DANIEL KANEVSKY, PAVEL CHISTYAKOV

LOMONOSOV MOSCOW STATE UNIVERSITY

THE USE OF THE EVOLUTIONARY ALGORITHM FOR THE AGGREGATION OF SOCIO-ECONOMIC INDICES (CONSIDERING THE ESTIMATE OF THE INNOVATIVE POTENTIAL OF RUSSIAN REGIONS)

ABSTRACT. Aggregating indicators are the numerical characteristics of objects and processes, reflecting their global properties, which often defy strict formalization. The problem of calculation of aggregating indicators arises in many branches of social science, economics, and geography. In this paper we introduce a new method, which uses several simple quantitative characteristics to construct a rating aggregating indicator. The evolutionary algorithm, underlying our method, doesn't use a provided formula or function to optimize, thus guaranteeing unbiased results. Moreover, the evolutionary algorithm takes into account modest effects, annihilated by factorial analysis. We illustrate the method calculating the ratings of innovation potential of Russian regions.

KEY WORDS: decision making, aggregating indicators, genetic algorithms, innovations, Russian regions.

INTRODUCTION

The indices aggregation is necessary for getting a synthetic idea of various processes and phenomena. The issue of the aggregation of socio-economic indices is relevant and unsettled. It is bound to the fact that the choice of mathematical aggregation modes (for example, a preference for various arithmetical operations over indices) is always subjective and the modes themselves turn the aggregation process into a mechanical operation. Also most effective methods don't take into account modest effects, annihilated by the factorial analysis. The aggregating index must not only represent a sum of major indications, but a new quality and it must be the most objective.

In this article a new approach to the indices aggregation, based on the evolutionary algorithm, is presented. Practical means of this approach are brought out using the example of calculation of the innovative potential of Russian regions.

EVOLUTIONARY ALGORITHMS

Evolutionary algorithms have only recently been used in calculus mathematics. They are regarded as the so-called "weak" methods, not often using the knowledge of the objective field. It makes them more universal but less effective. Evolutionary algorithms usually work relatively slowly and in some cases don't give satisfactory results. Their advantage is that they make it possible to tackle incorrectly posed and badly formalized tasks, and also very difficult polyvalent tasks, for which it is impossible to construct an adequate algorithm. In our opinion the aggregation of socio-economic indices is such a task. The evolutionary algorithm simulates the process of evolution, "the crown" of which will be the solution to the problem.

There are a lot of various empirical and analytical models of the development of a population. The differences are caused by the conditions of the evolutionary process (for example, limited resources) and by some specific features of the population (for example, the character of consumption). Not to concentrate on one particular case we take into consideration only the basic features: heredity, variability and natural selection. Let's present these notions by means of the mechanisms advancing the evolution. The mechanisms of recombination and mutation will ensure the variety of individuals (variants of problem solution) and inheritance of their useful features. The mechanism of selection, being the core of the algorithm, will contribute to the solution. The quality features of these processes are set according to the most general principles of the evolution.

It is worth mentioning that the evolutionary algorithm works like "the black box". Which means that there is no adequate mathematical model, allowing to optimize the set of parameters of the evolutionary algorithm from the point of view, let's say, working time or memory input.

CONSTRUCTING EVOLUTIONARY ALGORITHM FOR CALCULATION OF RANK AGGREAGTING INDICATOR

Now we can pass over to the setting of the evolutionary algorithm for the construction of rank aggregating indices. In this work its parameters are divided into the following five groups; the formalized description of various groups is presented in small print.

1. THE ENVIRONMENT OF THE POPULATION DEVELOPMENT AND ITS INITIAL STRUCTURE

The environment is set by means of an expert survey and is a selection of indications – initial indices for every territorial cell (region), and the weight of these indications. Every expert suggests their set of indices for the aggregation and their weights. If several experts coincide in their opinion and suggest identical indices, their weights are summed up.

In the evolutionary algorithm randomly set vectors of real numbers, which length equals the number of territorial cells, present "individuals" (later this notion will be used without inverted commas).

