An Information Geometry Problem in Mathematical Finance

Imre Csiszár[1] Thomas Breuer[2]

[1]Alfréd Rényi Institute of Mathematics, Hungarian Academy of Sciences, Hungary
csiszar@renyi.hu
Supported by the Hungarian National Science Foundation, Grant 105840

[2]PPE Research Centre, FH Vorarlberg, Austria
tb@fhv.at
Supported by the Christian Doppler Gesellschaft

GSI 2015
29 October 2015
Ecole Polytechnique, Paris-Saclay

The problem

Problem: Minimize the expectation $E_{\mathbb{P}}(X)$ of a real valued random variable X for distributions $\mathbb{P}$ in a "plausible set" Γ.

Motivation: The monetary payoff or utility of a financial act, such as a portfolio selection, is a function $X(\omega)$ of a collection of random risk factors whose distribution $\mathbb{P}$ is unknown but $\mathbb{P} \in \Gamma$ may be assumed.

A measure of risk of this act, in this multiple priors model, is the negative of the worst case expected payoff

$$\inf_{\mathbb{P}\in\Gamma} E_{\mathbb{P}}(X) = \inf_{\mathbb{P}\in\Gamma} \int X(\omega)\mathbb{P}(d\omega). \tag{1}$$

In the theory of preferences, the infimum (1) serves as a criterion by which decision makers may prefer one act to another.

How to chose the plausible set Γ

Intuition: the plausible distributions are those not deviating much from a default distribution $\mathbb{P}_0$.
Deviation measures: I-divergence (relative entropy) appears most versatile, larger classes are f-divergences and Bregman distances.

In this talk, we consider plausible sets consisting of distributions dominated by a given (σ-finite) measure μ, of form

$$\Gamma = \{\mathbb{P} : d\mathbb{P} = p d\mu,\ H(p) \leq k\} \quad (2)$$

where H is an entropy functional (convex integaral functional).
I-divergence balls, and general f-divergence or Bregman distance balls around a default distribution arise by specific choices of H.

Goal: Determine $\inf_{\mathbb{P}\in\Gamma} E_{\mathbb{P}}(X)$, the worst case distribution attaining the minimum (if attained), and the behavior of almost worst case distributions in general, our main contribution.

History sketch

Multiple prior models, risk measures, theory of preferences: Főllmer and Schied 2004, Hansen and Sargent 2008, Gilboa 2009. Plausible sets Γ: I-divergence balls Hansen and Sargent 2001, Ahmadi-Javid 2011; f-divergence balls Maccheroni, Marinacci, Rustichini 2006, Ben Tal and Teboulle 2007.

Axiomatic approach leading to specific divergences: In inference context to I-divergence, with f-divergences and Bregman distances as alternatives: Csiszár 1991. In mathematical finance, distinguishing I-divergence: Strzalecki 2011.

Moment problem: Geometric view goes back to Chentsov 1972, clustering of approximate solutions to Topsoe 1979, Csiszár 1984, convex duality approach to Borwein and Lewis 1991, 1993. General results relied upon in this talk: Csiszár and Matús 2012.

Basic framework used in this talk: Breuer and Csiszár 2013.

Formal definitions

σ-finite measure space $(\Omega, \mathcal{F}, \mu)$, nonnegative, finite valued (measurable) functions on Ω are denoted by p or q.
Equality $p = q$ in μ-a.e. sense.

Denote by $\mathbb{B}$ the class of functions $\beta(\omega, s)$ on $\Omega \times \mathbb{R}$ that are

- for each $s \in \mathbb{R}$, measurable in ω
- for each $\omega \in \Omega$, strictly convex, differentiable in s on $(0, +\infty)$, equal to $+\infty$ if $s < 0$, and $\beta(\omega, 0) = \lim_{s \downarrow 0} \beta(\omega, s)$.

For $\beta \in \mathbb{B}$ define the entropy functional $H = H_\beta$ by

$$H(p) = H_\beta(p) \triangleq \int_\Omega \beta(\omega, p(\omega))\mu(d\omega). \tag{3}$$

The functions $\beta \in \mathbb{B}$ are convex normal integrands, hence $\beta(\omega, p(\omega))$ and similar functions later on are measurable.
(Shannon differential entropy is $-H(p)$ with $\beta(\omega, s) = s \log s$.)

