LYAPUNOV FUNCTIONS AS THE DIVERGENCE FUNCTIONS FOR ENVIRONMENTAL ASSESSMENT: THEORETICAL BACKGROUND

YURI A. PYKH[1] & IRINA G. MALKINA-PYKH[2]
[1]St. Petersburg Mathematical Society, St. Petersburg, Russia
[2]St. Petersburg State Institute of Psychology and Social Work, St. Petersburg, Russia

ABSTRACT

Measures of divergence between two points that measure their discrepancy play a key role in many environmental problems. So, it has a great importance for more constructive and progressive approaches to ensure sustainability. In particular, for two probability distributions p and q, one can define a set of various divergence measures $D[p \| q]$. A divergence is not necessarily symmetric, that is, the relation $D[p \| q] = D[q \| p]$ does not generally exist, nor does it satisfy the triangular inequality. Usually the divergence function $D[p \| q]$ satisfies the following conditions: (1) $D[p \| q] \geq 0$, (2) $D[p \| q] = 0$ when and only when $p = q$. Recall that, the information geometry originated from the geometric study of the manifold of probability distribution, has been successfully applied to many fields of environmental assessment. It is well known from Amari's investigations that divergence function induced the information geometry and endowed it by Riemannian metric. On the other hand, Lyapunov–Meyer function for the special class of replicator system is the relative entropy function or the function of information divergence. Based on "The fundamental theorem of dynamical systems" by Conley and recent results of Barta et al. "Every ordinary differential equation with a strict Lyapunov function is a gradient system" we receive the Riemannian metric for dynamical pairwise systems on standard simplex and give for it a new example of divergence for different environmental problems.
Keywords: nonlinear pairwise interactions, replicator dynamics, Lyapunov–Meyer functions, distance measure, divergence, environmental assessment.

> "We always have to lay special emphasis on mathematical analogies because the concentration on them can promote the development of science"
>
> A.N. Kolmogorov (1957)

1 INTRODUCTION

In statistics and information geometry divergence or a contract function is a function that establishes the "distance" of one probability distribution to the other on a statistical manifold. The divergence is a weaker notion than that of the distance; in particular the divergence needn't be symmetric and needn't satisfy the triangle inequality.

Suppose σ_p^n is a space of all probability distributions $p(t) = (p_1(t), ..., p_n(t))$,

$$\sigma_p^n = \left(p \in \mathbb{R}^n : p_i > 0, i = 1, ..., n, \sum_{i=1}^{n} p_i = 1 \right).$$ Then a divergence on σ_p^n is a function

$D[p \| q] : \sigma_p^n \times \sigma_p^n \to \mathbb{R}$ satisfying [1]:

1. $D[p \| q] \geq 0$ for all $p, q \in \sigma_p^n$
2. $D[p \| q] = 0$ if and only if $p = q$.

It is easy to see that function D looks like Lyapunov functions for the systems of probability evolution equation. This is especially important in the view of the latest Conley's results [2]. Note that the concepts of information, entropy, energy, diversity, discrepancy and divergence occupy a fundamental position in various fields of mathematical sciences, e.g. statistical mechanics, information theory and especially in the environmental assessment.

2 LYAPUNOV FUNCTIONS

The study of Dynamical Systems crosses interdisciplinary boundaries from ecology to psychology to meteorology, from chemical kinetics to population genetics, from economics to mechanics. A dynamical system consists of three ingredients: a setting in which the dynamical behaviour takes place; a mathematical rule which specifies; and an initial condition or state from which the system starts.

Conley [2] presents some significant results about invariant sets, attractor–repeller pairs, chain recurrence and index theory in the setting of flows on compact metric spaces. These results can be stated very simply.

