Physica A 388 (2009) 4703–4712

Contents lists available at ScienceDirect

Physica A

journal homepage: www.elsevier.com/locate/physa

Nonlinear non-extensive approach for identification of structured information

Laura Rebollo-Neira [a,*], A. Plastino [b]

[a] Mathematics, Aston University, Birmingham, B4 7ET, UK

[b] IFLP-CCT-Conicet, Universidad National University La Plata, CC 727, 1900 La Plata, Argentina

ARTICLE INFO

Article history:
Received 4 June 2009
Received in revised form 21 July 2009
Available online 6 August 2009

Keywords:
Nonextensive nonlinear approximations
Oblique projections
Sparse representations

ABSTRACT

The problem of separating structured information representing phenomena of differing natures is considered. A structure is assumed to be independent of the others if can be represented in a complementary subspace. When the concomitant subspaces are well separated the problem is readily solvable by a linear technique. Otherwise, the linear approach fails to correctly discriminate the required information. Hence, a non-extensive approach is proposed. The resulting nonlinear technique is shown to be suitable for dealing with cases that cannot be tackled by the linear one.

1. Introduction

We consider the problem of discriminating information produced by phenomena of differing natures, via inverse methods. This involves the study of the physical state of a system by analyzing its response to some external interaction. We refer to the interactive carrier as *input signal* and to the system's reaction as *signal response*. Unfortunately, a particular response is not always directly available, as one may receive it 'disguised' by the interference with another independent phenomenon not being the focus of specific interest. In this paper we restrict our consideration to responses evoked by statistical systems. By this we understand systems which are fully characterized by a probability distribution indicating either the population of subsystems compressing the whole system, or the degree of uncertainty about the system being in one of its possible states. We regard both situations to be identical in the description and refer to subsystems as system's states.

In order to formulate the problem let us use the label 'i', ranging from 1 to M, to denote the i-th state of a system which is characterized by a probability p_i. Adopting Dirac's notation we indicate by a ket $|f_\nu\rangle$ the system's response to some input signal and by $|v_i\rangle$ the corresponding response of the i-th state. Consequently, the system's signal response satisfies

$$|f_\nu\rangle \propto \sum_{i=1}^{M} p_i |v_i\rangle.$$

This equation is transformed into an equality by simply relaxing the condition $\sum_{i=1}^{M} p_i = 1$, so that

$$|f_\nu\rangle = \sum_{i=1}^{M} c_i |v_i\rangle,$$

* Corresponding author.
E-mail address: rebollol@aston.ac.uk (L. Rebollo-Neira).

0378-4371/$ – see front matter
doi:10.1016/j.physa.2009.08.003

where the coefficients in the superposition are not necessarily normalized to unity. As already stated, we are interested in the problem of discriminating $|f_{\mathcal{V}}\rangle$ from a given signal $|f\rangle$ of which $|f_{\mathcal{V}}\rangle$ is a component. Out of the many situations involving this problem it immediately comes to our mind the intensity of X-rays produced simultaneously by dispersion and diffraction or an infrared emission spectrum superimposed on blackbody radiation. In order to model all relevant cases we assume that, rather than $|f_{\mathcal{V}}\rangle$, the available signal is $|f\rangle = |f_{\mathcal{V}}\rangle + |f_{\mathcal{W}^\perp}\rangle$, where $|f_{\mathcal{W}^\perp}\rangle$ is produced by an independent phenomenon. We focus on those cases ensuring a unique decomposition, i.e., we further assume that the subspaces hosting the components $|f_{\mathcal{V}}\rangle$ and $|f_{\mathcal{W}^\perp}\rangle$ are complementary. However, the focus of our interest refers to complementary subspaces being close enough together to move the problem of separating the components far away from the trivial one. Certainly, if the subspaces hosting the signal components are well separated, the problem is readily solvable by means of an oblique projection onto one of the subspaces and along the other [1,2]. Contrarily, if the subspaces are not well separated the construction of the necessary projector becomes ill posed and the problem needs to be tackled in an alternative way. In this Communication we address the matter by including a hypothesis upon the system producing the signal response. *We assume that the population of states is K-sparse in the sense that, out of the M possible states of the system, only $K < M$ of them are characterized by a significant probability.* Nevertheless, the hypothesis generates, in general, an intractable problem, because of course the populated states are unknown and the number of possibilities of populating K states out of M is a combinatorial number $\binom{M}{K}$. This makes the exhaustive search for the unknown states an impossible task for most values of M and K. In recent publications [3,4] a greedy strategy for making the search tractable has been proposed. In the present context, the proposal of that publications implies assuming a priori that no state is populated and looks for the populated ones in a stepwise manner. Here we investigate the possibility of addressing the problem from the opposite view point. Assuming a priori that all the states are equally populated, we will determine the actual population of each state via the minimization of the q-norm like quantity $\sum_{i=1}^M |c_i|^q$, $0 < q \leq 1$. The minimization of this quantity as an appropriate criterion for determining a sparse solution to an under-determined linear system is discussed in Refs. [5,6]. For nonnegative and normalized to unity coefficients $c_i, i = 1, \ldots, M$, this quantity is closely related to the non-extensive entropic measure broadly applied in physics [7–12] since Tsallis introduced it as the essential ingredient of his thermodynamic analysis framework [7]. In the present context the value of q plays a particular role. By choosing $0 < q \leq 1$ we introduce an assumption on the sought distribution. *We assume that not all the possible states in a system's model are significantly populated*. This assumption is meant to compensate for the actual overestimation of possibilities one usually makes when a system's signal response is modeled mathematically.