Special cases

- Let μ equal the default distribution $\mathbb{P}_0$, let $\beta(\omega, s) = f(s)$ be an autonomous convex integrand with $f(1) = 0$. Then $H(p)$ in (3) with $p = d\mathbb{P}/d\mu$ is the *f-divergence* $D_f(\mathbb{P} \,||\, \mathbb{P}_0) = \int f(\frac{d\mathbb{P}}{d\mathbb{P}_0})d\mathbb{P}_0$.
- Let μ and the default distribution $\mathbb{P}_0 \ll \mu$ be arbitrary, f a strictly convex differentiable function on $(0, +\infty)$, and for $s \geq 0$ let $\beta(\omega, s) = \Delta_f(s, p_0(\omega))$ where

$$\Delta_f(s,t) \triangleq f(s) - f(t) - f'(t)(s-t); \qquad (4)$$

  if f is steep ($f'(0) = -\infty$), assume that $p_0 > 0$ μ-a.e. Then $H(p)$ equals the Bregman distance $B_f(p, p_0)$.
- In the special case $f(s) = s \log s$, both examples above give for $H(p)$ in (3) with $p = d\mathbb{P}/d\mu$ the *I*-divergence $D(\mathbb{P}||\mathbb{P}_0) = \int p \log \frac{p}{p_0} d\mu$, . Then the plausible set Γ in (2) is the I-divergence ball $\{\mathbb{P} : D(\mathbb{P}||\mathbb{P}_0) \leq k\}$.

Standing Assumptions

- X is a real valued measurable function and $\mathbb{P}_0 \ll \mu$ a default distribution on Ω, with density p_0
- $-\infty \leq m < b_0 < M \leq +\infty$ where
$$m \triangleq \mu\text{-ess inf X}, \ M \triangleq \mu\text{-ess sup X}$$
$$b_0 \triangleq E_{\mathbb{P}_0}(X) = \int X(\omega) p_0(\omega) \mu(d\omega),$$
- $H(p) \geq H(p_0) = 0 \quad$ whenever $\int p d\mu = 1$.
- $0 < k < k_{\max} \triangleq \lim_{b \downarrow m} F(b)$ where
$$F(b) \triangleq \inf_{p: \int p d\mu = 1, \int X p d\mu = b} H(p). \tag{5}$$

The version of Problem (1) we address is

$$V(k) \triangleq \inf_{p: \int p d\mu = 1, H(p) \leq k} \int X p d\mu. \tag{6}$$

Main Lemma

Lemma (Main Lemma)

Under our standing assumptions, there exists a unique b with

$$F(b) = k, \qquad m < b < b_0 \tag{7}$$

and then $V(k) = b$. *A density* p *attains the minimum in* (6), *the definition of* $V(k)$, *if and only if it attains that in* (5) *for* b *in* (7).

Example

Autonomous integrand $\beta(\omega, s) = f(s) = -\log s$, let $\mu = \mathbb{P}_0$. Then $H(p)$ for $p = d\mathbb{P}/d\mathbb{P}_0$ is the reverse I-divergence $D(\mathbb{P}_0||\mathbb{P})$. Specifically, let $\Omega = (0,1)$, $X(\omega) = \omega$, let $\mu = \mathbb{P}_0$ have density 2ω with respect to the Lebesgue measure. Then

$$f^*(r) = -1 - \log(-r)\ (r < 0),\ \ K(\theta_1, \theta_2) = \int_0^1 [-1 - \log(-\theta_1 - \theta_2\omega)] 2\omega d\omega$$

$$\Theta = \mathrm{dom}\, K = \{(\theta_1, \theta_2) : \theta_1 \leq 0,\ \theta_1 + \theta_2 < 0\}.$$

To $\theta_2 < 0$ there exists θ_1 such that $p_{\theta_1,\theta_2}(\omega) = 1/(-\theta_2 - \theta_2\omega)$ is a μ-density ($\int p_{\theta_1,\theta_2}(\omega) 2\omega d\omega = 1$) if and only if $\theta_2 \geq -2$. Otherwise

$$G(\theta_2) = K(0, \theta_2) = \int_0^1 [-1 - \log(-\theta_2\omega)] 2\omega d\omega = -\log(-\theta_2) - 1/2.$$

Calculus gives that $\theta_2 < -2$ is a maximizer of $[k + G(\theta_2)]/\theta_2$ if $k = \log(-\theta_2) - 1/2 > \log 2 - 1/2$. Then $V(k) = e^{-(k+1/2)}$, and $q_k(\omega) = (\theta_2\omega)^{-1}$, not a density. If $k \leq \log 2 - 1/2$ then $\theta_2 \geq -2$ and $q_k = p_{\theta_1(\theta_2),\theta_2}$ is the worst case density (no explicit formulas).

