Integral indicators based on data and rank-scale expert estimations*

Vadim Strijov, Peter Letmathe
strijov@ccas.ru, peter.letmathe@uni-siegen.de
Computing Center of the Russian Academy of Sciences, Russia
Universität Siegen, Germany

Integral indicators play important role in decision making. To make a balanced decision one needs measured data and expert estimations. The expert estimations may contradict the data. Below we investigate a method of integral indicator construction. It uses rank-scaled expert estimations and resolves the possible contradiction between the estimations and the data.

Introduction

To compare objects or alternative decisions one must evaluate a value of quality or a measure of performance for each object. This real-valued scalar is called the integral indicator. Expert estimations of one expert or an expert group could be indicators [1, 2]. Also data could be used to calculate indicators as a convolution of object features [3, 4].

To construct an integral indicator one must perform the following steps. First, select a quality criterium or a comparison criterium for objects. Collect the set of comparable objects. Collect the set of features to describe the objects. Then fulfill the "object-feature" design matrix. We suppose the matrix has no outliers and empty values. The features of the design matrix have a unified linear scale and do not have significant cross-correlation coefficients [5, 6].

Let us state the problem of indicator construction. The design matrix $A = \{a_{ij}\}_{i,j=1}^{m,n},\ A \in \mathbb{R}^{m\times n}$ is given. An element a_{ij} of the matrix is j-th feature measurement for i-th object.

The integral indicator of the object is the linear combination

$$q_i = \sum_{j=1}^{n} w_j g_j(a_{ij}), \qquad (1)$$

where g_j is the normalizing function

$$g_j : a_{ij} \mapsto (-1)^{s_j} \frac{a_{ij} - \min\limits_i a_{ij}}{\max\limits_i a_{ij} - \min\limits_i a_{ij}} + s_j. \qquad (2)$$

This function keeps the principle "the bigger the better". According to this principle an object with bigger value of some feature has the better indicator. The modifier s_j is assigned to one if the optimal value of j-th feature must be minimal and to zero if the optimal value of the feature must be maximal. If the denominator of (2) equals zero for some j, then j-th feature must be withdrawn from the design matrix A.

If the condition (2) holds then the indicator $\mathbf{q} = A\mathbf{w}$, where the indicator $\mathbf{q} = \langle q_1, \dots, q_m\rangle^\mathsf{T}$ and the vector of the weights $\mathbf{w} = \langle w_1, \dots, w_n\rangle^\mathsf{T}$. Let $\mathbb{R}^m \ni \mathbf{q}$ be the space of objects and $\mathbb{R}^n \ni \mathbf{w}$ be the space of features. To construct the integral indicator (1) we must find the weights of given features.

List some obvious methods that use the model (1) and design matrix subject to (2).

1. The indicator q_i is the distance from i-th object to the object with the best features

$$q_i = \frac{1}{m}\left(\sum_{j=1}^{n}\Big(a_{ij} - \max_{\xi=1,\dots,m} a_{\xi j}\Big)^r\right)^{\frac{1}{r}},$$

where the parameter r defines the distance function.

2. The indicator $\mathbf{q}_{\mathrm{PCA}} = A\mathbf{w}_{\mathrm{1PC}}$ is the projection of the row vectors of the matrix A to the first principal component [7, 8], where $\mathbf{w}_{\mathrm{1PC}}$ is the first row of the matrix W defined by the singular values decomposition $A^\mathsf{T} A = W\Lambda^2 W^\mathsf{T}$.

3. The indicator $\mathbf{q}_1 = A\mathbf{w}_0$, is the linear combination of the columns of the matrix A, where $\mathbf{w}_0$ are the linear-scaled expert estimations of weights. The lower index 0 signals that the estimation was given by the expert.

4. The indicator $\mathbf{q}_{\mathrm{ESM}} = A\mathbf{w}_1$, where $\mathbf{q}_0$ are the linear-scaled expert estimations of indicators and

$$\mathbf{w}_1 = \arg\min_{\mathbf{w}\in\mathbb{R}^n} \|A\mathbf{w} - \mathbf{q}_0\|^2.$$

The solution of this optimization problem is $\mathbf{w}_1 = (A^\mathsf{T} A)^{-1} A^\mathsf{T} \mathbf{q}_0$. Denote by $\|\cdot\|$ the Euclidian norm.

