American Journal of Theoretical and Applied Statistics
2018; 7(6): 215-221
http://www.sciencepublishinggroup.com/j/ajtas
doi: 10.11648/j.ajtas.20180706.13
ISSN: 2326-8999 (Print); ISSN: 2326-9006 (Online)

National Innovation Systems Archetypal Analysis

Joseph Gogodze

Institute of Control Systems, Techinformi, Georgian Technical University, Tbilisi, Georgia

Email address:
jgogodze@gmail.com

To cite this article:
Joseph Gogodze. National Innovation Systems Archetypal Analysis. *American Journal of Theoretical and Applied Statistics.* Vol. 7, No. 6, 2018, pp. 215-221. doi: 10.11648/j.ajtas.20180706.13

Received: September 7, 2018; **Accepted**: September 18, 2018; **Published**: October 31, 2018

Abstract: The national innovation system (NIS) determines the innovation capability of a country, and its economic development. However, recently, very little is known regarding the determinants of NIS functioning in various countries. Probably the easiest way to obtain such an understanding is to begin with the structural representation of the NIS. Particularly, it is quite natural to assume that there exists several 'cornerstone type NIS' or 'archetypal NIS', and all the other types can be considered a mixture of them. The aim of this paper is to somewhat study the advances in the structural understanding of the NIS. For this purpose we conducted our study based on the data set from the Global Innovation Indexes' (GII) seven pillars and using archetypal analysis. It is also important to note that the concept of entropy was also naturally determined under archetypal analysis. We demonstrate that each NIS can be considered a mixture of three archetypical NISs, which are as follows: The first one is a prototype of a highly developed NIS (with a high level GII score and a low level of entropy); the second one is a prototype of an underdeveloped NIS (with a low level GII score and a low level of entropy); and the third one is an intermediate form of NIS (with a medium level GII score and a high level of entropy). Hence, we establish that such a multidimensional phenomenon, such as the NIS (described in this study as the 7-dimensional vector – GII pillars), with an acceptable level of the accuracy, essentially can be considered a 2-dimensional object; and the corresponding barycentric coordinates are a convenient means of describing NISs. We also introduce an important indicator – the NIS entropy – which characterises the level of the disorder or randomness in the NIS.

Keywords: Statistical Data Analysis, Archetypal Analysis, National Innovation System

1. Introduction

In today's world, fostering innovation is essential to improving economic growth and competitiveness across countries (see e.g. [1-6]). Given that an innovation activity can be efficient only in the appropriate enabling environment, namely, in the country's well-functioning national innovation system (NIS). A NIS determines the innovation capability of a country, and it can be seen as a socio-economic system where different actors along with, formal and informal institutions interact. A NIS necessarily exploits all available resources in a country. Moreover, it requires the generation and dissemination of knowledge, in addition to the utilisation of innovation. Finally, the results obtained by a NIS can aid in achieving economic development.

Current scientific literature seems to focus less on the determinants of a NIS functioning in different countries. Probably, the easiest way to obtain such an understanding is to begin with its structural representation. Especially, it is quite natural to assume that there exist several "cornerstone type NIS" or "archetypal NIS", and all other NIS can be considered as mixtures of them. This can be very simply interpreted, for instance, that the weight of a "cornerstone type NIS"/ "archetypal NIS" in a given NIS (i.e. the mixture) corresponds to the share of adherences of the either type nationally. Also, note that, generally "cornerstone type NIS"/"archetypal NIS" are not necessarily observed, and they may represent only some "ideal points".

The aim of this study is to identify a possible "cornerstone type NIS"/"archetypal NIS" and represent (at least approximately) all other NISs in the form of appropriate mixtures of archetypal NIS's. A convenient tool for conducting such a study is the archetypal analysis (AA), which was introduced in [7] (see also [8-10]). AA aims at finding a few extreme observations, also called archetypes, in a multivariate dataset, so that all other observations can be represented as convex combinations of archetypes. This means that we obtain the archetype convex hull for multivariate data representation. It is noticeable that data points lying inside the archetype convex hull are exact convex combinations of the archetypes, while data points lying outside are only approximated. The difference reflects a loss of information and is similar to the loss of information arising from truncating the number of eigenvectors used in principal component analysis (PCA).

In this study, we introduce the AA for investigating NIS. We extract extreme observations (archetypal NIS) based on the Global Innovation Index's (GII), [11], and the seven composition component (pillar) scores. In other words, we conduct the AA for the data set which is formed by the 7-dimensional vectors presented by the GII pillars for each country under consideration. Also note that the GII is a modern tool for measuring the NIS in its full multidimensional emergence and, hence, quite appropriate for our goals in this study.

