Metric analysis of estimation of territorial development

D.A. Gainanov
Institution of Russian Federation Academy of sciences
Institute of social and economic researches of Ufa scientific centre RAS
Ufa, Russia
e-mail: 2d2@inbox.ru

O.G. Kantor
Institution of Russian Federation Academy of sciences
Institute of social and economic researches of Ufa scientific centre RAS
Ufa, Russia
e-mail: o_kantor@mail.ru

Abstract[1]

The method of estimation of territorial development shown in this work is based on the objective data of metric analysis. This represents a solution of a multi-criteria problem of alternatives' regulation by means of minimization of the Euclid distances in indicator and criteria spaces.

1. Introduction

Development of territories – multigoal and multicreterial process, regardless of which object is implied under a term "territory" (region, city district, municipal education, etc.) In most general case the estimation of development of territories plugs in itself a few stages: determination of aims and criteria, calculation of values of criteria and implementation of procedure of direct estimation.

Social (improvement of education, feed and health protection, decline of level of poverty, making healthy of environment, etc.) and economic (increase of gross regional product, increase of volume of producible products, increase of profits, etc.) goals more frequently come forward as aims of development of territories. According to the aims of development of territories the system of criteria (descriptions of development) and indexes, which measure these criteria, is built. Depending on the aims of research procedure of estimation includes either watching the dynamics of change of separate indexes, either selection of homogeneous groups of the examined objects or organization of objects.

Research of dynamics of change of indexes allows conducting the detailed analysis, exposing tendencies, conformities to law and connections between them. A main lack of such approach is impossibility of complex estimation, which is most substantial at plenty of the investigated indexes, possessing multidirectional tendencies. For the selection of homogeneous groups of objects plenty of methods is developed on the basis of taxonomical and multidimensional statistical methods. As a result of application of such approach in the initial aggregate of the investigated objects the shallower are selected to the aggregate of objects: leaders, outsiders, etc. Thus a question is opened, as objects are well-organized into such groups. Procedure of organization of objects, as a rule, in practice is produced on the basis of integral criterion, the calculation of which is carried out on the base of methods: expert, a priori, "recognitions of patterns", factor and component analysis [1].

At the estimation of socio-economic development of separate territories it is expedient to select a few the most essential independent aims and to carry out monitoring of their achievement. Both territories, for which the results of comparisons serve as the indicators of accordance of their aims and reference points to the terms of external environment, and the state (in the face of organs of management of different levels) for adjustment of the policy and grounded acceptance of decisions, including distributing of facilities, funds, underbacks etc., need the results of such analysis.

For acceptance of the grounded administrative decisions the organs of power must own complete information about the dynamics of macroeconomic and territorial indicators of socio-economic and budget-financial development, and also to be in a position to carry out estimation, analysis of efficiency and effectiveness, design and prognostication on the indicated parameters.

In our opinion, the estimation of development of territories (on socio-economic and on other indexes), efficiency of the accepted administrative decisions and plenitude of their realization, is necessary to express with an integral index, because it will allow to give evident interpretation of results, simply to determine a difference in the estimation of levels of socio-economic development of territories and correct aims and methods of their achievement. In addition, the tasks of ranging of the investigated objects can effectively decide on the basis of integral index, that frequently is the primary purpose of the conducted estimation [2, 3], and tasks of allocation of resources (underbacks, funds, etc.)

Proceedings of the 11th International Workshop on Computer Science and Information Technologies CSIT'2009, Crete, Greece, 2009

2. Analysis of scientific reserve

The analysis of existent methods [4] stated that for the calculation of integral indexes (analogical names: generalized indexes, summary ratings, etc.) opinions of experts, which are expressed either in gravimetric coefficients or in ball or ratings estimations, are mainly used. Besides that an integral index, expected thus, initially is the reflection of subjective preferences of experts, quite often integral indexes, expected on different methods, sufficiently changes the grades of the investigated objects. Summarizing said, it is possible to draw conclusion, that the actual development of method of estimation of development of territories is being based on a calculation of objective integral index, realizing the obvious special purpose settings and the maximally deprived subjectivism.

The integral index of level of development of territories must accumulate in itself plenty enough of private indexes. Therefore task of determination of method of his calculation with subsequent organization of the examined aggregate of the investigated objects it is possible to interpret as a multicriterial problem of making decision.

