Asymmetric Topologies on Statistical Manifolds

Roman V. Belavkin

School of Science and Technology
Middlesex University, London NW4 4BT, UK

GSI2015, October 28, 2015

Sources and Consequences of Asymmetry

Method: Symmetric Sandwich

Results

Sources and Consequences of Asymmetry

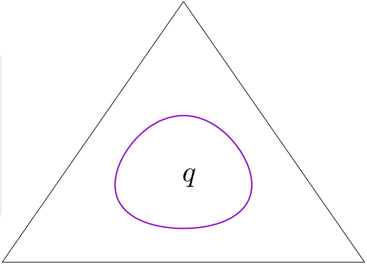
Method: Symmetric Sandwich

Results

Asymmetric Information Distances

Kullback-Leibler divergence

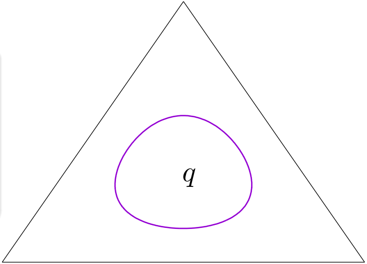
- $D[p, q] = \mathbb{E}_q\{\ln(p/q)\}$



Asymmetric Information Distances

Kullback-Leibler divergence

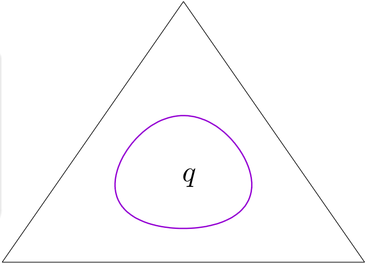
- $D[p, q] = \mathbb{E}_q\{\ln(p/q)\}$
- $D[p_1 \otimes p_2, q_1 \otimes q_2] = D[p_1, q_1] + D[p_2, q_2]$



Asymmetric Information Distances

Kullback-Leibler divergence

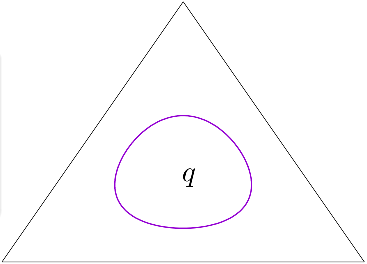
- $D[p,q] = \mathbb{E}_q\{\ln(p/q)\}$
- $D[p_1 \otimes p_2, q_1 \otimes q_2] = D[p_1, q_1] + D[p_2, q_2]$
- $\ln : (\mathbb{R}_+, \times) \to (\mathbb{R}, +)$



Asymmetric Information Distances

Kullback-Leibler divergence

- $D[p, q] = \mathbb{E}_q\{\ln(p/q)\}$
- $D[p_1 \otimes p_2, q_1 \otimes q_2] = D[p_1, q_1] + D[p_2, q_2]$
- $\ln : (\mathbb{R}_+, \times) \to (\mathbb{R}, +)$



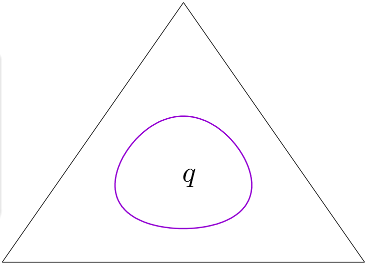
Asymmetry of the KL-divergence

- $D[p, q] \neq D[q, p]$

Asymmetric Information Distances

Kullback-Leibler divergence

- $D[p,q] = \mathbb{E}_q\{\ln(p/q)\}$
- $D[p_1 \otimes p_2, q_1 \otimes q_2] = D[p_1, q_1] + D[p_2, q_2]$
- $\ln : (\mathbb{R}_+, \times) \to (\mathbb{R}, +)$



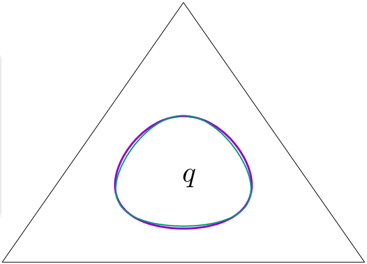
Asymmetry of the KL-divergence

- $D[p,q] \neq D[q,p]$
- $D[q + (p-q), q] \neq D[q - (p-q), q]$

Asymmetric Information Distances

Kullback-Leibler divergence

- $D[p, q] = \mathbb{E}_q\{\ln(p/q)\}$
- $D[p_1 \otimes p_2, q_1 \otimes q_2] = D[p_1, q_1] + D[p_2, q_2]$
- $\ln : (\mathbb{R}_+, \times) \to (\mathbb{R}, +)$



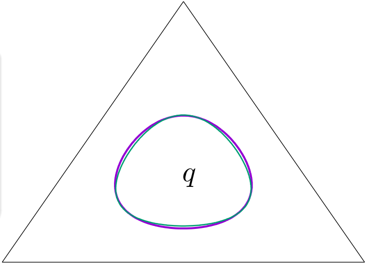
Asymmetry of the KL-divergence

- $D[p, q] \neq D[q, p]$
- $D[q + (p - q), q] \neq D[q - (p - q), q]$

Asymmetric Information Distances

Kullback-Leibler divergence

- $D[p,q] = \mathbb{E}_q\{\ln(p/q)\}$
- $D[p_1 \otimes p_2, q_1 \otimes q_2] = D[p_1, q_1] + D[p_2, q_2]$
- $\ln : (\mathbb{R}_+, \times) \to (\mathbb{R}, +)$



Asymmetry of the KL-divergence

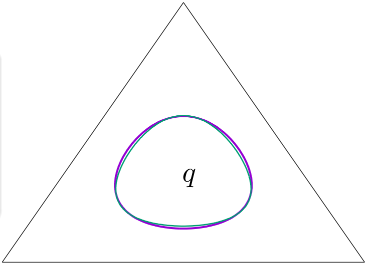
- $D[p,q] \neq D[q,p]$
- $D[q + (p-q), q] \neq D[q - (p-q), q]$

$$\|p - q| \quad = \quad \inf\{\alpha^{-1} > 0 : D[q + \alpha(p-q), q] \leq 1\}$$

