On minimization of entropy functionals under moment constraints

I. Csiszár (Budapest) F. Matúš (Prague)

Lecture at
LMS Durham Symposium
Mathematical Aspects of Graphical Models
June 30 – July 10, 2008

$(X, \mathcal{X}, \mu)$ ... a σ-finite measure space with μ nonzero

$(X, \mathcal{X}, \mu)$ ... a σ-finite measure space with μ nonzero

$\varphi = (\varphi_0, \ldots, \varphi_d) \colon X \to \mathbb{R}^{1+d}$ ...moment mapping

$(X, \mathcal{X}, \mu)$ ... a σ-finite measure space with μ nonzero

$\varphi = (\varphi_0, \ldots, \varphi_d)\colon X \to \mathbb{R}^{1+d}$ ...moment mapping

(a $(1+d)$-tuple of real-valued measurable functions on X)

$(X, \mathcal{X}, \mu)$ ... a σ-finite measure space with μ nonzero

$\varphi = (\varphi_0, \ldots, \varphi_d) \colon X \to \mathbb{R}^{1+d}$ ...moment mapping

(a $(1+d)$-tuple of real-valued measurable functions on X)

assuming $\varphi_0 \equiv 1$

$(X, \mathcal{X}, \mu)$ ... a σ-finite measure space with μ nonzero

$\varphi = (\varphi_0, \ldots, \varphi_d) \colon X \to \mathbb{R}^{1+d}$ ...moment mapping

(a $(1+d)$-tuple of real-valued measurable functions on X)

assuming $\varphi_0 \equiv 1$

The moment constraints

For $a = (a_0, \ldots, a_d) \in \mathbb{R}^{1+d}$

$$\mathcal{L}_a = \{ g \geqslant 0 \ \text{ measurable } : \int_X \varphi \, g \, d\mu = a \}$$

γ ... a strictly convex and differentiable function on $(0, +\infty)$

γ ... a strictly convex and differentiable function on $(0, +\infty)$ extended to $\gamma(0) = \lim_{t \downarrow 0} \gamma(t)$ and $\gamma(t) = +\infty$ for $t < 0$

γ ... a strictly convex and differentiable function on $(0, +\infty)$
extended to $\gamma(0) = \lim_{t \downarrow 0} \gamma(t)$ and $\gamma(t) = +\infty$ for $t < 0$

The entropy functional based on γ

For a measurable function $g \geqslant 0$ on X

$$J(g) = \int_X \gamma(g) \, d\mu$$

if the integral exists, finite of not, and $J(g) = +\infty$ otherwise.

γ ... a strictly convex and differentiable function on $(0, +\infty)$ extended to $\gamma(0) = \lim_{t \downarrow 0} \gamma(t)$ and $\gamma(t) = +\infty$ for $t < 0$

The entropy functional based on γ

For a measurable function $g \geqslant 0$ on X

$$J(g) = \int_X \gamma(g)\, d\mu$$

if the integral exists, finite of not, and $J(g) = +\infty$ otherwise.

Shannon functional: $\gamma(t) = t \ln t$

γ ... a strictly convex and differentiable function on $(0, +\infty)$ extended to $\gamma(0) = \lim_{t \downarrow 0} \gamma(t)$ and $\gamma(t) = +\infty$ for $t < 0$

The entropy functional based on γ

For a measurable function $g \geqslant 0$ on X

$$J(g) = \int_X \gamma(g)\, d\mu$$

if the integral exists, finite of not, and $J(g) = +\infty$ otherwise.

Shannon functional: $\gamma(t) = t \ln t$

Burg functional: $\gamma(t) = 1 - \ln t$

PROBLEM

For given $a \in \mathbb{R}^{1+d}$, minimize $J(g)$ subject to the moment constraints $g \in \mathcal{L}_a$.

PROBLEM

For given $a \in \mathbb{R}^{1+d}$, minimize $J(g)$ subject to the moment constraints $g \in \mathcal{L}_a$.

The value function

$$H(a) = \inf_{g \in \mathcal{L}_a} J(g)\,, \qquad a \in \mathbb{R}^{1+d}\,,$$

PROBLEM

For given $a \in \mathbb{R}^{1+d}$, minimize $J(g)$ subject to the moment constraints $g \in \mathcal{L}_a$.

The value function

$$H(a) = \inf_{g \in \mathcal{L}_a} J(g)\,, \qquad a \in \mathbb{R}^{1+d}\,,$$

ranges in $[-\infty, +\infty]$ and is convex.

$X = \mathbb{R}$, μ Lebesgue measure and $\varphi = (1, x, x^2)$

$X = \mathbb{R}$, μ Lebesgue measure and $\varphi = (1, x, x^2)$

For $a = (a_0, a_1, a_2)$

$$\mathcal{L}_a = \Big\{ g \geqslant 0 \colon \int_{\mathbb{R}} g(x)\, dx = a_0, \quad \int_{\mathbb{R}} x\, g(x)\, dx = a_1, \quad \int_{\mathbb{R}} x^2\, g(x)\, dx = a_2 \Big\}$$

$X = \mathbb{R}$, μ Lebesgue measure and $\varphi = (1, x, x^2)$

For $a = (a_0, a_1, a_2)$

$$\mathcal{L}_a = \Big\{ g \geqslant 0 \colon \int_{\mathbb{R}} g(x)\, dx = a_0 \quad \int_{\mathbb{R}} x\, g(x)\, dx = a_1 \quad \int_{\mathbb{R}} x^2\, g(x)\, dx = a_2 \Big\}$$

$\gamma(t) = t \ln t, \qquad t > 0$

$X = \mathbb{R}$, μ Lebesgue measure and $\varphi = (1, x, x^2)$

For $a = (a_0, a_1, a_2)$

$$\mathcal{L}_a = \Big\{ g \geqslant 0 \colon \int_{\mathbb{R}} g(x)\, dx = a_0 \quad \int_{\mathbb{R}} x\, g(x)\, dx = a_1 \quad \int_{\mathbb{R}} x^2\, g(x)\, dx = a_2 \Big\}$$

$\gamma(t) = t \ln t, \qquad t > 0$

$J(g) = \int_{\mathbb{R}} g(x) \ln g(x)\, dx = -$ the differential Shannon entropy

$X = \mathbb{R}$, μ Lebesgue measure and $\varphi = (1, x, x^2)$

For $a = (a_0, a_1, a_2)$

$$\mathcal{L}_a = \{g \geqslant 0\colon \int_{\mathbb{R}} g(x)\, dx = a_0 \quad \int_{\mathbb{R}} x\, g(x)\, dx = a_1 \quad \int_{\mathbb{R}} x^2\, g(x)\, dx = a_2\}$$

$\gamma(t) = t \ln t, \qquad t > 0$

$J(g) = \int_{\mathbb{R}} g(x) \ln g(x)\, dx = -$ the differential Shannon entropy