Theorem 1 [3]. *Any flow on a compact metric space decomposes into a chain recurrent part and a gradient-like part.*

The term gradient-like suggests a one-way behaviour, a flowing downhill mathematically quantifiable. The idea of gradient-like dynamical system is an extension from gradient flows of the idea of functions that decrease on solutions, called *Lyapunov functions.* A dynamical system is called *gradient-like* if there is some continuous real-valued function which is strictly decreasing on nonconstant solutions. A system is called *strongly gradient-like* "if the chain recurrent set is totally disconnected (and consequently equal to the rest point set" [2]. Next we'll use specific definition of Lyapunov function: A Lyapunov function is a continuous real valued function which is strictly decreasing (increasing) on orbits outside the chain recurrent set. This definition is an application of a refinement Uryson's Lemma [4]. So, we can see that definition of Lyapunov function is very similar to definition of divergence functions.

In 1968 Meyer [5] proved that for dynamical systems whose limit sets consist of only isolated rest points or cycles, i.e. for Morse–Smale systems, there always exists a Lyapunov function. So, we'll refer to them as Lyapunov–Meyer functions. Next result was proved in the article [6], and told us that if for the dynamical system a Lyapunov–Meyer function exists then there exists a Riemannian metric such that this dynamical system has a gradient form in this metric.

3 MATHEMATICAL MODEL

As shown in [7], the acceptance of two hypotheses about the structure and types of the interactions between objects in macrosystems leads to the following generalized replicator equations determining the evolution of the probability distributions $p(t)$

$$\dot{p} = D(f)\left(Wf - e\theta^{-1}(p)\langle f, Wf \rangle\right) \cdot \tag{1}$$

Here, we use the following notation: $f(p)$ is the vector $f=f(p)=(f_1(p_1),...,f_n(p_n))$ is a vector, where the f_i are nonlinear response functions (probability distribution functions) having at least continuous first derivatives and satisfying the conditions $f_i(0)=0$, $f_i(1)=1$,

$\dfrac{\partial f_i}{\partial p_i} > 0$ for $p_i > 0$, $\dfrac{\partial f_i}{\partial p_i} \geq 0$ for $p_i = 0$; $D(\mathrm{f}) = diag(f_1, f_2 ..., f_n)$; W is the matrix of interactions; and $\theta = \theta(\mathrm{p}) = \langle \mathrm{e}, \mathrm{f}(\mathrm{p}) \rangle$, where $\langle \cdot, \cdot \rangle$ is inner product.

Since $\langle \dot{\mathrm{p}}(t), \mathrm{e} \rangle \equiv 0$ and $f_i(0) = 0$ it follows that the simlex σ_p^n and each of its faces are invariant set for system (1). System (1), as well as the generalized Lotka–Volterra equations [8], determines the dynamics of objects with nonlinear pair interactions. The matrix W determines the structure of the interaction, and the response functions f_i determine their type. A brief survey of the best known systems with pair interactions is presented in [8].

The elements of the matrix W are generalized interaction strengths, i.e. by analogy with the thermodynamics of irreversible processes, are "reasons" causing changes in the speed of flows [7].

Let us rewrite eqn (1) as

$$\dot{\mathrm{p}} = \theta D(\mathrm{f})\left(\mathrm{Wf}\theta^{-1} - \mathrm{e}E(\mathrm{p})\right),\tag{2}$$

where $E(\mathrm{p}) = \theta^{-2}(\mathrm{p})\langle \mathrm{f}, \mathrm{Wf} \rangle$. Using the terminology of the theory of neural networks, we refer to $E(\mathrm{p})$ as the energy function of the macrosystem under consideration. System (2) and the energy function $E(\mathrm{p})$ naturally determine the introduction of new additional variables:

$$x_i(\mathrm{p}) = f_i(p_i)\theta^{-1}(\mathrm{p}) \qquad i = 1,...,n.\tag{3}$$

Obviously, $\mathrm{x} = (x_1, x_2 ..., x_n) \in \sigma_x^n = \left\{ \mathrm{x} \in \mathbb{R}^n : x_i \geq 0, \mathrm{e}^\mathrm{T} \mathrm{x} = 1 \right\}$ for $\mathrm{p} \in \sigma_p^n$.

