Using State Space Differential Geometry for Nonlinear Blind Source Separation

David N. Levin

University of Chicago
d-levin@uchicago.edu

Abstract. Given a time series of multicomponent measurements of an evolving stimulus, nonlinear blind source separation (BSS) usually seeks to find a "source" time series, comprised of statistically independent combinations of the measured components. In this paper, we seek a source time series that has a *phase-space* density function equal to the product of density functions of individual components. In an earlier paper, it was shown that the phase space density function induces a Riemannian geometry on the system's state space, with the metric equal to the local velocity correlation matrix of the data. From this geometric perspective, the vanishing of the curvature tensor is a necessary condition for BSS. Therefore, if this data-derived quantity is non-vanishing, the observations are not separable. However, if the curvature tensor is zero, there is only one possible set of source variables (up to transformations that do not affect separability), and it is possible to compute these explicitly and determine if they do separate the phase space density function. A longer version of this paper describes a more general method that performs nonlinear multidimensional BSS or independent subspace separation.

1 Introduction

Consider a set of data consisting of $\tilde{x}(t)$, a time-dependent multiplet of n measurements ($\tilde{x}_k$ for $k = 1, 2, \ldots, n$). The usual objectives of nonlinear BSS are: 1) to determine if these observations are instantaneous mixtures of n statistically independent source components $x(t)$

$$\tilde{x}(t) = f[x(t)] \tag{1}$$

where f is an unknown, possibly nonlinear, n-component mixing function, and, if so, 2) to compute the mixing function. In most approaches to this problem [1,2], the desired source components are required to be statistically independent in the sense that their state space density function $\rho(x)$ is the product of the density functions of the individual components. However, it is well known that this problem always has many solutions (see [3] and references therein). Specifically, any observed density function can be integrated in order to construct an entire family of functions f^{-1} that transform it into a separable (i.e., factorizable) form.

M.E. Davies et al. (Eds.): ICA 2007, LNCS 4666, pp. 65–72, 2007.

The observed trajectories of many classical physical systems [4] can be characterized by density functions in *phase space* (i.e., $(\tilde{x}, \dot{\tilde{x}})$-space). Furthermore, if such a system is composed of non-interacting subsystems, the state space variables can be chosen so that the system's phase space density function is separable (i.e., is the product of the phase space density functions of the subsystems). This fact motivates the approach to BSS described in this paper [5]: we search for a function of the state space variable $\tilde{x}$ that transforms the observed phase space density function $\tilde{\rho}(\tilde{x}, \dot{\tilde{x}})$ into a separable form. Unlike conventional BSS, this "phase space BSS problem" has a unique solution in the following sense: either the data are inseparable, or they can be separated by a mixing function that is unique, up to transformations that do not affect separability (translations, permutations, and possibly nonlinear rescaling of individual source components). This form of the BSS problem has a unique solution because separability in phase space is a stronger requirement than separability in state space. In other words, if a choice of variables x leads to a separable phase space density function, it also produces a separable state space density function; however, the converse is not true. In particular, the above-mentioned procedure of using integrals of the state space density function to transform it into separable form [3] cannot be used to separate the phase space density function.

It was previously demonstrated [6] that the phase space density function of a time series induces a Riemannian metric on the system's state space and that this metric can be directly computed from the local velocity correlation matrix of the data. In the following Section, we show how this differential geometry can be used to determine if there is a source coordinate system in which the phase space density function is separable and, if so, to find the transformation between the coordinate systems of the observed variables and the source variables. In a technical sense, the method is straight-forward. The data-derived metric is differentiated to compute the affine connection and curvature tensor on state space. If the curvature tensor does not vanish, the observed data are not separable. On the other hand, if the curvature tensor does vanish, there is only one possible set of source variables (up to translations, permutations, and transformations of individual components), and it is possible to compute these explicitly and determine if they do separate the phase space density function. A longer version of this paper [5] describes the solution of a more general BSS problem (sometimes called multidimensional independent component analysis [MICA] or independent subspace analysis) in which the source components can be partitioned into groups, so that components from different groups are statistically independent but components belonging to the same group may be dependent [7,8,9].