The function H admits an explicit formula, e.g. for $a = (1, 0, 1)$

$$H(a) = \min_{\mathcal{L}_a} J = -\tfrac{1}{2}\ln(2\pi) - \tfrac{1}{2} = J\Big(\tfrac{1}{\sqrt{2\pi}} e^{-x^2/2} \Big)$$

$X = \mathbb{R}$, μ Lebesgue measure and $\varphi = (1, x, x^2)$

For $a = (a_0, a_1, a_2)$

$$\mathcal{L}_a = \Big\{ g \geqslant 0 \colon \int_{\mathbb{R}} g(x)\, dx = a_0 \quad \int_{\mathbb{R}} x\, g(x)\, dx = a_1 \quad \int_{\mathbb{R}} x^2\, g(x)\, dx = a_2 \Big\}$$

$\gamma(t) = t \ln t, \qquad t > 0$

$J(g) = \int_{\mathbb{R}} g(x) \ln g(x)\, dx = -$ the differential Shannon entropy

The function H admits an explicit formula, e.g. for $a = (1, 0, 1)$

$$H(a) = \min_{\mathcal{L}_a} J = -\tfrac{1}{2} \ln(2\pi) - \tfrac{1}{2} = J\Big(\tfrac{1}{\sqrt{2\pi}} e^{-x^2/2} \Big)$$

the minimizer is unique,
Gaussian with the given moments

$(X, \mathcal{X}, \mu)$ probability space, φ arbitrary

$(X, \mathcal{X}, \mu)$ probability space, φ arbitrary

Shannon functional

$(X, \mathcal{X}, \mu)$ probability space, φ arbitrary

Shannon functional

For $a = (1, a_1, \ldots, a_2)$ if $g \in \mathcal{L}_a$

$(X, \mathcal{X}, \mu)$ probability space, φ arbitrary

Shannon functional

For $a = (1, a_1, \ldots, a_2)$ if $g \in \mathcal{L}_a$

any $g \in \mathcal{L}_a$ is the μ-density of a probability measure P

$(X, \mathcal{X}, \mu)$ probability space, φ arbitrary

Shannon functional

For $a = (1, a_1, \ldots, a_2)$ if $g \in \mathcal{L}_a$

any $g \in \mathcal{L}_a$ is the μ-density of a probability measure P

and $\int_X g \ln g \, d\mu = D(P\|\mu)$.

$(X, \mathcal{X}, \mu)$ probability space, φ arbitrary

Shannon functional

For $a = (1, a_1, \ldots, a_2)$ if $g \in \mathcal{L}_a$

any $g \in \mathcal{L}_a$ is the μ-density of a probability measure P

and $\int_X g \ln g \, d\mu = D(P\|\mu)$.

Thus, $\inf_{\mathcal{L}_a} J$ is the minimization of the divergence $D(P\|\mu)$

subject to $\int_X \varphi \, dP = a$.

$(X, \mathcal{X}, \mu)$ probability space, φ arbitrary

Shannon functional

For $a = (1, a_1, \ldots, a_2)$ if $g \in \mathcal{L}_a$

any $g \in \mathcal{L}_a$ is the μ-density of a probability measure P

and $\int_X g \ln g \, d\mu = D(P\|\mu)$.

Thus, $\inf_{\mathcal{L}_a} J$ is the minimization of the divergence $D(P\|\mu)$

subject to $\int_X \varphi \, dP = a$.

(Cs&M (2003) *IEEE Trans. IT*)

$X = \mathbb{R}$, $\mu = \sum_{n \geqslant 1} \frac{1}{n^2} \delta_n$ and $\varphi(x) = (1, x)$

$X = \mathbb{R}$, $\mu = \sum_{n \geqslant 1} \frac{1}{n^2} \delta_n$ and $\varphi(x) = (1, x)$

$\gamma(t) = e^{2/t}, \ t > 0$

$X = \mathbb{R}$, $\mu = \sum_{n \geqslant 1} \frac{1}{n^2} \delta_n$ and $\varphi(x) = (1, x)$
$\gamma(t) = e^{2/t}$, $t > 0$

$$J(g) = \int_X \gamma(g)\, d\mu = \sum_{n \geqslant 1} e^{2/g(n)} \frac{1}{n^2}, \quad g \geqslant 0\,.$$

$X = \mathbb{R}$, $\mu = \sum_{n\geqslant 1} \frac{1}{n^2}\,\delta_n$ and $\varphi(x) = (1, x)$
$\gamma(t) = e^{2/t}$, $t > 0$

$$J(g) = \int_X \gamma(g)\,d\mu = \sum_{n\geqslant 1} e^{2/g(n)}\,\frac{1}{n^2}\,, \quad g \geqslant 0\,.$$

If finite then $g(n) \geqslant \frac{1}{\ln n}$ eventually.

$X=\mathbb{R}$, $\mu=\sum_{n\geqslant 1} \frac{1}{n^2}\,\delta_n$ and $\varphi(x)=(1,x)$
$\gamma(t)=e^{2/t}$, $t>0$

$$J(g)=\int_X \gamma(g)\,d\mu=\sum_{n\geqslant 1} e^{2/g(n)}\frac{1}{n^2}\,,\quad g\geqslant 0\,.$$

If finite then $g(n)\geqslant \frac{1}{\ln n}$ eventually.
Therefore, $\int_X \varphi_1\, g\, d\mu=\sum_{n\geqslant 1} n\, g(n)\,\frac{1}{n^2}$ diverges.

$X = \mathbb{R}$, $\mu = \sum_{n \geqslant 1} \frac{1}{n^2} \delta_n$ and $\varphi(x) = (1, x)$
$\gamma(t) = e^{2/t}$, $t > 0$

$$J(g) = \int_X \gamma(g)\, d\mu = \sum_{n \geqslant 1} e^{2/g(n)} \frac{1}{n^2}, \quad g \geqslant 0\,.$$

If finite then $g(n) \geqslant \frac{1}{\ln n}$ eventually.
Therefore, $\int_X \varphi_1\, g\, d\mu = \sum_{n \geqslant 1} n\, g(n) \frac{1}{n^2}$ diverges.
This implies that g is not in the union of the families $\mathcal{L}_a$,

$X = \mathbb{R}$, $\mu = \sum_{n \geqslant 1} \frac{1}{n^2}\, \delta_n$ and $\varphi(x) = (1, x)$
$\gamma(t) = e^{2/t}$, $t > 0$

$$J(g) = \int_X \gamma(g)\, d\mu = \sum_{n \geqslant 1} e^{2/g(n)} \frac{1}{n^2}\,, \quad g \geqslant 0\,.$$

If finite then $g(n) \geqslant \frac{1}{\ln n}$ eventually.
Therefore, $\int_X \varphi_1\, g\, d\mu = \sum_{n \geqslant 1} n\, g(n)\, \frac{1}{n^2}$ diverges.
This implies that g is not in the union of the families $\mathcal{L}_a$,
thus $H \equiv +\infty$.