The indices x and p are used in the notation of simplexes in order to avoid confusion. Consider change (3) in more detail. If this is a diffeo-morphism, then it can be regarded not only as a simplifying change of variables customary in the theory of differential equations but also as the definition of a set of quantities with particular physical meaning.

To go further, we need the following assertion.

Statement 1. *System (1) is invariant with respect to the replacement of the interaction matrix* **W** *by a perturbed matrix* $\mathrm{W}_\zeta = \left(\mathrm{W} + \mathrm{e}\zeta^\mathrm{T}(\mathrm{p})\right)$, *where the components of the vector function* $\zeta(\mathrm{p}) = (\zeta_1(\mathrm{p}),...,\zeta_n(\mathrm{p})) : \sigma_p^n \to \mathbb{R}^n$ *are bounded on* σ_p^n.

This statement was proved in [9] for the case $\zeta(\mathrm{p}) = const$. The proof given in [9] carries over in an obvious way to the case of any bounded functions $\zeta_i(\mathrm{p})$ $i = 1, 2, ..., n$.

Now we state the main for our aim result.

Theorem 2 [10]. *If system (1) has a nontrivial equilibrium point* $\hat{\mathrm{p}} \in \mathrm{Int}\,\sigma_p^n$ *and the matrix* $\left(\mathrm{W}_\zeta^\mathrm{T} + \mathrm{W}_\zeta\right)$, $\forall \zeta \in \mathbb{R}^n$ *has* $n-1$ *negative characteristic numbers for any* $\zeta(\mathrm{p}) \in \mathbb{R}^n$, *then the function*

$$H(\mathrm{p}) = \sum_{i=1}^{n} \int_{\hat{p}_i}^{p_i} \frac{\hat{f}_i dx}{f_i(x)} + C,\tag{4}$$

where C is a constant, is a Lyapunov–Meyer function for system (1) on $\text{Int}\,\sigma_p^n$, and the energy function $E(\mathrm{p})$ of the system attains its maximum value $E(\hat{\mathrm{p}})$ as $\mathrm{t} \to \infty$.

Consider properties of the function $H(\mathrm{p})$. We begin with a quotation from [11]: "The most important information invariant of a pair of probability measures is relative entropy. This characteristic is asymmetric; therefore, we call it the information deviation of one distribution from the other...". Chentsov's book [11] was written in 1972. Since then, many terminological changes have occurred in this area; they are summarized in the recently published *Dictionary of Distances* [12]. We shall use the term relative entropy, or the function of information deviation.

The function $H(\mathrm{p})$ with $\mathrm{p} \in \text{Int}\,\sigma_p^n$ attains its maximum value at the point $\mathrm{p} = \hat{\mathrm{p}}$, where $H(\hat{\mathrm{p}}) = C$; i.e., $H(\mathrm{p}) < C$ for any $\mathrm{p} \neq \hat{\mathrm{p}}$, the covector of $H(\mathrm{p})$ is $\nabla H(\mathrm{p}) = \left(\hat{f}_i / f_i\,(p_i)\right)$, and the Hessian is $\mathcal{H}(H(\mathrm{p})) = -diag\left(\hat{f}_i \dfrac{\partial f_i}{\partial p_i} f_i^{-2}(p_i)\right)$. Since $\partial f_i / \partial p_i > 0$, it follows that the Hessian is negative definite, and the function $H(\mathrm{p})$ is concave on $\text{Int}\,\sigma_p^n$. Thus, $-H(\mathrm{p})$ can be regarded as relative entropy, or the function of information deviation between the distributions p and $\hat{\mathrm{p}}$.

The relative entropy $H(\mathrm{p})$ introduced by us depends on the interaction matrix only via the equilibrium point $\hat{\mathrm{p}}$ of the system and is largely determined by the form of the response functions. This allows us to use relation (4) for solving the "inverse" problem of finding the response functions corresponding to a given relative entropy function $H(\mathrm{p})$. Note that if $F: \mathbb{R}^1 \to \mathbb{R}^1$ is a monotonically increasing smooth function and $c_1 > 0, c_2 > 0$ and c_3 are the numbers than the function $H_F = c_1 F(c_2 H(\mathrm{p})) + c_3$ is also a Lyapunov–Meyer function for system (1), and this function is a separated function.