The article is organised as follows: In Section 2, we outline the material and methods used in our analysis. Particularly, we briefly present the data description and AA technique. Section 3 presents the results and, finally, Section 4 concludes the article.

2. Materials and Methods

2.1. Data

This study used the GII data for the period 2011–2015 at the pillar level. Note that we decided to use this data because it reflects the extensive experience of the previous studies and the current understanding of NISs' and the mechanisms behind their functioning. Besides, the GII is regularly published and contains detailed data on more than 100 countries. In addition, it uses well-defined measurement tools, and both the primary data and final indicators of the GII are subject to multiple external and internal tests. In particular, it is audited annually by the European Commission Joint Research Centre, which not only checks the conceptual and statistical consistency of its structure, but also the impact of the crucial modelling assumptions based on the GII scores and ranking.

The GII, [11], is built on a hierarchical basis and includes the following two sub-indices: Innovation Input, which is the composite (averaging) of five input indexes (pillars), and Innovation Output, which is the composite (averaging) of two output indexes (pillars). Each pillar is divided into sub-pillars, each of which is built using a number of relevant individual indicators. In this study, we refer to values of the GII pillars for the period 2011–2015.

Table 1. *A Short Description of the GII Pillars.*

Pillar	Short Description
I11	Institutions (Political Environment, Regulatory Environment, Business Environment)
I12	Human capital & research (Education, Tertiary Education, R&D)
I13	Infrastructure (ICT, General Infrastructure, Ecological Sustainability)
I14	Market Sophistication (Credit, Investment, Trade & Competition)
I15	Business Sophistication (Knowledge Workers, Innovation Linkages, Knowledge Absorption)
I21	Knowledge & technology outputs (Knowledge Creation, Knowledge Impact, Knowledge Diffusion)
I22	Creative outputs (Intangible Assets, Creative Goods & Services, Online Creativity)

Note: Pillar I21 "Knowledge and technology outputs" and Sub-pillar "Ecological sustainability" were named "Scientific outputs" and "Energy," respectively, in the GII 2011. Sub-pillar "Online creativity" was absent from the GII 2011.

The GII is the simple average of its Input and Output sub-indices. Moreover, the sub-indices are the simple average of their underlying pillar scores. Each pillar score is calculated as the weighted average of its sub-pillar scores, and each sub-pillar score is calculated as the weighted average of its individual indicators. The individual indicators (their numbers and compositions change from year to year, and they added up to 79–84 in the period 2011–2015) are obtained from various sources and scaled to be comparable across countries, by a division using the relevant scaling factor.

Individual indicators are also normalised to the [0, 100] range, with higher scores representing better outcomes. Such normalisation was obtained through the min-max method. Each year, the most recent value is used for each individual indicator. The details about their composition, data sources, processing techniques, and country selection methods can be obtained in [11]. For the readers' convenience, Table 1 represents short descriptions of the GII pillars.

Of course, the original dataset of GII pillars contain missing values from some country-years during the considered period (about 6% of country-years cases). We conducted data imputation by using the "first-last accessible value and linear interpolation method". Such an approach seems quite justified, because the considered indicators were characterised by considerable inertness. Therefore, we finally obtained a set of our primary data (seven indicators for 147 countries for the years 2011–2015), which will be called the GII primary dataset.

2.2. Methodology

AA was introduced in [7] as a dimensionality reduction statistical technique for the problem of multivariate data analysis. The general approach of the AA is to approximate each point of a data set as a convex combination of some (not necessarily observed) ideal types or archetypes. At the same time, archetypes must be mixtures of individual data points. Therefore, to identify archetypes, AA minimise the squared error in representing each data point as a mixture of archetypes. More precisely, let $X = [x_1, \dots, x_n] \in R^{m\times n}$ be a matrix of multivariate dataset, where each $x_i \in R^m (i = 1, \dots, n)$ represent some data point. The AA looks for a matrix $Z = [z_1, \dots, z_p] \in R^{m\times p}$ of archetypes $z_j \in R^m (j = 1, \dots, p)$ under the following requirements: each data vector $x_i (i = 1, \dots, n)$ should be well approximated by a convex combination of archetypes, and each archetype $z_j (j = 1, \dots, p)$ should be a convex combination of data points. Let us introduce the following notations: $A = [a_1, \dots, a_n] \in R^{m\times n}, B = [b_1, \dots, b_n] \in R^{m\times p}$, where $a_i \in \Delta_n (i = 1, \dots, n)$ named alpha coefficients, $b_i \in \Delta_p (j = 1, \dots, p)$ named beta coefficients, and for any natural number $N \geq 1$ simplex Δ_N defined by equality

$$\Delta_N = \left\{ \alpha = \in R^N | (\alpha_i \geq 0, i = 1, \dots, N) \; and \; \left(\sum_{i=1}^{N} \alpha_i = 1 \right) \right\}$$