Concordance of linear-scaled expert estimations

Let the indicators $\mathbf{q}_1 = A\mathbf{w}_0$ are obtained using the expert estimations of feature weights $\mathbf{w}_0$. Let the feature weights $\mathbf{w}_1 = A^+\mathbf{q}_0$ be obtained using the expert estimations of indicators $\mathbf{q}_0$. Obtain the pseudo-inverse linear operator A^+ using the singular values decomposition [9, 10] of the matrix A,

$$A^+ = W\Lambda U^\mathsf{T}.$$

The linear operator A maps the vector of expert estimations of weights $\mathbf{w}_0$ to the vector $\mathbf{q}_1$. The pseudo-inverse linear operator A^+ maps the vector of the expert estimation of indicators $\mathbf{q}_0$ to the vector $\mathbf{w}_1$. In the general case the estimated and the mapped vectors are different: $\mathbf{q}_1 \neq \mathbf{q}_0$ and $\mathbf{w}_1 \neq \mathbf{w}_0$.

The project is supported by RFBR grant 10-07-00422.

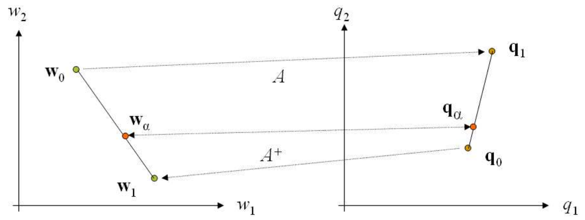

Fig. 1. Concordant vectors $\mathbf{w}_\alpha \in [\mathbf{w}_1, \mathbf{w}_0]$ and $\mathbf{q}_\alpha \in \in [\mathbf{q}_1, \mathbf{q}_0]$.

To resolve this contradiction we suggest to use the concordant estimations, see Fig. 1:

$$\mathbf{w}_\alpha \in [\mathbf{w}_0, \mathbf{w}_1] \text{ and } \mathbf{q}_\alpha \in [\mathbf{q}_1, \mathbf{q}_0]. \tag{3}$$

The vectors $\mathbf{w}_\alpha, \mathbf{q}_\alpha$, given α, are defined by the equation

$$\begin{aligned} \mathbf{w}_\alpha &= \alpha\mathbf{w}_0 + (1-\alpha)A^+\mathbf{q}_0; \\ \mathbf{q}_\alpha &= (1-\alpha)\mathbf{q}_0 + \alpha A\mathbf{w}_0. \end{aligned} \tag{4}$$

This equation brings the concordant expert estimations. They satisfy the condition (6). The structural parameter α defines preference of the expert estimations. When $\alpha \to 0$ one prefers the expert estimations of indicators. When $\alpha \to 1$ one prefers the expert estimations of weights. The structural parameter α could be assigned by the expert himself or could be defied as

$$\hat{\alpha} = \arg\min_{\alpha\in[0,1]} \frac{\|\mathbf{w}_0 - \mathbf{w}_\alpha\|}{n} + \frac{\|\mathbf{q}_0 - \mathbf{q}_\alpha\|}{m}. \tag{5}$$

Withdrawn the condition (3) and find the concordant estimations in the neighborhood of the vectors $\mathbf{w}_0, \mathbf{q}_0$ (see Fig. 2). Solve the optimization problem (5) with the regularizaion parameter γ^2. Rewrite this problem as

$$\mathbf{w}_\gamma = \arg\min_{\mathbf{w}\in\mathbb{R}^n} (\varepsilon^2 - \gamma^2\delta^2),$$

where $\varepsilon^2 = \|\mathbf{w} - \mathbf{w}_0\|^2$ and $\delta^2 = \|A\mathbf{w} - \mathbf{q}_0\|^2$. The concordant expert estimations are defined by the eqsssuation

$$\mathbf{w}_\gamma = (A^\top A + \gamma^2 I_n)^{-1}(A^\top\mathbf{q}_0 + \gamma^2\mathbf{w}_0)$$

and satisfy the condition (6). The parameter γ^2 defines preference of the expert estimations of indicators versus expert estimations of weights like the parameter α.

