Available online at www.sciencedirect.com
ScienceDirect
IFAC-PapersOnLine 48-11 (2015) 115–119

Lyapunov-Meyer functions and distance measure from generalized Fisher's equations

Yuri Pykh

Russian Academy of Sciences
Research Center for Interdisciplinary Environmental Cooperation (INENCO)
Saint-Petersburg, Russia
E-mail: inenco@mail.neva.ru

Abstract : The main goal of this report is to do the next step in the investigation of generalized Fisher's (replicator) equations. Recently Lyapunov-Meyer function was constructed by the author for above equation as relative entropy. In this paper we prove that negative relative entropy is a convex function for a probability space and receive new distance measure between two probability distributions. Also we use Legendre-Donkin-Fenchel transformation for dual coordinates. In particular it follows from these cross-disciplinary issue that nonlinear pairwise interactions is the origin of all known entropy functions.

Keywords : Lyapunov-Meyer function, replicator equation, relative entropy, Legendre transformation, distance measure, pair interactions.

1 INTRODUCTION

This report I dedicated to the memory of member-correspondent of RAS V.A.Yakubovich. In 1972 he supported my investigation of classical Fisher's equations which arised in the mathematical theory of the evolution of the genetical structure of population. My degree of candidate in physics and mathematics science I've received on his cathedra of theoretical cybernetics in Leningrad State University.

Recently I returned to the investigation in the field of generalized Fisher's equations. Now this equation is more known as generalized replicator equation.

The main goal of this report is to do the next step in the investigation of generalized Fisher's (replicator) equations. Recently Lyapunov-Meyer function was constructed by the author for above equation as relative entropy. In this paper we prove that negative relative entropy is a convex function for a probability space and receive new distance measure between two probability distributions. Also we use Legendre-Donkin-Fenchel transformation for dual coordinates. In particular it follows from these cross-disciplinary issue that nonlinear pairwise interactions is the origin of all known entropy functions.

2. GENERALIZED FISHER'S EQUATIONS

Consider a macrosystem formed by a sufficiently large number N of interacting objects. Suppose that, at a moment t, the macrosystem under consideration contains n different types of objects and the number of objects of type i is $x_i(t)$, where $i=1,\dots,n$, and $\sum_{i=1}^{n} x_i(t)=N(t)$. Consider the relative numbers $p_i(t)=\frac{x_i(t)}{N(t)}$ of objects of various types. Obviously, $\sum_{i=1}^{n} p_i(t)=1$, i.e., $\mathrm{p}(t)\in \sigma_p^n=\{\mathrm{p}\in\mathbb{R}^n : p_i\ge 0, i=1,2,\dots,n, \mathbf{e}^T\mathrm{p}=1\}$, where σ_p^n is the standard simplex in Euclidean n-space $\mathbb{R}^n$, and $\mathbf{e}$ is the vector of ones. Thus, the state of such a macrosystem at each moment t is determined by the vector $\mathrm{p}(t)=(p_i(t),\dots,p_n(t))$.

First, we consider only those systems in which only pair interactions occur during a short time interval $(t,t+\Delta t)$; in other words, those systems in which simultaneous interactions of more than two particles are impossible. Such a constraint is fairly natural and has found a recent confirmation in Morgan et al. (2003), Schneidman et al. (2006), Bealek et al. (2007).

We make the following two assumptions about interactions between objects in a macrosystem Pykh (2011).

Hypothesis 1. The interaction between objects of types i and j is characterized by the so-called interaction strength, which we denote by w_{ij} and regard as a quantitative characteristic of the effect of the interaction between two

2405-8963 © 2015, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2015.09.169

objects of types i and j on the rate of change of the relative number $p_i(t)$ of objects of type i.

Remark 1. Apparently, the notion of an interaction strength first appeared in mathematical ecology. A fairly detailed study of this notion and a survey of related results are contained in Laska et al. (1998).

Remark 2. Note that the definition given above does not imply that $w_{ij} = w_{ji}$. We also emphasize that the asymmetry of results of interaction plays an essential role in many cases.

Hypothesis 2. For each macrosystem under consideration, there exists a set of probability distribution functions $f_i(p_i)$, *where* $i = 1,...,n$, which determine the probability of the interaction of each object of type i with any other object in the macrosystems. Thus, the probability of the pairwise interaction between objects of types i and j is determined by the product $f_i(p_i) f_j(p_j)$.

