0)$ − a predefined first approximation, where $\lambda = te = (\lambda_1, \lambda_2, \ldots, \lambda_m)$,

$$h(\lambda) = h(t, e) - \sum_{j=1}^{J} \alpha_j^2 \left| (\lambda, \lambda) + 2(\lambda, r_j) - \beta_j^2 \right|.$$

Then for sequence $\lambda^{(0)}, \lambda^{(1)}, \ldots, \lambda^{(j)}, \ldots$,.

$$\forall j \geq 0 \; E(\lambda^{(j)}) > E(\lambda^{(j+1)}), \lim_{k \to \infty} E(\lambda^{(k)}) > E(\hat{\lambda})$$

which means that for $\hat{\lambda}$ either one of the conditions (20) is met, or the function $h(\lambda)$ decreases towards

$$\nu = -grad \; h(\lambda)\big|_{\lambda=\lambda^{(1)}}.$$

The results above represent a theoretical basis for an algorithm of visualisation and can be used for the development of software that will be introduce an interactive mode of visually-aided data mining in the process of which data mining a graphical output can be displayed to the user allowing the user to assess the quality of the mapping and swiftly adjust the variables as to improve the process of visualisation. Additionally, the explicit definition of the projection can be further used for visual analysis after the source data set is updated without the need for redundant reapplication of the initial procedure of transformation calculation.

It is worth mentioning that software implementation of the algorithm assumes development of a feature that would allow choosing "visually most preferable" (according to the user) mapping that would permit choosing standalone subsets as a basis for subsequent classification and identification of key factors within those subsets.

References / Список литературы

[1] A.N. Kolmogorov, S.V. Fomin. Elements of the theory of functions and functional analysis. M., Nauka, 1981, 623 p. (in Russian) / А.Н. Колмогоров, А.В. Фомин. Элементы теории функций и функционального анализа. 5-е изд. М., Наука, 1981 г., 623 стр.

[2] V.N. Volkova,V.A. Voronkov et al. Theory of systems and methods of system analysis in control and communication. M., Radio and Svyaz, 1983, 248 p. (in Russian) / В.Н. Волкова, В.А. Воронков и др. Теория систем и методы системного анализа в управлении и связи. М., Радио и связь, 1981 г., 248 стр.

[3] A.Yu. Terekhina. Data mining by the methods of multidimentional scaling. M., Nauka, 1986, 168 p. (in Russian) / А.Ю. Терёхина. Анализ методами многомерного шкалирования. М., Наука, 1986 г., 168 стр.

[4] G. Fukunaga. Introduction to Statistical Pattern Recognition. 2nd Edition. Academic Press, 1990, 626 p.

[5] B.N. Parlett. The Symmetric Eigenvalue Problem. Society for Industrial and Applied Mathematics, 1987, 416 p.

[6] M. Kendall. The Advanced Theory of Statistics, Vol. 3: Design and Analysis and Time-Series. 2nd Esitiion. Arnold, 1968, 567 p.

Информация об авторах / Information about authors

Ислам Александрович АЛЕКСАНДРОВ – кандидат технических наук, старший научный сотрудник. Сфера научных интересов: системный анализ, автоматизация, искусственные нейронные сети.

Islam Alexandrovich ALEXANDROV – Candidate of Technical Sciences, Senior Researcher. Research interests: system analysis, automation, artificial neural networks.

Владимир Жанович КУКЛИН – доктор технических наук, доцент, ведущий научный сотрудник. Сфера научных интересов: системный анализ, информационные системы, методы экспертного анализа данных, нейрокомпьютерные технологии.

Vladimir Zhanovich KUKLIN – Doctor of Technical Sciences, Associate Professor, Leading Researcher. Research interests: system analysis, information systems, methods of expert data analysis, neurocomputer technologies.

Александр Николаевич МУРАНОВ – кандидат технических наук, старший научный сотрудник. Сфера научных интересов: оптимизация, статистическая обработка данных, автоматизация.

Alexander Nikolaevich Muranov – Candidate of Technical Sciences, Senior Researcher. Research interests: optimization, statistical data processing, automation.

Аслан Адальбиевич ТАТАРКАНОВ – научный сотрудник. Сфера научных интересов: системный анализ, автоматизация, искусственные нейронные сети.

Aslan Adalbievich TATARKANOV – Researcher. Research interests: system analysis, automation, artificial neural networks.