These vectors are potential solutions to the task. Let's unite a selection of such randomly chosen individuals into a multitude – a population. Within the limits of this population we'll simulate the evolutionary process to arrive at the solution to the problem.

Let there be **n** objects and **m** indices. $\mathbf{V} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots \mathbf{v}_n\}$ will designate the multitude of objects, $\psi = \{\psi_1, \psi_2, \ldots, \psi_m\}$ will be the multitude of indices. Index ψ_i is set by a vector with length **n** $\mathbf{C}_i = (\mathbf{c}_{i1}, \mathbf{c}_{i2}, \ldots, \mathbf{c}_{in})$, where $\mathbf{c}_{ik}$ is a real number, equal to the value of i-th index for object $\mathbf{v}_k$. Besides, every ψ_i is confronted with its weight $\mathbf{w}_i$, determining the significance of ψ_i when searching the rank aggregating indicator. This information is provided by experts. It must undergo some extra processing to be used by the algorithm. Let's confront every vector $\mathbf{C}_i$ with a permutation of first **n** positive integers $\mathbf{S}_i$, where component $\mathbf{s}_{ik}$ is the number of an object, which has rank **k** according to index ψ_i (vector $\mathbf{C}_i$). In other words, we put vector $\mathbf{C}_i$ in descending order (the maximum value is in the first place), and then exchange each component of ordered $\mathbf{C}_i$ with the number of the corresponding object. The obtained succession of natural numbers represents permutation $\mathbf{S}_i$. Let's call $\mathbf{S}_i$ a rank permutation corresponding to $\mathbf{C}_i$. Thus, the input data for the algorithm are **m** pairs of type $(\mathbf{S}_i, \mathbf{w}_i)$ reflecting the information, set by the indices.

2. THE RULES OF BIRTH, MUTATION AND DEATH OF INDIVIDUALS.

As it was mentioned above, the recombination and mutation operations are the main source of new individuals in the evolutionary algorithm. These two mechanisms are probably one of the most important parameters of the evolutionary algorithms and they are subject to determination in every concrete case. Their choice is conditioned by the requirement of getting the maximum variety of individuals in a population, allowing to "stumble on the solution". For the development of the evolutionary process individual's death must depend on its adaptation. That's how the mechanism of natural selection joins in the model process. The most probable ("certain") death of the least adapted is the general rule of death modeling.

The birth of a new individual (recombination operation) is realized in the following way:

A pair of existing individuals – "a mother" and "a father" (hereafter without speech marks) is chosen. As the examined individuals are sexless, they are chosen randomly out of the population (mother and father must be different). A new vector is formed according to the following rule: every *i*-th position of "the child" has a value equiprobably "inherited" from mother or father (a value from the *i*-th position of mother or father). That's why on average half of "child's" values coincides with the values in the mother's positions, and the other half coincides with father's values. As in reality it isn't always so, "a child" can "resemble" his mother or father more.

Let's describe the rule of mutation. In this work pointed mutations are realized by means of random replace of one component of the chosen vector with another number. Both vector component, subject to the change, and the number, which replaces the chosen component (from the interval [0;1]) are chosen randomly.

The death in the population is simulated by deleting the vectors with the worst adaptation value (see below).

3. THE MEANS OF ESTIMATE OF THE ADAPTATION

This is the most important part of the evolutionary algorithm formation. The speed of achieving the acceptable solution and the fact of its achieving first of all depend on the means of the adaptation setting. Every new individual (a randomly chosen vector) is compared with all the indications, set by experts (the vectors of the statistic indices). It is necessary to get the quantitative estimate of the quality of all vectors in the population – the adaptation estimate of an individual: how close it is to the task solution. This estimate will allow to make the evolutionary process directed, providing the approach to the problem.