$H(a) = \inf_{g \in \mathcal{L}_a} J(g), \ \ a \in \mathbb{R}^{1+d}$ ... the primal problem

$H(a) = \inf_{g \in \mathcal{L}_a} J(g), \;\; a \in \mathbb{R}^{1+d}$ ... the primal problem

$H^*(\vartheta) = \sup_{a \in \mathbb{R}^{1+d}} \big[\langle \vartheta, a \rangle - H(a)\big], \;\; \vartheta \in \mathbb{R}^{1+d}$ ... the conjugate of H

$H(a) = \inf_{g \in \mathcal{L}_a} J(g), \;\; a \in \mathbb{R}^{1+d}$ ... the primal problem

$H^*(\vartheta) = \sup_{a \in \mathbb{R}^{1+d}} \left[\langle \vartheta, a \rangle - H(a)\right], \;\; \vartheta \in \mathbb{R}^{1+d}$ ... the conjugate of H

$\gamma^*(r) = \sup_{t>0} \left[rt - \gamma(t)\right], \;\; r \in \mathbb{R}$ ... the conjugate of γ

$H(a) = \inf_{g \in \mathcal{L}_a} J(g), \;\; a \in \mathbb{R}^{1+d}$ ... the primal problem

$H^*(\vartheta) = \sup_{a \in \mathbb{R}^{1+d}} \big[\langle \vartheta, a \rangle - H(a)\big], \;\; \vartheta \in \mathbb{R}^{1+d}$ ... the conjugate of H

$\gamma^*(r) = \sup_{t>0} \big[rt - \gamma(t)\big], \;\; r \in \mathbb{R}$ ... the conjugate of γ

Proposition (H^* is expressible through γ^*)

If $H \not\equiv +\infty$ then

$$H^*_\gamma(\vartheta) = \int_X \gamma^*(\langle \vartheta, \varphi \rangle)\, d\mu\,, \qquad \vartheta \in \mathbb{R}^{1+d}\,.$$

$H(a) = \inf_{g \in \mathcal{L}_a} J(g)$, $a \in \mathbb{R}^{1+d}$ ... the primal problem

$H^*(\vartheta) = \sup_{a \in \mathbb{R}^{1+d}} \big[\langle \vartheta, a \rangle - H(a)\big]$, $\vartheta \in \mathbb{R}^{1+d}$ ... the conjugate of H

$\gamma^*(r) = \sup_{t>0} \big[rt - \gamma(t)\big]$, $r \in \mathbb{R}$ ... the conjugate of γ

Proposition (H^* is expressible through γ^*)

If $H \not\equiv +\infty$ then

$$H^*_\gamma(\vartheta) = \int_X \gamma^*(\langle \vartheta, \varphi \rangle)\, d\mu\,, \qquad \vartheta \in \mathbb{R}^{1+d}\,.$$

(a full proof using ideas of Rockafellar 68)

$H(a) = \inf_{g \in \mathcal{L}_a} J(g)$, $a \in \mathbb{R}^{1+d}$ … the primal problem

$H^*(\vartheta) = \sup_{a \in \mathbb{R}^{1+d}} \big[\langle \vartheta, a \rangle - H(a)\big]$, $\vartheta \in \mathbb{R}^{1+d}$ … the conjugate of H

$\gamma^*(r) = \sup_{t>0} \big[rt - \gamma(t)\big]$, $r \in \mathbb{R}$ … the conjugate of γ

Proposition (H^* is expressible through γ^*)

If $H \not\equiv +\infty$ then

$$H^*_\gamma(\vartheta) = \int_X \gamma^*(\langle \vartheta, \varphi \rangle)\, d\mu\,, \qquad \vartheta \in \mathbb{R}^{1+d}\,.$$

(a full proof using ideas of Rockafellar 68)

$H^{**}(a) = \sup_{\vartheta \in \mathbb{R}^{1+d}} \big[\langle \vartheta, a \rangle - H^*(\vartheta)\big]$, $a \in \mathbb{R}^{1+d}$ … the dual problem

$H(a) = \inf_{g \in \mathcal{L}_a} J(g), \quad a \in \mathbb{R}^{1+d}$ ... the primal problem

$H^*(\vartheta) = \sup_{a \in \mathbb{R}^{1+d}} \big[\langle \vartheta, a \rangle - H(a)\big], \quad \vartheta \in \mathbb{R}^{1+d}$ ... the conjugate of H

$\gamma^*(r) = \sup_{t>0} \big[rt - \gamma(t)\big], \quad r \in \mathbb{R}$ ... the conjugate of γ

Proposition (H^* is expressible through γ^*)

If $H \not\equiv +\infty$ then

$$H^*_\gamma(\vartheta) = \int_X \gamma^*(\langle \vartheta, \varphi \rangle)\, d\mu\,, \qquad \vartheta \in \mathbb{R}^{1+d}\,.$$

(a full proof using ideas of Rockafellar 68)

$H^{**}(a) = \sup_{\vartheta \in \mathbb{R}^{1+d}} \big[\langle \vartheta, a \rangle - H^*(\vartheta)\big], \quad a \in \mathbb{R}^{1+d}$ ... the dual problem

$H^{**} \leqslant H$, with the equality at the points of lower semicontinuity

Shannon functional: $\gamma(t) = t \ln t$

Shannon functional: $\gamma(t) = t \ln t$

$\gamma^*(r) = \sup_{t>0} [rt - t \ln t] = e^{r-1}$

$$H^{**}(a) = \sup_{\vartheta \in \mathbb{R}^{1+d}} \left[\langle \vartheta, a \rangle - \int_X \exp(\langle \vartheta, \varphi \rangle - 1)\, d\mu \right]$$

Shannon functional: $\gamma(t) = t \ln t$

$\gamma^*(r) = \sup_{t>0} [rt - t \ln t] = e^{r-1}$

$$H^{**}(a) = \sup_{\vartheta \in \mathbb{R}^{1+d}} \left[\langle \vartheta, a \rangle - \int_X \exp(\langle \vartheta, \varphi \rangle - 1) \, d\mu \right]$$

where the bracket rewrites to

$$\vartheta_0 a_0 + \Sigma_{j=1}^d \vartheta_j a_j - e^{\vartheta_0 - 1} \int_X \exp\left(\Sigma_{j=1}^d \vartheta_j \varphi_j \right) d\mu \, .$$

Shannon functional: $\gamma(t) = t \ln t$

$\gamma^*(r) = \sup_{t>0} [rt - t \ln t] = e^{r-1}$

$$H^{**}(a) = \sup_{\vartheta \in \mathbb{R}^{1+d}} \left[\langle \vartheta, a \rangle - \int_X \exp(\langle \vartheta, \varphi \rangle - 1) \, d\mu \right]$$

where the bracket rewrites to

$$\vartheta_0 a_0 + \Sigma_{j=1}^d \vartheta_j a_j - e^{\vartheta_0 - 1} \int_X \exp\left(\Sigma_{j=1}^d \vartheta_j \varphi_j\right) d\mu \, .$$

