Strijov V.V., Letmathe P. Integral indicators based on data and rank-scale expert estimations // Intellectual Data Analysis: the International Scientific Conference Proceedings, 2010 : 107-110.

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

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

where the parameter r defines the distance function.

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

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 $\mathbf{q}_{\text{ESM}} = A\mathbf{w}_1$, 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.