It is necessary to determine which of the vectors is "better" and which is "worse" as the value of the aggregating indicator. The following informal statement underlies the suggested estimate: "The higher is the given input index for the aggregating indicator (the higher the weight shown by an expert), the more the new received indicator will "resemble" it". This statement is axiomatic, self-evident. Its acceptance is the logical principle of the work of the algorithm. Let's formalize this statement in order to get the adaptation estimate. For that we should define what we are going to understand by the notion "alike", meaning the index and aggregating indicator resemblance. As in this work the task of rank indicator setting is being considered, the “likeness” of indices will be defined as the “likeness” of orders, set by these indices on the multitude of objects. So the first step of formalization can be defined the following way: “alike” means “alike according to the set order”. As in this work the order, set by the index, is presented by means of rank permutation of the given index, the quantitative characteristic of the “likeness” of the indices is – informally – the measure of rank permutations closeness. Let’s pass over to the estimate of adaptation of an individual (a vector, representing a potential indicator).

Let vector V be an individual, which adaptation we need to estimate. Let’s build its rank permutation $\mathbf{S}_v$. Let’s examine the rank permutation of an arbitrary input index - $\mathbf{S}_i$. $(\mathbf{S}_i, \mathbf{S}_v)$ will designate the number of transpositions, necessary to obtain $\mathbf{S}_i$ from $\mathbf{S}_v$. $\mathbf{E}_i$ will designate the value equal to

$$\mathbf{E}_i = \mathbf{w}_i{*}(\mathbf{S}_i, \mathbf{S}_v),$$

which we will name the estimate relative to indication ψ_i. $\mathbf{E}_i$ is the measure of closeness of given individual V and index ψ_i, with regard to the weight of the latter. Now we can record the estimate of adaptation as

$$\mathbf{E} = \sum_i \mathrm{E}_i,$$

which means that the adaptation estimate is the sum of estimates concerning every indication. This formula is a formal representation of the statement given at the beginning of the paragraph. Thus, the less is the value of the estimate E, the better the adaptation of this individual.

We shall call value **E** the basic adaptation estimate of an individual.

3a. The multitude and the mechanism of exceptions. There are situations when an expert needs some correction of information, set by the indices. For example, some region has defects in the statistic form. If the indices of the additional information are not corrected, this region will have an inadequate rank in the aggregating indicator. In this case an expert’s task is to name the territory cells which statistic indices, in their opinion, are too high or too low in comparison with the real situation. Thus, the multitude of exceptions is part of the environment of population development, but exceptions must supplement and correct adaptation estimate of an individual. The instrument of estimate specification is called the mechanism of exceptions.

Let’s examine an arbitrary permutation **S** of the first numbers **n** of positive integers (fit in Fig. 1). (later on only these permutations will be considered and the specification will be omitted). Let’s divide **S** into **p** equal groups as it is shown in fig 1 (on the fig, **p** = 5). If we regard S as a rank permutation, the requirement of an object belonging to some definite groups specifies the diapason of possible ranks for the given object. In Fig. 1, where **p** = 5, these groups are named considering rank characteristics of the objects included – "maximum", "above the average", "average", "below the average", "minimum". On the account of this interpretation we'll call such groups – rank groups of permutation **S**.

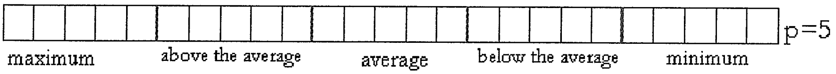

Fig. 1. Permutation division (Drawn up by authors)

Now let's give several definitions. Here and further on we'll identify an object with its number

("**k-th** object" and $\mathbf{v}_k$, are synonyms).

Definition 1: We'll say that object **k** belongs to some rank group of permutation **S**, if number **k** in this permutation belongs to this group.

Definition 2: We'll name one or several rank groups to which object **k** mustn't belong the exception multitude of object **k**.

Meaning that by setting the exception multitude it is possible to directly limit the diapason of ranks the object can have.

Definition 3: We'll say that permutation **S** has exception on object **k**, if object **k** in **S** belongs to its exception multitude.