For the next step let us introduce some auxiliary variables:

$$\varphi(\mathrm{p}) = \sum_{i=1}^{n} \varphi_i(p_i) = \sum_{i=1}^{n} \int_{0}^{p_i} f_i(x)dx. \tag{5}$$

It is evident that

$$\frac{\partial \varphi(p)}{\partial p_{\mathrm{i}}} = f_i(p_i) \qquad i = 1,....,n, \quad p \in \sigma_p^n.$$

So, our main eqn (1) maybe rewritten in the next form

$$\dot{\mathrm{p}} = P(\mathrm{p}) = D(\mathrm{f})\Pi \cdot W\text{grad}\,\varphi(\mathrm{p}) \qquad \mathrm{p} \in \text{Int}\,\sigma_p^n, \tag{6}$$

where

$$\Pi = (1 - \theta^{-1} e f^{\mathrm{T}}),$$

is a nonortogonal projection operator $\Pi: \sigma_p^n \to \sigma_p^n$, $W\text{grad}\,\varphi(\mathrm{p})$ is $\nabla_w \Phi$ gradient field on $\text{Int}\,\sigma_p^n$. So, we have that $D^{-1}(f)$ is a diagonal (Egorov's) Riemannian metric on $\text{Int}\,\sigma_p^n$. This diagonal Riemannian metric is hyperbolic Poincare type metric known in statistic as Fisher

metric [1]. Note, that system (6) gives us also the second Riemannian metric $\left\langle p, W^{-1}p \right\rangle$ on $Int\sigma_p^n$.

So, we have two different Riemannian metrics on $Int\sigma_p^n$ according to system (6). Moreover there is exists the relation between these two Riemannian metrics [13]:

$$\dot{H}(p,\hat{p}) = -\hat{\theta}\theta(p) \left(\frac{f_i(p_i)}{\theta(p)} - \frac{\hat{f_i}}{\hat{\theta}} \right)^T W \left(\frac{f_i(p_i)}{\theta(p)} - \frac{\hat{f_i}}{\hat{\theta}} \right). \tag{7}$$

Recall that due to Kingman's result [13] $\dot{H}(p,\hat{p}) > 0$ if $p \in Int\sigma_p^n$ and $p \neq \hat{p}$.

4 GEOMETRICAL PROPERTIES OF THE DIAGONAL RIEMANNIAN METRIC

As we see we get from the dynamical system (6) a diagonal Riemannian metric. Some properties of such metric was received in the article [14] for the general diagonal Riemannian metric

$$G(p) = diag\left(\frac{1}{g_1(x_1)^2}, \frac{1}{g_2(x_2)^2}, ..., \frac{1}{g_n(x_n)^2} \right),$$

defined on the positive octant $\mathbb{R}_{++}^n$ and its restriction to $Int\sigma_p^n$.

So, we have from [14]

a. Christofel's symbols

$$\Gamma_{ij}^m = -\frac{1}{g_i(p_i)} \frac{\partial g_i(p_i)}{\partial p_i} \delta_{im}\delta_{ij},$$

where δ_{ij} are Dirac measures.

b. Geodesic equation

$$\int \frac{1}{g_i(x_i)}dx_i = a_i t + b_i \qquad i = 1,...,n,$$

where a_i, b_i are real constants. It is easy to see that this relationship is very close to the well-known definition of straighten dynamical systems.

c. Null curvative

There are shown that Riemannian manifold σ_p^n endowed with the diagonal metric $G(p)$ has null curvative [13]. This means that σ_p^n is a flat statistical manifold and we can use all results from [1] and [15] about different divergence functions.