Maximizing over ϑ_0,

$$H^{**}(a) = a_0 \ln a_0 + \sup_{\vartheta_1, \ldots, \vartheta_d} \left[\Sigma_{j=1}^d \vartheta_j a_j - \int_X \exp\left(\Sigma_{j=1}^d \vartheta_j \varphi_j\right) d\mu \right] .$$

Shannon functional: $\gamma(t) = t \ln t$

$\gamma^*(r) = \sup_{t>0} [rt - t \ln t] = e^{r-1}$

$$H^{**}(a) = \sup_{\vartheta \in \mathbb{R}^{1+d}} \left[\langle \vartheta, a \rangle - \int_X \exp(\langle \vartheta, \varphi \rangle - 1)\, d\mu \right]$$

where the bracket rewrites to

$$\vartheta_0 a_0 + \Sigma_{j=1}^d \vartheta_j a_j - e^{\vartheta_0 - 1} \int_X \exp\left(\Sigma_{j=1}^d \vartheta_j \varphi_j \right) d\mu \,.$$

Maximizing over ϑ_0,

$$H^{**}(a) = a_0 \ln a_0 + \sup_{\vartheta_1, \dots, \vartheta_d} \left[\Sigma_{j=1}^d \vartheta_j a_j - \int_X \exp\left(\Sigma_{j=1}^d \vartheta_j \varphi_j \right) d\mu \right] .$$

... maximization of the normalized log-likelihood function in the exponential family based on μ and $(\varphi_1, \dots, \varphi_d)$.

Shannon functional: $\gamma(t) = t \ln t$

$\gamma^*(r) = \sup_{t>0} [rt - t \ln t] = e^{r-1}$

$$H^{**}(a) = \sup_{\vartheta \in \mathbb{R}^{1+d}} \left[\langle \vartheta, a \rangle - \int_X \exp(\langle \vartheta, \varphi \rangle - 1)\, d\mu \right]$$

where the bracket rewrites to

$$\vartheta_0 a_0 + \Sigma_{j=1}^d \vartheta_j a_j - e^{\vartheta_0 - 1} \int_X \exp\left(\Sigma_{j=1}^d \vartheta_j \varphi_j\right) d\mu\,.$$

Maximizing over ϑ_0,

$$H^{**}(a) = a_0 \ln a_0 + \sup_{\vartheta_1,\ldots,\vartheta_d} \left[\Sigma_{j=1}^d \vartheta_j a_j - \int_X \exp\left(\Sigma_{j=1}^d \vartheta_j \varphi_j\right) d\mu \right].$$

... maximization of the normalized log-likelihood function in the exponential family based on μ and $(\varphi_1, \ldots, \varphi_d)$.

(Cs&M (2008) *Probab. Th. Rel. F.*)

Theorem

Assume $a \in ri(dom(H_\gamma))$ and $H_\gamma(a) > -\infty$.

Theorem

Assume $a \in ri(dom(H_\gamma))$ and $H_\gamma(a) > -\infty$.

Theorem

Assume $a \in ri(dom(H_\gamma))$ *and* $H_\gamma(a) > -\infty$.
Then, $H_\gamma(a) = H_\gamma^{**}(a)$,

Theorem

Assume $a \in ri(dom(H_\gamma))$ and $H_\gamma(a) > -\infty$.

Then, $H_\gamma(a) = H_\gamma^{**}(a)$,

the dual value is attained by some $\vartheta \in \mathbb{R}^{1+d}$,

Theorem

Assume $a \in ri(dom(H_\gamma))$ and $H_\gamma(a) > -\infty$.
Then, $H_\gamma(a) = H_\gamma^{**}(a)$,
the dual value is attained by some $\vartheta \in \mathbb{R}^{1+d}$,
the function $g_a = \gamma^{*\prime}(\langle \vartheta, \varphi \rangle)$ does not depend on the choice of a maximizer ϑ,

Theorem

Assume $a \in ri(dom(H_\gamma))$ and $H_\gamma(a) > -\infty$.
Then, $H_\gamma(a) = H_\gamma^{**}(a)$,
the dual value is attained by some $\vartheta \in \mathbb{R}^{1+d}$,
the function $g_a = \gamma^{*\prime}(\langle \vartheta, \varphi \rangle)$ does not depend on the choice of a maximizer ϑ,
and for all $g \in \mathcal{L}_a$

$$J(g) = H(a) + B(g, g_a) + \int_X g \, |\gamma'(0) - \langle \vartheta, \varphi \rangle|_+ \, d\mu.$$

Theorem

Assume $a \in ri(dom(H_\gamma))$ and $H_\gamma(a) > -\infty$.
Then, $H_\gamma(a) = H_\gamma^{**}(a)$,
the dual value is attained by some $\vartheta \in \mathbb{R}^{1+d}$,
the function $g_a = \gamma^{*\prime}(\langle\vartheta, \varphi\rangle)$ does not depend on the choice of a maximizer ϑ,
and for all $g \in \mathcal{L}_a$

$$J(g) = H(a) + B(g, g_a) + \int_X g\,|\gamma'(0) - \langle\vartheta, \varphi\rangle|_+ \,d\mu.$$

(B … Bregman distance based on γ)

Theorem

Assume $a \in ri(dom(H_\gamma))$ and $H_\gamma(a) > -\infty$.
Then, $H_\gamma(a) = H_\gamma^{**}(a)$,
the dual value is attained by some $\vartheta \in \mathbb{R}^{1+d}$,
the function $g_a = \gamma^{*\prime}(\langle \vartheta, \varphi \rangle)$ does not depend on the choice of a maximizer ϑ,
and for all $g \in \mathcal{L}_a$

$$J(g) = H(a) + B(g, g_a) + \int_X g \, |\gamma'(0) - \langle \vartheta, \varphi \rangle|_+ \, d\mu.$$

(B ... Bregman distance based on γ)

The primal problem has a minimizer if and only if $g_a \in \mathcal{L}_a$.

Theorem

Assume $a \in ri(dom(H_\gamma))$ and $H_\gamma(a) > -\infty$.
Then, $H_\gamma(a) = H_\gamma^{**}(a)$,
the dual value is attained by some $\vartheta \in \mathbb{R}^{1+d}$,
the function $g_a = \gamma^{*\prime}(\langle \vartheta, \varphi \rangle)$ does not depend on the choice of a maximizer ϑ,
and for all $g \in \mathcal{L}_a$

$$J(g) = H(a) + B(g, g_a) + \int_X g \, |\gamma'(0) - \langle \vartheta, \varphi \rangle|_+ \, d\mu.$$

(B ... Bregman distance based on γ)

The primal problem has a minimizer if and only if $g_a \in \mathcal{L}_a$.
$\int_X$ vanishes when $\gamma'(0) = -\infty$ (γ is ess. smooth, or steep).

