Doclady Mathematics. V 72, N 2, 2005, pp.780-783.

CONSTRUCTION OF ENTROPY CHARACTERISTICS BASED ON REPLICATOR EQUATIONS WITH NONSYMMETRIC INTERACTION MATRICES

Yu. A. Pykh

Research Center for Interdisciplinary Environmental Cooperation, Russian Academy of Sciences, nab. Kutuzova 14, St. Petersburg, 191187 Russia
e-mail: inenco@mail.neva.ru

Presented by Academician A.B. Kurzhanskii February 4, 2005

Received February 10, 2005

In [1], a family of Lyapunov energy functions was constructed for generalized replicator equations with a symmetric interaction matrix and it was shown that virtually all the existing entropy characteristics and measures of the distance between probability distributions belong to this family. The symmetry of the interaction matrix is a rather severe constraint, and we will show below that it can be considerably weakened.

Generalized replicator equations govern the evolution of probability distributions $\mathbf{p}(t)=(p_1(t),...,p_n(t))\in\sigma$, where $\sigma=\{\mathbf{p}\in\mathbb{R}^n : p_i\geq 0, i=1,...,n, \mathbf{e}^T\mathbf{p}=1\}$ is the standard simplex in the n-dimensional Euclidean space $\mathbb{R}^n$ and $\mathbf{e}$ is a unit vector. These equations can be written as follows [2]:

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

Here, $\mathbf{f}(\mathbf{p})=(f_1(p_1),...,f_n(p_n))$, where f_i are nonlinear response functions obeying the conditions $f_i(0)=0$, $\partial f_i/\partial p_i>0$ for $p_i>0$, and $\partial f_i/\partial p_i\geqslant 0$ for $p_i=0$; $D(\mathbf{f})=diag(f_1,...,f_n)$; $\mathbf{W}=(w_{ij})$ is the interaction matrix; the function $h:\sigma\to(0,\infty]$ is specified in every particular problem; and $\theta(\mathbf{p})=\langle\mathbf{e},\mathbf{f}(\mathbf{p})\rangle$,, where $\langle\cdot,\cdot\rangle$ is the scalar product. Since $\langle\dot{\mathbf{p}}(t),\mathbf{e}\rangle\equiv 0$ and $f_i(0)=0$, it is obvious that σ and each of its faces are invariant sets of system (1). Note that, like the generalized Lotka--Volterra equations [3, 4], system (1) determines the dynamics of objects with nonlinear pair interactions. Here, the matrix $\mathbf{W}$ determines the structure of interactions, and the response function determines their type.

If $\mathbf{W}$ is nonsingular, system (1) has no more than one isolated equilibrium in $\mathrm{Int}\sigma$, which we call nontrivial.

Proposition 1 [2]. System (1) has a unique nontrivial equilibrium $\hat{\mathbf{p}}\in\mathrm{Int}\sigma$ if and only if the vector $\mathbf{W}^{-1}\mathbf{e}$ is either strictly positive or strictly negative.

The coordinates of this equilibrium are determined by the system of equations

$$\hat{\mathbf{f}} \diagup \langle \hat{\mathbf{f}}, \mathbf{e} \rangle = \mathbf{W}^{-1}\mathbf{e} \diagup \langle \mathbf{e}, \mathbf{W}^{-1}\mathbf{e} \rangle, \quad (2)$$

where $\hat{\mathbf{f}} = \left(f_1(\hat{p}_1), ..., f_n(\hat{p}_n) \right)$ with an additional constraint. Introducing the notation $\mathbf{W}^{-1}\mathbf{e} = \mathbf{a}$, we obtain the following system of $(n+1)$ equations for determining $\hat{\mathbf{p}}$: $\hat{p}_i = f_i^{-1}(\mu a_i)$ $i = 1, ..., n$, $\sum_{i=1}^{n} \hat{f}_i^{-1}(\mu a_i) = 1$, where μ is an auxiliary variable similar to a Lagrange multiplier. Obviously, this system has a unique solution if all a_i have the same sign.

Remark. It can easily be seen that the uniqueness of an equilibrium is substantially determined by the condition $\partial f_i / \partial p_i > 0$ for $p_i > 0$. If it is dropped and the picked functions f_i are, for example, unimodal or bimodal, then system (1) can have several equilibria in $Int\sigma$. This issue will be investigated elsewhere.
To state the basic theorem, we need the following definition.

Definition [5]. A continuous function on the phase space of a dynamical system is called a Lyapunov energy function for that system if it has a continuous time derivative, which is positive on the set of fixed points by virtue of the system.
Below is the main result.