5 DISTANCE MEASURES

Following Shennon information measure let us introduce $S(p)$ as relative entropy defined by [16]:

$$S(\mathrm{p}) = -H(\mathrm{p}) = -\sum_{i=1}^{n} \int_{\hat{p}_i}^{p_i} \frac{\hat{f}_i dx}{f_i(x)}, \quad \mathrm{p} \in \mathrm{Int}\sigma_p^n. \qquad (8)$$

It is easy to see that

$$\lambda \sum_{i=1}^{n} \int_{\hat{p}_i}^{p_i} dx = \lambda \sum_{i=1}^{n} (p_i - \hat{p}_i) \equiv 0 \quad \forall p, \hat{p} \in \sigma_p^n,$$

and using (8) we clearly have the following relationship

$$S(\mathrm{p}) = -\sum_{i=1}^{n} \int_{\hat{p}_i}^{p_i} \left(\frac{\hat{f}_i dx}{f_i(x)} + \lambda \right) dx \quad \forall \lambda \in \mathbb{R}^1.$$

Since, by the evident equality $\partial S(\mathrm{p}) / \partial p_i = 0$, with $p = \hat{p}$, we have that $\lambda = -1$. It is clear that λ is a type of Lagrange multiplier. Therefore:

$$S(\mathrm{p}) = \sum_{i=1}^{n} \int_{\hat{p}_i}^{p_i} \left(1 - \frac{\hat{f}_i}{f_i(x)} \right) dx \quad \mathrm{p} \in \mathrm{Int}\sigma_p^n. \qquad (9)$$

The gradient of $S(\mathrm{p})$ with $\mathrm{p} \in \mathrm{Int}\,\sigma_p^n$ is given by partial derivatives

$$\nabla S(\mathrm{p}) = \left(1 - \frac{\hat{f}_i}{f_i(p_i)} \right), \qquad (10)$$

that is strictly monotonically increasing for each $p_i > 0$ and Hessian of $S(\mathrm{p})$ on $\mathrm{Int}\,\sigma_p^n$ is $\mathcal{H}(\mathrm{p}) = diag\left(\hat{f}_i \frac{\partial f_i}{\partial p_i} f_i^{-2}(p_i) \right)$. Since $\partial f_i / \partial p_i > 0$ it follows that the Hessian is positive definite, and the function $S(\mathrm{p})$ according to definition is convex on $\mathrm{Int}\,\sigma_p^n$. As observed by Ball and Chen [17] entropy and convexity have played an important role in many areas of mathematics.

Another very well known definition of convexity Jensen [18], Hardy et al. [19] is the next inequality with $\mathrm{p}, \mathrm{q} \in \mathrm{Int}\,\sigma_p^n$

$$S(\mathrm{p}) - S(\mathrm{q}) - \langle \nabla_q S(\mathrm{q}), \mathrm{p} - \mathrm{q} \rangle \geq 0. \qquad (11)$$

It is evident that the expression from left-side inequality is the Bregman divergences denoted by $B_S(\mathrm{p}, \mathrm{q})$. This name was first given by Censor and Lent [20]. Bregman divergence or Bregman distance [21] is similar to a metric, but does not satisfy the triangle inequality nor symmetry. Using inequality (11) we can receive new weighted distance measure between two probability distributions

$$B_S(\mathrm{p}, \mathrm{q}) = \sum_{i=1}^{n} \left(\int_{p_i}^{q_i} \frac{\hat{f}_i dx}{f_i(x)} + \frac{\hat{f}_i}{f_i(q_i)}(p_i - q_i) \right) \geq 0. \qquad (12)$$

It is obvious that $B_S(\mathrm{p},\mathrm{q}) = 0$ for $\mathrm{p} = \mathrm{q}$, and $B_S(\mathrm{p},\hat{\mathrm{p}}) = S(\mathrm{p})$.

So, we can draw a conclusion that entropy-like function $S(\mathrm{p})$, energy-like function $\dot{\mathrm{H}}(\mathrm{p},\hat{\mathrm{p}})$ and Bregman-like function $B_S(\mathrm{p},\mathrm{q})$ may serve as divergence functions $D[p \| q]$.