Moment problem

Given a moment mapping $\phi : \Omega \to \mathbb{R}^d$, minimize $H(p)$ subject to the moment constraint $\int \phi p d\mu = a \ \ (a \in \mathbb{R}^d)$.
This moment problem arises in inferring a nonnegative function p (often a probability density) when only the moment vector $\int \phi p d\mu = a$ is known: one may adopt, as best guess, the minimizer of $H(p)$ subject to $\int \phi p d\mu = a$.

Extensively studied, particularly for entropy functional equal to I-divergence (thus for $\beta(\omega, s) = s \log s$). This has substantially contributed to the development of information geometry, including concepts like information projection and Pythagorean identities..

We will use results available on the moment problem, taken from Csiszár and Matúš 2012, with the choice $d = 2$, $\phi(\omega) = (1, X(\omega))$.

Differential geometry is often regarded a basic ingerdient of information geometry, but it will not be used here. On the other hand, as in the moment problem, convex duality will be a key tool.

Invoking moment problem results

For $\phi(\omega) = (1, X(\omega))$, consider minimization of $H(p) = H_\beta(p)$ subject to $\int \phi p d\mu \triangleq (\int p d\mu, \int X p d\mu) = (a, b) \in \mathbb{R}^2$:

$$J(a, b) \triangleq \inf_{p: \int p d\mu = a, \int X p d\mu = b} H(p). \tag{8}$$

Instances of CsM 2012, Theorem 1.1 and Lemma 6.6 :

- Convex conjugate $J^*(\theta_1, \theta_2) \triangleq \sup_{a,b}[\theta_1 a + \theta_2 b - J(a, b)]$ of the function (8) equals

$$K(\theta_1, \theta_2) \triangleq \int \beta^*(\omega, \theta_1 + \theta_2 X(\omega))\mu(d\omega). \tag{9}$$

  Convex conjugate and (later on) derivative of β are by its second variable.

- The interior of $\mathrm{dom} J \triangleq \{(a, b) : J(a, b) < +\infty\}$ is

$$\text{int dom } J = \{(a, b) : am < b < aM\}$$

Implications of convex duality

Since $F(b) = J(1,b)$, standard convex duality results give for $(1,b)$ not on the boundary of $\mathrm{dom} J$, i.e., for $b \neq m$, $b \neq M$, that

$$\begin{aligned} F(b) &= J(1,b) = J^{**}(1,b) = \sup_{\theta_1,\theta_2}[\theta_1 + \theta_2 b - K(\theta_1,\theta_2)] && (10) \\ &= \sup_{\theta_2}[\theta_2 b - G(\theta_2)] = G^*(b) && (11) \end{aligned}$$

where

$$G(\theta_2) \triangleq \inf_{\theta_1}[K(\theta_1,\theta_2) - \theta_1]. \quad (12)$$

Moreover, in (10),(11) the maximum is attained if $b \in (m, M)$. A maximizer (θ_1, θ_2) in (10) is equivalently a subgradient of J at $(1,b)$. Its component θ_2, a maximizer in (11), is the slope of a supporting line of the graph of F at b.

(11) implies that $F^* = G^{**}$. Stronger result: $F^* = G$.