Asymmetric Information Distances

Kullback-Leibler divergence

- $D[p, q] = \mathbb{E}_q\{\ln(p/q)\}$
- $D[p_1 \otimes p_2, q_1 \otimes q_2] = D[p_1, q_1] + D[p_2, q_2]$
- $\ln : (\mathbb{R}_+, \times) \to (\mathbb{R}, +)$



Asymmetry of the KL-divergence

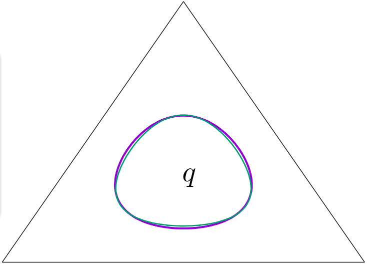
- $D[p, q] \neq D[q, p]$
- $D[q + (p - q), q] \neq D[q - (p - q), q]$

$$\begin{aligned} \|p - q| &= \inf\{\alpha^{-1} > 0 : D[q + \alpha(p - q), q] \leq 1\} \\ &\simeq \sup_x \{\mathbb{E}_{p-q}\{x\} : \mathbb{E}_q\{e^x - 1 - x\} \leq 1\} \end{aligned}$$

Asymmetric Information Distances

Kullback-Leibler divergence

- $D[p,q] = \mathbb{E}_q\{\ln(p/q)\}$
- $D[p_1 \otimes p_2, q_1 \otimes q_2] = D[p_1, q_1] + D[p_2, q_2]$
- $\ln : (\mathbb{R}_+, \times) \to (\mathbb{R}, +)$



Asymmetry of the KL-divergence

- $D[p,q] \neq D[q,p]$
- $D[q + (p-q), q] \neq D[q - (p-q), q]$

$$\begin{aligned} \|p-q\| &= \inf\{\alpha^{-1} > 0 : D[q + \alpha|(p-q)|, q] \leq 1\} \\ &\simeq \sup_x\{\mathbb{E}_{p-q}\{x\} : \mathbb{E}_q\{e^{|x|} - 1 - |x|\} \leq 1\} \end{aligned}$$

Functional Analysis in Asymmetric Spaces

Theorem (e.g. Theorem 1.5 in Fletcher and Lindgren (1982))

Every topological space with a countable base is quasi-pseudometrizable.

Functional Analysis in Asymmetric Spaces

Theorem (e.g. Theorem 1.5 in Fletcher and Lindgren (1982))

Every topological space with a countable base is quasi-pseudometrizable.

- An asymmetric seminormed space can be T_0, but not T_1 (and hence not Hausdorff T_2).

Functional Analysis in Asymmetric Spaces

Theorem (e.g. Theorem 1.5 in Fletcher and Lindgren (1982))

Every topological space with a countable base is quasi-pseudometrizable.

- An asymmetric seminormed space can be T_0, but not T_1 (and hence not Hausdorff T_2).
- Dual quasimetrics $\rho(x, y)$ and $\rho^{-1}(x, y) = \rho(y, x)$ induce two different topologies.

Functional Analysis in Asymmetric Spaces

Theorem (e.g. Theorem 1.5 in Fletcher and Lindgren (1982))

Every topological space with a countable base is quasi-pseudometrizable.

- An asymmetric seminormed space can be T_0, but not T_1 (and hence not Hausdorff T_2).
- Dual quasimetrics $\rho(x,y)$ and $\rho^{-1}(x,y) = \rho(y,x)$ induce two different topologies.
- There are 7 notions of Cauchy sequences: *left (right) Cauchy*, *left (right) K-Cauchy*, *weakly left (right) K-Cauchy*, *Cauchy*.

Functional Analysis in Asymmetric Spaces

Theorem (e.g. Theorem 1.5 in Fletcher and Lindgren (1982))

Every topological space with a countable base is quasi-pseudometrizable.

- An asymmetric seminormed space can be T_0, but not T_1 (and hence not Hausdorff T_2).
- Dual quasimetrics $\rho(x, y)$ and $\rho^{-1}(x, y) = \rho(y, x)$ induce two different topologies.
- There are 7 notions of Cauchy sequences: *left (right) Cauchy*, *left (right) K-Cauchy*, *weakly left (right) K-Cauchy*, *Cauchy*.
- This gives 14 notions of completeness (with respect to ρ or ρ^{-1}).

Functional Analysis in Asymmetric Spaces

Theorem (e.g. Theorem 1.5 in Fletcher and Lindgren (1982))

Every topological space with a countable base is quasi-pseudometrizable.

- An asymmetric seminormed space can be T_0, but not T_1 (and hence not Hausdorff T_2).
- Dual quasimetrics $\rho(x, y)$ and $\rho^{-1}(x, y) = \rho(y, x)$ induce two different topologies.
- There are 7 notions of Cauchy sequences: *left (right) Cauchy*, *left (right) K-Cauchy*, *weakly left (right) K-Cauchy*, *Cauchy*.
- This gives 14 notions of completeness (with respect to ρ or ρ^{-1}).
- Compactness is related to *outer precompactness* or *precompactness*, which are strictly weaker properties than *total boundedness*.

Functional Analysis in Asymmetric Spaces

Theorem (e.g. Theorem 1.5 in Fletcher and Lindgren (1982))

Every topological space with a countable base is quasi-pseudometrizable.

- An asymmetric seminormed space can be T_0, but not T_1 (and hence not Hausdorff T_2).
- Dual quasimetrics $\rho(x, y)$ and $\rho^{-1}(x, y) = \rho(y, x)$ induce two different topologies.
- There are 7 notions of Cauchy sequences: *left (right) Cauchy*, *left (right) K-Cauchy*, *weakly left (right) K-Cauchy*, *Cauchy*.
- This gives 14 notions of completeness (with respect to ρ or ρ^{-1}).
- Compactness is related to *outer precompactness* or *precompactness*, which are strictly weaker properties than *total boundedness*.
- An asymmetric seminormed space may fail to be a topological vector space, because $y \mapsto \alpha y$ can be discontinuous (Borodin, 2001).