As mentioned above, this paper exploits a stronger criterion of statistical independence than conventional approaches (i.e., separability of the phase space density function instead of separability of the state space density function). Furthermore, the new method differs from earlier approaches on the technical level. For example, the proposed method exploits statistical constraints on source time derivatives that are *locally* defined in the state space, in contrast to the usual criteria for statistical independence that are *global* conditions on the source time

series or its time derivatives [10]. Furthermore, the nonlinearities of the mixing function are unraveled by imposition of local second-order statistical constraints, unlike many conventional approaches that rely on higher-order statistics [1,2]. In addition, the constraints of statistical independence are used to construct the mixing function in a "deterministic" manner, without the need for parameterizing it (with a neural network architecture or other means) and without using probabilistic learning methods [11,12]. And, the new method is quite general, unlike some other techniques that are limited to the separation of post-nonlinear mixtures [13] or other special cases. Finally, the use of differential geometry in this paper should not be confused with existing applications of differential geometry to BSS. In our case, the observed measurement trajectory is used to derive *a metric on the system's state space*, and the vanishing of the curvature tensor is shown to be a necessary condition for separability of the data. In contrast, other authors [14] define *a metric on a completely different space, the search space of possible mixing functions,* so that "natural" (i.e., covariant) differentiation can be used to expedite the search for the function that optimizes the fit to the observed data.

2 Method

This Section describes how the phase space density function of the observed data induces a Riemannian geometry on the state space and shows how to compute the metric and curvature tensor of this space from the observed time series. Next, we show that, if curvature tensor is non-vanishing, the observed data are not separable. However, if the curvature tensor vanishes, we show how to determine whether the data are separable, and, if they are, we show how to find the mixing function, which is essentially unique.

Let $x = x(t)$ (x_k for $k = 1, 2, \ldots, n$) denote the trajectory of a time series. Suppose that there is a phase space density function $\rho(x, \dot{x})$, which measures the fraction of total time that the trajectory spends in each small neighborhood $dxd\dot{x}$ of $(x, \dot{x})$-space (i.e., phase space). As discussed in [6], most classical physical systems in thermal equilibrium with a "bath" have such a phase space density function: namely, the Maxwell-Boltzmann distribution [4]. Next, define $g^{kl}(x)$ to be the local second-order velocity correlation matrix [6]

$$g^{kl}(x) =< (\dot{x}_k - \bar{\dot{x}}_k)\,(\dot{x}_l - \bar{\dot{x}}_l) >_x \tag{2}$$

where the bracket denotes the time average over the trajectory's segments in a small neighborhood of x and where $\bar{\dot{x}} =< \dot{x} >_x$, the local time average of $\dot{x}$. In other words, g^{kl} is a combination of first and second moments of the local velocity distribution. Because this correlation matrix transforms as a symmetric contravariant tensor, it can be taken to be a contravariant metric on the system's state space. Furthermore, as long as the local velocity distribution is not confined to a hyperplane in velocity space, this tensor is positive definite and can be inverted to form the corresponding covariant metric g_{kl}. Thus, under these conditions, the time series induces a non-singular metric on state space. This metric

can then be used to compute the affine connection Γ^k_{lm} and Riemann-Christoffel curvature tensor $R^k{}_{lmn}$ of state space by means of the standard formulas of differential geometry [15]

$$\Gamma^k_{lm}(x) = \frac{1}{2} g^{kn} \left(\frac{\partial g_{nl}}{\partial x_m} + \frac{\partial g_{nm}}{\partial x_l} - \frac{\partial g_{lm}}{\partial x_n} \right) \tag{3}$$

and

$$R^k{}_{lmn}(x) = -\frac{\partial \Gamma^k{}_{lm}}{\partial x_n} + \frac{\partial \Gamma^k{}_{ln}}{\partial x_m} + \Gamma^k{}_{im} \Gamma^i{}_{ln} - \Gamma^k{}_{in} \Gamma^i{}_{lm} \tag{4}$$

where we have used the Einstein convention of summing over repeated indices.

