Mismatched Estimation in an Exponential Family

Subrahamanian Moosath K S and Harsha K V

Indian Institute of Space Science and Technology, Trivandrum, India ; Indian Institute of Technology Bombay, Mumbai, India
smoosath@iist.ac.in ; harshakv@iitb.math.ac.in

IGAIA 2016 conference at Liblice, Czech Republic. 12-17 June 2016

Introduction

- A mismatched estimation problem uses a mismatched model (unfaithful model) instead of the original model for estimation.
- Here we discuss the information geometric approach to the general estimation problem based on a mismatched model in an exponential family.

Exponential Family and Mismatched Estimator

Let $\mathcal{S} = \{p(x;\theta) = \exp\{\sum_{i=1}^{n} \theta^i x_i - \psi(\theta)\} \;/\; \theta \in E \subseteq \mathbb{R}^n\}$ be an n-dimensional exponential family, where $x = (x_1, \cdots, x_n)$ - set of random variables and $\theta = (\theta^1, \cdots, \theta^n)$ is the canonical coordinate. Since $\mathcal{S}$ is dually flat space, the dual coordinate $\eta = (\eta_1, \cdots, \eta_n)$ of θ is defined by $\eta_i = \int x_i \, p(x;\theta)\, dx$, [?].

Now let $M = \{q(x;u) \;/\; u = (u^a) \in \mathbb{R}^m\}$ be an m-dimensional curved exponential family. Consider a mismatched model $M^* = \{q'(x;u) \;/\; u = (u^a) \in \mathbb{R}^m\}$ corresponding to the original model M. Let the embedding functions of M and M^* in $\mathcal{S}$ be $\theta(u)$ and $\theta'(u)$ respectively. Let $\eta(u)$ and $\eta'(u)$ be the corresponding dual representations.

To estimate the parameter $u \in M$, we use M^* instead of M. Let $x_N = (x^1, \cdots, x^N)$ be N independent observations from $q(x;u) \in M$. Then the observed point $\bar{x} = (\bar{x}_1, \cdots, \bar{x}_n)$ defines a distribution in $\mathcal{S}$ whose η-coordinate is given by $\hat{\eta}_N = \bar{x}$. $\bar{x}$ is a sufficient statistic for M^*.

The estimator $\hat{u}'_N$ based on mismatched model M^* is represented as a mapping f'_N from $\mathcal{S}$ to M^*

$$f'_N : \mathcal{S} \longrightarrow M^* \quad \text{where} \quad \hat{\eta}_N \mapsto \hat{u}'_N = f'_N(\hat{\eta}_N) \tag{1}$$

The **ancillary manifold** or the **estimating submanifold** $A'_N(u)$ corresponding to the point $u \in M^*$ associated with f'_N is defined as

$$A'_N(u) = f'^{-1}_N(u) = \{\eta = (\eta_i) \in \mathcal{S} \;/\; f'_N(\eta) = u\} \tag{2}$$

That is, $A'_N(u)$ is the set of all points η in $\mathcal{S}$ which are mapped to $u \in M^*$ by the estimator f'_N.

Now we analyze the characteristics of an estimator $\{\hat{u}'_N, \; N = 1, 2, \cdots\}$ in M^* using the geometric properties of the ancillary submanifold $A'_N(u)$. Let

$$A'(u) = \lim_{N\to\infty} A'_N(u) \tag{3}$$

Also let $A_N(u)$ be the ancillary manifold corresponding to the point $u \in M$ and let

$$A(u) = \lim_{N\to\infty} A_N(u) \tag{4}$$

Consistency and efficiency of a general mismatched estimator

Theorem
An estimator $\{\hat{u}'_N, \; N = 1, 2, \cdots\}$ for $u \in M^*$ is consistent if and only if $\eta(u) \in M \subset \mathcal{S}$ is in the estimating submanifold $A'(u)$ attached to the point $u \in M^*$.

Theorem
A consistent estimator $\{\hat{u}'_N, \; N = 1, 2, \cdots\}$ for $u \in M^*$ is first order efficient if and only if $A'(u)$ is orthogonal to M at the intersecting point $\eta(u) \in M$.

Mismatched Maximum Likelihood Estimator (MLE)

Ozumi et al. [?] stated certain conditions for MLE based on a mismatched model to be consistent and efficient. We gave a theoretical formulation of these results and a detailed proof of the same.

Theorem
Let $\hat{u}'$ be the MLE in M^*. Then $\hat{u}'$ is a consistent estimator of u iff

$$q'(x;u) = \arg\min_{v\in M^*} D_{-1}(q(x;u), q'(x;v)) \tag{5}$$

Theorem

Let $\hat{u}'$ be the consistent MLE in M^*. Then $\hat{u}'$ is first order efficient iff

$$q(x;u) = \arg\min_{v\in M} D_{-1}(q'(x;u), q(x;v)) \tag{6}$$

where D_{-1} is the (-1)-divergence or the Kullback-Leibler divergence.

Corollary
Let $\hat{u}'$ be the MLE in M^*. Let γ be the (-1)-geodesic connecting $q(x;u) \in M$ and

$$q'(x;u) \in M^*.$$

Then

1. The MLE $\hat{u}'$ is consistent iff γ is orthogonal to M^*.
2. The consistent MLE $\hat{u}'$ is first order efficient iff γ is orthogonal to both M and M^*.

Conclusion

- An information geometric approach to the mismatched estimation problem in an exponential family is given.
- We proved a necessary and sufficient condition for an estimator based on a mismatched model to be consistent and efficient.
- Also we gave a theoretical formulation of the consistency and efficiency of the mismatched MLE.

We express our sincere gratitude to Prof. Shun-ichi Amari for the fruitful discussions.

References

- S.I. Amari, *Differential-Geometrical methods in Statistics*, Lecture Notes in Statistics, Springer-Verlab, New York, 28, 1985.
- Amari, S. and Nagaoka, H. (2000). *Methods of Information Geometry, Translations of Mathematical Monographs*. Oxford University Press: Oxford.
- Oizumi, M. Okada, M. and Amari, S. (2011). Information loss associated with imperfect observation and mismatched decoding. *Frontiers in Computational Neuroscience*, 5(9): 2011.