Fig. 2 illustrates the introduced notions (fit in Fig. 2). There are two permutations presented; one has exception on object **k**, the other doesn't. The exception multitude, corresponding to object **k** is marked black; the permutation component containing number **k** (object **k**) is marked gray.

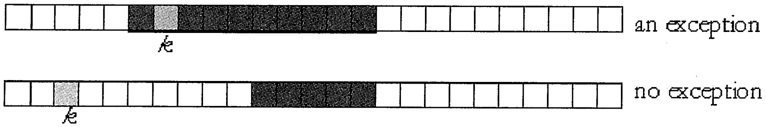

Fig. 2. Exceptions (Drawn up by authors)

Now let's get back to the task of constructing an aggregating rank indicator and we'll use these terms to add some information about the objects. The expert has an opportunity of setting an exception multitude in rank permutation of the aggregating indicator for every object, i.e. to limit the diapason of possible ranks of an object in the sought-for aggregating indicator. That is to say we make an extra demand: if vector **C** is a constructed aggregating indicator, **S** is its rank permutation, then **S** can't have exceptions on any of the objects.

Let's describe how the information set by exceptions is given. Let the permutation be divided into **p** groups for setting exceptions (in the current version of the algorithm **p** = 5, Fig. 1). Every object will be confronted by a vector with 0 or 1 length p, where 1 in the place **1** means that group **1** (with the numbering from left to right) belongs to the exception multitude of this object. This vector will be named the *object exception vector*. Out of the exception vectors for all the objects the *matrix of exceptions* is made, for which exception vectors are rows. Thus, apart from rank permutations of indicators and their weights, the considered algorithm gets the matrix of exceptions size $\mathbf{n}\times\mathbf{p}$, where n is a number of objects, **p** is a number of rank groups.

Now it is necessary to amplify the adaptation estimate to fulfill the requirement, imposed on the aggregating indicator by a set of exception multitudes. To register exceptions estimate **E** must be corrected so that for an individual with exceptions it would always be possible to point a more adaptable individual without exceptions. Then most probably the algorithm will "find" an individual, for which the requirements imposed by exception multitude are fulfilled.

Let vector V has an exception on object **k**. "To get rid" of the exception we can "move" number **k** to the left or to the right performing transpositions of number **k** with adjacent numbers. Let $\mathbf{D}_l$ be the minimum number of transpositions, necessary to get the permutation without the exception on **k**, if we move number **k** to the left. If it is impossible to get such a permutation, moving number **k** to the left, let D_l be equal to the infinity. Thus, $\mathbf{D}_l$ is a minimum "distance through the permutation" to the left up to the end of the exception multitude, corresponding to object **k**. Now we'll the same way define $\mathbf{D}_r$ for the movement to the right. Let **D** be equal the maximum of two numbers $\mathbf{D}_l$ and $\mathbf{D}_r$, on condition that neither of them is equal to the infinity. If one of them is equal to the infinity, let **D** be equal to the other. Thus **D** is "the distance to the father border of the exception multitude". This is illustrated in fig 3, which shows the permutation transformation of $\mathbf{S}_1$, which has exceptions, into permutation $\mathbf{S}_2$, which has no exceptions with the help of **D** transpositions. The exception multitude, corresponding to object **k** is marked black, the permutation component containing number **k** (object **k**) is marked gray (fit in Fig. 3).

Let's consider an estimate, set by exceptions, equal to

$$\Theta = 2*D*\sum_i w_i$$

The general adaptation estimate is the sum of the stated:

$$\Omega = E + \Theta.$$

Why is estimate Ω defined this way? Let S_1 be a permutation with exception on object **k**, S_2 is a permutation got from S_1 with the help of **D** transpositions (moving number **k** on **D** positions, Fig.3). Permutation S_1 has no exception on **k**. Let's evaluate the upper bound of main adaptation estimate **E** of permutation S_2:

$$E(S_2) \leq E(S_1) + D^* \sum_i w_i ,$$

as S_1 differs from S_2 in **D** transpositions.