Evaluation of $V(k)$

Fix $k \in (0, k_{\max})$.
By Main Lemma, $k = F(b)$ where $b = V(k) \in (m, b_0)$.
For this b the maximizer θ_2 in (11) is negative, thus

$$k = \max_{\theta_2 < 0} [\, \theta_2 V(k) - G(\theta_2) \,]. \tag{13}$$

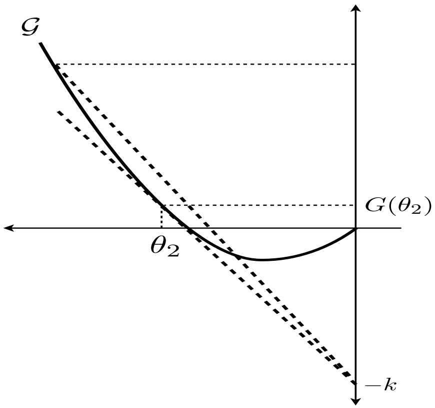
Consequence (extension of Ahmadi-Javid 2011, Theorem 5.1):

Theorem (1)

$$V(k) = \max_{\theta_2 < 0} \frac{k + G(\theta_2)}{\theta_2}. \tag{14}$$

Of course, the maximizers in (13) and (14) are the same.

Evaluation of $V(k)$



Generalized exponential family

Key concept for the moment problem, see CsM 2012. In our case of $\phi(\omega) = (1, X(\omega))$, it consists of the nonnegative functions

$$p_{\theta_1,\theta_2}(\omega) \triangleq (\beta^*)'(\omega, \theta_1 + \theta_2 X(\omega)), \quad (\theta_1, \theta_2) \in \Theta, \tag{15}$$

$$\Theta \triangleq \left\{(\theta_1, \theta_2) \in \mathrm{dom} K : \theta_1 + \theta_2 X(\omega) < \beta'(\omega, +\infty) \ \ \mu\text{-a.e.}\right\} \tag{16}$$

Instance of CsM 2012, Lemma 4.10: For $b \in (m, M)$, in the definition (5) of $F(b)$ the minimum is attained if and only if some p_{θ_1,θ_2} in (15) is a density with $\int X p_{\theta_1,\theta_2} d\mu = b$. Then

$$H(p_{\theta_1,\theta_2}) = F(b) = \theta_1 + \theta_2 b - K(\theta_1, \theta_2).$$

Via Main Lemma it follows: For $k \in (0, k_{\max})$, in the definition (6) of $V(k)$ the minimum is attained if and only if some p_{θ_1,θ_2} in (15) with $\theta_2 < 0$ is a density with $H(p_{\theta_1,\theta_2}) = k$. Then θ_2 is a minimizer in (14), and the minimizer in (6) (worst case density) is $p = p_{\theta_1,\theta_2}$.

Additional auxiliary results

If $K(\theta_1, \theta_2) < +\infty$ and $\bar{\theta}_1 < \theta_1$ then $(\bar{\theta}_1, \theta_2) \in \Theta$, hence $\mathrm{dom}\, K$ and Θ have the same projection to the θ_2 axis. Denote this projection by Θ_2 and its infimum by $\theta_{\min}$.

- $p_{\theta_1,\theta_2}(\omega) = 0$ if and only if $\theta_1 + \theta_2 X(\omega) \leq \beta'(\omega, 0)$. If $\beta'(\omega, 0) = -\infty$ μ-a.e then each p_{θ_1,θ_2} is positive μ-a.e.
- The generalized exponential family (15) contains the default density: $p_0 = p_{\theta_1,0}$ for some θ_1. In particular, $0 \in \Theta_2$.
- The standing assumption $k_{\max} > 0$ is equivalent to $\theta_{\min} < 0$, and implies that $F(b) > 0$ for each $b < b_0$.
- To each $\theta_2 \in \Theta_2$ there exists a unique $\theta_1 = \theta_1(\theta_2)$ with $G(\theta_2) = K(\theta_1, \theta_2) - \theta_1$. If $\int p_{\theta_1,\theta_2} d\mu = 1$ for some θ_1 then this θ_1 is unique and equals $\theta_1(\theta_2)$. Otherwise $\theta_1(\theta_2)$ is the largest θ_1 with $(\theta_1, \theta_2) \in \Theta$, and for this $\int p_{\theta_1,\theta_2} d\mu < 1$.

Bregman distance corresponding to any $\beta \in \mathbb{B}$:

$$B(p,q) = B_\beta(p,q) \triangleq \int \Delta_{\beta(\omega,.)}(p(\omega), q(\omega))\mu(d\omega), \qquad (17)$$

with $\Delta_\beta(\omega, .)$ defined as in (4), taking $f(s) = \beta(\omega, s)$ there.