Theorem

Assume $a \in ri(dom(H_\gamma))$ and $H_\gamma(a) > -\infty$.
Then, $H_\gamma(a) = H_\gamma^{**}(a)$,
the dual value is attained by some $\vartheta \in \mathbb{R}^{1+d}$,
the function $g_a = \gamma^{*\prime}(\langle\vartheta, \varphi\rangle)$ does not depend on the choice of a maximizer ϑ,
and for all $g \in \mathcal{L}_a$

$$J(g) = H(a) + B(g, g_a) + \int_X g\, |\gamma'(0) - \langle\vartheta, \varphi\rangle|_+ \, d\mu.$$

(B ... Bregman distance based on γ)

The primal problem has a minimizer if and only if $g_a \in \mathcal{L}_a$.
$\int_X$ vanishes when $\gamma'(0) = -\infty$ (γ is ess. smooth, or steep).
If $g_n \in \mathcal{L}_a$ and $J(g_n) \to H(a)$ then $B(g_n, g_a) \to 0$.

g_a .... generalized primal solution

$H(a) = \inf_{g \in \mathcal{L}_a} \int_{\mathbb{R}} g(x) \ln g(x)\, dx$, ... the primal problem

where $\mathcal{L}_a$, $a \in \mathbb{R}^3$, comes from the moments $1, x, x^2$

$H(a) = \inf_{g \in \mathcal{L}_a} \int_{\mathbb{R}} g(x) \ln g(x)\, dx$, ... the primal problem

where $\mathcal{L}_a$, $a \in \mathbb{R}^3$, comes from the moments $1, x, x^2$

the dual problem

$$H^{**}(a) = \sup_{\vartheta \in \mathbb{R}^3} \left[\vartheta_0 a_0 + \vartheta_1 a_1 + \vartheta_2 a_2 - \int_X \exp(\vartheta_0 + \vartheta_1 x + \vartheta_2 x^2 - 1) d\mu \right]$$

explicitly computable, finite on an open set in $\mathbb{R}^3$

$H(a) = \inf_{g \in \mathcal{L}_a} \int_{\mathbb{R}} g(x) \ln g(x)\, dx$, ... the primal problem

where $\mathcal{L}_a$, $a \in \mathbb{R}^3$, comes from the moments $1, x, x^2$

the dual problem

$$H^{**}(a) = \sup_{\vartheta \in \mathbb{R}^3} \left[\vartheta_0 a_0 + \vartheta_1 a_1 + \vartheta_2 a_2 - \int_X \exp(\vartheta_0 + \vartheta_1 x + \vartheta_2 x^2 - 1)\, d\mu \right]$$

explicitly computable, finite on an open set in $\mathbb{R}^3$

$H = H^{**}$ with the same open effective domain

$H(a) = \inf_{g \in \mathcal{L}_a} \int_{\mathbb{R}} g(x) \ln g(x)\, dx$, ... the primal problem

where $\mathcal{L}_a$, $a \in \mathbb{R}^3$, comes from the moments $1, x, x^2$

the dual problem

$$H^{**}(a) = \sup_{\vartheta \in \mathbb{R}^3} \left[\vartheta_0 a_0 + \vartheta_1 a_1 + \vartheta_2 a_2 - \int_X \exp(\vartheta_0 + \vartheta_1 x + \vartheta_2 x^2 - 1) d\mu \right]$$

explicitly computable, finite on an open set in $\mathbb{R}^3$

$H = H^{**}$ with the same open effective domain

For a in the open domain

$H(a) = \inf_{g \in \mathcal{L}_a} \int_{\mathbb{R}} g(x) \ln g(x)\, dx$, ... the primal problem

where $\mathcal{L}_a$, $a \in \mathbb{R}^3$, comes from the moments $1, x, x^2$

the dual problem

$$H^{**}(a) = \sup_{\vartheta \in \mathbb{R}^3} \left[\vartheta_0 a_0 + \vartheta_1 a_1 + \vartheta_2 a_2 - \int_X \exp(\vartheta_0 + \vartheta_1 x + \vartheta_2 x^2 - 1) d\mu \right]$$

explicitly computable, finite on an open set in $\mathbb{R}^3$

$H = H^{**}$ with the same open effective domain

For a in the open domain

g_a has the form $\exp(\vartheta_0 + \vartheta_1 x + \vartheta_2 x^2 - 1)$, thus is proportional to a Gaussian density

$H(a) = \inf_{g \in \mathcal{L}_a} \int_{\mathbb{R}} g(x) \ln g(x)\, dx$, ... the primal problem

where $\mathcal{L}_a$, $a \in \mathbb{R}^3$, comes from the moments $1, x, x^2$

the dual problem

$$H^{**}(a) = \sup_{\vartheta \in \mathbb{R}^3} \left[\vartheta_0 a_0 + \vartheta_1 a_1 + \vartheta_2 a_2 - \int_X \exp(\vartheta_0 + \vartheta_1 x + \vartheta_2 x^2 - 1) d\mu \right]$$

explicitly computable, finite on an open set in $\mathbb{R}^3$

$H = H^{**}$ with the same open effective domain

For a in the open domain

g_a has the form $\exp(\vartheta_0 + \vartheta_1 x + \vartheta_2 x^2 - 1)$, thus is proportional to a Gaussian density

adjusting the moments, g_a is the unique primal solution

$X = [0,1]$, $d\mu = 2x\,dx$ and $\varphi = (1, x)$

$X = [0,1]$, $d\mu = 2x\,dx$ and $\varphi = (1, x)$

Burg functional: $\gamma(t) = 1 - \ln t, \quad t > 0$

$X = [0,1]$, $d\mu = 2x\,dx$ and $\varphi = (1, x)$

Burg functional: $\gamma(t) = 1 - \ln t, \quad t > 0$

$\gamma^*(r) = -\ln(-r), \quad r < 0$

$X = [0, 1]$, $d\mu = 2x\, dx$ and $\varphi = (1, x)$

Burg functional: $\gamma(t) = 1 - \ln t, \quad t > 0$

$\gamma^*(r) = -\ln(-r), \quad r < 0$

$\gamma^{*\prime}(r) = -1/r, \quad r < 0$

$X = [0, 1]$, $d\mu = 2x\,dx$ and $\varphi = (1, x)$

Burg functional: $\gamma(t) = 1 - \ln t, \quad t > 0$

$\gamma^*(r) = -\ln(-r), \quad r < 0$

$\gamma^{*\prime}(r) = -1/r, \quad r < 0$

the primal problem $H(a) = \inf_{g \in \mathcal{L}_a} \int_0^1 \, [1 - \ln g(x)]\, 2x\, dx$

$X = [0,1]$, $d\mu = 2x\,dx$ and $\varphi = (1, x)$

Burg functional: $\gamma(t) = 1 - \ln t, \quad t > 0$

$\gamma^*(r) = -\ln(-r), \quad r < 0$

$\gamma^{*\prime}(r) = -1/r, \quad r < 0$

the primal problem $H(a) = \inf_{g \in \mathcal{L}_a} \int_0^1 [1 - \ln g(x)]\, 2x\, dx$

the dual problem

$$H^{**}(a) = \sup_{\vartheta \in \mathbb{R}^2} \left[\vartheta_0 a_0 + \vartheta_1 a_1 + \int_0^1 \ln(-\vartheta_0 - \vartheta_1 x)\, 2x\, dx \right]$$