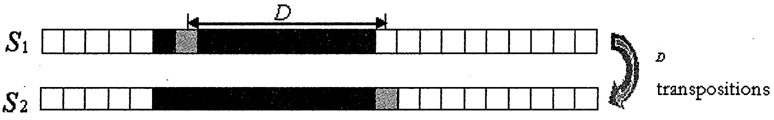

Fig. 3. Getting rid of the exceptions. (Drawn up by authors)

Adding number Θ to $E(S_1)$ according to the formula above makes the general estimate S_2 less than the general estimate S_1, thus it makes an individual without any exception on **k** more adapted than the one with this exception.

4. RULES OF THE DEVELOPMENT OF A POPULATION

We should set a number of arbitrary values, corresponding to the number of the dead/born in this "population", the number of mutations in the "generation" etc. (a generation is viewed as a population at the given moment of the process). There are also some accompanying questions, such as: "Should we limit the size of the population and if we should then by what number and by means of what mechanism?" "What should be the order of the changes (deaths, births, mutations) at the given moment of the process?"

These and other parameters set the general view of the development of the population, which determines not only the time of the work of the algorithm, but also the quality of the final solution.

It was decided to limit the size of the population as an unlimited size will lead to enormous expenses of the calculating resources and doesn't give any considerable improvement of the result. At every moment of the process at first a necessary quantity of births and mutations with the adaptation estimate is brought into the world and only then a necessary quantity of less adapted individuals is removed. Thus, even less adapted individuals in this generation are able to give birth to the individuals which can turn out to be more adapted.

The number limiting the size of the population is determined by the number of input indices and is equal to $A = C^*m$, where **C** is a constant, **m** is a number of input indices. The death limitation in a population is performed with the help of the parameters of random processes, determining the number of births and deaths (see below). The size of the population gradually increases with the slowing down of the tempos of the adaptation improvement. It is so because, the current size of the population is enough for a remarkable improvement when the adaptation changes are considerable. When the tempos of improvement of the adaptation (the approach to the solution) slow down, it is necessary to raise the resources of the process by means of increasing the size of the population.

Random variables, characterizing the number of births, deaths and mutations, are considered to be Gaussian. Their expectation and standard deviation depend on the number of individuals in the generation. Let's consider these parameters of distributions. Let **N** be the current size of the population. Then:

— the number of births:
  expectation = $\sqrt{\mathbf{N}}$,
  mean square deviation $= \mathbf{1 + \lg (1 + N)}$
— the number of deaths:
  expectation = $\mathbf{N/\sqrt{A}}$,
  mean square deviation $= \mathbf{1 + \lg (1 + N)}$
— the number of mutations:
  expectation = $\sqrt{\mathbf{N}}$,
  mean square deviation $= \mathbf{1 + \ln (1 + N)/\ln(A)}$

Here **A** is the limitation of the population size, determined in the previous point. Now it is clear why **A** is the limitation of the number of individuals: the expectation of the number of deaths is equal to the expectation of the number of births when $\mathbf{N = A}$: $\sqrt{\mathbf{N}} = \mathbf{N/\sqrt{A}}$. when $\mathbf{N<A}$ on average the births outnumber the deaths, when $\mathbf{N>A}$ there are more deaths. Low and identical variance of both arbitrary values guarantees inconsiderable oscillations around **A**. Thus, the size of the population increases approximately up to **A**, and then remains close to this number with some inconsiderable oscillation.

5. THE WORK TERMINATION CRITERION

As the exact criteria of the determination whether the solution is achieved, are unknown, it is necessary to work out an indirect criterion with the help of which it will be possible to determine that the given individual can be regarded as the final solution and that the evolutionary process needs to be terminated. Thus, answering the criterion we "suppose" that we've got "the crown of creation", and we won't get any better results or we don't have an opportunity (resources) for the continuation of the work.

