Available online at www.sciencedirect.com
ScienceDirect
Physica D 234 (2007) 70–80
PHYSICA D
www.elsevier.com/locate/physd

Jacobi fields on statistical manifolds of negative curvature

C. Cafaro*, S.A. Ali

Department of Physics, University at Albany-SUNY, 1400 Washington Avenue, Albany, NY 12222, USA

Received 8 October 2006; accepted 1 July 2007
Available online 5 July 2007

Communicated by C.K.R.T. Jones

Abstract

Two entropic dynamical models are considered. The geometric structure of the statistical manifolds underlying these models is studied. It is found that in both cases, the resulting metric manifolds are negatively curved. Moreover, the geodesics on each manifold are described by hyperbolic trajectories. A detailed analysis based on the Jacobi equation for geodesic spread is used to show that the hyperbolicity of the manifolds leads to chaotic exponential instability. A comparison between the two models leads to a relation among statistical curvature, stability of geodesics and relative entropy-like quantities. Finally, the Jacobi vector field intensity and the entropy-like quantity are suggested as possible indicators of chaoticity in the ED models due to their similarity to the conventional chaos indicators based on the Riemannian geometric approach and the Zurek–Paz criterion of linear entropy growth, respectively.

Keywords: Inductive inference; Information geometry; Statistical manifolds; Entropy; Nonlinear dynamics and chaos

1. Introduction

Entropic Dynamics (ED) [1] is a theoretical framework constructed on statistical manifolds to explore the possibility that laws of physics, either classical or quantum, might be laws of inference rather than laws of nature. It is known that thermodynamics can be obtained by means of statistical mechanics which can be considered as a form of statistical inference [2] rather than a pure physical theory. Indeed, even some features of quantum physics can be derived from principles of inference [3]. Finally, recent research considers the possibility that Einstein's theory of gravity is derivable from general principles of inductive inference [4]. Unfortunately, the search for the correct variables that encode relevant information about a system is a major obstacle in the description and understanding of its evolution. The manner in which relevant variables are selected is not straightforward. This selection is made, in most cases, on the basis of intuition guided by experiment. The Maximum relative Entropy (ME) method [5–7] is used to construct ED models. The ME method is designed to be a tool of inductive inference. It is used for updating from a prior to a posterior probability distribution when new information in the form of constraints becomes available. We use known techniques [1] to show that this principle leads to equations that are analogous to equations of motion. Information is processed using ME methods in the framework of Information Geometry (IG) [8] that is, Riemannian geometry applied to probability theory. In our approach, probability theory is a form of generalized logic of plausible inference. It should apply in principle, to any situation where we lack sufficient information to permit deductive reasoning.

In this paper, we focus on two special entropic dynamical models. In the first model (ED1), we consider a hypothetical system whose microstates span a 2D space labelled by the variables $x_1 \in \mathbb{R}^+$ and $x_2 \in \mathbb{R}$. We assume that the only testable information pertaining to the quantities x_1 and x_2 consists of the expectation values $\langle x_1 \rangle$, $\langle x_2 \rangle$ and the variance Δx_2. In the second model (ED2), we consider a 2D space of microstates labelled by the variables $x_1 \in \mathbb{R}$ and $x_2 \in \mathbb{R}$. In this case, we assume that the only testable information pertaining to the quantities x_1 and x_2 consists the expectation values $\langle x_1 \rangle$ and $\langle x_2 \rangle$ and of the variances Δx_1 and Δx_2. Our models

* Corresponding author.
E-mail addresses: carlocafaro2000@yahoo.it (C. Cafaro), alis@alum.rpi.edu (S.A. Ali).

0167-2789/$ - see front matter
doi:10.1016/j.physd.2007.07.001

may be extended to more elaborate systems (highly constrained dynamics) where higher dimensions are considered. However, for the sake of clarity, we restrict our considerations to the above relatively simple cases. Given two known boundary macrostates, we investigate the possible trajectories of systems on the manifolds. The geometric structure of the manifolds underlying the models is studied. The metric tensor, Christoffel connections coefficients, Ricci and Riemann curvature tensors are calculated in both cases and it is shown that in both cases the dynamics takes place on negatively curved manifolds. The geodesics of the dynamical models are hyperbolic trajectories on the manifolds. A detailed study of the stability of such geodesics is presented using the equation of geodesic deviation (Jacobi equation). The notion of statistical volume elements is introduced to investigate the asymptotic behavior of a one-parameter family of neighboring geodesics. It is shown that the behavior of geodesics on such manifolds is characterized by exponential instability that leads to chaotic scenarios on the manifolds. These conclusions are supported by the asymptotic behavior of the Jacobi vector field intensity. Finally, a relation among entropy-like quantities, instability and curvature in the two models is presented.

2. Curved statistical manifolds

In the case of ED1, a measure of distinguishability among the states of the system is achieved by assigning a probability distribution $p(\vec{x}|\vec{\theta})$ to each state defined by expected values $\theta_1^{(1)}, \theta_1^{(2)}, \theta_2^{(2)}$ of the variables x_1, x_2 and $(x_2 - \langle x_2 \rangle)^2$. In the case of ED2, one assigns a probability distribution $p(\vec{x}|\vec{\theta})$ to each state defined by expected values $\theta_1^{(1)}, \theta_2^{(1)}, \theta_1^{(2)}, \theta_2^{(2)}$ of the variables x_1, $(x_1 - \langle x_1 \rangle)^2$, x_2 and $(x_2 - \langle x_2 \rangle)^2$. The process of assigning a probability distribution to each state provides the statistical manifolds of the ED models with a metric structure. Specifically, the Fisher–Rao information metric [9–12] defined in (7) is used to quantify the distinguishability of probability distributions $p(\vec{x}|\vec{\theta})$ that live on the manifold (the family of distributions $\{p^{(\text{tot})}(\vec{x}|\vec{\theta})\}$ is as a manifold, each distribution $p^{(\text{tot})}(\vec{x}|\vec{\theta})$ is a point with coordinates θ^i where i labels the macrovariables). As such, the Fisher–Rao metric assigns an IG to the space of states.