Theorem 1. If system (1) has a nontrivial equilibrium $\hat{\mathbf{p}} \in Int\sigma$ and if the matrix $\left(\mathbf{W}^T + \mathbf{W}\right)$ has exactly $(n-1)$ negative eigenvalues, then the function

$$H(\mathbf{p}) = \sum_{i=1}^{n} \int_{\hat{p}_i}^{p_i} \frac{\hat{f}_i dx}{f_i(x)} \quad (3)$$

is a Lyapunov energy function for system (1).

Proof. The derivative of $H(\mathbf{p})$ on the trajectories of system (1) can be expressed as

$$\dot{H}(\mathbf{p}) = h(\mathbf{p}) \hat{\mathbf{f}}^T \left(\mathbf{W}\mathbf{f} - \mathbf{e}\theta^{-1} \langle \mathbf{f}, \mathbf{W}\mathbf{f} \rangle \right). \quad (4)$$

Consider a set of values $x_i(\mathbf{p})$ defined as

$$x_i(\mathbf{p}) = f_i(p_i) / \theta(\mathbf{p}).$$

The vector $\mathbf{x} = (x_1, ..., x_n)$ is known as an escort distribution [2] and satisfies the obvious inclusion $\mathbf{x}(t) \in \sigma$. In terms of the escort distribution, relation (4) can be written as

$$\dot{H}(\mathbf{p}) = h(\mathbf{p}) \hat{\theta} \theta^{-1} \left(\hat{\mathbf{x}}^T \mathbf{W}\mathbf{x} - \langle \mathbf{x}, \mathbf{W}\mathbf{x} \rangle \right) \quad (5)$$

Here, we took into account $\hat{\mathbf{x}}^T \mathbf{e} = 1$. Define the auxiliary variables $y_i = x_i - \hat{x}_i$, $i = 1, ..., n$. Obviously, the vector $\mathbf{y} = (y_1, ..., y_n)$ belongs to the shifted simplex

$$\sigma^0 = \left\{ \mathbf{y} : \mathbf{e}^T \mathbf{y} = 0, -\hat{x}_i \leqslant y_i \leqslant 1 - \hat{x}_i \right\}.$$

Rewriting formula (5) in terms of the new variables and collecting similar terms yields

$$\dot{H}(\mathbf{p}) = -h(\mathbf{p}) \hat{\theta} \theta^{-1} \langle \mathbf{y}, \mathbf{W}\mathbf{y} \rangle \tag{6}$$

By the theorem condition, the matrix $\left(\mathbf{W} + \mathbf{W}^T\right)$ has exactly $(n-1)$ negative eigenvalues. Therefore, the quadratic form $\langle \mathbf{y}, \mathbf{W}\mathbf{y} \rangle$ is negative definite on the simplex σ^0 [6], which proves the theorem.

Before discussing the theorem, we consider the properties of $H(\mathbf{p})$. A simple analysis shows that this function reaches its maximum at the point $\mathbf{p} = \hat{\mathbf{p}}$ $H(\hat{\mathbf{p}}) = 0$. At all the remaining points $\mathbf{p} \in \sigma$, the function $H(\mathbf{p})$ is negative and may be unbounded from below on the boundaries of σ. As was indicated in [7], the loss of boundedness from below for the characteristic entropy is fairly natural and is exemplified by the Shannon entropy. Thus, the theorem has the following obvious consequence.

Corollary 1. If the conditions of Theorem 1 are fulfilled, then the equilibrium $\hat{\mathbf{p}}$ of system (1) is globally stable in $Int\sigma$.

Note that formula (6) formally defines the production of entropy in the system. Therefore, it is of interest to find the conditions on $\mathbf{W}$ under which this formula can be rewritten in terms of energy [2], i.e., in terms of $E(\mathbf{p}) = \theta^{-2} \langle \mathbf{f}, \mathbf{W}\mathbf{f} \rangle$.

Corollary 2. If the conditions of Theorem 1 are fulfilled and $\mathbf{W}$ is such that $\mathbf{W}^T \mathbf{W}^{-1}$ is a stochastic matrix, i.e.,

$$\mathbf{W}^T \mathbf{W}^{-1} \mathbf{e} = \mathbf{e}, \tag{7}$$

then the entropy production is defined by the formula

$$\dot{H}(\mathbf{p}) = h(\mathbf{p}) \theta \hat{\theta} \left(E(\hat{\mathbf{p}}) - E(\mathbf{p}) \right) \geqslant 0. \tag{8}$$

Proof. The corollary is proved by noting that $\mathbf{W}^T \mathbf{W}^{-1} \mathbf{e} = \mathbf{e}$, combined with (2), implies that

$$\mathbf{W}^T \hat{\mathbf{f}} = \mathbf{e} \left(\hat{\theta} / \left\langle \mathbf{e}, \mathbf{W}^{-1} \mathbf{e} \right\rangle \right),$$

and by using the formula for $E(\mathbf{p})$.