To define the work termination criterion we should determine the individual we regard the solution, in other words, fundamentally the best. As the exact solution is unknown, the widespread criterion is the absence of the improvement of the adaptation within a definite number of generations. This very criterion is chosen in this work. To be more exact: if the value of the best adaptation doesn't change within **K** populations (**K** has the order of 1,000), and an individual with a better adaptation has no exceptions on any of the indications, this individual is regarded to be the solution and the work of the algorithm is completed. As the achieved indices differ even with identical input, the choice of the final solution can be made by several launches of the algorithm and averaging (or choosing the best from) the partial solutions.

Thus, the evolutionary algorithm, used to construct rank aggregating indicators, is completely set. In conclusion let's once again analyze the work of the algorithm. The rank permutations of the input indices (suggested by the experts) and weights corresponding to them are the input data. Besides, there is an opportunity to set multitude of exception for any object (not presented on the scheme not to overburden the picture). The algorithm represents a cycle of births, mutations and deaths of less adapted individuals. It results in the determination of the best adapted individual, not dominated within a large number of generations. (fit in Fig. 4).

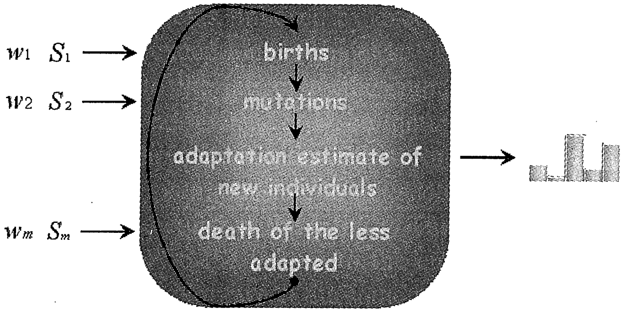

Fig. 4. A general scheme of the work of the algorithm (drawn up by authors)

Summing up the description of the method, the evolutionary algorithm can be presented as an infinite loop of births, mutations and deaths of individuals, which simulates the evolution of the population, going on according to the set rules of its development. The loop exit is performed when the situation described in point 6 arises, that is when the condition of the population corresponds to the work termination criterion.

DISCUSSION AND RESULTS

The obvious advantage of the method is that it doesn't give "the exact" results, but only sets the interval, to which this or that region should belong.

The interval is computed by means of the analysis of the whole population, consisting only of the crowns of the evolution, got after repeated launches of the algorithm. By the intervals we can judge the break in the value of the index, i.e. form the types. This rank is more adequate for the regions, countries or cities, getting one and the same rank resulted from a repeated use of the evolutionary algorithm. Whereas to define more exactly the rank of the regions, which have a considerable dispersion in the results, some additional information is necessary. Such regions most likely have some divergence from the general pattern.

The approbation of the suggested method was carried out by means of calculating the innovative potential of the regions of Russia (Table 1). On the basis of the poll of 5 experts 16 indices were selected; they determined the trend of the evolutionary process. Among them the most highly influential were the indices of the educational level, the spread of the Internet in the region, the part of city dwellers (fit in Fig 5.).

Considerable differences in the socio-economic level of development of the regions are typical for Russia. It is bound up with the spread of the tide of innovation in the vast territories of the country. A distinct division of the regions into the innovative core, subcore, subperiphery and periphery has always been obvious. And today, concerning the innovative potential, Moscow leaves far behind other regions and even stands out in the innovative core.

The regions which belong to the type “the core” aren’t a territorial aggregating formation. These are creative regions and strong acceptors, where the quickest transformation of novations into innovations takes place. The creative regions of the Sub core adjoin “the core” regions, which the innovations from the nearest core reach quickly. Subperipheral regions don’t practically produce innovations, but their introduction is faster and easier than in peripheral regions. The table shows that less than a quarter of the regions of Russia belong to the core and subcore zone, which indicates a considerable concentration of the innovative potential. Subperipheral regions adjoin the core and sub core regions, introducing the last trait to the image of nuclear territory systems. Khabarovskiy and Primorskiy regions, which don’t border on the creative regions of Russia, also belong to this type. On the one hand, they form some far eastern innovative sub center of the country; on the other hand, they are influenced by the innovations, spreading from the countries of the Asian-Pacific region (fit in Table 1).