Functional Analysis in Asymmetric Spaces

Theorem (e.g. Theorem 1.5 in Fletcher and Lindgren (1982))

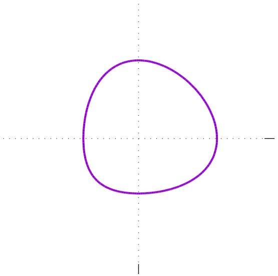
Every topological space with a countable base is quasi-pseudometrizable.

- An asymmetric seminormed space can be T_0, but not T_1 (and hence not Hausdorff T_2).
- Dual quasimetrics $\rho(x, y)$ and $\rho^{-1}(x, y) = \rho(y, x)$ induce two different topologies.
- There are 7 notions of Cauchy sequences: *left (right) Cauchy*, *left (right) K-Cauchy*, *weakly left (right) K-Cauchy*, *Cauchy*.
- This gives 14 notions of completeness (with respect to ρ or ρ^{-1}).
- Compactness is related to *outer precompactness* or *precompactness*, which are strictly weaker properties than *total boundedness*.
- An asymmetric seminormed space may fail to be a topological vector space, because $y \mapsto \alpha y$ can be discontinuous (Borodin, 2001).
- Practically all other results have to be reconsidered (e.g. Baire category theorem, Alaoglu-Bourbaki, etc). (Cobzas, 2013).

Random Variables as the Source of Asymmetry

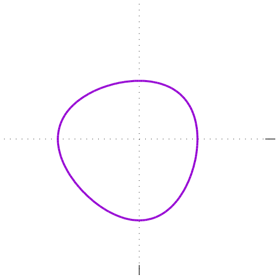
$M^\circ := \{x : \langle x, y\rangle \leq 1,\ \forall\, y \in M\}$

M

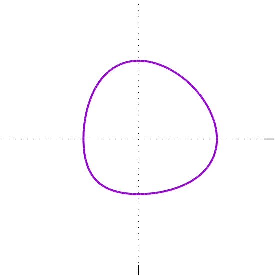


Random Variables as the Source of Asymmetry

$M^\circ := \{x : \langle x, y\rangle \leq 1,\ \forall\, y \in M\}$

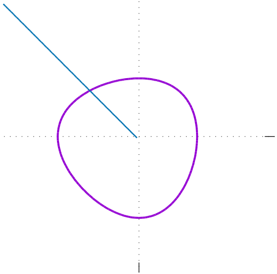


M

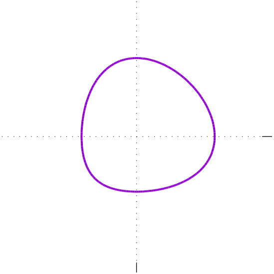


Random Variables as the Source of Asymmetry

$M^\circ := \{x : \langle x, y\rangle \leq 1,\ \forall\, y \in M\}$



M

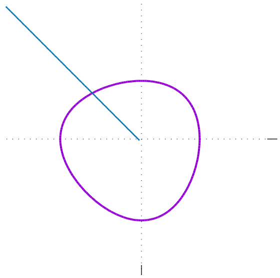


- Minkowski functional:

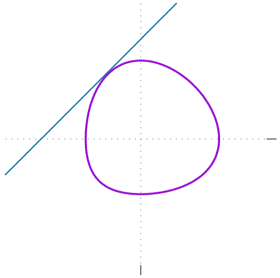
$$\mu M^\circ(x) = \inf\{\alpha > 0 : x/\alpha \in M^\circ\}$$

Random Variables as the Source of Asymmetry

$M^\circ := \{x : \langle x, y\rangle \leq 1,\ \forall\, y \in M\}$



M



- Minkowski functional:

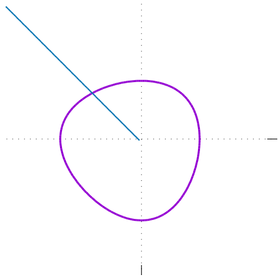
  $\mu M^\circ(x) = \inf\{\alpha > 0 : x/\alpha \in M^\circ\}$

- Support function

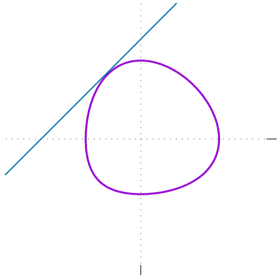
  $= sM(x) = \sup\{\langle x, y\rangle : y \in M\}$

Random Variables as the Source of Asymmetry

$M^\circ := \{x : \langle x, y\rangle \leq 1,\ \forall\, y \in M\}$



M



- Minkowski functional:

  $\mu M^\circ(x) = \inf\{\alpha > 0 : x/\alpha \in M^\circ\}$

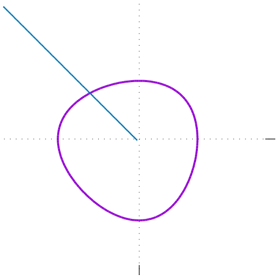
- Support function

  $= sM(x) = \sup\{\langle x, y\rangle : y \in M\}$

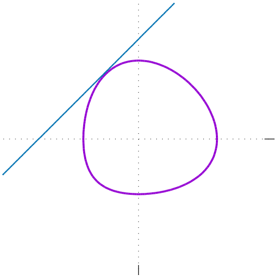
- $M = \{u : D[(1+u)z, z] \leq 1\}$

Random Variables as the Source of Asymmetry

$M^\circ := \{x : \langle x, y\rangle \leq 1,\ \forall\, y \in M\}$



M



- Minkowski functional:

  $$\mu M^\circ(x) = \inf\{\alpha > 0 : x/\alpha \in M^\circ\}$$

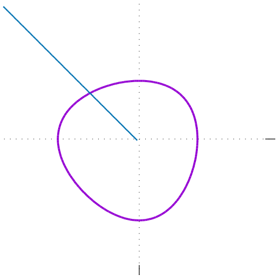
- Support function

  $$= sM(x) = \sup\{\langle x, y\rangle : y \in M\}$$

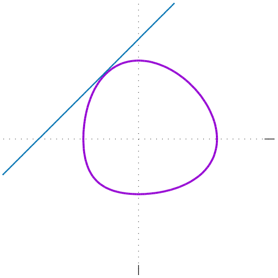
- $M = \{u : D[(1+u)z, z] \leq 1\}$
- $D = \langle (1+u)\ln(1+u) - u, z\rangle$