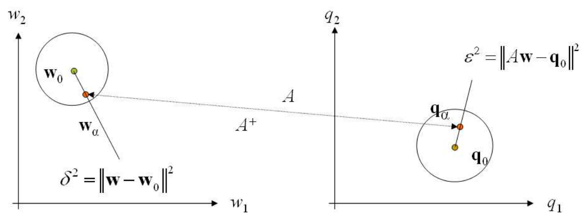

Fig. 2. The concordant vectors $\mathbf{w}_\gamma$ and $\mathbf{q}_\gamma$ are in the neighborhood of the vectors $\mathbf{w}_0$ and $\mathbf{q}_0$.

Concordance for rank-scaled expert estimations

We suppose hereby that an expert marks his estimations in the rank scales. As it was described in the previous section, there are two types of indicators: estimated $\mathbf{q}_0$ and calculated $\mathbf{q}_1 = A\mathbf{w}_0$. In the general case these indicators are different, $\mathbf{q}_0 \neq \mathbf{q}_1$. The goal is to make a concordant indicator, which accounts data and rank-scaled expert estimations as well.

We call a *concordant pair* a vector $\mathbf{q}$ of indicators and a vector $\mathbf{w}$ of weights such that

$$\begin{aligned} \mathbf{q} &= A\mathbf{w}; \\ \mathbf{w} &= A^+\mathbf{q}, \end{aligned} \tag{6}$$

where A^+ is a linear map, pseudo-inverse for A: $AA^+A = A$, $A^+AA^+ = A^+$ and $(AA^+)^\top = AA^+$ $+$, $(A^+A)^\top = A^+A$. The main problem we will solve in this paper is to modify expert estimations to hold (6).

So the expert estimations $\mathbf{q}_0, \mathbf{w}_0$ are given. Any monotonic transformation could be applied to the elements of the vectors. The design matrix $A \in \mathbb{R}^{m\times n}$ is given. It satisfies the condition (2).

Without loss of generality let us consider the following order relation:

$$q_1 \geqslant \cdots \geqslant q_m \geqslant 0 \text{ and } w_1 \geqslant \cdots \geqslant w_n \geqslant 0. \tag{7}$$

To hold this condition one must rearrange elements of the vector $\mathbf{q}_0$ and corresponding columns of the matrix A. Similar rearrangement must be performed for elements of the vector $\mathbf{w}_0$ and corresponding rows of the matrix A.

Rewrite the condition (7) as the following system of linear inequalities (show indicators only)

$$\begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \\ 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ \cdots \\ q_{m-1} \\ q_m \end{pmatrix} \geqslant \mathbf{0}.$$

Denote the bidiagonal matrix by J and rewrite (7) as

$$J_m\mathbf{q} \geqslant \mathbf{0} \text{ and } J_n\mathbf{w} \geqslant \mathbf{0}.$$

The number of rows in the square matrix J equals the number of ineqiations in the system. The number of columns in J equals the number of elements in $\mathbf{q}$.

Denote by $\mathcal{Q}$ and by $\mathcal{W}$ the expert-given cones for the objects and for the features respectively.

$$\begin{aligned} \mathcal{Q} &= \{\mathbf{q} \,|\, J_m\mathbf{q} \geqslant \mathbf{0}\}; \\ \mathcal{W} &= \{\mathbf{w} \,|\, J_n\mathbf{w} \geqslant \mathbf{0}\}. \end{aligned} \tag{8}$$

Let the vectors $\mathbf{q}$ and $\mathbf{w}$ be elements of arbitrary cones $\mathcal{Q}$ and $\mathcal{W}$. The linear operator A maps the cone $\mathcal{W} \ni \mathbf{w}_0$ of the expert estimations of features (8) to the cone $A\mathcal{W} = \mathcal{P} \ni \mathbf{q}_1$, see Fig. 3:

$$A : \mathcal{W} \to \mathcal{P}; \quad A : \mathbf{w}_0 \mapsto \mathbf{q}_1.$$

Two variants of the mapping are possible.

1. The cones $\mathcal{P}$ and $\mathcal{Q}$ intersect; in this case the expert estimations are concordant: there exists a pair $\mathbf{q}_p \in \mathcal{P} \cup \mathcal{Q}, \mathbf{w}_p = A^+\mathbf{q}_p \in \mathcal{W}$, which satisfies (6).
2. The intersection of $\mathcal{P}$ and $\mathcal{Q}$ is empty, the concordance procedure is needed.

These variants are described in the following sections.