$X = [0,1]$, $d\mu = 2x\,dx$ and $\varphi = (1, x)$

Burg functional: $\gamma(t) = 1 - \ln t, \quad t > 0$

$\gamma^*(r) = -\ln(-r), \quad r < 0$

$\gamma^{*\prime}(r) = -1/r, \quad r < 0$

the primal problem $H(a) = \inf_{g \in \mathcal{L}_a} \int_0^1 [1 - \ln g(x)]\, 2x\, dx$

the dual problem

$$H^{**}(a) = \sup_{\vartheta \in \mathbb{R}^2} \left[\vartheta_0 a_0 + \vartheta_1 a_1 + \int_0^1 \ln(-\vartheta_0 - \vartheta_1 x)\, 2x\, dx \right]$$

$H = H^{**}$, the domain open $a_0 > 0$, $a_1 < a_0$

$X = [0,1]$, $d\mu = 2x\,dx$ and $\varphi = (1, x)$

Burg functional: $\gamma(t) = 1 - \ln t, \quad t > 0$

$\gamma^*(r) = -\ln(-r), \quad r < 0$

$\gamma^{*\prime}(r) = -1/r, \quad r < 0$

the primal problem $H(a) = \inf_{g \in \mathcal{L}_a} \int_0^1 \, [1 - \ln g(x)]\, 2x\, dx$

the dual problem

$$H^{**}(a) = \sup_{\vartheta \in \mathbb{R}^2} \left[\vartheta_0 a_0 + \vartheta_1 a_1 + \int_0^1 \, \ln(-\vartheta_0 - \vartheta_1 x)\, 2x\, dx \right]$$

$H = H^{**}$, the domain open $a_0 > 0$, $a_1 < a_0$

unique maximizer $\vartheta = (0, -1/a_1)$ when $a_0 \geqslant 2a_1$,

$X = [0,1]$, $d\mu = 2x\,dx$ and $\varphi = (1, x)$

Burg functional: $\gamma(t) = 1 - \ln t, \quad t > 0$

$\gamma^*(r) = -\ln(-r), \quad r < 0$

$\gamma^{*\prime}(r) = -1/r, \quad r < 0$

the primal problem $H(a) = \inf_{g \in \mathcal{L}_a} \int_0^1 [1 - \ln g(x)]\, 2x\, dx$

the dual problem

$$H^{**}(a) = \sup_{\vartheta \in \mathbb{R}^2} \left[\vartheta_0 a_0 + \vartheta_1 a_1 + \int_0^1 \ln(-\vartheta_0 - \vartheta_1 x)\, 2x\, dx\right]$$

$H = H^{**}$, the domain open $a_0 > 0$, $a_1 < a_0$

unique maximizer $\vartheta = (0, -1/a_1)$ when $a_0 \geqslant 2a_1$,

generalized primal solution $\quad g_a(x) = a_1/x$

$X = [0,1]$, $d\mu = 2x\,dx$ and $\varphi = (1, x)$

Burg functional: $\gamma(t) = 1 - \ln t, \quad t > 0$

$\gamma^*(r) = -\ln(-r), \quad r < 0$

$\gamma^{*\prime}(r) = -1/r, \quad r < 0$

the primal problem $H(a) = \inf_{g \in \mathcal{L}_a} \int_0^1 [1 - \ln g(x)]\, 2x\, dx$

the dual problem

$$H^{**}(a) = \sup_{\vartheta \in \mathbb{R}^2} \left[\vartheta_0 a_0 + \vartheta_1 a_1 + \int_0^1 \ln(-\vartheta_0 - \vartheta_1 x)\, 2x\, dx \right]$$

$H = H^{**}$, the domain open $a_0 > 0$, $a_1 < a_0$

unique maximizer $\vartheta = (0, -1/a_1)$ when $a_0 \geqslant 2a_1$,

generalized primal solution $\quad g_a(x) = a_1/x$

if $a_0 = 1$, $a_1 < 1/2$ it is not a density: $\int_0^1 \frac{a_1}{x}\, 2x\, dx = 2a_1 < a_0$

$X = [0,1]$, $d\mu = 2x\,dx$ and $\varphi = (1, x)$

Burg functional: $\gamma(t) = 1 - \ln t, \quad t > 0$

$\gamma^*(r) = -\ln(-r), \quad r < 0$

$\gamma^{*\prime}(r) = -1/r, \quad r < 0$

the primal problem $H(a) = \inf_{g\in\mathcal{L}_a} \int_0^1 [1 - \ln g(x)]\, 2x\,dx$

the dual problem

$$H^{**}(a) = \sup_{\vartheta\in\mathbb{R}^2} \left[\vartheta_0 a_0 + \vartheta_1 a_1 + \int_0^1 \ln(-\vartheta_0 - \vartheta_1 x)\, 2x\,dx\right]$$

$H = H^{**}$, the domain open $a_0 > 0$, $a_1 < a_0$

unique maximizer $\vartheta = (0, -1/a_1)$ when $a_0 \geqslant 2a_1$,

generalized primal solution $\quad g_a(x) = a_1/x$

if $a_0 = 1$, $a_1 < 1/2$ it is not a density: $\int_0^1 \frac{a_1}{x}\, 2x\,dx = 2a_1 < a_0$

NO primal solution! (a variation on Borwein & Lewis (1993))

$cc_{\varphi}(\mu) \subseteq \mathbb{R}^{1+d}$ ... convex core of the φ-image of μ, intersection of all convex Borel sets $B \subseteq \mathbb{R}^{1+d}$ s.t. $\mu(\varphi^{-1}(\mathbb{R}^{1+d} \setminus B)) = 0$.

$cc_\varphi(\mu) \subseteq \mathbb{R}^{1+d}$ ... convex core of the φ-image of μ, intersection of all convex Borel sets $B \subseteq \mathbb{R}^{1+d}$ s.t. $\mu(\varphi^{-1}(\mathbb{R}^{1+d} \setminus B)) = 0$.

$cn_\varphi(\mu) = \{ta \colon t \geqslant 0\,,\ a \in cc_\varphi(\mu)\}$ ... φ-cone of μ

$cc_\varphi(\mu) \subseteq \mathbb{R}^{1+d}$ ... convex core of the φ-image of μ, intersection of all convex Borel sets $B \subseteq \mathbb{R}^{1+d}$ s.t. $\mu(\varphi^{-1}(\mathbb{R}^{1+d} \setminus B)) = 0$.

$cn_\varphi(\mu) = \{ta\colon\, t \geqslant 0\,,\; a \in cc_\varphi(\mu)\}$ ... φ-cone of μ

Lemma: The set $\mathcal{L}_a$ is nonempty if and only if $a \in cn_\varphi(\mu)$.

$cc_\varphi(\mu) \subseteq \mathbb{R}^{1+d}$ ... convex core of the φ-image of μ, intersection of all convex Borel sets $B \subseteq \mathbb{R}^{1+d}$ s.t. $\mu(\varphi^{-1}(\mathbb{R}^{1+d} \setminus B)) = 0$.

