

An approach to design a composite index of economic development and identifying the bounds of its levels

Pavlo Hryhoruk
Department of Automated Systems and
Modeling in Economics
Khmelnytskyi National University
Khmelnytskyi, Ukraine
violete@ukr.net

Nila Khrushch
Department of Finance, Banking and
Insurance
Khmelnytskyi National University
Khmelnytskyi, Ukraine
nila.ukr@gmail.com

Svitlana Grygoruk
Department of Software Engineering
Khmelnytskyi National University
Khmelnytskyi, Ukraine
grygoruk.svitlana@gmail.com

Abstract— Paper considers the task of assessment of regions' economic development level using composite index assessment technology. It is proposed to design composite index by block convolution of two groups of partial indicators. First group contains indicators measured in absolute scale. Second group consists of relative measures of dynamics. For each group we proposed own procedure for data normalization and own method for data convolution in partial composite index. The procedure for identification of intervals' bounds on the composite index's scale for each level of economic development is proposed to the aim of classification of Ukraine's regions. Presented means for identification of intervals' bounds of data grouping may be used to solve other tasks independently from the way of design of composite index and set of initial indicators. The results calculated may be used in the decision-making system as source data for further calculations, as well as a criterion for determining the socio-economic development directions.

Keywords — Composite index assessment technology, regions' economic development, economic information, data normalization, convolution, bounds of the level of economic development.

I. INTRODUCTION

Solving problems of socio-economic development has always been in the category of strategic and most important issues for any country. Priorities of such development in the Government regulation system are determined by the concept of socio-economic development, which is implemented in the strategy and appropriate programmes of development. In particular, for Ukraine, such programs are reflected in [1-4]. For strategic management of the region as a socio-economic system, all its components must be considered: social, ecological and economic, the most important of which, in our opinion, is the latter. Studying the state of socio-economic development of the regions should be based on analytical approaches that comprehensively and adequately take into account changes in the processes of their development and allow to reveal its dynamics and asymmetry.

The key to making informed managerial decisions is economic information. The main indicator for identifying the individual countries development level as a whole and the comparing different countries by the economic development level traditionally considers GDP, which is as a universal measure of development and welfare [5, 6]. For interregional comparisons within one country such indicator is GRP. However, GDP as a single measure of development was subject to broad criticism [7, 8, 9, 10]. On the other hand, in our opinion, it is not worthwhile to completely abandon this

indicator and indicators associated with it. Human Development Index, Index of Economic Freedom, Index of Sustainable Economic Welfare and Global Competitiveness Index are considered as the most well-known alternatives to the traditional use of GDP [11]. In Ukraine, the level of region's socio-economic development is assessed by the measuring regional human development index [12]. It is calculated for 33 indicators, divided into six groups. In our opinion, the use of the indicated indices has its disadvantages: they are characterized by the complexity of obtaining initial information and the difficulty in visualizing the results obtained.

The idea of using a composite index assessment technology for assessing the level of regions development is reflected in the writings of many scholars [11, 13-21]. Their results are a different by a set of initial indicators for assessing the regional development level, the area of application and interpretation. At the same time, the identification and justification of the composite indicator's bounds for the established levels of development favor as a separate issue. The papers [13, 17, 21] offer their own scales for measuring the level of the regions socio-economic development and set the bounds for levels of development. In our view, the regions' socio-economic development is characterized by natural nonlinearity, while the proposed approaches involve the use of equal intervals for each level.

A rather common approach to establishing the correspondence between the quantitative indicator's measure and a qualitative characteristic is the use of the desirability functions of and corresponding verbal-numerical scales. Usually such functions are non-linear and deprived of the above-mentioned disadvantage. Paper [18] presents the Harrington's scale modification for this purpose. Such approaches are fairly widespread, but usually they are oriented towards the use of linguistic interpretation of the intervals of the desirability scale without taking into account the transformation in accordance with Harrington' function, which determines the relationship between the initial indicator and the desirability scale.

In this paper, we propose an approach to design composite index of regions' economic development and iterative procedure for identification of the interval bounds for each level.