Random Variables as the Source of Asymmetry

$M^\circ := \{x : \langle x, y\rangle \leq 1,\ \forall\, y \in M\}$



M



- Minkowski functional:

  $\mu M^\circ(x) = \inf\{\alpha > 0 : x/\alpha \in M^\circ\}$

- $M^\circ \simeq \{y : D^*[x, 0] \leq 1\}$

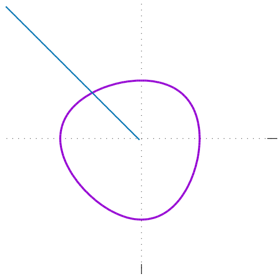
- Support function

  $= sM(x) = \sup\{\langle x, y\rangle : y \in M\}$

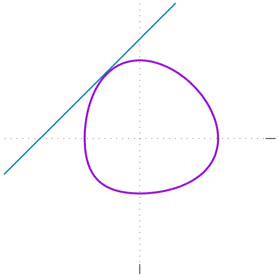
- $M = \{u : D[(1 + u)z, z] \leq 1\}$
- $D = \langle (1 + u)\ln(1 + u) - u, z\rangle$

Random Variables as the Source of Asymmetry

$M^\circ := \{x : \langle x, y\rangle \leq 1,\ \forall\, y \in M\}$



M



- Minkowski functional:

$$\mu M^\circ(x) = \inf\{\alpha > 0 : x/\alpha \in M^\circ\}$$

- Support function

$$= sM(x) = \sup\{\langle x, y\rangle : y \in M\}$$

- $M^\circ \simeq \{y : D^*[x, 0] \leq 1\}$
- $D^*[x, 0] = \langle e^x - 1 - x, z\rangle$
- $M = \{u : D[(1 + u)z, z] \leq 1\}$
- $D = \langle (1 + u)\ln(1 + u) - u, z\rangle$

Examples

Example (St. Peterbourgh lottery)

- $x = 2^n$, $q = 2^{-n}$, $n \in \mathbb{N}$.

Examples

Example (St. Peterbourgh lottery)

- $x = 2^n$, $q = 2^{-n}$, $n \in \mathbb{N}$.
- $\mathbb{E}_q\{x\} = \sum_{n=1}^{\infty}(2^n/2^n) \to \infty$

Examples

Example (St. Peterbourgh lottery)

- $x = 2^n$, $q = 2^{-n}$, $n \in \mathbb{N}$.
- $\mathbb{E}_q\{x\} = \sum_{n=1}^{\infty}(2^n/2^n) \to \infty$
- $\mathbb{E}_p\{x\} < \infty$ for all biased $p = 2^{-(1+\alpha)n}$, $\alpha > 0$.

Examples

Example (St. Peterbourgh lottery)

- $x = 2^n$, $q = 2^{-n}$, $n \in \mathbb{N}$.
- $\mathbb{E}_q\{x\} = \sum_{n=1}^{\infty}(2^n/2^n) \to \infty$
- $\mathbb{E}_p\{x\} < \infty$ for all biased $p = 2^{-(1+\alpha)n}$, $\alpha > 0$.
- $2^n \notin \operatorname{dom}\mathbb{E}_q\{e^x\}$, $-2^n \in \operatorname{dom}\mathbb{E}_q\{e^x\}$

Examples

Example (St. Peterbourgh lottery)

- $x = 2^n$, $q = 2^{-n}$, $n \in \mathbb{N}$.
- $\mathbb{E}_q\{x\} = \sum_{n=1}^{\infty}(2^n/2^n) \to \infty$
- $\mathbb{E}_p\{x\} < \infty$ for all biased $p = 2^{-(1+\alpha)n}$, $\alpha > 0$.
- $2^n \notin \operatorname{dom} \mathbb{E}_q\{e^x\}$, $-2^n \in \operatorname{dom} \mathbb{E}_q\{e^x\}$
- $0 \notin \operatorname{Int}(\operatorname{dom} \mathbb{E}_q\{e^x\})$

Examples

Example (St. Peterbourgh lottery)

- $x = 2^n$, $q = 2^{-n}$, $n \in \mathbb{N}$.
- $\mathbb{E}_q\{x\} = \sum_{n=1}^{\infty}(2^n/2^n) \to \infty$
- $\mathbb{E}_p\{x\} < \infty$ for all biased $p = 2^{-(1+\alpha)n}$, $\alpha > 0$.
- $2^n \notin \operatorname{dom}\mathbb{E}_q\{e^x\}$, $-2^n \in \operatorname{dom}\mathbb{E}_q\{e^x\}$
- $0 \notin \operatorname{Int}(\operatorname{dom}\mathbb{E}_q\{e^x\})$

Example (Error minimization)

- Minimize $x = \frac{1}{2}\|a - b\|_2^2$ subject to $D_{KL}[w, q \otimes p] \leq \lambda$, $a, b \in \mathbb{R}^n$.

Examples

Example (St. Peterbourgh lottery)

- $x = 2^n$, $q = 2^{-n}$, $n \in \mathbb{N}$.
- $\mathbb{E}_q\{x\} = \sum_{n=1}^{\infty}(2^n/2^n) \to \infty$
- $\mathbb{E}_p\{x\} < \infty$ for all biased $p = 2^{-(1+\alpha)n}$, $\alpha > 0$.
- $2^n \notin \operatorname{dom}\mathbb{E}_q\{e^x\}$, $-2^n \in \operatorname{dom}\mathbb{E}_q\{e^x\}$
- $0 \notin \operatorname{Int}(\operatorname{dom}\mathbb{E}_q\{e^x\})$

Example (Error minimization)

- Minimize $x = \frac{1}{2}\|a - b\|_2^2$ subject to $D_{KL}[w, q \otimes p] \leq \lambda$, $a, b \in \mathbb{R}^n$.
- $\mathbb{E}_w\{x\} < \infty$ minimized at $\overline{w} \propto e^{-\beta x} q \otimes p$.