Denoting the sum $\sum_{i=1}^{n} f_i(p_i)$ by $\theta(\mathrm{p})$, we obtain the following system of differential equations determining the evolution of the probability distribution $\mathrm{p}(t)$, Pykh (2011):

$$\dot{p}_i = f_i(p_i)\left(\sum_{j=1}^{n} w_{ij} f_j(p_j) - \theta^{-1}(\mathrm{p}) \sum_{ij}^{n} w_{ij} f_i(p_i) f_j(p_j)\right), \quad (1)$$
$$i = 1,...,n$$

In what follows, it is convenient to pass to the matrix form. System (1) is written in this form as

$$\dot{\mathrm{p}} = D(\mathrm{f})\left(\mathrm{Wf} - \mathbf{e}\theta^{-1}(\mathrm{p})\langle \mathrm{f}, \mathrm{Wf}\rangle\right) \quad (2)$$

Here, $\mathrm{f}(\mathrm{p})$ is the vector $(f_1(p_1),...,f_n(p_n))$, where the f_i are probability distribution functions (nonlinear response functions) satisfying the conditions $f_i(0) = 0$, $\partial f_i / \partial p_i > 0$ for $p_i > 0$, $\frac{\partial f_i}{\partial p_i} \geqslant 0$ for $p_i = 0$ and $f_i(1) = 1$; $D(\mathrm{f}) = diag(f_1, f_2 ..., f_n)$; $\mathrm{W} = (w_{ij})$ is the matrix of interactions; and $\theta(\mathrm{p}) = \langle \mathbf{e}, \mathrm{f}(\mathrm{p})\rangle$, where $\langle \cdot,\cdot \rangle$ denotes inner product. Obviously, since $\langle \dot{\mathrm{p}}(t), \mathbf{e}\rangle \equiv 0$ and $f_i(0) = 0$ it follows that the simplex σ_p^n and each of its faces are invariant sets for system (1).

3. PRELIMINARY RESULTS FROM PYKH (2014)

Let us rewrite equations (2) as

$$\dot{\mathrm{p}} = \theta D(\mathrm{f})\left(\mathrm{Wf}\theta^{-1} - \mathbf{e}\mathrm{E}(\mathrm{p})\right), \quad (3)$$

where $\mathrm{E}(\mathrm{p}) = \theta^{-2}(\mathrm{p})\langle \mathrm{f}, \mathrm{Wf}\rangle$. Using the terminology of the theory of neural networks, we refer to $\mathrm{E}(\mathrm{p})$ as the energy function of the macrosystem under consideration. System (3) and the energy function $\mathrm{E}(\mathrm{p})$ naturally determine the introduction of new additional variables – escort distributions:

$$x_i(\mathrm{p}) = f_i(p_i)\theta^{-1}(\mathrm{p}) \qquad i = 1,...,n. \quad (4)$$

Obviously,
$\mathrm{x} = (x_1, x_2 ..., x_n) \in \sigma_x^n = \{\mathrm{x} \in \mathbb{R}^n : x_i \geq 0, \mathbf{e}^{\mathrm{T}}\mathrm{x} = 1\}$ 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 (4) 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.

Theorem 1. *Under the conditions customarily used for the response functions* f_i, *for* $\mathrm{p} \in \sigma_p^n$, *there exists a one-to-one inverse mapping to* (4), *which is defined by*

$$p_i = f_i^{-1}(x_i) \Big/ \sum_{j=1}^{n} f_j^{-1}(x_j), \quad i = 1, 2..., n, \quad (5)$$

where $f_i^{-1}(\cdot)$ *denotes the function inverse to* $f_i(\cdot)$.

Remark 1. Recall that the notation $f_i^{-1}(\cdot)$ is used for both functions inverse in the sense of function theory and functions inverse in the algebraic sense. It is always clear from the context what is meant.

Statement 1. *System* (1) *is invariant with respect to the replacement of the interaction matrix* W *by a perturbed matrix* $\mathrm{W}_\zeta = (\mathrm{W} + \mathrm{e}\,\zeta^T(\mathrm{p}))$, *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.

The further considerations are largely based on the stability postulate stated by Chetaev (1936).

Stability postulate. *Stability, which is a fundamentally general phenomenon, apparently, must manifest itself in basic laws of nature in some way. If knowledge is constructed from the requirement of small deviations from nature, then scientific thinking must (or can) rely on some Lyapunov function V. Certainly, this function always exists according to the stability postulate.*

In 1968, this postulate was stated mathematically by Meyer, who 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, which increases on the set of wandering points of the system. Meyer himself suggested the term energy functions for such functions; however, it is more natural and convenient to refer to them as Lyapunov-Meyer functions. One of the best known examples of such functions is entropy in Boltzmann's H-theorem, De Roeck et al. (2006). Following the stability postulate, we state the following theorem, which extends a result obtained in Pykh (1983, 2001).

Theorem 2. *If there exists a vector* ζ, *for which* $\mathrm{W}_\zeta = \mathrm{W}_\zeta^T$, *then the energy function of the system*

$$E_\zeta(\mathrm{p}) = \langle \mathrm{f}(\mathrm{p}), \mathrm{W}_\zeta \mathrm{f}(\mathrm{p})\rangle \theta^{-2}(\mathrm{p}) \quad (6)$$

is its Lyapunov-Meyer function.