Also, note that for any $\alpha = (\alpha_1, \dots, \alpha_N) \in \Delta_N$ can be defined entropy $E(\alpha) = -\frac{1}{N} \sum_{i=1}^{N} \alpha_i ln\alpha_i$.

Now, the aforementioned requirements can be formulated as follows: each vector x_i should be close to $Za_i (i = 1, \dots, n)$and each z_j has representation $z_j = Xb_j (j = 1, \dots, p)$ for the same $b_i \in \Delta_p (j = 1, \dots, p)$. Hence, we arrive at the following nonlinear programming problem, which is reflected in the essence of:

$$\begin{cases} RSS = ||X - XBA||_F^2 \to inf \\ a_i \in \Delta_n (i = 1, \dots, n), \\ b_i \in \Delta_p (j = 1, \dots, p). \end{cases}$$

where $||\cdot||_F$ is Frobenius norm of a matrix, i.e. for a matrix $A \in R^{m\times n}$, $||A||_F = (\sum_{i=1}^{n} \sum_{k=1}^{m} |a_{ik}|^2)^{1/2}$. Therefore, we must minimise the residual sum of squares (RSS) under the corresponding simplex-constraints. Respectively, the archetypes are represented by equality $Z = XB$.

Thus, the implementation of AA consists of two of the following components: The procedures for determining the number of archetypes, p and the method of solving the aforementioned nonlinear programming problem. Unfortunately, there is no clear and unified rule for determining the correct number of archetypes p which is necessary to conduct the AA for different p to analyse the behaviour of the RSS (as a function of p) and achieve an appropriate estimation of the archetypes number. Also, note that a simple and common method in practical statistics is to use the “elbow criterion” on the RSS (as a function of p), where a flattening of the curve indicates the correct value of p. However, solving the AA nonlinear programming problem is also associated with difficulties, because the algorithm requires restarting several times with different initial points to achieve the “global minimum” of the RSS.

For conducting the AA, the publicly available R package archetypes [1] was used in this study. Moreover, for result visualization, we used the R packages rworldmap[2], which is also publicly available.

3. Results

3.1. Descriptive Statistics

As stated before, the seven indicators (GII pillars) for the 147 countries for the years 2011-2015 compiled a set of primary data, which was at our disposal. With the goal of making our further analysis more robust, we decided to conduct a time-averaging of the data. Furthermore, we named this (averaged) dataset as the GII-dataset.

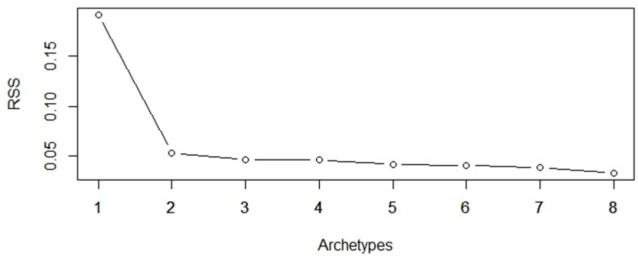

Figure 1. *RSS scree plot for the GII-dataset.*

Table 2 represents the simple descriptive statistic for the GII-dataset. As we can see, the GII-dataset distributions of all the pillars are to a certain extent right-tailed and, at the same time, distributions of all the pillars are platykurtic, i.e. their distributions produce less extreme outliers as compared to the normal distribution.

3.2. Archetypal Analysis

Generally speaking, given the seven indicators for the 147 countries, it is not clear how many archetypes are reasonable for the existing data description. Following the recommendations given in section 2.2., we performed a numerical analysis of the nonlinear programming problem for the RSS. Specifically, we considered the RSS with up to eight potential archetypes and solved the corresponding nonlinear programming problems each time with three repetitions of the different initial points to achieve the “global minimum”. Figure 1 shows a corresponding scree plot for the best model in each case. Accounting for the essential flattening on the RSS graph that was observed for three archetypes (see Figure 1), we can conclude, in accordance with the “elbow criterion”, that three (maybe) is the best number of archetypes in in this case.