6 EXAMPLES

A lot of different examples one can be found in the article [22]. In [23] we show how the logistic relative entropy works also in the case of the estimation of wastewater treatment. Different classical divergence functions are presented in the article [15, p. 350]. Now we add only few additional examples:

1. **Mahalanobis squared distance** [24] we can receive if we use (3) and if W is positively definite matrix then from (7) it follows:

$$\mathrm{D}_\mathrm{M}[x \| \hat{x}] = (x - \hat{x})^\mathrm{T} \mathrm{W}(x - \hat{x}) > 0 \qquad x, \hat{x} \in \sigma_x^n, x \neq \hat{x}.$$

2. **Kagan divergence** [25]. Let us put $\mathrm{W} = \mathrm{D}^{-1}(\hat{f}/\hat{\theta})$ in (7). If we take into account (3) then we receive

$$\mathrm{D}_\mathrm{K}[x \| \hat{x}] = \sum \frac{(x_i - \hat{x}_i)^2}{\hat{x}_i} \qquad x, \hat{x} \in \sigma_x^n.$$

3. Now let us consider as an example **Bregman divergence** (12). If we put $f_i(p_i) = p_i$ and $\hat{p}_i = \frac{1}{n}$ then taking into account the results on Kullback–Leibler relative entropy from [26] we receive

$$\mathrm{D}_\mathrm{B}[p \| q] = \frac{1}{n} \sum_{i=1}^{n} (\ln \frac{q_i}{p_i} + \frac{p_i - q_i}{q_i}).$$

4. Well-known case with **normal distributions** [27]. In this case the response function is the probability integral (Laplace function):

$$f_i(p_i) = \int_0^{p_i} e^{x^2} dx. \qquad (13)$$

It is obvious that $f_i(0) = 0$, $\partial f_i(p_i)/\partial p_i > 0$. Since the integral in (13) cannot be express through the elemental functions, then all the meanings of (13) one can find in the corresponding tables. The more full analysis of this issue is presented in [27].

7 CONCLUSIONS

In this article we discuss the most serious problems affecting sustainable development, i.e. the problems of different distance measures between any two objects. The concept of divergence for random phenomena occupies a fundamental position in various fields of environmental sciences. It's clear that this problem is very important in the goal of environmental assessment by the evaluation and implementation of sustainable development alternatives that reflects both environmental constraints and opportunities [28]. In this context we establish the set of next important links: nonlinear pairwise interaction → generalized

Fisher (replicator) equations → Lyapunov–Meyer functions → relative entropy → different distance measures → divergence $D[p \| q]$.

In particular it follows from these links that nonlinear pairwise interaction is the origin of all known relative entropy functions and different divergences.

REFERENCES

[1] Ay, N. & Amari, S., A novel approach to canonical divergences within information geometry. *SFI Working Papers*, 2015-10-038.

[2] Conley, C., Isolated invariant sets and the Morse Index. *CBMS Regional Conference Series in Mathematics,* **38**, American Mathematical Society: Providence, R.I., 1978.

[3] Norton, D.E., The fundamental theorem of dynamical systems. *Comment. Math. Univ. Carolin.*, **36**, pp. 585–597, 1995.

[4] Patrao, M., *Morse Decomposition of Semiflows on Topological Spaces*, 2005.

[5] Meyer, K., Energy functions for Morse–Smale systems. *American Journal of Mathematics,* **90**, pp. 1031–1040, 1968.

[6] Barta, T., Chill, R. & Fasangova, E., Every ordinary differential equation with a strict Lyapunov function is a gradient system. *Monatshefte für Mathematik*, **166**(1), pp. 57–72, 2012.

[7] Pykh, Y.A., Entropy characteristics of macrosystems with nonlinear pair interactions. *Doklady Mathematics*, **84**, pp. 669–671, 2011.

