Information Geometry in Multiple Priors Models, Worst Case and Almost Worst Case Distributions

Imre Csiszár
Alfréd Rényi Institute of Mathematics
Hungarian Academy of Sciences, Hungary
e-mail: csiszar.imre@renyi.mta.hu

Thomas Breuer
PPE Research Centre
FH Vorarlberg, Austria
e-mail: thomas.breuer@fhv.at

Abstract

Minimisation of the expectation $E_{\mathbb{P}}(X)$ of a random variable X over a family Γ of plausible distributions $\mathbb{P}$ is addressed, when Γ is a level set of some entropy functional. It is shown, using a generalized Pythagorean identity, that whether or not a worst case distribution (minimising $E_{\mathbb{P}}(X)$ subject to $\mathbb{P} \in \Gamma$) exists, the almost worst case distributions cluster around an explicitely specified, perhaps incomplete distribution. It is also analysed how the existence of a worst case distribution depends on the threshold defining the set Γ.

1 Entropy functionals

An *entropy functional* over nonnegative measurable functions pm a set Ω, equipped with a (finite or σ-finite) measure μ, is

$$H(p) = H_\beta(p) := \int_\Omega \beta(\omega, p(\omega))\mu(d\omega),$$

where β belongs to the collection $\mathbb{B}$ of functions $\beta(\omega, s)$ on $\Omega \times \mathbb{R}^+$, measurable in ω for each $s \in \mathbb{R}^+$, and strictly convex and differentiable in s on $(0, +\infty)$ for each $\omega \in \Omega$, with $\beta(\omega, 0) = \lim_{s\downarrow 0} \beta(\omega, s)$.

Entropy functionals are of main interest for densities $p = \frac{d\mathbb{P}}{d\mu}$ of distributions $\mathbb{P}$. Best known is I-divergence or relative entropy, also familiar are other f*-divergences*

$$D_f(\mathbb{P}||\mu)) = \int f(p(\omega))\mu(d\omega) \quad \text{if } \mu(\Omega) = 1,$$

and *Bregman distances*

$$B_f(p, q) = \int \Delta_f(p(\omega), q(\omega))\mu(d\omega).$$

Here $\Delta_f(s, r) = f(s) - f(r) - f'(r)(s - r)$. More general Bregman distances, with any $\beta \in \mathbb{B}$ in the role of f, will also be needed.

2 The problem

In Mathematical Finance, the monetary payoff or utility of some action, e.g. of a portfolio choice, is a function $X(\omega)$ of a collection ω of random risk factors governed by a distribution not known exactly that can often be assumed to belong to a set Γ of "plausible" distributions. Then the negative of the worst case expected payoff

$$\inf_{\mathbb{P}\in\Gamma} E_{\mathbb{P}}(X) = \inf_{\mathbb{P}\in\Gamma} \int_{\Omega} X(\omega)\mathbb{P}(d\omega)$$

is a measure of the risk of this action. Familiar choices for Γ are an I-divergence ball or some other f-divergence ball around a "default distribution".

More gererally, $\Gamma = \{\mathbb{P} : d\mathbb{P} = p d\mu,\ H(p) \leq k\}$ will be considered with any entropy functional H and threshold k, addressing the corresponding infimum $V(k)$. This problem motivated by Mathematical Finance appears of independent interest, and the results obtained are expected to be relevant also in other fields. Study of the closely related problem of minimising entropy functionals subject to moment constraints has substantially contributed to the development of Information Geometry. Results of joint work of the first author with F. Matúš on that problem, including a generalised Pythagorean identity, will be essentially used in this talk.

3 Sketch of results

First, a known expression of $V(k)$ when H is an f-divergence is extended to general entropy functionals.

A main result (new even for I-divergence) is that the densities p with $H(p)$ close to k and $E_p(X)$ close to $V(k)$ cluster in Bregman distance around an explicitly specified function that equals the worst case density if it exists, but otherwise may even have integral less than 1.

Next, the function $V(k)$ is analysed, it is shown differentiable in typical cases but not always. Finally, the dependence on k of the existence of a worst case density is addressed. In case of f-divergences, it is shown to either exist for all k, or to exist/do not exist for k less/larger than a critical value. A conjecture is formulated about how this result might extend to general entropy functionals.

Acknowledgement

Imre Csiszár is supported by the Hungarian National Science Foundation, Grant 105840. Thomas Breuer is supported by the Josef Ressel Centre for Scientific Computing in Finance, Logistics, and Energy. Parts of this paper were presented at ISIT 2013 in Istanbul and at GSI 2015 in Paris. The full paper has been submitted to IEEE Transactions on Information Theory.