Theorem (2: Worst case localiser)

For $k \in (0, k_{\max})$), let $\theta_2 < 0$ maximize $[k + G(\theta_2)]/\theta_2$, and let $\theta_1 = \theta_1(\theta_2)$. Then each density p with $H(p) < +\infty$ satisfies

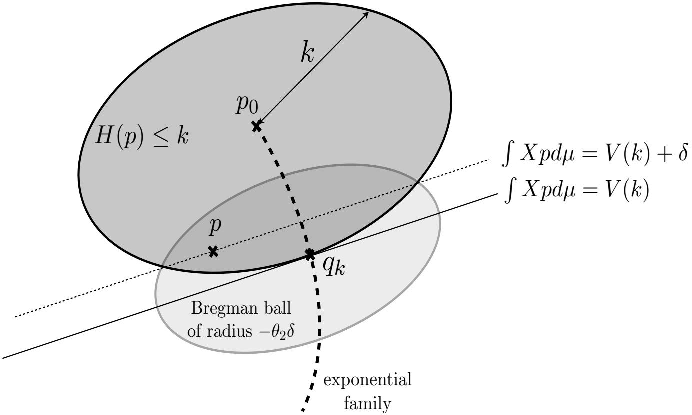
$$B(p, p_{\theta_1,\theta_2}) \leq H(p) - k - \theta_2 \left(\int X p d\mu \; - \; V(k) \right). \qquad (18)$$

Consequently, each sequence of densities p_n with

$$H(p_n) \to k, \quad \int X p_n d\mu \to V(k) \qquad (19)$$

converges to p_{θ_1,θ_2} locally in measure.

Illustration of Main result



Worst case localiser

- By Theorem 2, almost worst case densities, i.e., densities with $H(p)$ close to k and $\int Xpd\mu$ close to $V(k)$, cluster in Bregman distance around the function p_{θ_1,θ_2}.
- This function p_{θ_1,θ_2} is uniquely determined by $k \in (0, k_{\max})$. It will be called worst case localiser (WCL), denoted by q_k.
- The parameters θ_1, θ_2 of the WCL $q_k = p_{\theta_1,\theta_2}$ need not be unique, but they are if $q_k > 0$ μ-a.e.
- If a worst case density (attaining the minimum in (6)) exists, it equals the WCL. However, the clustering property also holds when no worst case density exists.
- A sufficient condition for $q_k = p_{\theta_1,\theta_2}$ to be the worst case density is $(\theta_1, \theta_2) \in \text{int dom } K$.
- q_k fails to be a density if and only if θ_2 in Theorem 2 is such that $\int p_{\theta_1,\theta_2} d\mu < 1$ for each θ_1 with $(\theta_1, \theta_2) \in \Theta$.

Proof of Theorem 2

Instance of CsM 2012, Lemma 4.15 (proven by simple algebra):

$$
\begin{aligned}
H(p) &= \theta_1 + \theta_2 \int X p d\mu - K(\theta_1, \theta_2) + B(p, p_{\theta_1,\theta_2}) \\
&+ \int |\beta'(\omega, 0) - \theta_1 - \theta_2 X(\omega)|_+ p(\omega) \mu(d\omega). \qquad (20)
\end{aligned}
$$

For θ_1, θ_2 in Theorem 2, $\theta_2 V(k) = k + K(\theta_1, \theta_2) - \theta_1$, hence

$$
\begin{aligned}
H(p) &= k - \theta_2 V(k) + \theta_2 \int X p d\mu + B(p, p_{\theta_1,\theta_2}) \\
&+ \int |\beta'(\omega, 0) - \theta_1 - \theta_2 X(\omega)|_+ p(\omega) \mu(d\omega). \qquad (21)
\end{aligned}
$$

This proves (18), whence (19) follows as $B(p_n, q) \to 0$ implies $p_n \to q$ locally in measure (CsM 2012, Corollary 2.14).