Examples

Example (St. Peterbourgh lottery)

- $x = 2^n$, $q = 2^{-n}$, $n \in \mathbb{N}$.
- $\mathbb{E}_q\{x\} = \sum_{n=1}^{\infty}(2^n/2^n) \to \infty$
- $\mathbb{E}_p\{x\} < \infty$ for all biased $p = 2^{-(1+\alpha)n}$, $\alpha > 0$.
- $2^n \notin \operatorname{dom}\mathbb{E}_q\{e^x\}$, $-2^n \in \operatorname{dom}\mathbb{E}_q\{e^x\}$
- $0 \notin \operatorname{Int}(\operatorname{dom}\mathbb{E}_q\{e^x\})$

Example (Error minimization)

- Minimize $x = \frac{1}{2}\|a - b\|_2^2$ subject to $D_{KL}[w, q \otimes p] \leq \lambda$, $a, b \in \mathbb{R}^n$.
- $\mathbb{E}_w\{x\} < \infty$ minimized at $\overline{w} \propto e^{-\beta x} q \otimes p$.
- Maximization of x has no solution.

Examples

Example (St. Peterbourgh lottery)

- $x = 2^n$, $q = 2^{-n}$, $n \in \mathbb{N}$.
- $\mathbb{E}_q\{x\} = \sum_{n=1}^{\infty}(2^n/2^n) \to \infty$
- $\mathbb{E}_p\{x\} < \infty$ for all biased $p = 2^{-(1+\alpha)n}$, $\alpha > 0$.
- $2^n \notin \operatorname{dom} \mathbb{E}_q\{e^x\}$, $-2^n \in \operatorname{dom} \mathbb{E}_q\{e^x\}$
- $0 \notin \operatorname{Int}(\operatorname{dom} \mathbb{E}_q\{e^x\})$

Example (Error minimization)

- Minimize $x = \frac{1}{2}\|a - b\|_2^2$ subject to $D_{KL}[w, q \otimes p] \leq \lambda$, $a, b \in \mathbb{R}^n$.
- $\mathbb{E}_w\{x\} < \infty$ minimized at $\overline{w} \propto e^{-\beta x} q \otimes p$.
- Maximization of x has no solution.
- $\frac{1}{2}\|a - b\|_2^2 \notin \operatorname{dom} \mathbb{E}_{q \otimes p}\{e^x\}$, $-\frac{1}{2}\|a - b\|_2^2 \in \operatorname{dom} \mathbb{E}_{q \otimes p}\{e^x\}$

Examples

Example (St. Peterbourgh lottery)

- $x = 2^n$, $q = 2^{-n}$, $n \in \mathbb{N}$.
- $\mathbb{E}_q\{x\} = \sum_{n=1}^{\infty}(2^n/2^n) \to \infty$
- $\mathbb{E}_p\{x\} < \infty$ for all biased $p = 2^{-(1+\alpha)n}$, $\alpha > 0$.
- $2^n \notin \operatorname{dom} \mathbb{E}_q\{e^x\}$, $-2^n \in \operatorname{dom} \mathbb{E}_q\{e^x\}$
- $0 \notin \operatorname{Int}(\operatorname{dom} \mathbb{E}_q\{e^x\})$

Example (Error minimization)

- Minimize $x = \frac{1}{2}\|a - b\|_2^2$ subject to $D_{KL}[w, q \otimes p] \leq \lambda$, $a, b \in \mathbb{R}^n$.
- $\mathbb{E}_w\{x\} < \infty$ minimized at $\overline{w} \propto e^{-\beta x} q \otimes p$.
- Maximization of x has no solution.
- $\frac{1}{2}\|a - b\|_2^2 \notin \operatorname{dom} \mathbb{E}_{q\otimes p}\{e^x\}$, $-\frac{1}{2}\|a - b\|_2^2 \in \operatorname{dom} \mathbb{E}_{q\otimes p}\{e^x\}$
- $0 \notin \operatorname{Int}(\operatorname{dom} \mathbb{E}_{q\otimes p}\{e^x\})$

Sources and Consequences of Asymmetry

Method: Symmetric Sandwich

Results

Method: Symmetric Sandwich

- $s[-A \cap A] \leq sA \leq s[-A \cup A]$

Method: Symmetric Sandwich

- $s[-A \cap A] \leq sA \leq s[-A \cup A]$
- $\mu\mathrm{co}\,[-A^\circ \cup A^\circ] \leq \mu A^\circ \leq \mu[-A^\circ \cap A^\circ]$

Method: Symmetric Sandwich

- $s[-A \cap A] \leq sA \leq s[-A \cup A]$
- $\mu\mathrm{co}\,[-A^\circ \cup A^\circ] \leq \mu A^\circ \leq \mu[-A^\circ \cap A^\circ]$
- $s[-A \cap A] = s(-A)\mathrm{co} \wedge sA = \inf\{sA(z) + sA(z-y) : z \in Y\}$

Method: Symmetric Sandwich

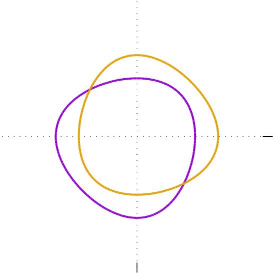
- $s[-A \cap A] \leq sA \leq s[-A \cup A]$
- $\mu\mathrm{co}\,[-A^\circ \cup A^\circ] \leq \mu A^\circ \leq \mu[-A^\circ \cap A^\circ]$
- $s[-A \cap A] = s(-A)\mathrm{co} \wedge sA = \inf\{sA(z) + sA(z-y) : z \in Y\}$
- $s[-A \cup A] = s(-A) \vee sA$