Now, assume that the data are separable and that x represents a set of source variables; i.e., assume that the phase space density function ρ is equal to the product of density functions of each component of x. It follows from definition (2) that the metric $g^{kl}(x)$ is diagonal and has positive diagonal elements, each of which is a function of the corresponding coordinate component. Therefore, the individual components of x can be transformed in order to create a new state space coordinate system in which the metric is the identity matrix and the curvature tensor (4) vanishes. It follows that the curvature tensor must vanish in every coordinate system, including the coordinate system $\tilde{x}$ defined by the observed data

$$\tilde{R}^k{}_{lmn}(\tilde{x}) = 0 \tag{5}$$

In other words, the vanishing of the curvature tensor is a necessary consequence of separability. Therefore, if this data-derived quantity does not vanish, the data cannot be transformed so that their phase space density function is separable.

On the other hand, if the data do satisfy (5), there is only one possible separable coordinate system (up to transformations that do not affect separability), and it can be explicitly constructed from the observed data $\tilde{x}(t)$. To see this, first note that, on a flat manifold (e.g., (5)) with a positive definite metric, it is always possible to explicitly construct a "Euclidean" coordinate system for which the metric is the identity matrix. Furthermore, if a coordinate system has a diagonal metric with positive diagonal elements that are functions of the corresponding coordinate components, it can be derived from this Euclidean one by means of an n-dimensional rotation, followed by transformations that do not affect separability (i.e., translations, permutations, and transformations of individual components). Therefore, because every separable coordinate system must have a diagonal metric with the aforementioned properties, all possible separable coordinate systems can be found by constructing a Euclidean coordinate system and then finding all rotations of it that are separable. The first step is to construct a Euclidean coordinate system in the following manner: 1) at some arbitrarily-chosen point $\tilde{x}_0$, select n small vectors $\delta\tilde{x}_{(i)}$ $(i = 1, 2, \ldots, n)$ that are orthonormal with respect to the metric at that point (i.e., $\tilde{g}_{kl}(\tilde{x}_0)\delta\tilde{x}_{(i)k}\delta\tilde{x}_{(j)l} = \delta_{ij}$, where δ_{ij} is the Kronecker delta); 2) starting at $\tilde{x}_0$, use the affine connection to repeatedly parallel transfer all $\delta\tilde{x}$ along $\delta\tilde{x}_{(1)}$; 3) starting at each point along the resulting geodesic path, repeatedly parallel transfer these vectors along $\delta\tilde{x}_{(2)}$; ... continue the parallel transfer process along other directions ... n+1) starting

at each point along the most recently produced geodesic path, parallel transfer these vectors along $\delta\tilde{x}_{(n)}$. Finally, each point is assigned the geodesic coordinate s $(s_k, k = 1, 2, \ldots, n)$, where s_k represents the number of parallel transfers of the vector $\delta\tilde{x}_{(k)}$ that was required to reach it. Differential geometry [15] guarantees that the metric of a flat, positive definite manifold will be the identity matrix in a geodesic coordinate system constructed in this way. We can now transform the data into this Euclidean coordinate system and examine the separability of all possible rotations of it. The easiest way to do this is to compute the second-order correlation matrix

$$\sigma_{kl} =< (s_k - \bar{s}_k)(s_l - \bar{s}_l) > \tag{6}$$

where the brackets denote the time average over the entire trajectory and $\bar{s} =< s >$. If this data-derived matrix is not degenerate, there is a unique rotation that diagonalizes it, and the corresponding rotation of the s coordinate system is the only candidate for a separable coordinate system (up to transformations that do not affect separability). Its separability can be determined by explicitly computing the data's phase space density function in order to see if it factorizes in this rotated coordinate system. Alternatively, we can use higher-order statistical criteria to see if the rotated s components are truly independent.

