Max-Planck-Institut
für Mathematik
in den Naturwissenschaften
Leipzig

C^1-statistical manifolds are statistical models

by

Hông Vân Lê

Preprint no.: 77 2003

C^1-statistical manifolds are statistical models

Hông Vân Lê
hvle@mis.mpg.de
Max-Planck-Institute for Mathematics in Sciences
Inselstraße 22-26
D-04103 Leipzig

August 19, 2003

Abstract

We prove that any C^1-statistical manifold can be embedded into the space of all probability measures on a finite set. As a result we get a positive answer to the Lauritzen question on realization of C^1-statistical manifolds as statistical models.

1 Introduction

According to Lauritzen [Lau1987] a statistical manifold is a Riemannian manifold (M, g) with a 3-symmetric tensor T on M. His motivation examples are statistical models provided with the Fisher metric g^F and the Amari-Centsov tensor T^{AC}. Let us remind that statistical models are families $\{P_x, x \in M\}$ of probability measures P_x on a measurable space $(\Omega, \mathcal{A})$. It is natural in statistical problems to assume that M and Ω are C^k-manifolds, moreover $\{P_x\}$ is a family of absolutely muatually continuous Borel measures on Ω [CeMo1990]. Thus we shall write $p_x = p(x, \omega)d\omega$, where $p(x, \omega)$ is a non-negative function and $d\omega$ is a fixed Borel measure on Ω. The Fisher metric g^F and the Amari-Centsov tensor T^{AC} are defined as follows

$$g^F(V, W)_x := \int_\Omega (\partial_V \ln p(x, \omega))(\partial_W \ln p(x, \omega))p(x, \omega)d\omega,$$

$$T^{AC}(V,W,Z)_x := \int_\Omega (\partial_V \ln p(x,\omega))(\partial_W \ln p(x,\omega))(\partial_Z \ln p(x,\omega))p(x,\omega)d\omega.$$

The Fisher metric and the Amari-Censov tensor are the main geometric ingredients of the information geometry [Am2002] where one studies statistical estimation problems with geometric methods.

Lauritzen asked [Lau1987] how large is the class of statistical models inside the class of all statistical manifolds. In [Le2003] we proved that if $k \geq 3$ then any C^k-statistical manifold is a statistical model on a space Ω^N of $N < \infty$ elementary events. Here we mean C^k-statistical manifold as a smooth manifold M provided with C^k Riemannian metric g and C^k tensor T. This statement is equivalent to the statement on the existence of isostatistical embedding of the corresponding statistical manifolds, since the isostatistical embedding induces the probability potential $p(x,\omega)$ (see also [Le2003] for more details).

Here we say that an immersion $f : (M,g,T) \to (\bar{M},\bar{g},\bar{T})$ is called **isostatistical**, if $f^*(\bar{g}) = g$ and $f^*(\bar{T}) = T$.

We also notice that this embedding problem is close to the embedding problem of Riemannian manifolds, since any Riemannian manifold (M,g) is a statistical manifold with $T = 0$. Unlike the Riemannian case where there is no C^1-invariant of a Riemannian structure, we have constructed in [Le2003] a class of C^0 and C^1-invariants of statistical structures. Many of these invariants represent obstructions to embeddings of statistical manifolds, even in the case that the dimension of the target statistical space is arbitrary larger than the dimension of the domain statistical space. So there are serious differences in the embedding problems of Riemannian manifolds and of statistical manifolds.

Main Theorem. *Any C^1-statistical manifold is a statistical model. More precisely for any statistical manifold (M^m,g,T) there exists an isostatistical embedding $f : (M^m,g,T) \to (\Omega^{N(m)},g^F,T^{AC})$ with $N(m) = 4(m+1)[(m+1)(m+4)+2]$.*

In the second section we give the proof of our main theorem whose main ingriedents are the Nash C^1-embedding theorem, the Gromov C^1-embedding theorem for a manifold with a symmetric form of degree 3 and an embedding theorem for a linear statistical manifold which is a preciser (better) version of Lemma 5.1 in [Le2003].

2 Proof of the main theorem.

Let us first recall three embedding theorems we shall use in our proof of Main Theorem. On $\mathbf{R}^n$ we denote by g_0 the Euclidean metric and by T_0 the standard symmetric form of degree 3: $T_0 = \sum_{i=1}^n (dx_i)^3$.