Pythagorean identities

For any $b \in (m.M)$, (20) gives for densities p with $\int Xpd\mu = b$ and θ_1, θ_2 attaining $F(b) = \theta_1 + \theta_2 b - K(\theta_1, \theta_2)$ that

$$H(p) = F(b) + B(p, p_{\theta_1,\theta_2}) + \text{correction term.} \tag{22}$$

The correction term is the last integral in (20). When in the definition $F(b) = \inf_{p:\int pd\mu=1, \int Xpd\mu=b} H(p)$ the minimum is attained (necessarily by p_{θ_1,θ_2}), and the correction term vanishes (specifically, when $p_{\theta_1,\theta_2} > 0$ μ-a.e.), this reduces to

$$H(p) = H(p_{\theta_1,\theta_2}) + B(p, p_{\theta_1,\theta_2}),$$

a familiar Pythagorean identity. The general version of identity (22) (for arbitrary noment mapping ϕ) appears in CsM 2012 as generalized Pythagorean identity.
The novel feature of the proof of Theorem 2 is that identity (20) is applied also to densities p with $\int Xpd\mu \neq b$.

Example

Autonomous integrand $\beta(\omega, s) = f(s) = -\log s$, let $\mu = \mathbb{P}_0$. Then $H(p)$ for $p = d\mathbb{P}/d\mathbb{P}_0$ is the reverse I-divergence $D(\mathbb{P}_0||\mathbb{P})$. Specifically, let $\Omega = (0,1)$, $X(\omega) = \omega$, let $\mu = \mathbb{P}_0$ have density 2ω with respect to the Lebesgue measure. Then

$$f^*(r) = -1 - \log(-r)\ (r<0),\ K(\theta_1, \theta_2) = \int_0^1 [-1 - \log(-\theta_1 - \theta_2\omega)] 2\omega d\omega$$

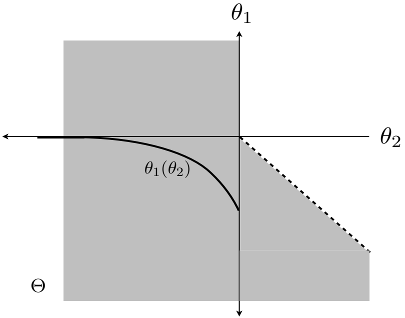
$$\Theta = \mathrm{dom}\, K = \{(\theta_1, \theta_2) : \theta_1 \leq 0,\ \theta_1 + \theta_2 < 0\}.$$

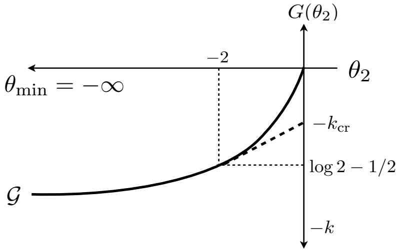
To $\theta_2 < 0$ there exists θ_1 such that $p_{\theta_1,\theta_2}(\omega) = 1/(-\theta_2 - \theta_2\omega)$ is a μ-density ($\int p_{\theta_1,\theta_2}(\omega) 2\omega d\omega = 1$) if and only if $\theta_2 \geq -2$. Otherwise

$$G(\theta_2) = K(0, \theta_2) = \int_0^1 [-1 - \log(-\theta_2\omega)] 2\omega d\omega = -\log(-\theta_2) - 1/2.$$

Calculus gives that $\theta_2 < -2$ is a maximizer of $[k + G(\theta_2)]/\theta_2$ if $k = \log(-\theta_2) - 1/2 > \log 2 - 1/2$. Then $V(k) = e^{-(k+1/2)}$, and $q_k(\omega) = (\theta_2\omega)^{-1}$, not a density. If $k \leq \log 2 - 1/2$ then $\theta_2 \geq -2$ and $q_k = p_{\theta_1(\theta_2),\theta_2}$ is the worst case density (no explicit formulas).

Illustration of Example





A pathological example

With Ω, μ, X in the previous example, now let $\mathbb{P}_0$ be the uniform distribution on $\Omega = (0,1)$, that has μ-density $p_0(\omega) = \frac{1}{2\omega}$.
Let $H(p)$ be the Bregman distance $B_{f,\mu}(p, p_0)$ with $f(s) = -\log s$, formally the functional $H_\beta(p)$ with

$$\beta(\omega, s) \triangleq \Delta_f(s, p_0(\omega)) = -\log s - \log(2\omega) + 2\omega(s - \frac{1}{2\omega}).$$