The major instrument of decision of multicriterial tasks is principle of Pareto, which consists of the selection of so called set of Pareto, consisting of not dominating alternatives (under an alternative any object from an analyzable aggregate will be implied), that alternatives which can not be improved on any criterion. This set plugs in itself most "contrasting" alternatives, difficult for comparison. In case when Pareto set consists of plenty enough of alternatives, each of which in same queue has estimations on a few of criteria, comparison is possible to carry out only on the basis of the special methods. Determination of Pareto set is the important stage of decision of task of choice of the best alternative, because in most existent methods the proof of belonging of the decision to the Pareto set is an obligatory procedure for the decision.

It is traditionally accepted to distinguish three basic tasks of making decision [5].

1. Regulating of alternatives.
2. Distributing the alternatives on the classes of decisions.
3. Selection of the best alternative.

For the decision of tasks plenty of methods which can be divided into three groups are developed:

1. *axiomatic methods*, to the number of which the Multicriterial Utility Theory (MAUT-methods) [6] and methods of Prospect Theory, taking into account the real lines of human conduct in tasks with subjective probabilistic estimations;
2. *structural methods*: Analytic Hierarchy Process (AHP) and ELECTRE methods (group of methods, based on development of indexes of bi-comparison of alternatives);
3. *heuristic methods* (method of the weighed sum of estimations of criteria, method of indemnification and other).

The main lack of axiomatic methods (above all things MAUT-methods) is uncheckable character of axioms that means a requirement to take on a trust the rules of rational conduct, effluent from one or another theory in practice.

Constructible methods suppose the use of straightly defined sequence of action. They are based on the subjective estimations of decision-maker (DM) and therefore, can be examined as the ones joining heuristic. These methods had found a wide practical application because of their simplicity and evidentness.

The dignity of all heuristic methods is simplicity and comfort, and basic failing is that all of them do not have a scientific ground. Indeed, these methods are based either on giving the alternatives estimations, either on the bi-comparison of alternatives on each criterion. Both procedures are based on DM opinion, which means that the result priori is subjective.

The basic idea of most methods related to the analysis of alternatives applied in practice consists of possibility of construction relatively by simple facilities of well-organized estimations of their comparative position that allows deciding any of the tasks of making decision listed above. Therefore, two stages are obligatory:

- determination of the system of initial indexes which alternatives are compared on the basis of;
- construction of the system of comparable indicators.

On the first stage, on the basis of opinion of a DM, the array of indexes, this must be full enough from point of decision of the put tasks. These indexes are often called criteria, although, in our view, such name is justified only in case that DM makes decision exceptionally on the basis of direct values of initial indexes. If DM , realizing the having a special purpose settings, enters in consideration functions the arguments of which are initial indexes, then exactly these objective functions more logical to name criteria. On the second stage on the basis of the formed array of indexes the system of comparable indicators, main property of which must be the maximal comparableness, is built, that must eliminate various dimensions and various directions of initial indexes.

Realization of the second stage is provided by application of procedure of normalization, in basis of which correlation of actual values of indexes is fixed with proper components of "ideal vector", that vector with "ideal" values of indexes. Components of this vector more frequent than all are either normative values, either average (for example, on the analyzable aggregate of subjects) or standard (as which the best values can be chosen on the examined aggregate of subjects). There are other ways of task "ideal vector" (for example, by means of account of maximally possible variation of the proper indexes), but in the real work they are not examined. As a result of application of procedure of normalization from the actual values of criteria passing is carried out to the dimensionless sizes (to the indicators), which have an identical range of changes, as a rule, consilient with a segment [0,1]. Thus the "ideal" values of indicators must correspond to the values of indexes, which are equal 1, and worst – 0. We will carry out the formed record of procedure of normalization.

Let M present alternatives $A^1,\dots,A^M$, each of which is characterized by the set of values N criteria. Thus, there is a set of initial indexes $\left\{x_j^i, j=\overline{1,N}, i=\overline{1,M}\right\}$. For each of the indexes we find:

- "ideal" values of indexes x_j^*, $j=\overline{1,N}$;
- values x_j^- and x_j^+, $j=\overline{1,N}$, being accordingly the worst and the best values on every index. These values can be got on the basis of indexes of the investigated aggregate of alternatives (the worst and the best values accordingly), or can be set a priori, according to sense of criteria (for example, scopes of turn-down of values of index).