$cn_\varphi(\mu) = \{ta\colon\, t \geqslant 0\,,\, a \in cc_\varphi(\mu)\}$ ... φ-cone of μ

Lemma: The set $\mathcal{L}_a$ is nonempty if and only if $a \in cn_\varphi(\mu)$.

Corollary: $H = +\infty$ outside $cn_\varphi(\mu)$.

$cc_\varphi(\mu) \subseteq \mathbb{R}^{1+d}$ ... convex core of the φ-image of μ, intersection of all convex Borel sets $B \subseteq \mathbb{R}^{1+d}$ s.t. $\mu(\varphi^{-1}(\mathbb{R}^{1+d} \setminus B)) = 0$.

$cn_\varphi(\mu) = \{ta\colon t \geqslant 0\,,\ a \in cc_\varphi(\mu)\}$ ... φ-cone of μ

Lemma: The set $\mathcal{L}_a$ is nonempty if and only if $a \in cn_\varphi(\mu)$.

Corollary: $H = +\infty$ outside $cn_\varphi(\mu)$.

Theorem

$cc_\varphi(\mu) \subseteq \mathbb{R}^{1+d}$ ... convex core of the φ-image of μ, intersection of all convex Borel sets $B \subseteq \mathbb{R}^{1+d}$ s.t. $\mu(\varphi^{-1}(\mathbb{R}^{1+d} \setminus B)) = 0$.

$cn_\varphi(\mu) = \{ta\colon t \geqslant 0\,,\ a \in cc_\varphi(\mu)\}$ ... φ-cone of μ

Lemma: The set $\mathcal{L}_a$ is nonempty if and only if $a \in cn_\varphi(\mu)$.

Corollary: $H = +\infty$ outside $cn_\varphi(\mu)$.

Theorem

If μ is finite and $\gamma(0) = +\infty$ then dom(H) equals ri$(cn_\varphi(\mu))$ or $\emptyset$.

$cc_\varphi(\mu) \subseteq \mathbb{R}^{1+d}$ ... convex core of the φ-image of μ, intersection of all convex Borel sets $B \subseteq \mathbb{R}^{1+d}$ s.t. $\mu(\varphi^{-1}(\mathbb{R}^{1+d} \setminus B)) = 0$.

$cn_\varphi(\mu) = \{ta \colon t \geqslant 0\,,\ a \in cc_\varphi(\mu)\}$ ... φ-cone of μ

Lemma: The set $\mathcal{L}_a$ is nonempty if and only if $a \in cn_\varphi(\mu)$.

Corollary: $H = +\infty$ outside $cn_\varphi(\mu)$.

Theorem

If μ is finite and $\gamma(0) = +\infty$ then dom(H) equals ri$(cn_\varphi(\mu))$ or $\emptyset$.

If μ is finite and $\gamma(0)$ finite then dom$(H) = cn_\varphi(\mu)$.

$cc_\varphi(\mu) \subseteq \mathbb{R}^{1+d}$ ... convex core of the φ-image of μ, intersection of all convex Borel sets $B \subseteq \mathbb{R}^{1+d}$ s.t. $\mu(\varphi^{-1}(\mathbb{R}^{1+d} \setminus B)) = 0$.

$cn_\varphi(\mu) = \{ta \colon t \geqslant 0\,,\ a \in cc_\varphi(\mu)\}$ ... φ-cone of μ

Lemma: The set $\mathcal{L}_a$ is nonempty if and only if $a \in cn_\varphi(\mu)$.

Corollary: $H = +\infty$ outside $cn_\varphi(\mu)$.

Theorem

If μ is finite and $\gamma(0) = +\infty$ then dom(H) equals ri$(cn_\varphi(\mu))$ or $\emptyset$.

If μ is finite and $\gamma(0)$ finite then dom$(H) = cn_\varphi(\mu)$.

If μ is infinite and $\gamma(0) = 0$ then dom$(H) = cn_\varphi(\mu)$.

$cc_\varphi(\mu) \subseteq \mathbb{R}^{1+d}$ ... convex core of the φ-image of μ, intersection of all convex Borel sets $B \subseteq \mathbb{R}^{1+d}$ s.t. $\mu(\varphi^{-1}(\mathbb{R}^{1+d} \setminus B)) = 0$.

$cn_\varphi(\mu) = \{ta\colon t \geqslant 0\,,\ a \in cc_\varphi(\mu)\}$ ... φ-cone of μ

Lemma: The set $\mathcal{L}_a$ is nonempty if and only if $a \in cn_\varphi(\mu)$.

Corollary: $H = +\infty$ outside $cn_\varphi(\mu)$.

Theorem

If μ is finite and $\gamma(0) = +\infty$ then $dom(H)$ equals $ri(cn_\varphi(\mu))$ or $\emptyset$.

If μ is finite and $\gamma(0)$ finite then $dom(H) = cn_\varphi(\mu)$.

If μ is infinite and $\gamma(0) = 0$ then $dom(H) = cn_\varphi(\mu)$.

If μ is infinite and $\gamma(0) > 0$ then $dom(H) = \emptyset$.

$cc_\varphi(\mu) \subseteq \mathbb{R}^{1+d}$ ... convex core of the φ-image of μ, intersection of all convex Borel sets $B \subseteq \mathbb{R}^{1+d}$ s.t. $\mu(\varphi^{-1}(\mathbb{R}^{1+d} \setminus B)) = 0$.

$cn_\varphi(\mu) = \{ta\colon\, t \geqslant 0\,,\; a \in cc_\varphi(\mu)\}$ ... φ-cone of μ

Lemma: The set $\mathcal{L}_a$ is nonempty if and only if $a \in cn_\varphi(\mu)$.

Corollary: $H = +\infty$ outside $cn_\varphi(\mu)$.

Theorem

If μ is finite and $\gamma(0) = +\infty$ then $dom(H)$ equals $ri(cn_\varphi(\mu))$ or $\emptyset$.

If μ is finite and $\gamma(0)$ finite then $dom(H) = cn_\varphi(\mu)$.

If μ is infinite and $\gamma(0) = 0$ then $dom(H) = cn_\varphi(\mu)$.

If μ is infinite and $\gamma(0) > 0$ then $dom(H) = \emptyset$.

If μ is infinite and $\gamma(0) < 0$ then $H = -\infty$ on $dom(H) = cn_\varphi(\mu)$.

Theorem

Assume $a \in \mathbb{R}^{1+d}$ with $H_\gamma(a)$ finite.

Theorem

Assume $a \in \mathbb{R}^{1+d}$ with $H_\gamma(a)$ finite.