---

1 Comprehensive R Archive Network (CRAN): https://cran.r-project.org/web/packages/archetypes.

2 Comprehensive R Archive Network (CRAN): https://cran.r-project.org/web/packages/rworldmap.

Table 2. *Descriptive statistics for the GII-dataset.*

	I11	I12	I13	I14	I15	I21	I22
Mean	61,30	33,44	33,29	45,43	36,05	28,05	33,41
Median	59,60	31,38	32,5	43,22	33,62	24,50	32,58
Standard Deviation	16,75	13,91	11,23	11,78	10,79	12,21	12,11
Kurtosis	-0,65	-0,59	-0,43	0,75	0,21	0,44	-0,04
Skewness	0,18	0,52	0,47	0,91	0,70	0,96	0,16
Minimum	20,54	10,26	13,30	23,04	9,24	2,76	2,88
Maximum	94,30	66,70	62,92	83,96	71,00	65,76	64,26
Count	147	147	147	147	147	147	147

Now, we must perform a suitable analysis to attain an understanding of what the obtained archetypes represent in the ideal sense. Apparently, the best way to achieve this goal is a graphical representations of the archetypes, , for instance Figure 2 (also refer to the numerical representations of the archetypes vectors in Table 3).

Table 3. *The archetypes composition by pillars.*

Pillar	A1	A2	A3
I11	91,24238	69,36485	36,03577
I12	62,00475	30,54890	16,49293
I13	51,17822	45,78206	16,68858
I14	69,93281	46,00967	32,58522
I15	61,61020	34,46357	24,26086
I21	58,85278	13,94363	18,00570
I22	54,55286	44,12805	14,34609

We begin our visual inspection of the archetypes from the bar plot in Figure 2, where archetypes are presented as follows: archetype A1 is dark grey, archetype A2 is grey, and archetype A3 is light grey. As we can see, archetype A1 has superior values for all the pillars compared to the archetype A2; on the other hand, archetype A3 has the lowest values for all the pillars in comparison with the archetype A2, except Pillar I21 (Knowledge & Technology outputs) for which the values in archetypes A2 and A3 are comparable in magnitude, but with lesser superiority in archetype A3.

Consider now the position of the archetypes in relation to the set of primary data represented by the GII-dataset. For this purpose, let us consider the parallel coordinate plot in Figure 3. Note that the representatives of the GII-dataset are visualised in grey. Also, depicted are the parallel coordinate plots of archetypes A1 (solid red line), A2 (dotted red line), and A3 (long dash red line). Figure 3 confirms the previous conclusions as a whole; but, at the same time, allows us to consider somewhat more deeply the issue connected with the aforementioned anomaly of behaviour of the archetypes A2 and A3 on the pillar I21 (Knowledge & Technology outputs).