Using the denotations formulas entered higher for the calculation of the system of indicators $\left\{p_j^i, j=\overline{1,N}, i=\overline{1,M}\right\}$ can be written as

$$p_j^i = 1 - \left| \frac{x_j^* - x_j^i}{max\left\{\left|x_j^* - x_j^-\right|, \left|x_j^* - x_j^+\right|\right\}} \right|, \quad (1)$$

$$j=\overline{1,N}, \; i=\overline{1,M},$$

if $x^* \in \left(x^-; x^+\right)$, or $x^* \in \left(x^+, x^-\right)$;

$$p_j^i = \frac{x_j^i - x_j^-}{x_j^+ - x_j^-}, \; j=\overline{1,N}, \; i=\overline{1,M}, \quad (2)$$

if $x^* = x^+ > x^-$, or $x^* = x^+ < x^-$;

$$p_j^i = \frac{x_j^i - x_j^-}{x_j^* - x_j^-}, \; j=\overline{1,N}, \; i=\overline{1,M}, \quad (3)$$

if $x^* > x^+ > x^-$, or $x^* < x^+ < x^-$.

A formula (1) corresponds to the method of normalization, when in "ideal" quality gets out, for example, average, and in quality x^- and x^+ - accordingly the least and most values (or vice versa) of the examined index. Formula (2) used, when it is possible only to assert for the examined index, that than anymore (less than) his value, so much the better. Therefore in "ideal" quality index used, accordingly, most or the least value of index from a number present. A formula (3) answers a situation, when there is a normative value on the examined index, but actual values of index (on all aggregate of alternatives) "does not drag as far as" to him.

Obviously, that expected on formulas (1) and (2) sizes p_j^i take on values in a range from 0 to 1, and in the case of application of formula (3) $p_j^i \in \left[0, p_j^0\right]$, $p_j^0 < 1$, and than nearer value of initial index x_j^i to "ideal", the nearer to 1 a proper value p_j^i. For sure, there are other ways of normalization systems of indexes, which let us to get range changes of indicators, different from $[0,1]$, and which are not necessarily based on linear dependence of indicators on the values of indexes. But, basic principle is necessarily observed in any case "what nearer value of initial index x_j^i to "ideal", the nearer to maximal the proper value of p_j^i indicator". Not limiting community of reasoning, we will suppose that all expected indicators $\left\{p_j^i, j=\overline{1,N}, i=\overline{1,M}\right\}$ change in a range from 0 to 1.

In future for definiteness will talk only about a sequencing of alternatives problem, because on its basis can be decided task of distributing of alternatives on classes, and task of determination of alternative, possessing the best descriptions.

In practice number of criteria *N* can suffice great, because of what application of methods MAUT and AHP can be it is either very labored or it is impossible in principle. Therefore more frequent than all use simple methods, based on the calculation of the generalized index relatively, in spite of the fact that these methods possess the row of failings. The most widespread methods of calculation of the generalized index are calculations middle arithmetic, middle geometrical, middle quadratic, middle harmonic, additive and multiplicative aggregative functions. Sometimes for the aims of research most suitable is an index characterizing variation of values from averages. In this case as the generalized index use mean quadratic deviation.

Application for the calculation of formulas middle arithmetic, middle geometrical, middle quadratic, middle harmonic, additive and multiplicative aggregative functions allows to get the large values of summary indexes with the increase of values of indicators (the best are consider alternatives with the greatest values of all indicators).

Basic dignity of additive aggregation function

$$p^i = \alpha_1 p_1^i + \dots + \alpha_N p_N^i, \; i=\overline{1,M}, \quad (4)$$

where $\alpha_1,\dots,\alpha_N$ - weight coefficients, $\sum_{j=1}^{N}\alpha_j = 1$, is that the necessary and sufficient terms of optimality are related to it on Pareto: from one side, for any set of the set weight coefficients $\alpha_1,\dots,\alpha_N$ alternative with the most value of the generalized index, expected on a formula (4), will belong to the great number of Pareto. From the other side, for any alternative, from the great number of Pareto it is possible to pick up weight coefficients $\alpha_1,\dots,\alpha_N$ so that a result of additive aggregation function (4) for this alternative was maximal. The calculation of mean quadratic deviation is used on purpose carry out organization so that the best were alternatives with the least variation of values of indicators (that priority gives oneself up to the alternatives with the approximately identical values of all indicators).