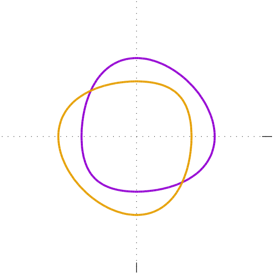
Method: Symmetric Sandwich

- $s[-A \cap A] \leq sA \leq s[-A \cup A]$
- $\mu\text{co}\,[-A^\circ \cup A^\circ] \leq \mu A^\circ \leq \mu[-A^\circ \cap A^\circ]$
- $s[-A \cap A] = s(-A)\text{co} \wedge sA = \inf\{sA(z) + sA(z-y) : z \in Y\}$
- $s[-A \cup A] = s(-A) \vee sA$

Method: Symmetric Sandwich

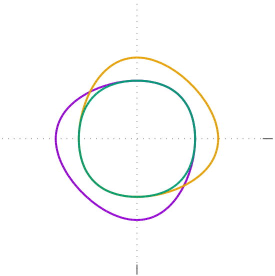
- $s[-A \cap A] \leq sA \leq s[-A \cup A]$
- $\mu\mathrm{co}\,[-A^\circ \cup A^\circ] \leq \mu A^\circ \leq \mu[-A^\circ \cap A^\circ]$
- $s[-A \cap A] = s(-A)\mathrm{co} \wedge sA = \inf\{sA(z) + sA(z-y) : z \in Y\}$
- $s[-A \cup A] = s(-A) \vee sA$



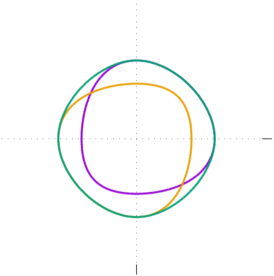


Method: Symmetric Sandwich

- $s[-A \cap A] \leq sA \leq s[-A \cup A]$
- $\mu\mathrm{co}\,[-A^\circ \cup A^\circ] \leq \mu A^\circ \leq \mu[-A^\circ \cap A^\circ]$
- $s[-A \cap A] = s(-A)\mathrm{co} \wedge sA = \inf\{sA(z) + sA(z-y) : z \in Y\}$
- $s[-A \cup A] = s(-A) \vee sA$



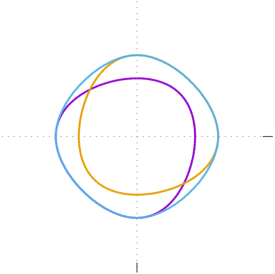
$\mu M^\circ \leq \mu(-M^\circ) \vee \mu M^\circ$



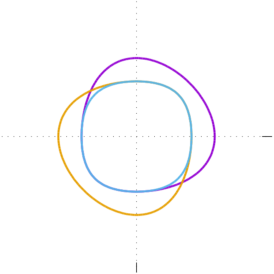
$\mu(-M)\mathrm{co} \wedge \mu M \leq \mu M$

Method: Symmetric Sandwich

- $s[-A \cap A] \leq sA \leq s[-A \cup A]$
- $\mu\mathrm{co}\,[-A^\circ \cup A^\circ] \leq \mu A^\circ \leq \mu[-A^\circ \cap A^\circ]$
- $s[-A \cap A] = s(-A)\mathrm{co} \wedge sA = \inf\{sA(z) + sA(z-y) : z \in Y\}$
- $s[-A \cup A] = s(-A) \vee sA$

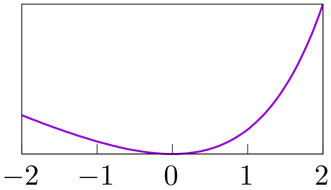


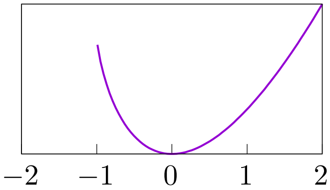
$$\mu(-M^\circ)\mathrm{co} \wedge \mu M^\circ \leq \mu M^\circ$$



$$\mu M \leq \mu(-M) \vee \mu M$$

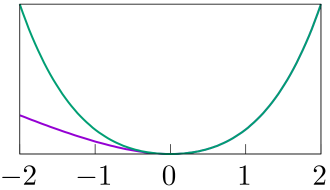
Lower and upper Luxemburg (Orlicz) norms

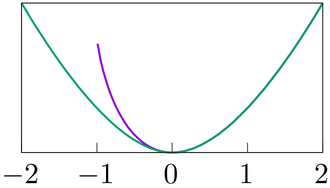




$$\varphi^*(x) = e^x - 1 - x \qquad \varphi(u) = (1+u)\ln(1+u) - u$$

Lower and upper Luxemburg (Orlicz) norms

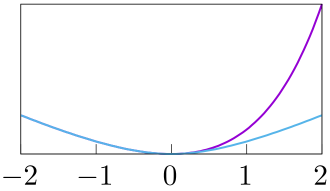


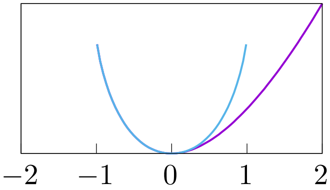


$$\begin{aligned}\varphi^*(x) &= e^x - 1 - x\\ \varphi^*_+(x) &= \varphi^*(|x|) \notin \Delta_2\end{aligned}$$

$$\begin{aligned}\varphi(u) &= (1+u)\ln(1+u) - u\\ \varphi_+(u) &= \varphi(|u|) \in \Delta_2\end{aligned}$$

Lower and upper Luxemburg (Orlicz) norms

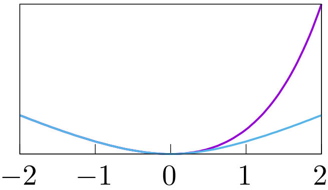


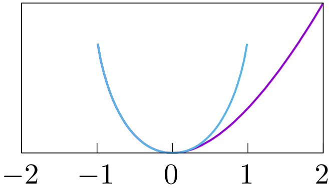


$$\begin{aligned}\varphi^*(x) &= e^x - 1 - x \\ \varphi^*_+(x) &= \varphi^*(|x|) \notin \Delta_2 \\ \varphi^*_-(x) &= \varphi^*(-|x|) \in \Delta_2\end{aligned}$$

$$\begin{aligned}\varphi(u) &= (1+u)\ln(1+u) - u \\ \varphi_+(u) &= \varphi(|u|) \in \Delta_2 \\ \varphi_-(u) &= \varphi(-|u|) \notin \Delta_2\end{aligned}$$