Then

$$\beta^*(\omega, r) = \log 2\omega - \log(-r + 2\omega),\ (\beta^*)'(\omega, r) = 1/(-r + 2\omega)\,(r < 2\omega)$$

and $\Theta_\beta = \mathrm{dom} K_\beta$ consists of those (θ_1, θ_2) for which $(\theta_1, \theta_2 - 2)$ belongs to the set Θ of the previous example. For $(\theta_1, \theta_2) \in \Theta_\beta$ the function $p_{\theta_1,\theta_2}(\omega) = 1/[-\theta_1 - (\theta_2 - 2)\omega)]$ coincides with $p_{\theta_1,\theta_2-2}(\omega)$ of the previous example, never a density if $\theta_2 < 0$.

Hence, in this example, no worst case density exists for any $k > 0$.

When WCL is a density

Theorem (3)

(i) For $k \in (0, k_{\max})$, if the WCL q_k is a density, it fails to be the worst case density only if $\theta_{\min} \in \Theta_2$, $G'(\theta_{\min}) > -\infty$ and

$$k > k_{\mathrm{cr}} \triangleq -G(\theta_{\min}) + \theta_{\min} G'(\theta_{\min}). \tag{23}$$

(ii) q_k is always a density if $K(\theta_1, 0) < +\infty$ for all $\theta_1 \in \mathbb{R}$.

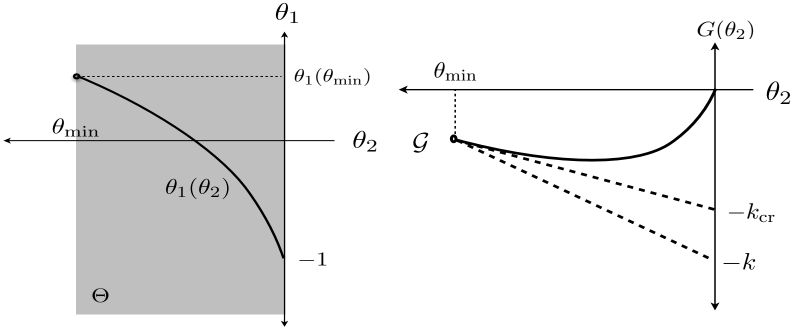
Proof.

$(\theta_1, \theta_2) \in \Theta$ with $q_k = p_{\theta_1,\theta_2}$ maximizes $\theta_1 + \theta_2 V(k) - K(\theta_1, \theta_2)$. Taking directional derivative gives for $(\tilde{\theta}_1, \tilde{\theta}_2) \in \mathrm{dom}\, K$

$$(\tilde{\theta}_1 - \theta_1)(1 - \int p_{\theta_1,\theta_2} d\mu) + (\tilde{\theta}_2 - \theta_2)(V(p) - \int X p_{\theta_1,\theta_2} d\mu) \leq 0.$$

If p_{θ_1,θ_2} is a density, it follows that $\int X p_{\theta_1,\theta_2} d\mu) \leq V(k)$, with equality if $\theta_1 > \theta_{\min}$. Under the hypothesis of (ii) one can take $\tilde{\theta}_2 = 0$ and $\tilde{\theta}_1$ arbitrarily large, this rules out $\int p_{\theta_1,\theta_2} d\mu < 1$. □

Illustration of Theorem



Critical value

The condition of Theorem 3 (ii) holds, e.g., if $H(p)$ is an f-divergence defined by a cofinite f (i.e., $f'(+\infty) = +\infty$). Then the worst case density either exists for each $k \in (0, k_{\max})$ or it exists if and only if k does not excceed a critical value.

For f-divergences with non-cofinite f a similar result can be proved (using that in that case the standing assumption $k_{\max} > 0$ implies $m > -\infty$, and one may assume $m = 0$).

It remains open whether the functional H_β has this property also for each $\beta \in \mathbb{B}$. Recalling the pathological example, this talk is concluded by the

Conjecture: The set of those $\theta_2 < 0$ for which there exists θ_1 with $(\theta_1, \theta_2) \in \Theta$, $\int p_{\theta_1,\theta_2} d\mu = 1$ (and hence $q_k = p_{\theta_1,\theta_2}$ is the worst case density for some $k \in (0, k_{\max})$) is, if nonempty, an interval with right endpoint 0.