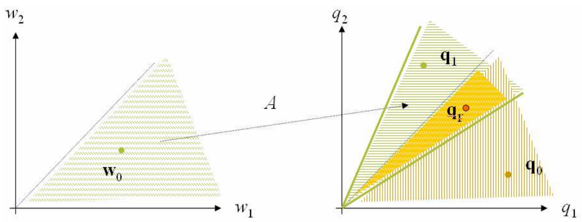

Fig. 3. The linear operator A maps the convex polyhedral cone of feature estimations to the space of indicators.

Mapping and intersection of convex polyhedral cones

Note some statements on the cone properties to justify the algorithms described below. The set of vectors $\mathcal{Q}$ in $\mathbb{R}^m$ is called a *cone*, if for any vector $\mathbf{q} \in \mathcal{Q}$ the vector $\lambda\mathbf{q}$ is also in $\mathcal{Q}$. *A convex polyhedral cone* is an intersection of the finite number of half-spaces that have a point in common. This point is called *the top of the cone.*

The convex polyhedral cone with the top in the origin of coordinates is a set of solutions of the system of linear inequalities:

$$\begin{cases} a_{11}w_1 + \dots + a_{1n}w_n \geqslant 0; \\ \dots \quad \dots \quad \dots \quad \dots \\ a_{m1}w_1 + \dots + a_{mn}w_n \geqslant 0. \end{cases}$$

Thus the system of linear inequalities $J\mathbf{w} \geqslant \mathbf{0}$ defines the convex polyhedral cone.

List the following statements about mapping and intersection of convex polyhedral cones.

Statement 1. *Let two convex polyhedral cones that have the common top. The intersection of these cones is the convex polyhedral cone. It is defined by the system of linear homogenic inequalities with the matrix*

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}.$$

Statement 2. *The set $\mathcal{W}$ of all vectors $\mathbf{w} = \langle w_1, \dots, w_n \rangle$, which satisfy the condition $w_1 \geqslant w_2 \geqslant \dots \geqslant w_n \geqslant 0$ is the cone.*

Statement 3. *The geometric locus defined by the image of the map $A : \mathcal{W} \to \mathcal{Q}$ is the cone.*

Thus, if the set $\mathcal{W}$ is the convex polyhedral cone, then the linear operator A maps it to the convex polyhedral cone $\mathcal{P} = A\mathcal{W}$. The corresponding pseudo-inverse linear operator A^+ maps the cone $\mathcal{W}$ to the cone $A^+\mathcal{W}$.

Statement 4. *If the cones are defined in the space of indicators by the systems of linear inequalities $B_1\mathbf{q} \geqslant \mathbf{0}$ and $B_2\mathbf{q} \geqslant \mathbf{0}$ intersect, their preimages in the feature space intersect, too.*

In the context of the investigated problem, if the convex polyhedral cones $\{\mathbf{q}|J_m\mathbf{q} \geqslant \mathbf{0}\}$ and $\{\mathbf{w}|AJ_n\mathbf{w} \geqslant \mathbf{0}\}$ intersect in the space of indicators, then their preimages in the feature space $\{\mathbf{q} : A^+J_m\mathbf{q} \geqslant \mathbf{0}\}$ and $\{J_n\mathbf{w} \geqslant \mathbf{0}\}$ intersect, too. Denote by $\mathcal{W}_p = \mathcal{W} \cup A^+\mathcal{Q}$ and $\mathcal{Q}_p = \mathcal{Q} \cup A\mathcal{W}$ the intersections, mentioned above.

Statement 5. *If the cone $\mathcal{Q}_p \neq \varnothing$ then $\mathcal{W}_p \neq \varnothing$; otherwise both cones are empty.*

This statement is equivalent to the following. For each vector $\mathbf{w}_p \in \mathcal{W}_p$ there exists some concordant vector $\mathbf{q}_p \in \mathcal{Q}_p$ such that the condition (6) holds.

To find the intersection $\mathcal{Q}_p$ describe the corresponding sets with the system of linear inequalities. The set $\mathcal{Q}_p \ni \mathbf{q}_p$ is the solution of the system of linear inequalities

$$\begin{cases} J_m\mathbf{q}_0 \geqslant 0; \\ AJ_n\mathbf{w}_0 \geqslant 0. \end{cases} \tag{9}$$

The solution $\mathcal{Q}_p$ is the cone; it could be trivial.