Lower and upper Luxemburg (Orlicz) norms

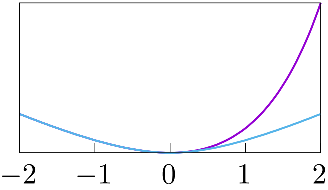


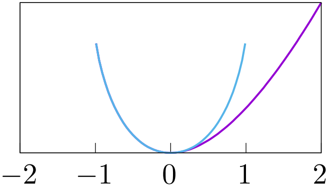


$$
\begin{aligned}
\varphi^*(x) &= e^x - 1 - x \\
\varphi^*_+(x) &= \varphi^*(|x|) \notin \Delta_2 \\
\varphi^*_-(x) &= \varphi^*(-|x|) \in \Delta_2 \\
\|x|^*_\varphi &= \mu\{x : \langle \varphi^*(x), z \rangle \leq 1\}
\end{aligned}
$$

$$
\begin{aligned}
\varphi(u) &= (1+u)\ln(1+u) - u \\
\varphi_+(u) &= \varphi(|u|) \in \Delta_2 \\
\varphi_-(u) &= \varphi(-|u|) \notin \Delta_2 \\
\|u|_\varphi &= \mu\{u : \langle \varphi(u), z \rangle \leq 1\}
\end{aligned}
$$

Lower and upper Luxemburg (Orlicz) norms





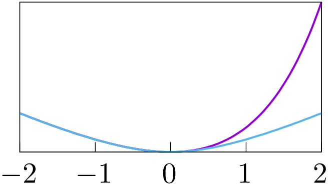
$$\begin{aligned}
\varphi^*(x) &= e^x - 1 - x \\
\varphi^*_+(x) &= \varphi^*(|x|) \notin \Delta_2 \\
\varphi^*_-(x) &= \varphi^*(-|x|) \in \Delta_2 \\
\|x|^*_\varphi &= \mu\{x : \langle \varphi^*(x), z \rangle \leq 1\}
\end{aligned}$$

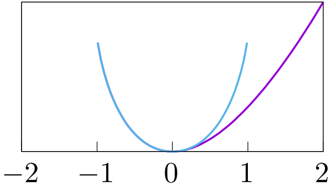
$$\begin{aligned}
\varphi(u) &= (1+u)\ln(1+u) - u \\
\varphi_+(u) &= \varphi(|u|) \in \Delta_2 \\
\varphi_-(u) &= \varphi(-|u|) \notin \Delta_2 \\
\|u|_\varphi &= \mu\{u : \langle \varphi(u), z \rangle \leq 1\}
\end{aligned}$$

Proposition

- $\|\cdot\|^*_{\varphi+}$, $\|\cdot\|^*_{\varphi-}$ *are Luxemburg norms and* $\|x\|^*_{\varphi-} \leq \|x|^*_\varphi \leq \|x\|^*_{\varphi+}$
- $\|\cdot\|_{\varphi+}$, $\|\cdot\|_{\varphi-}$ *are Luxemburg norms and* $\|u\|_{\varphi+} \leq \|u|_\varphi \leq \|u\|_{\varphi-}$

Lower and upper Luxemburg (Orlicz) norms





$$\begin{aligned} \varphi^*(x) &= e^x - 1 - x \\ \varphi^*_+(x) &= \varphi^*(|x|) \notin \Delta_2 \\ \varphi^*_-(x) &= \varphi^*(-|x|) \in \Delta_2 \\ \|x|^*_\varphi &= \mu\{x : \langle \varphi^*(x), z\rangle \le 1\} \end{aligned}$$

$$\begin{aligned} \varphi(u) &= (1+u)\ln(1+u) - u \\ \varphi_+(u) &= \varphi(|u|) \in \Delta_2 \\ \varphi_-(u) &= \varphi(-|u|) \notin \Delta_2 \\ \|u|_\varphi &= \mu\{u : \langle \varphi(u), z\rangle \le 1\} \end{aligned}$$

Proposition

- $\|\cdot\|^*_{\varphi+}$, $\|\cdot\|^*_{\varphi-}$ *are Luxemburg norms and* $\|x\|^*_{\varphi-} \le \|x|^*_\varphi \le \|x\|^*_{\varphi+}$
- $\|\cdot\|_{\varphi+}$, $\|\cdot\|_{\varphi-}$ *are Luxemburg norms and* $\|u\|_{\varphi+} \le \|u|_\varphi \le \|u\|_{\varphi-}$

Sources and Consequences of Asymmetry

Method: Symmetric Sandwich

Results

KL Induces Hausdorff (T_2) Asymmetric Topology

Theorem

$(Y, \|\cdot|_\varphi)$ *(resp.* $(X, \|\cdot|^*_\varphi)$*) is Hausdorff.*

KL Induces Hausdorff (T_2) Asymmetric Topology

Theorem

$(Y, \| \cdot |_\varphi)$ *(resp.* $(X, \| \cdot |^*_\varphi)$*) is Hausdorff.*

Proof.

$\|u\|_{\varphi+} \leq \|u|_\varphi$ (resp. $\|x\|_{\varphi-} \leq \|x|_\varphi$) implies $(Y, \| \cdot |_\varphi)$ (resp. $(X, \| \cdot |^*_\varphi)$) is finer than normed space $(Y, \| \cdot \|_{\varphi+})$ (resp. $(X, \| \cdot \|^*_{\varphi-})$). □

Separable Subspaces

Theorem

$(Y, \|\cdot\|_{\varphi+})$ *(resp.* $(X, \|\cdot\|^*_{\varphi-})$*) is a separable Orlicz subspace of* $(Y, \|\cdot|_{\varphi})$ *(resp.* $(X, \|\cdot|^*_{\varphi})$*).*

Separable Subspaces

Theorem

$(Y, \|\cdot\|_{\varphi+})$ *(resp.* $(X, \|\cdot\|^*_{\varphi-})$*) is a separable Orlicz subspace of* $(Y, \|\cdot|_\varphi)$ *(resp.* $(X, \|\cdot|^*_\varphi)$*).*

Proof.