2.1. The Nash C^1-embedding theorem. [Na1954] *Any C^1-Riemannian manifold (M^n, g) can be isometriccally embedded into $(\mathbf{R}^{n+1}, g_0)$.*

2.2. The Gromov C^1-embedding theorem. [Gr 1986, 2.4.3'] *Suppose that M^m is given with a C^1 symmetric 3-form T. Then there exists an embedding $f : M^m \to \mathbf{R}^{(m+1)(m+2)+m}$ such that $f^*(T_0) = T$.*

We denote by Cap_+^N the interior part of Cap^N. Is is well-known [A-N 2000], [Le2003] that the statistical manifold (Cap^N, g^F, T^{AC}) can be relialized as the positive quadrant $S_+^{N-1}(2)$ of the sphere of radius 2 in the positive quadrant $\mathbf{R}_+^n \subset \mathbf{R}^n$ which is provided with the Eucidean metric g_0 and following symmetric 3-form T_S

$$T_S(x_1, \cdots, x_n) = \sum_{i=1}^n \frac{(dx_i)^3}{x_i}, \tag{2.3}$$

In other words the metric g^F on S_+^N is the standard metric of positive constant curvature and the tensor T^{AC} on S_+^{N-1} is the restriction of T_S to $S_+^{N_1}$.

2.4. Embedding theorem for linear statistical manifolds. [Le2003, Lemma 5.17]*For any A and any open bounded domain $D^n \subset \mathbf{R}^n$ there exists an isostatstical embedding $i_s : (D^n, g_0, A \cdot T_0) \to (Cap^{4n}+, g^F, T^{AC})$ i.e.*

$$i_s(g^F) = g_0,\ i_s(A \cdot T^{AC}) = T_0.$$

In [Le 2003] our argument yields actually a little stronger statement. Here is a precise version of Lemma 5.17 in [Le2003]. This version has been used (inexplicitely) in our proof of [Le2003, Theorem 5.2] for the unbounded case.

We denote by $S^{4n-1}(B)$ the sphere of radius B in $\mathbf{R}^{4n}$ and by i the standard embedding of this sphere into $\mathbf{R}^{4n}$.

2.4.a. Proposition. Embbedding theorem for linear statistical manifolds. *Given positive numbers A, B and for any open bounded domain $D \subset \mathbf{R}^n$ and any closed subset $U \subset S_+^{4n-1}(B)$ there exists an isostatistical embedding*

$$i_s : (D^n, g_0, A \cdot T_0) \to (S_+^{4n-1}(B) \setminus U, i^*(g_0), i_*(T_S)).$$

In the first step of the proof of Main Theorem we shall show

2. 5. Theorem. *Suppose that (M^m, g, T) is a compact C^1-statistical manifold. Then there exists a positive number A and an isostatistical embedding $f : (M^m, g, T) \to (\mathbf{R}^{l(m)}, g_0, A \cdot T_0)$ with $l(m) = (m+1)(m+4) + m$.*

We fisrt note that Theorem 2.5 does not hold for non-compact statistical manifolds, (unlike the Riemanian case), since in this case the commass $\mathcal{M}^1(M)$ (see [Le2003] for a definition) can be equal to infinity.

Proof of Theorem 2.5. First we shall take an embedding $f_1 : (M^m, g, T) \to (\mathbf{R}^{(m+1)(m+2)+m}, g_0, T_0)$ such that

$$f_1^*(T_0) = T.$$

The existence of f_1 follows from the Gromov embedding thereom.

Then we choose a positive number A^{-1} such that

$$g - A^{-1}(f_1^*(g_0)) = g_1$$

is a Riemannian metric on M. Such a number A exists, since M is compact.

Now we shall choose an isometric embedding $f_2 : (M^m, g_1) \to (\mathbf{R}^{m+1}, g_0)$. The existence of f_2 follows from the Nash C^1-embedding theorem.

Next we choose a linear isometric embedding $L_{m+1} : \mathbf{R}^{m+1} \to |R^{2m+2}$ such that $L_{n+1}(T_0) = 0$. The existence of such a linear mapping is obvious, see e.g. [Le2003, Theorem 3.13.b].

Finally we take an embedding

$$f_3 : M^m \to \mathbf{R}^{(m+1)(m+2)+m} \oplus \mathbf{R}^{2m+2}$$

as follows.

$$f_3(x) = A \cdot f_1(x) \oplus (L_{n+1} \circ f_2).$$

Clearly f_3 is an isostatistical embedding $(M, g, T) \to (\mathbf{R}^{l(m)}, g_0, A \cdot T_0)$.□

Proof of Main Theorem. If M is compact, then the existence of an isostatistical embedding $i_s(M, g, T) \to (Cap^{4l(m)}, g^F, T^{AC})$ follows from Theorem 2.5 and Proposition 2.4.a. To find an isostatistical embedding as in our main Theorem we shall repeat the argument in [Le2003] for a non-compact M. For the convenience of the reader (and to convice her/him) we shall write down here a proof with more details than in [Le2003].