II. PROBLEM DESCRIPTION AND METHODOLOGY

A. The model for designing composite index

Generally, the design of a composite index is carried out in such a way that its values ranged from 0 to 1. This

simplifies the meaningful interpretation of the results obtained. It is also assumed that there is a positive direction of the result: the higher the value of the composite index corresponds to the higher level of economic development. This implies the submission of all primary indicators as incentives. In addition, it is necessary to eliminate the influence of measurement units of the initial indicators on the result. Compliance with these requirements provides a procedure for normalizing the initial data, that is, transformation of their values to an interval from 0 to 1. This procedure is carried out according to the rule:

$$u_{ij} = 1 - \left| x_{ij} - x_j^* \right| / (x_{j\max} - x_{j\min}), \quad (1)$$

where u_{ij} are normalized values of indicators, x_{ij} – initial values of indicators, $x_{j\min} = \min_i x_{ij}$, $x_{j\max} = \max_i x_{ij}$, $i=1..m$, $j=1..n$, m – number of objects observed (regions), n – number of initial indicators.

Values x_j^* are identified by the formula:

$$x_j^* = \begin{cases} x_{j\max}, & \text{when } X_j \text{ is a incentive;} \\ x_{j\min}, & \text{when } X_j \text{ is a disincentive;} \end{cases} \quad (2)$$

In such cases, when indicators X_j are relative measures of dynamics, it is expedient to make transformation according to the formula:

$$u_{ij} = \begin{cases} x_{ij} / x_{j\max}, & \text{when } X_j \text{ is a incentive;} \\ x_{j\min} / x_{ij}, & \text{when } X_j \text{ is a disincentive;} \end{cases} \quad (3)$$

There are other rules to normalize data depending on the nature of the indicator and its behavior.

Usually a linear additive or multiplicative convolution is used to shape composite index Q :

$$Q_{Ai} = \sum_{j=1}^n w_j u_{ij} \quad (4)$$

$$Q_{Mi} = \prod_{j=1}^n u_{ij}^{w_j} \quad (5)$$

where Q_{Ai} , Q_{Mi} are the values of composite indexes for additive or multiplicative convolution appropriately, $i=1..m$; w_j – weight coefficients of initial indicators, $j=1..n$,

We note, there are other approaches to determine synthesizing function for composite index. There are no formal rules to use on practice concrete method of convolution. But we consider, when initial indicators X_j are measured by the absolute scale it is expedient to use additive convolution like formula (4). If initial indicators are relative measures of dynamic, we propose to use multiplicative convolution like formula (5). In that case, when among the selected indicators there are both absolute and relative, it is proposed to use block convolution according to the formula:

$$Q_i = \alpha_A Q_i^{Abs} \cdot \alpha_R Q_i^{Rel}, \quad (6)$$

where Q_i – final value of composite index, Q_i^{Abs} – value of partial composite index which is calculated by the formula like (4) for absolute measured initial indicators, Q_i^{Rel} – value of partial composite index which is calculated by the formula like (5) for relative measures of dynamics, α_A , α_R – weight coefficients of partial composite indexes.

Also, it is expedient to use block convolution in cases where the initial set of indicators contains several subsets of indicators grouped according to certain functional attributes.

Weight coefficients should meet the condition:

$$\sum_{j=1}^n w_j = 1 \quad (7)$$

A similar condition should also be fulfilled for the weight coefficients of partial composite indexes in the case of block convolution.

B. An approach to identify economic development levels bounds on the scale of composite index

We proposed an approach to identify economic development levels bounds based on the use of training set of data. Assume that the composite index is designed for studying the levels of economic development of m regions (objects of observation), its values are Q_1, Q_2, \dots, Q_m . In the future, we will consider them as points on the axis of the integral index. Let us have the location of regions into classes for similar characteristics of economic development. Each class corresponds to its level of economic development. Let's denote the number of classes through s . Consequently, the scale of the composite index is divided into s intervals. Let's indicate the initial bounds of the intervals for the levels through $I_0^{(0)}, I_1^{(0)}, \dots, I_{s-1}^{(0)}, I_s^{(0)}$, where $I_{i-1}^{(0)}$ and $I_i^{(0)}$ are bounds of i -th interval corresponds to i -th class (i -th level of economic development); $i=1..s$; $I_0^{(0)}=0, I_s^{(0)}=1$. Let's denote through $x_i^{(0)}$ – the extreme right point belonging to the i -th interval (it represents the largest value of the composite index for an object that belongs to the i -th level of the investigated quality), and through $y_{i+1}^{(0)}$ – the extreme left point belonging to the next, $i+1$ -th the interval (it represents the smallest value of the integral index for an object that belongs to the $i+1$ th level of economic development). The recalculation of the intervals' bounds will be carried out by an iterative process, which is determined by the formula:

$$I_i^{(k)} = \left(y_{i+1}^{(k-1)} \cdot n_{i+1}^{(k-1)} + x_i^{(k-1)} \cdot n_i^{(k-1)} \right) / \left(n_{i+1}^{(k-1)} + n_i^{(k-1)} \right) \quad (8)$$

where k is the step number of iterative process, $k=1, 2, \dots$; $n_i^{k-1}, n_{i+1}^{(k-1)}$ – number of point that are belonged to i -th and $i+1$ -th intervals respectively, $i=1..s-1$. After that, redistribution of points between the new intervals is carried out. The iterative process will be completed if the number of points for each interval does not change at that, that is, the condition will meet:

$$n_i^{k-1} = n_i^{(k)}, i = 1..s. \quad (9)$$

In accordance to formula (8), the more points belong to the interval, the more clearly defined its bounds. So, the boundary values that determine the interval's bounds, will be closer located to the extreme points of this interval.

When organizing the iterative process by the formula (8), a situation may arise when at the one of the next steps of the conversion there will be no points into some intervals. In this case, we propose don't recalculate bounds of these intervals in the next steps, but left unchanged.

The accuracy of calculations can be increased if the training set of data will contain samplings over several periods. In this case, recalculation of the intervals of intervals at each step of the iterative process should be performed for each such sampling individually, and the final values of the interval's bounds in this step are obtained by calculating the average values according to the data of each sampling. If at the next step of the iterative process for one of the sampling one or more intervals are empty (doesn't contain any points), the calculation of the average value for intervals' bounds is proposed to be made according to the results of calculations for other samplings.

Other approaches to calculating intervals' bounds can be proposed, in particular, for this aim we can use the average value of the extreme points of the corresponding intervals; centers of intervals.

III. RESULTS AND DISCUSSIONS

Consider the construction of the composite index of economic development of the Ukraine's regions and identify the bounds of the intervals of economic development levels. The analysis of scientific publications in this area [13, 17, 18, 22, 23], as well as own authors' studies, make it possible to conclude that as the initial indicators determining the economic development of regions, it is expedient to select such indicators: X_1 – capital investment, mil. UAH; X_2 – export volume, mil. USD; X_3 – employment rate, %; X_4 – unemployment rate after the methodology of ILO, %; X_5 – volume index of GRP in prices of the previous year, %; X_6 – industrial output index, % to the previous year; X_7 – agriculture output index, % to the previous year. To provide calculation we form training set of data including indicators values for 2015-2017 [24]. First four indicators are values measured on absolute scale, so for their normalization we use formula like (1). Indicators $X_5 - X_7$ are relative measures of dynamics, therefore we use formula like (3) to normalize them. In this case, the extreme (maximum and minimum) values for normalization of each indicator will be calculated not on a separate sampling in one year, but on all data in three years. Let's also take into account the fact that the indicator X_4 is a disincentive, and the rest of the indicators are incentives. We calculate the partial composite index for the first five indicators by the formula like (4) and for the last three - by the formula like (5).

Since we don't have any information on the significance of the selected initial indicators, we will consider them to be equivalent. Therefore, the values of weight coefficients will be considered the same. Accordingly, for the first group of indicators values are $w_j = 1/4, j = 1..4$; for the second group $w_j = 1/3, j = 1..3$.

The final values of the composite index of economic development are calculated by the formula (6). In this case, we also consider values of weight coefficients for partial composite indexes calculated by the data of each group are equal: $\alpha_A = \alpha_R = 1/2$. Values of the composite index of economic development are shown in the Table I.