The cones does not intersect: the concordance is needed

If the cones $\mathcal{Q}_p = \mathcal{Q} \cap A\mathcal{W} = \varnothing$ and $\mathcal{W}_p = \mathcal{W} \cap A^+\mathcal{Q} = \varnothing$, then we will use a modified α-concordance procedure. Consider two rays, defined by the vectors $\mathbf{q} \in \mathcal{Q}$ and $\mathbf{p} \in \mathcal{P} = A\mathcal{W}$. Find the nearest rays in the edges or the faces of the cones $\mathcal{Q}, \mathcal{P}$ as

$$\cos(\mathbf{q}, \mathbf{p}) = \frac{\mathbf{q}^\mathsf{T}\mathbf{p}}{\|\mathbf{q}\|\|\mathbf{p}\|} \to \max.$$

Then apply the procedure (4) to the result vectors $\hat{\mathbf{q}}$ and $\hat{\mathbf{p}}$. To find these vectors one must solve the following optimization problem:

$$\begin{aligned} \text{maximize} \quad & \mathbf{q}^\mathsf{T}\mathbf{p} \\ \text{subject to} \quad & \mathbf{q}^\mathsf{T}\mathbf{q} = 1, \quad \mathbf{p}^\mathsf{T}\mathbf{p} = 1, \\ & J_n\mathbf{q} \geqslant \mathbf{0}, \quad AJ_m\mathbf{p} \geqslant \mathbf{0}. \end{aligned}$$

Describe an algorithm to find the vectors $\mathbf{q}^{(2k)}$ and $\mathbf{p}^{(2k+1)}$ on the even and odd step of its iterations. The vectors $\mathbf{x} = \mathbf{q}^{(2k)}$ and $\mathbf{y} = \mathbf{p}^{(2k+1)}$ will be the solutions of two sequentially-solved optimization problems. Let the arbitrary vector $\mathbf{p}^{(0)} \in \mathcal{P}$ on the step $k=0$.

Problem $2k$:		Problem $2k+1$:	
maximize	$\mathbf{x}^\mathsf{T}\mathbf{p}^{(2k)}$	maximize	$\mathbf{q}^{T(2k+1)}\mathbf{y}$
subject to	$\mathbf{x}^\mathsf{T}\mathbf{x}=1,$	subject to	$\mathbf{y}^\mathsf{T}\mathbf{y}=1,$
	$J_n\mathbf{x} \geqslant 0.$		$AJ_m\mathbf{y} \geqslant 0.$

For each step of the algorithm the constants $\mathbf{p}^{(2k)}$ and $\mathbf{q}^{(2k+1)}$ equal the corresponding solutions $\mathbf{x}$ and $\mathbf{y}$ of the previous step. Since the maximizing functions and the constrains of both problems are convex, then the solution will be found in the finite number of steps. We recommend to use in the practice the following convex optimization methods [11, 12].

Obtain the solution of the optimization problem, vectors $\hat{\mathbf{p}}, \hat{\mathbf{q}}$ and execute the α-concordance procedure

$$\mathbf{q}_\alpha = (1-\alpha)\hat{\mathbf{p}} + \alpha\hat{\mathbf{q}},$$

if there exists a nontrivial solution $\mathbf{q}_\alpha$, that is $\hat{\mathbf{p}}^\mathsf{T}\hat{\mathbf{q}} \neq -1$. The vector $\mathbf{q}_\alpha$ and the vector $\mathbf{w}_\alpha = A^+\mathbf{q}_\alpha$ satisfy the condition (6). These vectors define the cones $\mathcal{W}$ and $\mathcal{Q}$ in the corresponding spaces and the intersection $\mathcal{W}_p = A\mathcal{W} \cap \mathcal{Q}$ is not empty. As for α-concordance procedure, when $\alpha \to 0$ we prefer expert estimations of objects. When $\alpha \to 1$ we prefer expert estimations of features.

The next section shows a way to obtain linear-scaled indicators using rank-scaled expert estimations.

The cones intersect: disturb the design matrix to get robust indicators

Consider the obtained intersection cone $\mathcal{Q}_p$ and the design matrix A. Disturb elements of the matrix $A = A + \Delta$. The normal distribution hypothesis is $\Delta = \delta I,\ \delta \sim \mathcal{N}(0, \sigma^2)$. The image of the linear map $\mathbf{q} = (A+\Delta)\mathbf{w}$ will have also the normal distribution. According to this hypothesis, take the indicator $\mathbf{q}_p$ as robust to a small disturbance of the matrix A when it has the maximal distance to the all faces of the cone $\mathcal{Q}_p$ subject to $\|\mathbf{q}_p\| = 1$. We call the Chebyshev point the vector $\mathbf{q}_p$ in the center of the inscribed sphere of the cone $\mathcal{Q}_p$.