The paper is organized as follows: Section 2 introduces the mathematical setting of the problem and discusses the construction of oblique projectors. Section 3 stresses the need for nonlinear approaches to separate signal components living in subspaces which are 'theoretically' complementary, but close enough to prevent the components discrimination being realized by a linear operation. The proposed strategy, based on the minimization of $\sum_{i=1}^M |c_i|^q$ subject to recursively selected constraints, is discussed in Section 4 and illustrated in the same section by a numerical simulation. The numerical experiment is especially designed to highlight the robustness of the proposed approach against significant errors in the data. The conclusions are presented in Section 5.

2. Mathematical setting of the problem

As already mentioned, adopting Dirac's notation we represent the response of a statistical system to some external interaction as $|f_{\mathcal{V}}\rangle$, which is expressible in the form

$$|f_{\mathcal{V}}\rangle = \sum_{i=1}^{M} c_i |v_i\rangle. \tag{1}$$

Since the kets are elements of an inner product space, their square norm is induced by the inner product, i.e., $\| |f_{\mathcal{V}}\rangle \|^2 = \langle f_{\mathcal{V}} | f_{\mathcal{V}}\rangle$.

The problem we are concerned with entails 'rescuing' a ket response $|f_{\mathcal{V}}\rangle$ from an available mixture $|f\rangle = |f_{\mathcal{V}}\rangle + |f_{\mathcal{W}^\perp}\rangle$, where $|f_{\mathcal{W}^\perp}\rangle$ is produced by an independent phenomenon (e.g. a structured interference that one would call *background* referring to a persistent effect out of the focus of the main interest).

Denoting $\mathcal{V} = \text{span}\{|v_i\rangle\}_{i=1}^M$ and assuming that the subspace $\mathcal{W}^\perp$ such that $|f_{\mathcal{W}^\perp}\rangle \in \mathcal{W}^\perp$ is known, we restrict considerations to the case $\mathcal{V} \cap \mathcal{W}^\perp = \{0\}$ so as to ensure the uniqueness of the decomposition $|f\rangle = |f_{\mathcal{V}}\rangle + |f_{\mathcal{W}^\perp}\rangle$. Such a problem has a straightforward 'theoretical' solution. Certainly, from $\hat{E}_{\mathcal{V}\mathcal{W}^\perp}$, the oblique projector onto $\mathcal{V}$ along $\mathcal{W}^\perp$, one immediately has

$$\hat{E}_{\mathcal{V}\mathcal{W}^\perp} |f\rangle = \hat{E}_{\mathcal{V}\mathcal{W}^\perp}(|f_{\mathcal{V}}\rangle + |f_{\mathcal{W}^\perp}\rangle) = |f_{\mathcal{V}}\rangle.$$

However, as will be discussed in the next section, when the subspaces $\mathcal{V}$ and $\mathcal{W}^\perp$ are not well separated the numerical construction of $\hat{E}_{\mathcal{V}\mathcal{W}^\perp}$ becomes ill posed, thus preventing the signal separation from being correctly realized.