2.1. *The statistical manifold $\mathcal{M}_{S_1}$*

Consider a hypothetical physical system evolving over a 2D space. The variables $x_1 \in \mathbb{R}^+$ and $x_2 \in \mathbb{R}$ label the 2D space of microstates of the system. We assume that all information relevant to the dynamical evolution of the system is contained in the probability distributions. For this reason, no other information (such as external fields) is required. We assume that the only testable information pertaining to the quantities x_1 and x_2 consists of the expectation values $\langle x_1 \rangle$, $\langle x_2 \rangle$ and the variance Δx_2. Therefore, these three expected values define the 3D space of macrostates $\mathcal{M}_{S_1}$ of the ED1 model. Each macrostate may be thought as a point of a 3D statistical manifold with coordinates given by the numerical values of the expectations $\theta_1^{(1)}, \theta_1^{(2)}, \theta_2^{(2)}$. The available information can be written in the form of the following constraint equations,

$$\begin{aligned}
\langle x_1 \rangle &= \int_0^{+\infty} \mathrm{d}x_1 x_1 p_1\left(x_1|\theta_1^{(1)}\right), \\
\langle x_2 \rangle &= \int_{-\infty}^{+\infty} \mathrm{d}x_2 x_2 p_2\left(x_2|\theta_1^{(2)}, \theta_2^{(2)}\right), \\
\Delta x_2 &= \sqrt{\left\langle (x_2 - \langle x_2 \rangle)^2 \right\rangle} \\
&= \left[\int_{-\infty}^{+\infty} \mathrm{d}x_2 \left(x_2 - \langle x_2 \rangle\right)^2 p_2\left(x_2|\theta_1^{(2)}, \theta_2^{(2)}\right)\right]^{\frac{1}{2}},
\end{aligned} \tag{1}$$

where $\theta_1^{(1)} = \langle x_1 \rangle$, $\theta_1^{(2)} = \langle x_2 \rangle$ and $\theta_2^{(2)} = \Delta x_2$. The probability distributions p_1 and p_2 are constrained by the conditions of normalization,

$$\begin{aligned}
&\int_0^{+\infty} \mathrm{d}x_1 p_1\left(x_1|\theta_1^{(1)}\right) = 1, \\
&\int_{-\infty}^{+\infty} \mathrm{d}x_2 p_2\left(x_2|\theta_1^{(2)}, \theta_2^{(2)}\right) = 1.
\end{aligned} \tag{2}$$

Information theory identifies the exponential distribution as the maximum entropy distribution if only the expectation value is known. The Gaussian distribution is identified as the maximum entropy distribution if only the expectation value and the variance are known. ME methods allow us to associate a probability distribution $p^{(\text{tot})}(\vec{x}|\vec{\theta})$ to each point in the space of states. The distribution that best reflects the information contained in the prior distribution $m(\vec{x})$ updated by the constraints $(\langle x_1 \rangle, \langle x_2 \rangle, \Delta x_2)$ is obtained by maximizing the relative entropy

$$\begin{aligned}
&\left[S\left(\vec{\theta}\right)\right]_{\text{ED1}} \\
&\quad = -\int_0^{+\infty}\int_{-\infty}^{+\infty} \mathrm{d}x_1 \mathrm{d}x_2 p^{(\text{tot})}(\vec{x} \mid \vec{\theta}) \log\left[\frac{p^{(\text{tot})}(\vec{x}|\vec{\theta})}{m(\vec{x})}\right],
\end{aligned} \tag{3}$$

where $m(\vec{x}) \equiv m$ is the uniform prior probability distribution. The prior $m(\vec{x})$ is set to be uniform since we assume the lack of initial available information about the system (postulate of equal *a priori* probabilities). Upon maximizing (3), given the constraints (1) and (2), we obtain

$$\begin{aligned}
p^{(\text{tot})}(\vec{x}|\vec{\theta}) &= p_1\left(x_1|\theta_1^{(1)}\right) p_2\left(x_2|\theta_1^{(2)}, \theta_2^{(2)}\right) \\
&= \frac{1}{\mu_1}\mathrm{e}^{-\frac{x_1}{\mu_1}} \frac{1}{\sqrt{2\pi\sigma_2^2}} \mathrm{e}^{-\frac{(x_2-\mu_2)^2}{2\sigma_2^2}},
\end{aligned} \tag{4}$$

where $\theta_1^{(1)} = \mu_1$, $\theta_1^{(2)} = \mu_2$ and $\theta_2^{(2)} = \sigma_2$. The probability distribution (4) encodes the available information concerning the system and $\mathcal{M}_{s_1}$ becomes,

$$\begin{aligned}
\mathcal{M}_{s_1} = \Bigg\{ & p^{(\text{tot})}(\vec{x}|\vec{\theta}) = \frac{1}{\mu_1}\mathrm{e}^{-\frac{x_1}{\mu_1}} \frac{1}{\sqrt{2\pi\sigma_2^2}} \\
& \times \mathrm{e}^{-\frac{(x_2-\mu_2)^2}{2\sigma_2^2}} : \vec{x} \in \mathbb{R}^+ \times \mathbb{R} \text{ and } \vec{\theta} \equiv (\mu_1, \mu_2, \sigma_2) \Bigg\}.
\end{aligned} \tag{5}$$