The calculation of the generalized index by means of each of the transferred formulas of middle, and also additive and multiplicative aggregates possesses both the dignities and failings. Formulas middle arithmetic and middle quadratic are simple and well interpreted, but the best alternative, certain in obedience to these formulas, can possess extremely low (and sometimes and in general the worst) values on the row of criteria, that does not allow to examine it in its best quality. Middle geometrical formula in case if some alternative has a worst value even on one criterion (that the proper value of indicator is equal 0), automatically will put a zero (worst) value of the generalized index in accordance to this alternative, even if on other criteria an alternative can be considered a leader. An alternative with the worst values of all indexes in obedience to the formula of mean quadratic deviation will be considered the best, by virtue of that will possess a zero variation of values of indicators. Application of formula middle harmonic, if even one of indicators is equal to the zero, it is impossible. The lacks of formulas of additive and multiplicative aggregation function ensue from that they are more general cases of formulas middle arithmetic and middle geometrical accordingly.

Besides the indicated lacks of the transferred formulas there is another problem, inherent each of methods of calculation, essence of which consists in that not clear what from two alternatives to consider the best, if the values of the proper generalized indexes are identical. It ensues from all foregoing, that actual is development of the special methods of realization of multicriterial tasks of making decision, which must combine in itself not only relative simplicity of process of receipt of decision and possibility of evident interpretation of results but also possibility to level the situation of multiplicity of alternatives with identical final estimations.

3. Metric analysis: logic and algorithm

A method, based on taking of initial multicriterial problem of making decision to the bicriterial problem in space of values of criteria, realizing the having a special purpose settings of DM, is offered in the real work.

We will consider single N-measured cube E^N, each of his ribs equated with the proper indicator. In this case it is possible to talk that E^N is space of indicators. To every alternative $i=\overline{1,M}$ in E^N the unique point p^i with co-ordinates $\{p^i_j, j=\overline{1,N}\}$ corresponds. The aggregate of all such points $\{p^i, i=\overline{1,M}\}$ will be designated as G. Obviously, that by virtue of construction of indicators on formulas (1)-(3), a point $e(1,1,\dots,1)$ (with all co-ordinates equal 1) is "standard" in that sense, that this point corresponds to the alternative (more frequent hypothetical) with the best values on all criteria. We will name by a "diagonal" E^N cube's segment, connecting points $\theta(0,0,\dots,0)$ (all co-ordinates are zeros) and e. Any point, lying on a diagonal, is characterized by equality of all co-ordinates, and the proper alternative – by equality of all indicators.

Obviously, that than nearer to "standard" a point p^i beds, the higher the proper alternative is estimated. Therefore logical is introduction to consideration of criterion:

$$K_1' = \sqrt{\sum_{i=1}^{N}(1-p_i)^2}\ , \qquad (5)$$

characterizing Euclidean distance from the examined point to "standard". From two alternatives the best will be considered the one for which the value of criterion K_1' will be the least. Minimum value, equal to the zero, a criterion K_1' takes on an in the unique point - e, and maximal, equal to $\sqrt{N}$, - also in the unique point θ.

It is possible to prove that point from a set G, in which a criterion K_1' takes on a minimum value, belongs to the great number of points, optimum on Pareto. In general, a criterion (5) runs to a minimum not in the unique point. Therefore a question is opened, how to conduct organization of alternatives which the values of criterion K_1' coincide for.

We will assume that for two different alternatives A^1 and A^2 their proper points p^1 and p^2 in E^N are identically remote from e. Then, if all indexes which alternatives are put in order on do equivalent, that having identical priority, the best will count an alternative, possess less variation of values of the proper indicators. (A case with different priorities of criteria in the real work is not examined.) Such approach responds to the principle of "identity of indexes", which the following is implied under. An alternative is characterized by identical indexes, if all its indicators are equal. Such alternative has a zero variation of values of indicators. To every alternative with identical indexes in space E^N corresponds to a point, belonging to the diagonal of cube.