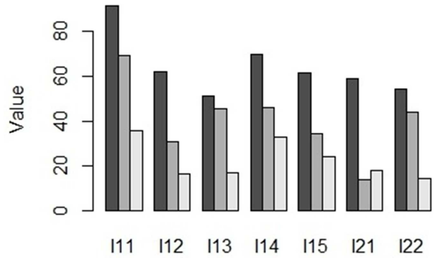

Figure 2. *Visualizing archetypes by bar plots.*

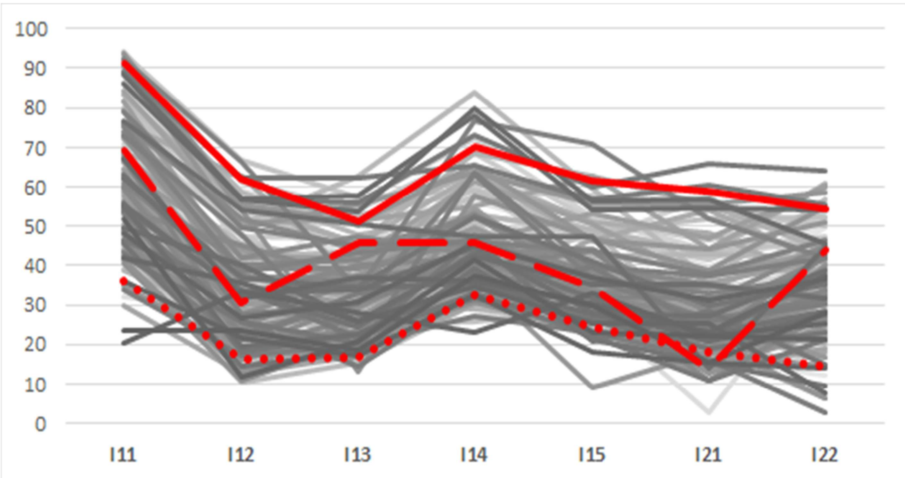

Figure 3. *Parallel plot for the GII-dataset and archetypes.*

Regardless of not being in possession of a clear and unambiguous explanation of this fact, we see from Figure 3 that the archetypes reflect the characteristic behaviour for the respective groups of countries. Furthermore, we can assume that this anomaly is connected to the poor quality of information provided by the countries that are "close" to archetype A3 (Table 4, represents archetypes composition by countries). In other words, the countries which are close to archetype A3, in most cases, overestimate the primary indicators constituting the pillar "Knowledge and Technology outputs" (e.g. domestic resident patent applications, new businesses, and so on, [11]). Therefore, we are inclined to believe that the value of the pillar I21 for archetype A3 is no greater than for archetype A2 in reality.

Table 4. *The archetypes composition by countryes.*

Archetype	Country Code/ Wight (nonzero beta coefficient)						Sum
A1	DNK	FIN	SGP	SWE	CHE		0,99985
	0,04891	0,23080	0,24400	0,05798	0,41817		
A2	BTN	ISL	ARE	URY			1,00014
	0,56716	0,17244	0,10880	0,15173			
A3	BDI	GIN	LAO	MMR	SDN	SWZ	0,99997
	0,09574	0,19028	0,22680	0,37671	0,07639	0,03406	

Based on the above considerations, we are also inclined to believe that the obtained archetypes represent the base structures of the NIS and also reflect the typical approaches to its functioning. To explicitly show the relationship of the archetypes with the special groups of countries, we have provided another visualisation of the GII-dataset through their barycentric coordinates (alpha coefficients). The barycentric (triangle) plot for GII-dataset is presented in Figure 4, where each archetype corresponds to the vertex of an equilateral triangle. We also highlighted in Figure 4 the top ten and the bottom ten NISs[3], according to the GII-dataset.

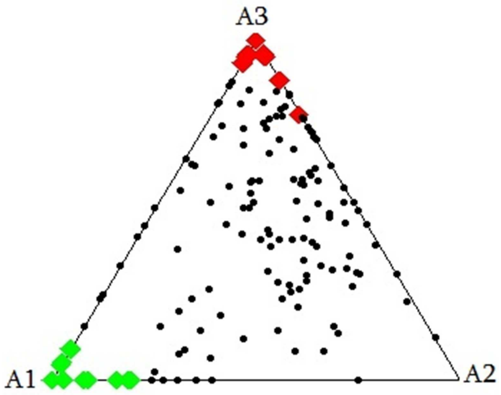

Figure 4. *Triangle plot for the GII-dataset and archetypes.*

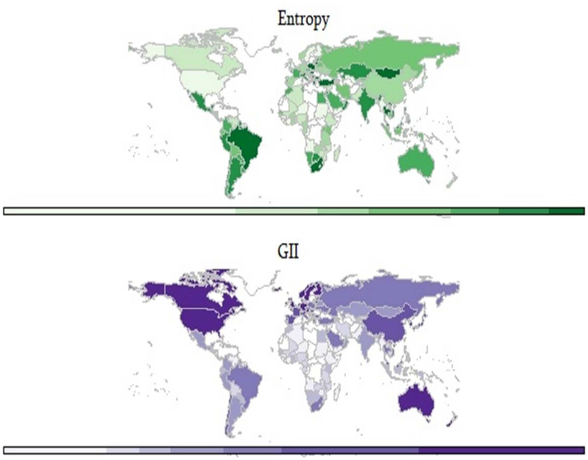

Figure 5. *Geospatial distributions of the GII and NIS Entropy.*

3 Top ten NIS: CHE, SWE, GBR, SGP, NLD, FIN, USA, HKG, DNK, IRL; Bottom ten NIS: AGO, DZA.GIN, NER, BDI, LAO, TGO, YEM, MMR, SDN.

As we can see in Figure 4, the NISs (countries) with the highest GII scores are concentrated near the vertex corresponding to archetype A1 and the NISs (countries) with the lowest GII scores are concentrated near the vertex corresponding to archetype A3. The other NISs (countries) occupy some intermediate position and archetype A2 is not attractive.