In summary, the BSS problem can be solved by the following procedure:

1. Use the data $\tilde{x}(t)$ to compute the metric, affine connection, and curvature of the state space [(2-4)].
2. If the curvature does not vanish at each point, the data are not separable.
3. If the state space curvature does vanish:
   (a) Compute the transformation to a Euclidean coordinate system s and transform the data into it.
   (b) Find the rotation that diagonalizes the second-order correlation matrix σ and transform to the corresponding rotation of the s coordinate system.
   (c) Compute the phase space density function of the data in the rotated s coordinate system.
   (d) If the density function factorizes, the data are separable, and the rotated s coordinates are the unique source variables (up to translations, permutations, and transformations of individual components). If the density function does not factorize, the data are not separable.

3 Discussion

This paper outlines a new approach to nonlinear BSS that is based on a notion of statistical independence, which is characteristic of a wide variety of classical noninteracting physical systems. Specifically, the new method seeks to determine if the observed data are mixtures of source variables that have a *phase-space* density function equal to the product of density functions of individual components. This criterion of statistical independence is stronger than that of conventional approaches to BSS, in which only the *state-space* density function is required to be separable. Because of the relative strength of this requirement, the new

approach to BSS produces a unique solution in each case (i.e., data are either inseparable or are separable by a unique mixing function), unlike the conventional approach that always finds an infinite number of mixing functions. Given a time series of observations in a measurement-defined coordinate system ($\tilde{x}$) on the system's state space, the basic problem is to determine if there is another coordinate system (a source coordinate system x) in which the density function is factorizable. The existence (or non-existence) of such a source coordinate system is a coordinate-system-independent property of the time series of data (i.e., an intrinsic or "inner" property). This is because, in *all* coordinate systems, there either is or is not a transformation to such a source coordinate system. In general, differential geometry provides mathematical machinery for determining whether a manifold has a coordinate-system-independent property like this. In the case at hand, we can induce a geometric structure on the state space by identifying its metric with the local second-order correlation matrix of the data's velocity [6]. Then, a necessary condition for BSS is that the curvature tensor vanishes in all coordinate systems (including the measurement coordinate system). Therefore, if this data-derived quantity is non-vanishing, the observations are not separable. However, if the curvature tensor is zero, the data are separable if and only if the density function is seen to factorize in a coordinate system that can be explicitly constructed from the data-derived affine connection. In that case, these coordinates are the unique source variables (up to transformations that do not affect separability).

A longer version of this paper [5] describes the solution of a more general BSS problem (sometimes called multidimensional ICA or independent subspace analysis) in which the source components are only required to be partitioned into groups that are statistically independent of one another but contain statistically interdependent variables [7,8,9]. The possible separable coordinate systems are a subset of all coordinate systems in which the metric is *block*-diagonal (instead of fully diagonal as in this paper). All of these "block-diagonal coordinate systems" can be derived from geodesic coordinate systems constructed from geodesics along a finite number of special directions in state space, and these special directions can be computed from algebraic equations involving the curvature tensor. Thus, it is possible to construct every block-diagonal coordinate system and then explicitly determine if the density function is separable in it. An exceptional situation arises if the metric can be transformed into a block-diagonal form with two or more one-dimensional blocks. In this case, there is an unknown rotation on this two-dimensional (or higher dimensional) subspace that is not determined by the requirement of metric block-diagonality. However, much as in Sect. 2, this rotation can be determined by applying other statistical requirements of separability, such as block diagonality of the second-order state variable correlation matrix or block-diagonality of higher-order local velocity correlation functions. In reference [5], this procedure for performing multidimensional ICA is described in detail, and it is illustrated with analytic examples, as well as with a detailed numerical simulation of an experiment.

