Minimization of entropy functionals under moment constraints

Imre Csiszár

Rényi Mathematical Institute

Based on joint work with F. Matúš, Prague, we address minimization of *convex integral functionals*

$$H_\beta(p) \triangleq \int_\Omega \beta(\omega, p(\omega)\mu(d\omega) \tag{1}$$

of non-negative functions p on a σ-finite measure space $(\Omega, \mathcal{F}, \mu)$, subject to the constraint that p has a specified *moment vector* $\int_\Omega \varphi p \, d\mu$ for a given vector-valued function $\varphi \colon \Omega \to \mathbb{R}^d$. Here β is a convex normal integrand, with $\beta(\omega, t)$, $\omega \in \Omega$, strictly convex for $t > 0$ and equal to $+\infty$ for $t < 0$. Typical examples of $H_\beta(p)$ are negative Shannon or Burg entropy, L^2-norm square, Kullback I-divergence, general f-divergences and Bregman distances.

Let $\mathcal{P}_a$ denote the class of the non-negative functions with moment vector $a = (a_1, \dots, a_d)$. If $\phi_1 \equiv 1$ and $a_1 = 1$ then $\mathcal{P}_a$ consists of probability densities.

Convex duality techniques are applied, studying the *primal problem*

$$J_\beta(a) \triangleq \inf\nolimits_{p \in \mathcal{P}_a} H_\beta(p) \, , \qquad a \in \mathbb{R}^d \, , \tag{2}$$

via the *dual problem* (where * denotes convex conjugation)

$$J_\beta^{**}(a) \triangleq \sup\nolimits_{\vartheta \in \mathbb{R}^d} \left[\langle \vartheta, a \rangle - J_\beta^*(\vartheta) \right] , \qquad a \in \mathbb{R}^d \, ; \tag{3}$$

in particular, an integral representation of $J_\beta^*(\vartheta)$ is given.

A minimizer/maximizer in (2)/(3) is called a primal/dual solution. A special case of problem (3) is maximum likelihood estimation in exponential families, and the concept of exponential family admits an extension to the general case.

A general *Pythagorean identity* will be established, which implies in general that even if no primal solution exists, a dual solution gives rise to a *generalized primal solution* which is a specified member of the exponential family in the extended sense.

Other results include a geometric characterization of the set

$$\mathrm{dom}(J_\beta) \triangleq \{a : J(a) < +\infty\},$$

and a recipe to extend results known under the *constraint qualification* that a is in the relative interior of $\mathrm{dom}(J_\beta)$, to all $a \in \mathrm{dom}(J_\beta)$.