$\varphi_+(u) = (1 + |u|)\ln(1 + |u|) - |u| \in \Delta_2$ (resp. $\varphi^*_-(x) = e^{-|x|} - 1 + |x| \in \Delta_2$). Note that $\varphi_- \notin \Delta_2$ and $\varphi^*_+ \notin \Delta_2$. □

Completeness

Theorem

$(Y, \|\cdot|_\varphi)$ *(resp.* $(X, \|\cdot|^*_\varphi)$*) is*

1. *Bi-Complete:* ρ^s*-Cauchy* $y_n \xrightarrow{\rho^s} y$.

Completeness

Theorem

$(Y, \|\cdot|_\varphi)$ *(resp.* $(X, \|\cdot|^*_\varphi)$*) is*

1. *Bi-Complete:* ρ^s*-Cauchy* $y_n \xrightarrow{\rho^s} y$.
2. ρ*-sequentially complete:* ρ^s*-Cauchy* $y_n \xrightarrow{\rho} y$.

Completeness

Theorem

$(Y, \| \cdot |_\varphi)$ (resp. $(X, \| \cdot |_\varphi^*)$) is

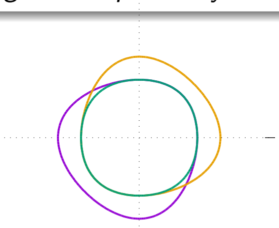
1. *Bi-Complete: ρ^s-Cauchy $y_n \stackrel{\rho^s}{\to} y$.*
2. *ρ-sequentially complete: ρ^s-Cauchy $y_n \stackrel{\rho}{\to} y$.*
3. *Right K-sequentially complete: right K-Cauchy $y_n \stackrel{\rho}{\to} y$.*

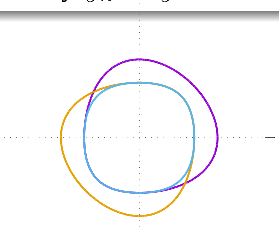
Completeness

Theorem

$(Y, \|\cdot|_\varphi)$ *(resp.* $(X, \|\cdot|^*_\varphi)$*) is*

1. *Bi-Complete:* ρ^s*-Cauchy* $y_n \xrightarrow{\rho^s} y$.
2. ρ*-sequentially complete:* ρ^s*-Cauchy* $y_n \xrightarrow{\rho} y$.
3. *Right K-sequentially complete: right K-Cauchy* $y_n \xrightarrow{\rho} y$.





Proof.

$\rho^s(y,z) = \|z-y|_\varphi \vee \|y-z|_\varphi \le \|y-z\|_{\varphi-}$, where $(Y, \|\cdot\|_{\varphi-})$ is Banach.
Then use theorems of Reilly et al. (1982) and Chen et al. (2007). □

Summary and Further Questions

- Topologies induced by asymmetric information divergences may not have the same properties as their symmetrized counterparts (e.g. Banach spaces), and therefore many properties have to be re-examined.

Summary and Further Questions

- Topologies induced by asymmetric information divergences may not have the same properties as their symmetrized counterparts (e.g. Banach spaces), and therefore many properties have to be re-examined.
- We have proved that topologies induced by the KL-divergence are:

Summary and Further Questions

- Topologies induced by asymmetric information divergences may not have the same properties as their symmetrized counterparts (e.g. Banach spaces), and therefore many properties have to be re-examined.
- We have proved that topologies induced by the KL-divergence are:
    - Hausdorff.

Summary and Further Questions

- Topologies induced by asymmetric information divergences may not have the same properties as their symmetrized counterparts (e.g. Banach spaces), and therefore many properties have to be re-examined.
- We have proved that topologies induced by the KL-divergence are:
  - Hausdorff.
  - Bi-complete, ρ-sequentially complete and right K-sequentially complete.

Summary and Further Questions

- Topologies induced by asymmetric information divergences may not have the same properties as their symmetrized counterparts (e.g. Banach spaces), and therefore many properties have to be re-examined.
- We have proved that topologies induced by the KL-divergence are:
  - Hausdorff.
  - Bi-complete, ρ-sequentially complete and right K-sequentially complete.
  - Contain a separable Orlicz subspace.

Summary and Further Questions

- Topologies induced by asymmetric information divergences may not have the same properties as their symmetrized counterparts (e.g. Banach spaces), and therefore many properties have to be re-examined.
- We have proved that topologies induced by the KL-divergence are:
  - Hausdorff.
  - Bi-complete, ρ-sequentially complete and right K-sequentially complete.
  - Contain a separable Orlicz subspace.
- Total boundedness, compactness?

Summary and Further Questions

- Topologies induced by asymmetric information divergences may not have the same properties as their symmetrized counterparts (e.g. Banach spaces), and therefore many properties have to be re-examined.
- We have proved that topologies induced by the KL-divergence are:
  - Hausdorff.
  - Bi-complete, ρ-sequentially complete and right K-sequentially complete.
  - Contain a separable Orlicz subspace.
- Total boundedness, compactness?
- Other asymmetric information distances (e.g. Renyi divergence).

Sources and Consequences of Asymmetry

Method: Symmetric Sandwich

Results

Borodin, P. A. (2001). The Banach-Mazur theorem for spaces with asymmetric norm. *Mathematical Notes*, *69*(3–4), 298–305.

Chen, S.-A., Li, W., Zou, D., & Chen, S.-B. (2007, Aug). Fixed point theorems in quasi-metric spaces. In *Machine learning and cybernetics, 2007 international conference on* (Vol. 5, p. 2499-2504). IEEE.

Cobzas, S. (2013). *Functional analysis in asymmetric normed spaces*. Birkhäuser.

Fletcher, P., & Lindgren, W. F. (1982). *Quasi-uniform spaces* (Vol. 77). New York: Marcel Dekker.

Reilly, I. L., Subrahmanyam, P. V., & Vamanamurthy, M. K. (1982). Cauchy sequences in quasi-pseudo-metric spaces. *Monatshefte für Mathematik*, *93*, 127–140.