TABLE I. THE VALUES OF COMPOSITE INDEX

Region	The values of composite index Q		
	2015	2016	2017
Vinnitsia	0.58	0.59	0.55
Volyn	0.48	0.42	0.35
Dnipro	0.81	0.81	0.85
Donetsk	0.37	0.47	0.48
Zhytomyr	0.48	0.52	0.54
Zakarpattia	0.52	0.52	0.51
Zaporizhzhia	0.60	0.59	0.61
Ivano-Frankivsk	0.53	0.53	0.58
Kyiv	0.68	0.76	0.75
Kyrovohrad	0.43	0.45	0.43
Luhansk	0.24	0.42	0.30
Lviv	0.59	0.63	0.68
Mykolaiv	0.59	0.62	0.58
Odesa	0.63	0.68	0.69
Poltava	0.49	0.50	0.52
Rivne	0.54	0.52	0.49
Sumy	0.50	0.54	0.58
Ternopil	0.37	0.42	0.41
Kharkiv	0.63	0.70	0.71
Kherson	0.50	0.49	0.51
Khmelnitskyi	0.45	0.53	0.56
Cherkasy	0.51	0.53	0.53
Chernivtsi	0.48	0.52	0.55
Chernihiv	0.49	0.50	0.52

We determine five levels of economic development with appropriate identification of five groups of regions. First group contains regions with a relatively high level of economic development; second group contains regions with a level of economic development above the average value; third group consists of regions with an average level of economic development, fourth group contains regions with a level of economic development below the average value and fifth group consists of regions with a relatively low level of economic development.

To make calculations, we assume that each group contains at least one region at least one-year data. Also, we set the initial values of intervals' bounds so that the value of the composite index, equal to 0.5, is covered by the third interval corresponding to the group of regions with the average level of economic development. Initial values of intervals' bounds are presented in the Table II, and final bounds obtained after 10 iterations are shown in Table III.

TABLE II. THE INITIAL GROUPS' BOUNDS OF COMPOSITE INDEX

Bounds	Group number				
	<i>first</i>	<i>second</i>	<i>third</i>	<i>fourth</i>	<i>fifth</i>
Left bound	0.80	0.65	0.45	0.31	0.00
Right bound	1.00	0.79	0.64	0.44	0.30

TABLE III. THE FINAL GROUPS' BOUNDS OF COMPOSITE INDEX

Bounds	Group number				
	<i>first</i>	<i>second</i>	<i>third</i>	<i>fourth</i>	<i>fifth</i>
Left bound	0.69	0.55	0.45	0.34	0.00
Right bound	1.00	0.68	0.54	0.44	0.33

Results presented in table 3 show that third interval has a small width, which is entirely consistent with its economic interpretation: it corresponds to average level of regions' economic development. Note, that first and fifth intervals are quite broad, which somewhat does not correspond to the economic interpretation of the appropriate economic development level. This can be explained by the fact that the corresponding intervals initially contained a small number of points.

In order to reduce the width of these intervals, we propose to use two artificial points of the composite index with values $Q_0=0$ and $Q_{m+1}=1$ (in conditions that such points do not belong to the studied set of values of the composite index). First point gets into the first interval, the second – into the last one. Accordingly, the number of points for these intervals will increase by 1. In this case, the bounds will be wider for those intervals that contain more points. The results of calculations by this method are shown in Table IV.

TABLE IV. THE GROUPS' BOUNDS OF COMPOSITE INDEX AFTER RECALCULATION

Bounds	Group number				
	<i>first</i>	<i>second</i>	<i>third</i>	<i>fourth</i>	<i>fifth</i>
Left bound	0.78	0.64	0.45	0.32	0.00
Right bound	1.00	0.77	0.63	0.44	0.31

Results presented in the Table IV show that width of the first interval was reduced, in so doing width of the third interval was increased. The width of the remaining intervals has practically not changed. In relation to the fifth interval, this result can be explained by the small number of points in the training set of data that initially got into this interval.

In our view, obtained outcome is also quite acceptable and can be used in the future for practical application in determining the levels of economic development of the regions in subsequent periods.

IV. CONCLUSION

In this paper, an approach to design the composite index of regions economic development is presented and the means for identification of the bounds of its intervals based on the training set of data are described. The results obtained are consistent with the economic interpretation of the identified bounds of economic development levels. The designed scale

can be used for practical application in determining the levels of economic development of the regions. The proposed approach to identifying the intervals' bounds may be applied to any scale regardless of the measure units, the number of determined levels of investigated quality and the training set of data. The results calculated may be used in the decision-making system as source data for further calculations, as well as a criterion for determining the socio-economic development directions.