Therefore, the analysis of the empirical data carried out shows that a multidimensional phenomenon such as the NIS (note that in this study we use 7-dimensional vector representations for the NISs) is essentially a 2-dimensional object. Moreover, barycentric coordinates are a convenient means of describing NISs. Accounting for the fact that the alpha and beta coefficients belong to the corresponding simplexes, we can define the entropies for each country and each of the archetypes. In particular, the archetypes entropies are: E(A1) = 0.27250, E(A2) = 0.23088, E(A3) = 0.31183.

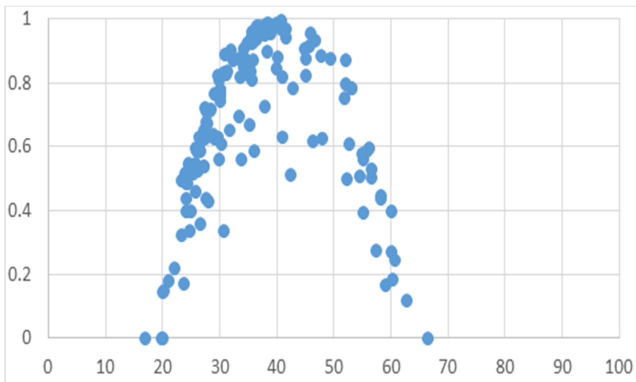

Figure 6. *Entropy vs GII score for NISs.*

Note:Vertical axis: entropy, Horizontal axis: GII score (refer to main text for explanation)

The NISs' barycentric coordinates and entropies are presented in Table 5, and Figure 5 illustrates the corresponding geospatial distribution of these indicators. Figure 6 shows the empirically established relation between the NIS GII score and NIS entropy. This illustration indicates that the NIS entropy is an important indicator and deserves a more detailed study.