What are the limitations of the applicability of this method? It is certainly critical that there be a well-defined metric on state space. However, this will be the case if the measurement time series is described by a phase space density function, a requirement that is satisfied by the trajectories of a wide variety of physical systems [6]. In practical applications, the measurements must cover state space densely enough to be able to compute the metric, as well as its first and second derivatives (required to calculate the affine connection and curvature tensor). In the numerical simulation in [5], approximately 8.3 million short trajectory segments (containing a total of 56 million points) were used to compute the metric and curvature tensor on a three-dimensional state space. Of course, if the dimensionality of the state space is higher, even more data will be needed. So, a relatively large amount of data may be required in order to be able to determine their separability. There are few other limitations on the applicability of the technique. For example, computational expense is not prohibitive. The computation of the metric is the most CPU-intensive part of the method. However, it can be distributed over multiple processors by dividing the observed data into "chunks" corresponding to different time intervals, each of which is sent to a different processor where its contribution to the metric (2) is computed. As additional data are accumulated, they can be processed separately and then added into the time average of the data that were used to compute the earlier estimate of the metric. Thus, the earlier data need not be processed again, and only the latest observations need to be kept in memory.

Acknowledgments

The author is indebted to Michael A. Levin for insightful comments and for a critical reading of the manuscript. Aapo Hyvärinen and Shun-ichi Amari also read the manuscript, and their comments on the uniqueness issue are appreciated. This work was performed with the partial support of the National Institute of Deafness and Communicative Disorders.

References

1. Hyvärinen, A., Karhunen, J., Oja, E.: Independent Component Analysis. Wiley, New York (2001)
2. Jutten, C., Karhunen, J.: Advances in nonlinear blind source separation. In: Proceedings of the 4$^{\text{th}}$ International Symposium on Independent Component Analysis and Blind Signal Separation (ICA 2003), Nara, Japan (April 2003)
3. Hyvärinen, A., Pajunen, P.: Nonlinear independent component analysis: existence and uniqueness results. Neural Networks 12, 429–439 (1999)
4. Sears, F.W.: Thermodynamics, the Kinetic Theory of Gases, and Statistical Mechanics, 2nd edn. Addison-Wesley, Reading, MA (1959)
5. Levin, D.N.: Using state space differential geometry for nonlinear blind source separation (2006), `http://arxiv.org/abs/cs/0612096`
6. Levin, D.N.: Channel-independent and sensor-independent stimulus representations. J. Applied Physics 98, 104701 (2005), Also, see the papers posted at `http://www.geocities.com/dlevin2001/`

7. Cardoso, J-F.: Multidimensional independent component analysis. In: Proc. 1998 IEEE Int. Conf. on Acoustics, Speech and Signal Processing (ICASSP'98), Seattle, pp. 1941–1944. IEEE Computer Society Press, Los Alamitos (1998)
8. Bingham, E., Hyvärinen, A.: A fast fixed-point algorithm for independent component analysis of complex-valued signals. Int. J. of Neural Systems 10, 1–8 (2000)
9. Nishimori, Y., Akaho, S., Plumbley, M.D.: Riemannian optimization method on the flag manifold for independent subspace analysis. In: Rosca, J., Erdogmus, D., Príncipe, J.C., Haykin, S. (eds.) ICA 2006. LNCS, vol. 3889, pp. 295–302. Springer, Heidelberg (2006)
10. Lagrange, S., Jaulin, L., Vigneron, V., Jutten, C.: Analytic solution of the blind source separation problem using derivatives. In: Puntonet, C.G., Prieto, A.G. (eds.) ICA 2004. LNCS, vol. 3195, pp. 81–88. Springer, Heidelberg (2004)
11. Haykin, S.: Neural Networks - A Comprehensive Foundation, 2nd edn. Prentice Hall, New York (1998)
12. Yang, H., Amari, S.I., Cichocki, A.: Information-theoretic approach to blind separation of sources in non-linear mixture. Signal Processing 64, 291–300 (1998)
13. Taleb, A., Jutten, C.: Source separation in post nonlinear mixtures. IEEE Trans. on Signal Processing 47, 2807–2820 (1999)
14. Amari, S.: Natural gradient works efficiently in learning. Neural Computation 10, 251–276 (1998)
15. Weinberg, S.: Gravitation and Cosmology - Principles and Applications of the General Theory of Relativity. Wiley, New York (1972)