Therefore the diagonal of cube E^N can be examined as a reflection of great number of alternatives with identical indexes. This principle in practice can be realized by the calculation of distance from points from a great number G to the diagonal of cube E^N, that predetermines introduction of criterion, characterizing this distance in same queue. Using the formulas of analytical geometry, this criterion can be written down as follows:

$$K_2' = \sqrt{\frac{(N-1)\sum_{i=1}^{N} p_i^2 - 2\sum_{i=1}^{N}\sum_{j>i}^{N} p_i p_j}{N}}. \qquad (6)$$

It's easy, using the necessary and sufficient terms of extremum of function of many variables, to prove that K_2' takes on a minimum (the best) value, equal to the zero, in any point for which $p_1 = p_2 = \ldots = p_N$, so that in any point, belonging diagonal of cube E^N. Formula (6) in the case of calculation of criterion K_2' for the arbitrary top of cube E^N, different from e and θ, it is possible to simplify:

$$K_2' = \sqrt{k\frac{N-k}{N}}, \qquad (7)$$

where k – is the number of co-ordinates of top, equal to 1.

Obviously, the expression $k\frac{N-k}{N}$ takes on a maximal value at $k = N/2$. Therefore if N is even, then the maximal (worst) value of criterion K_2' is equal $\sqrt{N}/2$, and if N is odd, then a criterion K_2' arrives at a maximal value, equal to $\sqrt{\frac{N^2-1}{4N}}$, at k equal $(N+1)/2$ and $(N-1)/2$. Criterion K_2' comes to the smallest value on the great number of tops of cube E^N, different from e and θ, at $k=1$ and $k=N-1$. Thus, from two alternatives A^1 and A^2, mentioned before, the best will be considered the one for which the value of criterion K_2' will be the least.

Being based on foregoing, the problem of set regulating of alternatives can be formulate as a bicriterial task on a discrete set G - to put in order alternatives in accordance with a criterion K, characterizing the closeness of alternatives to the standard e and the set of alternatives with identical indexes in space of indicators E^N:

$$K = \left\{ \begin{array}{l} K_1 = \frac{1}{\sqrt{N}} K_1' \\ K_2 = \begin{cases} \frac{2}{N} K_2', N - even, \\ \frac{2}{\sqrt{N^2-1}} K_2', N - odd \end{cases} \end{array} \right\} \to min. \qquad (8)$$

In order that both the quotients of criterion K_1 and K_2 in (8) were not only identically directed but also had an identical range of values from 0 to 1, criteria K_1' and K_2' were rationed.

Thus each of private criteria K_1 and K_2 is possible to examine as a metrics in space of indicators.

On a fig. 1 graphic illustration of location of optimum decisions on private criteria K_1 and K_2 in E^2 is presented:

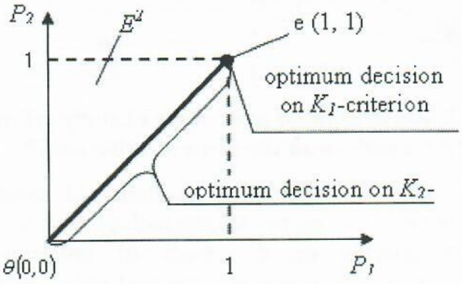

Fig. 1. Optimum decisions on private criteria in the space of indicators (the case when $N = 2$)

From the shown above picture it is seen that sets of optimum decisions on criteria K_1 and K_2 do not eliminate each other, that criteria are not contradictory. And if a task was put on all E^N set, but not on a subset $G \subset E^N$, then the decision would be banal (point e). But a problem of regulating of alternatives a priori is discrete, therefore its formalization on the basis of criterion (8) supposes development of clear algorithm of procedure of decision.

After introduction of criteria K_1 and K_2, every alternative is put in accordance with a steam of numbers $\{k_1, k_2\}$, essence of which are the values of criteria K_1 and K_2 expected for them. Thus, it is possible to consider space of values of criteria K_1 and K_2, which by virtue of ranges of their possible values coincides with space E^2. In this space the unique point corresponds every alternative with co-ordinates (k_1, k_2). The set of all such points will be designated as V. Taking into account an orientation of each of these criteria, it is possible to assert that a "standard" point in the entered space is a point with zero co-ordinates $(0,0)$. Therefore as generalized logically, in our view, is to consider a criterion

$$K=\sqrt{k_1^2+k_2^2}, \qquad (9)$$

which characterizes Euclidean distance from the points of set V to standard $(0,0)$. A function K can also be examined as a metrics in space of values of criteria K_1 and K_2.