REFERENCES

- [1] <http://zakon3.rada.gov.ua/laws/show/385-2014-%D0%BF#n11>
- [2] <http://www.slideshare.net/KyivSchoolofEconomics/ss-38253878>.
- [3] <http://zakon.rada.gov.ua/laws/show/5/2015>
- [4] http://www.ua.undp.org/content/dam/ukraine/docs/SDGreports/UNDP_Strategy_v06-optimized.pdf
- [5] M. D. Oțil, A. Miculescu, and L. M. Cismaș, "Disparities in Regional Economic Development in Romania", *Scientific Annals of Economics and Business*, vol. 62(s1), pp. 37-51, 2015
- [6] Tb. Ai. Munandar, R. Wardoyo, "Fuzzy-Klassen Model for Development Disparities Analysis based on Gross Regional Domestic Product Sector of a Region", *International Journal of Computer Applications*, vol.123 (7), pp. 17-22, 2015
- [7] E. Dhimitri, A. Gabeta, M. Bello, "Regional Development Theories And Models, A Comparative Analyse. Challenge Of Regional Development In Albania", *The USV Annals of Economics and Public Administration*, vol. 15(2(22)), pp. 51-59, 2015
- [8] [http://www.europarl.europa.eu/RegData/etudes/BRIE/2016/577953/EPRS_BRI\(2016\)577953_EN.pdf](http://www.europarl.europa.eu/RegData/etudes/BRIE/2016/577953/EPRS_BRI(2016)577953_EN.pdf)
- [9] http://www.foreurope.eu/fileadmin/documents/pdf/Workingpapers/WWWforEurope_WPS_no056_MS211.pdf
- [10] A. F. Ivković, "Limitations of the GDP as a measure of progress and well-being", *Review of Contemporary Business, Entrepreneurship and Economic*, iss. 29(1), pp. 257-272, 2016
- [11] A. Jindrova, "Composite indicator in evaluation of the regional disparities in quality of life in the Czech Republic", *Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*, vol. 61(7). 7, pp. 2239–2248, 2013. doi: <https://doi.org/10.11118/actaun201361072239>
- [12] http://www.uazakon.com/documents/date_1a/pg_ibcnog/index.htm
- [13] H.V., Azarova, "Integral Evaluation of the Level of Socio-Economic Development of Ukraine's Regions", *Economics and Law. Series "Economics"*, vol. 1, pp. 61-65, 2015. (in Ukrainian)
- [14] V. Babenko, M. Pasmor, Ju. Pankova, and M. Sidorov, "The place and perspectives of Ukraine in international integration space", *Problemy i Perspektivy v Menedzhmenti (Problems and Perspectives in Management)*, vol. 15 (1), pp. 80-92, 2017
- [15] E. Poledniková, "Comparing regions' ranking by MCDM methods: the case of visegrad countries", *WSEAS Transactions on Business and Economics*, vol. 11(1), pp. 496-507, 2014
- [16] M. Staničková, and L. Melecký, "Multidimensional Approach to Evaluation of Visegrad Countries' Competitiveness in Comparison with Austria and Germany", *Journal of Competitiveness*, vol. 4(3), pp. 58-76, 2012
- [17] I. Parshyn, "Integral Assessment of the Regional Economic Development Efficiency", *Economics of development*, vol. 1, pp. 27-33, 2014. (in Ukrainian)
- [18] U. S. Savkiv, "Integral Estimates of Sustainable Development of the Region", *Business Inform*, vol. 1, pp. 45-50, 2012. (in Ukrainian)
- [19] <http://www.economy.nayka.com.ua/?op=1&z=4896>
- [20] D. F. Meyer, J. De Jongh, and N. Meyer. "The Formulation of a Composite Regional Development Index", *International Journal of Business and Management Studies*, vol. 8(1), pp. 100-116, 2016
- [21] D. F. Meyer, and J. De Jongh, "An Alternative Multi-Dimensional Regional Economic Development Index: A Provincial Application In South Africa", *International Journal of eBusiness and eGovernment Studies*, vol. 10(1), pp. 97-113, 2018
- [22] E. Minarčíková, "Evaluation of Regional Disparities in Visegrad Four based on Selected MCDM Methods", *In Proceedings of Online Conference: 9th International Days of Statistics and Economics*, pp. 1128-1137, 2015
- [23] Tb. Ai Munandar, E. Winarko, "Regional Development Classification Model using Decision Tree Approach", *International Journal of Computer Application*, vol. 114 (8), pp. 28-33, 2015
- [24] <http://www.ukrstat.gov.ua/>