Theorem

Assume $a \in \mathbb{R}^{1+d}$ with $H_\gamma(a)$ finite.
Denote by F the face of the convex cone $cn_\varphi(\mu)$ with $a \in ri(F)$

Theorem

Assume $a \in \mathbb{R}^{1+d}$ with $H_\gamma(a)$ finite.
Denote by F the face of the convex cone $cn_\varphi(\mu)$ with $a \in ri(F)$
Then, the adjusted dual problem

$$\tilde{H}(a) = \sup_{\vartheta \in \mathbb{R}^{1+d}} \left[\langle \vartheta, a \rangle - \int_{\varphi^{-1}(cl(F))} \gamma^*(\langle \vartheta, \varphi \rangle) \, d\mu \right].$$

has a maximizer $\vartheta \in \mathbb{R}^{1+d}$,

Theorem

Assume $a \in \mathbb{R}^{1+d}$ with $H_\gamma(a)$ finite.
Denote by F the face of the convex cone $cn_\varphi(\mu)$ with $a \in ri(F)$
Then, the adjusted dual problem

$$\tilde{H}(a) = \sup_{\vartheta \in \mathbb{R}^{1+d}} \left[\langle \vartheta, a \rangle - \int_{\varphi^{-1}(cl(F))} \gamma^*(\langle \vartheta, \varphi \rangle) \, d\mu \right].$$

has a maximizer $\vartheta \in \mathbb{R}^{1+d}$,
$H_\gamma(a) = \tilde{H}(a) + \gamma(0) \cdot \mu(X \setminus \varphi^{-1}(cl(F))),$

Theorem

Assume $a \in \mathbb{R}^{1+d}$ with $H_\gamma(a)$ finite.
Denote by F the face of the convex cone $cn_\varphi(\mu)$ with $a \in ri(F)$
Then, the adjusted dual problem

$$\tilde{H}(a) = \sup_{\vartheta \in \mathbb{R}^{1+d}} \left[\langle \vartheta, a \rangle - \int_{\varphi^{-1}(cl(F))} \gamma^*(\langle \vartheta, \varphi \rangle)\, d\mu \right].$$

has a maximizer $\vartheta \in \mathbb{R}^{1+d}$,
$H_\gamma(a) = \tilde{H}(a) + \gamma(0) \cdot \mu(X \setminus \varphi^{-1}(cl(F)))$,
the function $g_a = \gamma^{*\prime}(\langle \vartheta, \varphi \rangle) 1_{\varphi^{-1}(F)}$ does not depend on its choice

Theorem

Assume $a \in \mathbb{R}^{1+d}$ with $H_\gamma(a)$ finite.
Denote by F the face of the convex cone $cn_\varphi(\mu)$ with $a \in ri(F)$
Then, the adjusted dual problem

$$\tilde{H}(a) = \sup_{\vartheta \in \mathbb{R}^{1+d}} \left[\langle \vartheta, a \rangle - \int_{\varphi^{-1}(cl(F))} \gamma^*(\langle \vartheta, \varphi \rangle)\, d\mu \right].$$

has a maximizer $\vartheta \in \mathbb{R}^{1+d}$,
$H_\gamma(a) = \tilde{H}(a) + \gamma(0) \cdot \mu(X \setminus \varphi^{-1}(cl(F)))$,
the function $g_a = \gamma^{*\prime}(\langle \vartheta, \varphi \rangle) 1_{\varphi^{-1}(F)}$ does not depend on its choice and for all $g \in \mathcal{L}_a$

$$J(g) = H(a) + B(g, g_a) + \int_X g\, |\gamma'(0) - \langle \vartheta, \varphi \rangle|_+ \, d\mu.$$

Theorem

Assume $H > -\infty$ and $a \in dom(H^{**})$.

Theorem

Assume $H > -\infty$ *and* $a \in dom(H^{**})$.

Theorem

Assume $H > -\infty$ and $a \in dom(H^{**})$.
Then, there exists a unique nonnegative function h_a such that

$$H^{**}(a) - \left[\langle \vartheta, a \rangle - \int_X \gamma^*(\langle \vartheta, \varphi \rangle)\, d\mu\right] \geqslant B(h_a, \gamma^{*\prime}(\langle \vartheta, \varphi \rangle)) + \int_X h_a \, |\gamma'(0) - \langle \vartheta, \varphi \rangle|_+ \, d\mu$$

for $\vartheta \in dom(H^*_\gamma)$ satisfying $\langle \vartheta, \varphi \rangle < \gamma'(+\infty)$, μ-a.e.

Theorem

Assume $H > -\infty$ and $a \in dom(H^{**})$.

Then, there exists a unique nonnegative function h_a such that

$$H^{**}(a) - \left[\langle \vartheta, a \rangle - \int_X \gamma^*(\langle \vartheta, \varphi \rangle)\, d\mu \right] \geqslant B(h_a, \gamma^{*\prime}(\langle \vartheta, \varphi \rangle)) + \int_X h_a \, |\gamma'(0) - \langle \vartheta, \varphi \rangle|_+ \, d\mu$$

for $\vartheta \in dom(H^*_\gamma)$ satisfying $\langle \vartheta, \varphi \rangle < \gamma'(+\infty)$, μ-a.e.

If $H_\gamma(a) = H^{**}(a)$ then $h_a = g_a$.

Theorem

Assume $H > -\infty$ and $a \in dom(H^{**})$.

Then, there exists a unique nonnegative function h_a such that

$$H^{**}(a) - \left[\langle \vartheta, a \rangle - \int_X \gamma^*(\langle \vartheta, \varphi \rangle) \, d\mu \right] \geqslant B(h_a, \gamma^{*\prime}(\langle \vartheta, \varphi \rangle)) + \int_X h_a \, |\gamma'(0) - \langle \vartheta, \varphi \rangle|_+ \, d\mu$$

for $\vartheta \in dom(H^*_\gamma)$ satisfying $\langle \vartheta, \varphi \rangle < \gamma'(+\infty)$, μ-a.e.

If $H_\gamma(a) = H^{**}(a)$ then $h_a = g_a$.

For $\gamma(t) = t \ln t$, this is MLE in EF; an explicit construction of h_a is available in Cs&M (2008) *Probab. Th. Rel. F.*

Theorem

Assume $H > -\infty$ and $a \in dom(H^{**})$.
Then, there exists a unique nonnegative function h_a such that

$$H^{**}(a) - \left[\langle \vartheta, a \rangle - \int_X \gamma^*(\langle \vartheta, \varphi \rangle)\, d\mu\right] \geqslant B(h_a, {\gamma^*}'(\langle \vartheta, \varphi \rangle)) + \int_X h_a \, |\gamma'(0) - \langle \vartheta, \varphi \rangle|_+ \, d\mu$$

for $\vartheta \in dom(H_\gamma^*)$ satisfying $\langle \vartheta, \varphi \rangle < \gamma'(+\infty)$, μ-a.e.
If $H_\gamma(a) = H^{**}(a)$ then $h_a = g_a$.

For $\gamma(t) = t \ln t$, this is MLE in EF; an explicit construction of h_a is available in Cs&M (2008) *Probab. Th. Rel. F.*

The talk is based on a contribution to *Proc. IEEE ISIT*, Toronto, being published this week.