The regions not shown in this table represent the innovative periphery, occupying a vast territory of Russia. The innovative potential of the regions, with little exception, is determined by the innovative potential of the regional center and other big cities. Within the region the innovative tides spread as a rule in the hierarchical order of the settlements, not within the territory.

Table 1. The ranking of the RF oblasts according to the innovative potential, calculated with the help of the evolutionary algorithm (made up by authors)

THE CORE	THE SUBCORE	
1. The city of Moscow	16. Leningradskaya oblast	30. Orenburgskaya oblast
2. Saint-Petersburg	17. Volgogradskaya oblast	31. Stavropol Territory
3. Moscovskaya oblast	18. Voronezhskaya oblast	32. Yaroslavskaya oblast
4. Sverdlovskaya oblast	19. Krasnoyarsk Territory	33. Udmurtskaya oblast
5. The republic of Tatarstan	The Subperiphery	34. Kurskaya oblast
6. Samarskaya oblast	20. Irkutskaya oblast	35. Penzenskaya oblast
7. Rostovskaya oblast	21. Kemerovskaya oblast	36. Vladimirskaya oblast
8. The Republic of Bashkortostan	22. Altai Territory	37. Tverskaya oblast
9. Novosibirskaya oblast	23. Omurskaya oblast	38. Orlovskaya oblast
10. Nizjegorodskaya oblast	24. Tulskaya oblast	39. Lipetskaya oblast
11. Krasnodar Territory	25. Khabarovsk Territory	40. Kirovskaya oblast
12. Saratovskaya oblast	26. Primorski Krai	41. Vologodskaya oblast
13. Tomskaya oblast	27. Tumenskaya oblast	
14. Permskaya oblast	28. Ulyanovskaya oblast	
15. Chelyabinskaya oblast	29. Riazanskaja oblast	

Source: Authors' elaboration based on using evolutionary algorithm

As the first table shows, the results of the indices aggregation turned out to be quite adequate and conceptually similar with the results of the research using other approaches, mainly at the quality level ([1]). But these results have considerable advantages over the results, achieved using the standard methods of aggregation (means the selection of indices according to the factorial analysis with later addition. The method of rank addition gave even less adequate results).

As one of the striking disadvantages we'll mention that Saint-Petersburg took only the sixth place in the aggregating rating according to the standard methods, being inferior to Bashkirya and Nizjegorodskaya region. To our mind, the first three places in the rating must be constantly taken by Moscow, Saint-Petersburg and Moscovskaya region. The largest number of universities, planning institutes, Scientific Research Institutes and experimental productions are concentrated in the two capitals. Moscovskaya region, with the most powerful generators of innovations – Dubna, Reutov, Troitsk, Chernogolovka, Pushino, Protvino, took only the 12th place.

Service and cultural innovations are the two spheres which are the quickest to root in the metropolitan area of Moscow and Saint-Petersburg. A constant growth of economical, political and institutional potential in these regions contributes to the increase of creativity and sensitivity to innovations.

Tomskaya region was another one which took only the 21st place according to the standard methods in comparison with the aggregation, using the evolutionary algorithm. Tomsk is the oldest Siberian scientific center, where the process of creating of technoparks with the succession from science to production is now taking place as well as in Novosibirsk. We can hardly agree with the 55th place of Leningradskaya region according to the standard methods. No doubt, in the course of the historical process Moscovskaya region had an opportunity to accumulate its economical, political, cultural and innovative potential, but with the regard of immediate proximity of the largest innovative center – Saint-Petersburg, the 16th place and reference to the subcore is more or less admissible.

The largest dispersion between the maximum and minimum ranks, resulted from a repeated use of the evolutionary algorithm, can be observed in Arkhangelskaya region (from 44 to 57 place), Kaluzhskaya region (from 45 to 55 place) and Brianskaya region (from 47 to 57).