Note that formula (8) is similar in structure to relation (6) in [1], which was obtained for a symmetric interaction matrix. However, in the case under consideration, $E(\mathbf{p})$ is not necessarily monotonically increasing on the trajectories of system (1). We can only state that it will reach its maximum value on σ with time elapsed. In addition, we note that (7) is obviously satisfied for symmetric and twice quasi-stochastic matrices. At the same time, an analysis based on Maple V shows that, even in the case $n = 3$, there are 12 types of matrices satisfying condition (7). Thus, the problem arises of describing all the types of real matrices satisfying (7).

The following result is useful in the analysis of the properties of system (1).

Proposition 2. System (1) is invariant under the replacement of the interaction matrix $\mathbf{W}$ with the matrix $\mathbf{W}_\zeta = \left(\mathbf{W} + \mathbf{e}\boldsymbol{\zeta}^T\right)$, where $\boldsymbol{\zeta}$ is an arbitrary vector from $\mathbb{R}^n$.

Proof. The parenthesized expression on the right-hand side of system (1) with the matrix $\left(\mathbf{W} + \mathbf{e}\boldsymbol{\zeta}^T\right)$ can be represented in the form

$$\left(\left(\mathbf{W} + \mathbf{e}\boldsymbol{\zeta}^{\mathrm{T}}\right)\mathbf{f} - \mathbf{e}\theta^{-1}\left(\mathbf{f}^{\mathrm{T}}\left(\mathbf{W} + \mathbf{e}\boldsymbol{\zeta}^{\mathrm{T}}\right)\mathbf{f}\right)\right) =$$
$$= \left(\mathbf{Wf} + \mathbf{e}\boldsymbol{\zeta}^{\mathrm{T}}\mathbf{f} - \mathbf{e}\theta^{-1}\left(\mathbf{f}^{\mathrm{T}}\mathbf{Wf}\right) - \mathbf{e}\theta^{-1}\mathbf{f}^{\mathrm{T}}\mathbf{e}\boldsymbol{\zeta}^{\mathrm{T}}\mathbf{f}\right).$$

Since $\mathbf{f}^T\mathbf{e} = \theta$, this expression can be rewritten as

$$\left(\mathbf{Wf} - \mathbf{e}\theta^{-1}\langle\mathbf{f}, \mathbf{Wf}\rangle\right),$$

which completes the proof of the proposition.

In certain cases, Proposition 2 substantially simplifies the analysis of the properties of system (1). For example, if all the diagonal elements of $\mathbf{W}$ are either the largest or smallest in their columns, then $\boldsymbol{\zeta}$ can be chosen so that all the diagonal elements of $\mathbf{W}_\zeta$ have the same sign, while the off-diagonal elements have an opposite sign. Therefore, by the Hawkins--Simon theorem [8], system (1) has a nontrivial equilibrium $\hat{\mathbf{p}} \in \mathrm{Int}\sigma$. Another example is as follows. In certain cases, $\boldsymbol{\zeta}$ can be picked so that $\mathbf{W}_\zeta$ is symmetric and, hence, all the results of [1] can be applied to system (1). In particular, this can always be done if $\mathbf{W}$ is tridiagonal. Another important observation is that $\boldsymbol{\zeta}$ can always be chosen so that $\mathbf{W}_\zeta^T$ is quasi-stochastic, i.e., satisfies

$$\mathbf{W}_\zeta^T\mathbf{e} = \alpha\mathbf{e} \qquad \text{where } \alpha > 0 \; .$$

In particular, this means that formula (8) is also valid for quasi-stochastic matrices.

Now, we use Theorem 1 to construct entropy characteristics and "distances" between probability characteristics. First, we note that, if $F : \mathbb{R}^1 \to \mathbb{R}^1$ is a monotonically increasing smooth function and C and C_1 are two numbers, then the function

$$H_F = F\left(H\left(\mathbf{p}\right) + C\right) + C_1 \tag{9}$$

is also a Lyapunov energy function for system (1).

Clearly, the domain of F coincides with the range of $\left(H\left(\mathbf{p}\right) + C\right)$, $\mathbf{p} \in \sigma$.