We can deal with this case by using the compact decomposition of M^m as Nash did for the isometric embedding for the smooth case[Nash1956]. Namely we cover M^m by disk neighborhoods N_i^j, $j = \overline{1,m}$ in such a way. For each j let

$$C^j := \cup_i N_i^j. \tag{2.6}$$

Then we require that the union in (2.6) is a disjoint union, i.e. $N_i^j \cap N_k^j = \emptyset$, if $i \neq k$. We also require that each N_i^j overlaps only a finite number of other N_l^k. Now we "compactify" N_i^j via an surjective smooth mapping : $\phi_i : N_i^j \to S_i^j$, where S_i^j is a sphere of the same dimension m. The map ϕ_i^j can be extended to the whole M^m, since it maps the boundary of N_i^j into the north point of the sphere. On the other hand, this map ϕ_i^j is injective in a large (enough) subdomain $\bar{N}_i^j \subset N_i^j$. We can furthermore use the unity partition function to define a C^1 statistical structure on each S_i^j such that the (sum of) pull back via ϕ_i^j is the given statistical structure on M. In other words we can consider the C^1 statistical structure on M^m as induced from (infinitely many) spheres S_i^j via the smooth mapping ϕ_i^j.

Now let

$$S^j = \cup_i S_i^j.$$

Using Theorem 2.4.a and Theorem 2.5 we can find isostatistical embedding

$$\psi^j : S^j \to (S_+^{4l(m)-1}(\frac{2}{\sqrt{m+1}}), i^*(g_0), i^*(T_S))$$

inductively, since each S_i^j is compact.

Now let us consider the map

$$I_j : M^m \to (S_+^{4l(m)-1}(\frac{2}{\sqrt{m+1}}), i^*(g_0), i^*(T_S)),$$

as a composition of the map $\phi_i^j : M \to S_i^j$ and the map ψ_j.

Finally the product mapping

$$I_0 = I_1 \times \cdots \times I_{m+1}$$

is the desired embedding, since the statistical structure on $S^{N(m)-1}$ is induced from the (non-linear) statistical structure on $\mathbf{R}^{N(m)}$. We can see this easily by noticing that both symmetrric forms g_0 and T_S are decomposable w.r.t. the embedding $i : \mathbf{R}^n \to \mathbf{R}^{n+l}$ for any $n, l > 0$, i.e. $g_0(\mathbf{R}^n) = i_*(g_0)(\mathbf{R}^{n+l})$, $T_S(R^n_+) = i^*(T_S)(\mathbf{R}^{n+l}_+)$. $\square$

2.6. Remark. Theorem 2.2 has been proved by Gromov [Gr1986] by using the condition of "hyper regularity" and the triangular form of the corresponding PDE. This technique is a part of the convex integration method which Gromov generalized from the Nash C^1-embedding theorem. We note that the PDE for an isostatistical embedding $(M, g, T) \to (R^N, g_0, T_0)$ cannot have triangular form, since it cannot be solved even locally without an extra condition on the comass $\mathcal{M}^1(M, g, T)$. This condition has been taken into account in our Theorem 2.5 by adding a positive constant A.

Acknowledgement. I thank Nihat Ay, Jürgen Jost for their interest and helpful comments.

References

[Ay2002] N. Ay, *An information-geometric approach to a theory of pragmatic structuring*, Ann. of Prob., v.30, N. 1, 416-436 (2002).

[A-N2002] S. Amari and H. Nagaoka, *Methods of Information Geometry*, Trans. of Math. Monograph (2000).

[Am1997] S. Amari, *Information geometry*, Contemporary Math. vol 203, 81-95, (1997).

[C1972] N.N. Chentsov, *Statistical decision rules and optimal inference*, AMS., (1982), (originally published in Russian, Nauka, Moscow , 1972).

[CeMo1990] E.A. Morozova and N.N. Chentsov, *Markov invariant geometry on manifolds of states* (in Russian), Itogi Nauki i Techniki 36, 69-102, (1990).

[Gr1986] M. Gromov, *Partial Differential Relations,* Springer-Verlag, (1986).

[Jost2002] J. Jost, *Information geometry*, preprint MPI 2002.

[Lau1987] S. Lauritzen, *Statistical manifolds*, In : Differential geometry in Statistical Inference, IMS Lecture Notes, Mongraph Serie 10., Inst. of Math. Stat. Hayward, California, (1987).

[Le2003] H.V. Le *Monotonen invariants and embeddings of statistical manifolds*, preprint of MPI http:www.mis.mpg.de/.

[Nash1954] J. Nash, C^1*-isometric imbeddings*, Ann. Math. 60, 383-396, (1954).