Table 5. *Barycentric coordinate and Entropy for NISs.*

Country	Alpha1	Alpha2	Alpha3	E	Country	Alpha1	Alpha2	Alpha3	E
ALB	0,08045	0,44025	0,47926	0,83418	LVA	0,37478	0,48126	0,14395	0,90915
DZA	0,00423	0,21369	0,78207	0,49618	LBN	0,27855	0,21625	0,50522	0,93948
AGO	0,04675	0,00000	0,95329	0,17185	LSO	0,00000	0,50021	0,49974	0,63094
ARG	0,16099	0,43267	0,40632	0,93067	LTU	0,28965	0,49526	0,21506	0,94430
ARM	0,24472	0,23177	0,52352	0,93039	LUX	0,72783	0,27227	0,00000	0,53289
AUS	0,63140	0,27834	0,09032	0,78595	MKD	0,25486	0,31496	0,43018	0,97865
AUT	0,60877	0,39128	0,00000	0,60922	MDG	0,00000	0,27196	0,72803	0,53268
AZE	0,09476	0,32962	0,57560	0,82563	MWI	0,12990	0,05614	0,81400	0,54096
BHR	0,26256	0,47017	0,26725	0,96357	MYS	0,53957	0,22342	0,23707	0,91842
BGD	0,02572	0,16955	0,80473	0,51871	MLI	0,07114	0,10930	0,81958	0,53982
BRB	0,49514	0,00000	0,50494	0,63086	MLT	0,57858	0,35587	0,06559	0,78550
BLR	0,34143	0,02262	0,63602	0,67396	MUS	0,23187	0,47855	0,28956	0,95617
BEL	0,65126	0,27291	0,07589	0,75493	MEX	0,15781	0,44970	0,39247	0,92647
BLZ	0,10020	0,32073	0,57906	0,82978	MDA	0,35157	0,12598	0,52250	0,88081
BEN	0,00000	0,22943	0,77057	0,49024	MNG	0,20670	0,43927	0,35401	0,96015
BTN	0,00000	0,87664	0,12329	0,33996	MNE	0,28141	0,44029	0,27828	0,97755
BOL	0,08989	0,20248	0,70763	0,71423	MAR	0,08855	0,39205	0,51937	0,83926
BIH	0,33335	0,03352	0,63319	0,70032	MOZ	0,18728	0,00000	0,81278	0,43892
BWA	0,13783	0,33566	0,52649	0,88960	MMR	0,00000	0,00000	1,00000	0,00000
BRA	0,23616	0,35086	0,41298	0,97718	NAM	0,05168	0,48635	0,46193	0,78322
BRN	0,09602	0,58646	0,31745	0,82123	NPL	0,00000	0,29221	0,70779	0,54989
BGR	0,31059	0,37013	0,31928	0,99722	NLD	0,84047	0,15965	0,00000	0,39959
BFA	0,06187	0,16049	0,77764	0,60200	NZL	0,69209	0,30799	0,00000	0,56201
BDI	0,00154	0,04710	0,95138	0,18324	NIC	0,00000	0,36761	0,63237	0,59865
CPV	0,00000	0,68874	0,31118	0,56443	NER	0,06553	0,00000	0,93450	0,22018
KHM	0,10537	0,13726	0,75739	0,65551	NGA	0,00000	0,26537	0,73462	0,52667
CMR	0,04718	0,17651	0,77631	0,58873	NOR	0,66393	0,00000	0,33617	0,58110
CAN	0,75807	0,24204	0,00000	0,50367	OMN	0,17273	0,48943	0,33781	0,92812
CHL	0,23202	0,63154	0,13638	0,82006	PAK	0,00000	0,16103	0,83898	0,40175
CHN	0,58063	0,00000	0,41949	0,61903	PAN	0,01512	0,75273	0,23206	0,56087
COL	0,12592	0,59986	0,27416	0,83948	PRY	0,11313	0,29569	0,59117	0,83521
CRI	0,26669	0,33844	0,39487	0,98857	PER	0,11408	0,55916	0,32672	0,85397
CIV	0,00000	0,15889	0,84111	0,39852	PHL	0,13291	0,19766	0,66945	0,78037
HRV	0,28709	0,45240	0,26049	0,97171	POL	0,30625	0,40808	0,28566	0,98859
CYP	0,49437	0,31539	0,19027	0,93566	PRT	0,36114	0,55315	0,08569	0,82458
CZE	0,56174	0,29121	0,14710	0,87853	QAT	0,29133	0,62296	0,08567	0,78703
DNK	0,80587	0,19425	0,00000	0,44805	ROM	0,28503	0,30268	0,41230	0,98741
DOM	0,00908	0,59279	0,39806	0,65479	RUS	0,40792	0,03283	0,55932	0,73085
ECU	0,00000	0,47791	0,52204	0,63006	RWA	0,02400	0,45222	0,52373	0,71648
EGY	0,06277	0,33309	0,60412	0,76862	SAU	0,23149	0,61007	0,15840	0,84841
SLV	0,00000	0,54386	0,45608	0,62743	SEN	0,00147	0,42943	0,56905	0,63118
EST	0,58904	0,20871	0,20231	0,87570	SRB	0,25201	0,30507	0,44293	0,97416
ETH	0,00000	0,22571	0,77428	0,48611	SYC	0,15170	0,52354	0,32471	0,90126
FJI	0,24869	0,03593	0,71543	0,64187	SGP	0,94852	0,00000	0,05165	0,18494
FIN	0,91041	0,08972	0,00000	0,27469	SVK	0,31856	0,44449	0,23694	0,97032
FRA	0,65751	0,18667	0,15590	0,79987	SVN	0,47073	0,40898	0,12031	0,88759
GAB	0,08258	0,14667	0,77076	0,62641	ZAF	0,31879	0,25017	0,43106	0,97744
GMB	0,04924	0,17409	0,77668	0,59065	ESP	0,54792	0,00000	0,45218	0,62673
GEO	0,23925	0,21008	0,55070	0,90887	LKA	0,06373	0,32279	0,61346	0,76479
DEU	0,73651	0,24068	0,02288	0,59573	SDN	0,00000	0,00000	1,00000	0,00000
GHA	0,20965	0,04004	0,75035	0,61160	SWZ	0,18268	0,00000	0,81738	0,43272
GRC	0,22266	0,46784	0,30948	0,95832	SWE	0,97149	0,02867	0,00000	0,11826