To obtain the maximal distance from the desired vector $\mathbf{q}_p$ to a nearest face $\mathbf{s}$ of the cone solve the optimization problem

$$\hat{\mathbf{q}}_p = \arg\max_{\mathbf{q}_p \in \mathcal{Q}_p} \big\{\|\mathbf{q}_p - \mathbf{s}\|^2, \text{where } \mathbf{s} \in \mathbb{R}^m \setminus \mathcal{Q}_p \text{ and } \|\mathbf{q}_p\| \leqslant 1\big\}.$$

Take in account the system of linear inequalities (9), which define the cone $\mathcal{Q}_p$. Denote by $\mathbf{s}_\ell$ the normal vector corresponding to the line number ℓ of this system. The dot product $\mathbf{x}^\mathsf{T}\mathbf{s}_\ell = 0$ defines a plane passing through the origin of coordinates in the space of indicators. The distance d from the vector $\mathbf{q}_p$ to this plane equals

$$d(\mathbf{q}_p, \mathbf{s}_\ell) = \frac{\mathbf{q}_p^\mathsf{T}\mathbf{s}_\ell}{\|\mathbf{s}_\ell\|}.$$

This is the convex optimization problem

$$\begin{aligned} \text{maximize} \quad & \inf_{\ell=1,\dots,L} \left(\mathbf{x}^\mathsf{T}\mathbf{s}_\ell\|\mathbf{s}_\ell\|^{-1}\right) \\ \text{subject to} \quad & \mathbf{x}^\mathsf{T}\mathbf{x} = 1, \\ & J_n\mathbf{x} \geqslant \mathbf{0}, \\ & AJ_m\mathbf{x} \geqslant \mathbf{0}. \end{aligned}$$

See geometry of the convex optimization methods in [11, 13]. The solution of the problem is the robust vector $\hat{\mathbf{x}} = \hat{\mathbf{q}}_p$. Also calculate the vector of the weights as $\mathbf{w}_p = A^+\hat{\mathbf{q}}_p$. These two vectors makes the concordant pair (6).

Thus, the linear-scaled indicator $\hat{\mathbf{q}}_p$ is obtained using the rank-scaled expert estimations.

References

[1] *Saari D. G.* Mathematics and voting // Notices of the American Mathematical Society. — 2008. — Vol. 55, No. 4. — Pp. 448–455.

[2] *Schiefer J., Fischer C.* The gap between wine expert ratings and consumer preferences: measures, determinants and marketing implications // International journal of wine business research. — 2008. — Vol. 20, No. 4. — Pp. 335–351.

[3] *Strijov V., Granic G., Juric Z., Jelavic B., S. A. M.* Integral indicator of ecological footprint for croatian power plants // HED Energy Forum "Quo Vadis Energija in Times of Climate Change Zagreb, Croatia. — 2009.

[4] *Strijov V., Shakin V.* Index construction: the expert-statistical method // Environmental research, engineering and management. — 2003. — Vol. 26, No. 4. — Pp. 51–55.

[5] *Draper N. R., Smith H.* Appied regression analysis. — John Wihley and Sons, 1998.

[6] *Belsley D. A.* Conditioning diagnostics: collinearity and weak data in regression. — New York: John Wiley and Sons, 1991.

[7] *Jolliffe I. T.* Principal component analysis. — Springer, 2002.

[8] *Isenmann A. J.* Modern multivariate statistical techniques. — Springer, 2008.

[9] *Golub G., Van Loan C.* Matrix computations. — Johns Hopkins Studies in Mathematical Sciences, 1996.

[10] *Hogben L.* Handbook of linear algebra. — CRC Press, 2007.

[11] *Boyd S., Vandenberghe L.* Convex optimization. — Cambridge University Press, 2009.

[12] *Minoux M.* Mathematical programming: theory and algorithms. — Wiley, 1986.

[13] *Dattorro J.* Convex optimization & Euclidean Distance Geometry. — Meboo Publishing, 2005.