Entropy characteristics. Note that Theorem 1 can be used in two ways.

1. The response functions can be found for previously known entropy characteristics. Relevant examples for the Shannon and Tsallis entropy were given in [1]. It is easy to show that this approach yields response functions for all the entropy characteristics proposed in [9].
2. New entropy characteristic can be derived from some functions obeying the conditions stated for response functions. As an example, we consider the logistic function, which is widely applied in mathematical ecology and economics. This function is defined as

$$f(x) = \frac{c}{(b+c)} \frac{\left(1-e^{-\alpha x}\right)}{\left(b+ce^{-\alpha x}\right)},$$

where $b > 0,\ c > 0,\ \alpha > 0$.

It is obvious that $f(0) = 0$, $f(+\infty) = c / b(b+c)$. Evaluating the corresponding integral and taking into account that arbitrary parameters can be picked in (9) (to get rid of the factors and summands), we derive the following expression for the logistic entropy $H_l(\mathbf{p})$:

$$H_l(\mathbf{p}) = \sum_{i=1}^{n} \ln\left(1-e^{-\alpha p_i}\right).$$

Recall that $H_l(\mathbf{p})$ is a Lyapunov energy function for system (1) if and only if the conditions of Theorem 1 are satisfied and the equilibrium $\hat{\mathbf{p}}$ is a uniform distribution; i.e., $\hat{p}_i = n^{-1}$. Since all the response functions are identical in the case under study, it is necessary and sufficient that $\mathbf{W}$ be quasi-stochastic. Note that the interest in objects described by noncanonical distributions has increased noticeably over resent years [7]. It is this type of distribution that can be derived from formula (9) suggested for entropy characteristics. For further use of these characteristics, it is natural to apply Jaynes' principle of maximum information entropy [10].

Distances between distributions. First, we underline that distance functions are defined in the case where the equilibrium distribution in system (1) is not uniform. As in the previous case, Theorem 1 can be used in two ways.

1. The response functions can be determined for previously known distance functions. Such an example with response functions obtained for the Kullback--Leibler relative entropy can be found in [1].

2. New distance functions between distributions can be derived from given response functions. As before, we consider the logistic function as an example. Let the conditions of Theorem 1 be satisfied. Evaluating the integral in (3) and using formula (9) gives the function

$$H_l(\mathbf{p}, \hat{\mathbf{p}}) = \sum_{i=1}^{n} \hat{f}_i \ln\left(\frac{1-e^{-\alpha p_i}}{1-e^{-\alpha \hat{p}_i}}\right),$$

which, by analogy, is naturally called the relative logistic entropy.

To conclude, we note that the approach suggested for constructing entropy characteristics and distances between probability distributions based on the direct Lyapunov method applied to kinetic replicator equations (i.e., equations determining the dynamics of systems with nonlinear pair interactions) provides new strategies in the macroscopic analysis of complex open objects in various areas of the natural sciences [11]. It should also be emphasized that this approach makes it possible to introduce mixed entropy, i.e., an entropy characteristic arising when the interacting subsystems have different response functions.

REFERENCES

1. Yu. A. Pykh, Dokl. Akad. Nauk 396, pp.162-165 (2004) [Dokl. Math. 69, pp. 355-358 (2004)].
2. Yu. A. Pykh, Proceedings of International Conference on Physics and Control (IFEE, St. Petersburg 2003), Vol. 1, pp. 271-276.
3. Yu. A. Pykh, Proceedings of 5th IFAC Symposium on Nonlinear Control Systems (IFAC, St. Petersburg, 2001), pp. 1655-1660.
4. Yu. A. Pykh, Equilibrium and Stability in Population Dynamics Models (Nauka, Moscow, 1983) [in Russian], 184 p.
5. K. Meyer, Am. J. Math. Vol 90 (N4), 1031-1040 (1968).
6. J. F. C. Kingman, Proc. Cambridge Phil. Soc. 57 (N1), 574-582 (1961).
7. Yu. G. Rudoi, Teor. Mat. Fiz. 135 (N1), 3-54 (2003) [in Russian].
8. D. Hawkins and H. A. Simon, Econometrica, V. 17 (N3) (1949).
9. M. D. Esteban and D. Morales, Kybernetica Vol. 31, pp.337-346 (1995).
10. E. T. Jaynes, Phys. Rev. (1957) V.106, p. 620; V.108, p. 171 (1957).
11. J. Hofbauer and K. Sigmund, Bull. Am. Math. Soc. Vol.40 (N4), pp.479-519 (2003).