Probably such dispersion can be explained by the demolition of a previously considerable innovative potential. These regions belong to the depressive type, accordingly, such major centers as Severodvinsk (the center of submarine production in Arkhangelskaya region), Obninsk (scientific center in the north of Kaluzjskaja region) experience this depression. Tomskaya region, which suffered a serious economic crisis in 1990s, has a wide rank dispersion. Thus, the use of the evolutionary algorithm provided us with a rather accurate picture of the innovative potential of the regions of Russia, allowed to reveal both territorial and hierarchical spread of innovative tides.

CONCLUSION

In this paper we suggested a new method for calculation of rank aggregating indicators, based on evolutionary algorithm. These are the main theses and results:

1. The aggregating indicator to be calculated is based on the set of quantitative characteristics, suggested by a group of qualified experts, thus providing unbiased input data. Moreover, it takes into account different influence of the input indices, using the weights, confronted to each index, suggested by an expert.
2. The statistic pattern, produced by the set of input indices, may be corrected in order to achieve more adequate and up-to-date results. This correction is made using an additional set of parameters, called exceptions, and is realized using a special estimate correction, that we called exceptions mechanism.
3. The evolutionary algorithm, underlying the method, doesn't use a provided formula. This property of evolutionary algorithms allows us to speak about impartiality of the resulting indicator. Furthermore, the black-box structure

of the algorithm provides slightly different output for identical input data, which gives additional information for analysis. For example, if the rank of an object considerably changes from run to run, it points to the lack of information about this object. The analysis of the outputs of many runs of algorithm points to intervals, to which the ranks of the objects belong. This information allows forming the clusters or types of objects.

4. The method was tested on the task of calculation of innovation potential of Russian regions. The input data, results and analysis of the test are presented in this paper. All the results are absolutely adequate and considerably better, than several standard methods with the same input data. A number of runs of the algorithm allowed determining innovative core, subcore, subperiphery and periphery of Russian regions.

A future research will be concentrated on the improvement of the algorithm in order to count not only ranks, but also the concrete values of the aggregating index. Besides, the search of the optimal and adaptive set of parameters of the evolutionary algorithm itself is obviously an important task and a field of work in the future.

REFERENCES

Bäck, Th. *Optimization by Means of Genetic Algorithms*, http://citeseer.ist.psu.edu/79660.html

Bäck, Th. and Hoffmeister F. *Global Optimization by Means of Evolutionary Algorithms*, http://citeseer.ist.psu.edu/219671.html

Hinterding, R., Gielewski, H., and Peachey T.C. 1995: The nature of mutation in genetic algorithms. *Proceedings of the Sixth International Conference on Genetic Algorithms*, L.J. Eshelman, ed. San Francisco: Morgan Kaufmann. pp.65-72

Holland, J.H. 1975: *Adaptation in natural and artificial systems*, The University of Michigan Press, Ann Arbor.

Kostenko, V.A. 2002: *The principles of constructing genetic algorithms and their use in the solving of optimization tasks*. The works of the IVth International conference Discrete models in the theory of operating systems, pp. 49-55.

Redko, V.G. 2001: *Evolutionary Cybernetics*, Moscow, Science.

Strizjov V.V. 2002: *The agreement of expert estimates when constructing aggregating indicators*, author's abstract of Ph. D. thesis, Moscow.

CORRESPONDENCE TO:

Daniel Kanevsky
Calculating Math and Cybernetics
Department of Mathematical Methods of Forcasting
Lomonosov Moscow State University, Russia

Leninskie Gory 1
+ 7 095 9394765
[e-mail: lpg@geogr.msu.su]

Pavel Chistyakov
Department of Economic and Social Geography of Russia
Faculty of Geography
Lomonosov Moscow State University, Russia
Leninskie Gory 1
+ 7 095 9394765
[e-mail: chipal@yandex.ru]