[8] Pykh, Y.A., Pairwise interactions origin of entropy functions. Preprint. arXiv: 1506.05731v1 [nlin.AO], 18 Jun. 2015.

[9] Pykh, Y.A., *Equilibrium and Stability in the Models of Population Dynamics,* Nauka: Moscow, 1983.

[10] Pykh, Y.A., Energy Lyapunov function for generalized replicator equations. *Proceedings of International Conference Physics and Control,* pp. 271–276, 2003.

[11] Chentsov, N.N., *Statistical and Decision Rules and Optimal Conclusions,* Nauka: Moscow, 1972.

[12] Deza, E. & Deza, M.M., *Dictionary of Distances,* Elsevier: Amsterdam, 2006.

[13] Pykh, Y.A., Escort distribution, energy and relative entropy of macrosystems with nonlinear pair interactions. *Doklady Mathematics*, **90**, pp. 710–714, 2014.

[14] Quiroz, E.A.P. & Oliveira, P.R., New results on linear optimization through diagonal metrics and Riemannian geometry. *XXXVI-SPBO*, pp. 1938–1949, 2004.

[15] Eguchi, S., A differential geometric approach to statistical inference on the basis of contrast functional. *Hiroshima Math. J.,* **15**, pp. 341–291, 1985.

[16] Pykh Y.A., Lyapunov–Meyer functions and distance measure from generalized Fisher's equations. *IFAC-PapersOnLine*, **48–11**, pp. 115–119, 2015.

[17] Ball, J.M. & Chen, G., Entropy and convexity for nonlinear partial differential equations. *Phil. Trans. R. Soc. A.*, **371**, 20120340, 8 pp., 2013.

[18] Jensen, G.L.W.V., Sur les fonctions convexes et les ineqalites entre les valenrs moyennes. *Acta Matematica*, **30**(1), pp. 175–193, 1906.

[19] Hardy, G.H., Littlewood, J.E. & Polya, G., *Inequalities*, Cambridge University Press: Cambridge, 1934.

[20] Censor, Y. & Lent, A., An iterative row-action method for interval convex programming. *Journal of Optimization Theory and Applications,* **34**(3), pp. 321–353, 1981.

[21]   Bregman, L.M., The relaxation method of finding the common points of convex sets and its application to the solution of problems in convex programming. *Computational Mathematics and Mathematical Physics*, **17**(3), pp. 200–217, 1967.

[22]   Pykh, Y.A. & Malkina-Pykh, I.G., Replicator equation, response functions and entropy measures in science: Mathematical background. *Proceedings Ecosystems and Sustainable Development V*, WIT Press: Southampton and Boston, pp. 15–22, 2005.

[23]   Pykh, Y.A. & Malkina-Pykh, I.G., Distance measures from replicator systems with nonlinear pairwise interactions for environmental risk management. *WIT Transactions on Ecology and the Environment*, vol. 237, WIT Press: Southampton and Boston, pp. 163–174, 2019.

[24]   Mahalanobis, P.C., On the generalized distance in statistics. *Proc. Nat. Inst. of Sciences of India*, **2**, pp. 49–55, 1936.

[25]   Kagan, A.M., On the theory of Fisher's amount of information. *Dokl. Akad. Nauk SSSR*, **151**, pp. 277–278, 1963.

[26]   Pykh, Y.A., Construction of entropy characteristics based on Lyapunov energy functions. *Doklady Mathematics*, **69**, pp. 355–358, 2004.

[27]   Costa, S.I.R., Santos, S.A. & Strapasson, J.T., Fisher information distance: A geometrical reading. *Discrete Applied Mathematics,* **197**, pp. 59–69, 2015.

[28]   Schultink, G., Environmental indices and public policy: A systems perspective on impact assessment and development planning. *Environmental Indices: Systems Analysis Approach (INDEX-97)*, eds Y.A. Pykh, D.E. Hyatt & R.E. Lenz, EOLSS Publishers: Oxford, pp. 169–181, 1999.