On the basis of formula (9) for every alternative the values of the generalized criterion can be counted in accordance with possible direct realization of procedure of regulating of alternatives.

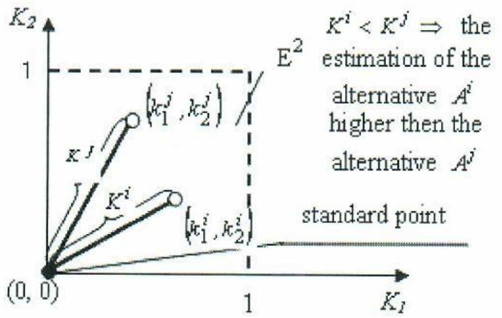

Fig. 2. Illustration of procedure of comparison of alternatives on the basis of criterion (9)

The algorithm of decision of multicriterial regulating problem is below given for M alternatives on a set of N equivalent indexes on the basis of criterion (8), characterizing a closeness to the standard and to the set of alternatives with identical indexes in space of indicators.

Stage 1. Formation of array of indexes, which will be the basis comparison of $\{x_j^i, j=\overline{1,N}, i=\overline{1,M}\}$ alternatives.

Stage 2. Construction of the system of comparable indicators $\{p_j^i, j=\overline{1,N}, i=\overline{1,M}\}$, applying formulas (1)-(3).

Stage 3. Calculation of values of criteria K_1 and K_2 for each *i*-ouch alternative, $i=\overline{1,M}$ on the proper set of indicators $\{p_j^i, j=\overline{1,N}\}$, in obedience to formulas (5), (6), (9), (10). As a result a steam of numbers $\{k_1^i, k_2^i\}$ for each i-ouch alternative is put.

Stage 4. Calculation of values of the generalized criterion K^i for each i-ouch alternative, $i=\overline{1,M}$, according to a formula (9).

Stage 5. Regulation of alternatives $\{A^i, i=\overline{1,M}\}$ in order from smallest to greatest of values $\{K^i, i=\overline{1,M}\}$. Thus, the best alternatives will be corresponded to the least grades.

4. Conclusion

1. As a result of application of the offered method both the task of determination of the best alternative (obviously, that the best alternative has a grade, equal to 1) and the task of distributing the alternatives can be decided on classes (by means of breaking up on the certain number of intervals the set of values of the generalized criterion K).

2. Basic dignities of the offered method it is been, at first, bicriterial going near formalization of the having a special purpose settings, which the degree of closeness of values of indexes to the standard sizes and variation of values of indexes is taken into account at; secondly, obvious principle of minimization of Euclidean distances (birth-certificates) in space of indicators and space of criteria, fixed in basis of comparison of alternatives; and, thirdly, absence of subjectivism.

3. The offered method of estimation of level of development of territories on the basis of application of metric analysis for the calculation of objective integral index gives possibility to range territories and carry out the complex analysis of the investigated objects, that is necessary for acceptance of the grounded decisions, especially in part of distributing of facilities and adjustment of aims of development. Use of the offered method for the of the same type groups of indexes, for example, characterizing the sphere of health protection, educations, profits of population and other, allows to expect the generalized group estimation directly for analysable direction, that in same queue does possible not only the lead through of more detailed comparative analysis but also determination of spheres, requiring priority attention.

References

1. Raiskaya N.N., Sergienko Y.V., Frenkel A.A. "Reiting of regions according to integral index of investment seductiveness". *Voprosy statistiki,* 2009; 1: 56-60.
2. Decree of the President of Russian Federation of 28.06.2007, № 825.
3. Decree of the President of Russian Federation of 28.04.2008, № 607.
4. Tulicheva L.D. "Use of the comparative research for the determination of the strategic priorities". GUAP, Saint-Petersburg, 2007.
5. Larichev O.I. "Theory and methods of decision making". Logos, Moscow, 2003.
6. Keeney R., Raiffa H. "Decisions with Multiple Objectives: Preferences and Value Tradeoffs". Radio i svyaz', Moscow, 1981.