Country	Alpha1	Alpha2	Alpha3	E	Country	Alpha1	Alpha2	Alpha3	E
GTM	0,09107	0,31711	0,59181	0,81272	CHE	1,00000	0,00000	0,00000	0,00000
GIN	0,00000	0,11568	0,88434	0,32607	SYR	0,00000	0,25540	0,74459	0,51719
GUY	0,22122	0,19520	0,58360	0,88013	TJK	0,13470	0,00000	0,86535	0,35971
HND	0,04304	0,33394	0,62299	0,72498	TZA	0,00000	0,28083	0,71915	0,54045
HKG	0,81146	0,18872	0,00000	0,44075	THA	0,34527	0,19833	0,45644	0,95213
HUN	0,48005	0,25163	0,26836	0,95803	TGO	0,00000	0,03788	0,96215	0,14666
ISL	0,64905	0,35102	0,00000	0,58987	TTO	0,19844	0,21224	0,58933	0,87522
IND	0,21887	0,19446	0,58669	0,87730	TUN	0,08834	0,54809	0,36351	0,82995
IDN	0,05391	0,35808	0,58799	0,76227	TUR	0,17561	0,40607	0,41830	0,94301
IRN	0,10403	0,15500	0,74098	0,67951	UGA	0,06757	0,25272	0,67970	0,72102
IRL	0,90901	0,00000	0,09114	0,27766	UKR	0,34762	0,00000	0,65246	0,58794
ISR	0,84200	0,00000	0,15814	0,39728	ARE	0,25096	0,74900	0,00000	0,51285
ITA	0,49944	0,11628	0,38435	0,87790	GBR	0,92297	0,07718	0,00000	0,24730
JAM	0,07738	0,43233	0,49025	0,82834	USA	0,95510	0,00000	0,04504	0,16704
JPN	0,76223	0,00000	0,23789	0,49931	URY	0,09187	0,59681	0,31126	0,81070
JOR	0,20522	0,38169	0,41308	0,96288	UZB	0,12190	0,00000	0,87814	0,33739
KAZ	0,13799	0,38706	0,47493	0,90506	VEN	0,08992	0,05332	0,85679	0,45998
KEN	0,18601	0,12103	0,69300	0,74874	VNM	0,28273	0,14725	0,57007	0,87348
KOR	0,75071	0,00000	0,24942	0,51120	YEM	0,00000	0,00000	1,00000	0,00000
KWT	0,26283	0,23013	0,50705	0,94087	ZMB	0,04032	0,16270	0,79698	0,55139
KGZ	0,15905	0,09807	0,74291	0,67442	ZWE	0,02375	0,12325	0,85301	0,43916
LAO	0,03877	0,00000	0,96128	0,14925					

4. Conclusion

In this study, using the methods of AA and the GII-dataset, we demonstrate that each NIS can be considered as mixture of three basic or archetypical NISs. In other words, such multidimensional phenomenon, such as the NIS (described in this study as the 7-dimensional vector) is essentially a 2-dimensional object. We also introduced important indicator – the NIS entropy– which characterises the level of disorder or randomness in the system (NIS). It is established that the first one from the above-mentioned three basic archetypes is a prototype of a highly developed NIS (with a high level of GII score and a low level of entropy), the second one is a prototype an underdeveloped NIS (with a low level of GII score and a low level of entropy), and the third one is an intermediate form of NIS (with a medium level of GII score and a high level of entropy). Hence, it would be interesting to undertake an in-depth investigation of the issues related to NIS classification and the dynamics of their development in the context of the archetypical representations of NISs.

References

[1] Verspagen, B., 2005. Innovation and economic growth, in: Fagerberg, J., Mowery, D. C., Nelson R. R. (Eds.), The Oxford Handbook of Innovation, Oxford University Press, pp. 487–513.

[2] Fagerberg, J., Srholec, M., 2008. National innovation systems, capabilities and economic Development. Research Policy, 37, 1417–1435.

[3] Acemoglu, D., 2015. Localised and biased technologies: Atkinson and Stiglitz's new view, induced innovations, and directed technological change. The Economic Journal, 125 (583), 2015, pp.443-463.

[4] Kogan, L., Papanikolaou, D., Seru, A. and Stoffman, N., Technological innovation, resource allocation, and growth. The Quarterly Journal of Economics, 132 (2), 2017, pp.665-712.

[5] Hidalgo, C. A., Economic complexity: From useless to keystone. Nature Physics, 2018, 14 (1), p.9.

[6] Thompson, M., Social capital, innovation and economic growth. Journal of behavioral and experimental economics, 73, 2018, pp.46-52.

[7] Cutler A., Breiman L., 1994. Archetypal Analysis. Technometrics, 36(4), 338–347.

[8] Eugster, M. J. and Leisch, F., Weighted and robust archetypal analysis. Computational Statistics & Data Analysis, 55 (3), 2011, pp.1215-1225.

[9] Kaufmann, D., Keller, S. and Roth, V., Copula archetypal analysis. In German Conference on Pattern Recognition, Springer, 2015, pp. 117-128.

[10] Vinué, G., Anthropometry: An R package for analysis of anthropometric data. Journal of Statistical Software, 77 (6), 2017, pp.1-39.

[11] INSEAD. Global innovation index 2011; Global innovation index 2012; Global innovation index 2013; Global innovation index 2014; Global innovation index 2015. Fontainebleau: INSEAD.