Partition identification for general distributions using multi-armed bandits

Sandeep Juneja

Tata Institute
Mumbai, India

Parts jointly with Subhashini Krishnasamy and Shubhada Agarwal, TIFR

PAAP, ICTS

August 17, 2019

Finding the correct partition set containing given vector of distributions

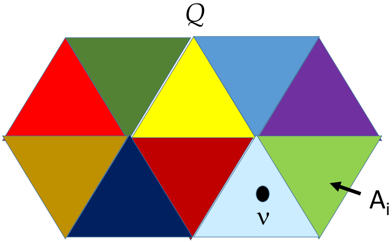
- $\mathcal{Q}$ is a collection of vectors $\nu = (\nu_1, \ldots, \nu_K)$ where each ν_i is a probability distribution. Can sample independently from each arm ν_i.

Finding the correct partition set containing given vector of distributions

- $\mathcal{Q}$ is a collection of vectors $\nu = (\nu_1, \ldots, \nu_K)$ where each ν_i is a probability distribution. Can sample independently from each arm ν_i.
- $\mathcal{Q} = \cup_{i=1}^{p} A_i$ where the A_i are disjoint

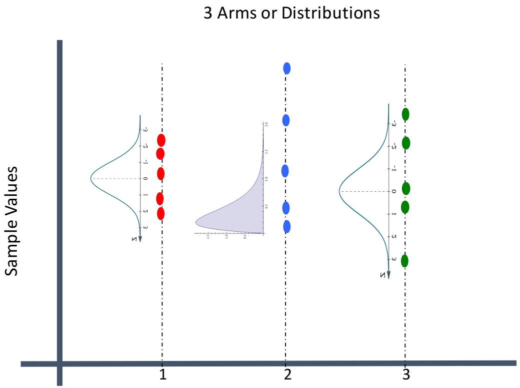
Finding the correct partition set containing given vector of distributions

- $\mathcal{Q}$ is a collection of vectors $\nu = (\nu_1, \ldots, \nu_K)$ where each ν_i is a probability distribution. Can sample independently from each arm ν_i.
- $\mathcal{Q} = \cup_{i=1}^{p} A_i$ where the A_i are disjoint
- Given a $\nu \in \mathcal{Q}$ need a δ- correct algorithm that finds the A_i it belongs to.

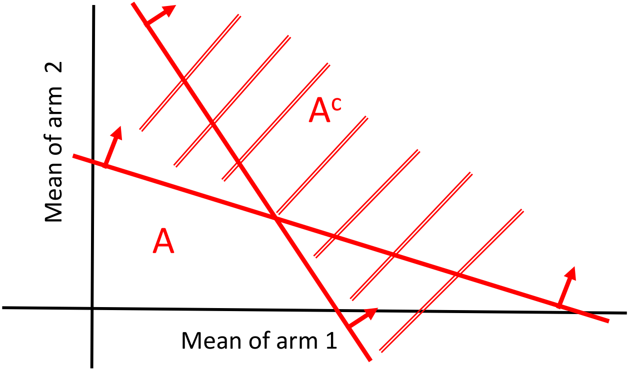


Classical Monte Carlo: Finding a distribution or best *arm* - arm with the largest mean

$\nu = (\nu_1, \nu_2, \nu_3)$. Need to find arm with the largest mean



Typically, $\mathcal{Q}$ comprises two sets: $A \cup A^c$. We consider finding if vector of means lies in a convex set (e.g., a polytope), or its complement



Agenda

- Develop methodology for computing lower bounds on computational effort for δ - correct algorithms.

Agenda

- Develop methodology for computing lower bounds on computational effort for δ - correct algorithms.
- This involves exploiting the geometry of the problem structure; use of duality or minimax theorem.

Agenda

- Develop methodology for computing lower bounds on computational effort for δ - correct algorithms.
- This involves exploiting the geometry of the problem structure; use of duality or minimax theorem.
- Develop δ-correct algorithms with matching computational bounds in general settings including the half space problem, the convex and the complement of convex set.

Agenda

- Develop methodology for computing lower bounds on computational effort for δ - correct algorithms.
- This involves exploiting the geometry of the problem structure; use of duality or minimax theorem.
- Develop δ-correct algorithms with matching computational bounds in general settings including the half space problem, the convex and the complement of convex set.
- The results are first discussed for single parameter exponential family of distributions. Later we discuss generalizations to

$$\mathcal{L} \triangleq \{\kappa \in \mathcal{P}(\Re) : E_{X\sim\kappa} f(|X|) \leq B\}$$

where f is non-negative, non-decreasing, convex and $\frac{f(x)}{x} \to \infty$ as $x \to \infty$. e.g., $f(x) = |x|^{1+\epsilon}$, $\epsilon > 0$.

Brief Literature Review

- Ranking and selection (Best arm pure exploration problem)

Brief Literature Review

- Ranking and selection (Best arm pure exploration problem)
  - Statistics: Bechhofer et. al. (1968) Uniform sampling, Paulson (1964) - elimination based. Earlier Chernoff (1959), Albert (1961)

Brief Literature Review

- Ranking and selection (Best arm pure exploration problem)
  - Statistics: Bechhofer et. al. (1968) Uniform sampling, Paulson (1964) - elimination based. Earlier Chernoff (1959), Albert (1961)
  - Simulation: Bechhofer, Goldsman, Nelson and others 90's, 2000's, Ho et. al. (1990), Dai (1996), Chen et al (2000), Glynn and J (2004)

Brief Literature Review

- Ranking and selection (Best arm pure exploration problem)
  - Statistics: Bechhofer et. al. (1968) Uniform sampling, Paulson (1964) - elimination based. Earlier Chernoff (1959), Albert (1961)
  - Simulation: Bechhofer, Goldsman, Nelson and others 90's, 2000's, Ho et. al. (1990), Dai (1996), Chen et al (2000), Glynn and J (2004)
  - Computer Science - Evan-Dar et. al. (2006), Bubeck, Audibert (2010), Kaufmann, Cappe, Garivier (2016), Garivier, Kaufmann (2016), Russo (2016)

Brief Literature Review

- Ranking and selection (Best arm pure exploration problem)
  - Statistics: Bechhofer et. al. (1968) Uniform sampling, Paulson (1964) - elimination based. Earlier Chernoff (1959), Albert (1961)
  - Simulation: Bechhofer, Goldsman, Nelson and others 90's, 2000's, Ho et. al. (1990), Dai (1996), Chen et al (2000), Glynn and J (2004)
  - Computer Science - Evan-Dar et. al. (2006), Bubeck, Audibert (2010), Kaufmann, Cappe, Garivier (2016), Garivier, Kaufmann (2016), Russo (2016)
- Regret minimization - Lai and Robbins (85), Bubeck, Auer, Audibert .... , Cappe, Garivier, Maillard, Munos, Stoltz (2013), Agarwal and Goyal (2011, 2012), Honda and Takemura (2010), Magureanu, Combes, Proutiere (2014)

We develop optimal δ-correct algorithms for the partition identification problem

- Given a vector of K arms or probability distributions, an algorithm specifies
  - an adaptive sampling strategy
  - a stopping time τ, and finally
  - a recommendation (a subset from the partition)

We develop optimal δ-correct algorithms for the partition identification problem

- Given a vector of K arms or probability distributions, an algorithm specifies
  - an adaptive sampling strategy
  - a stopping time τ, and finally
  - a recommendation (a subset from the partition)
- *An algorithm is said to be δ-correct,*

We develop optimal δ-correct algorithms for the partition identification problem

- Given a vector of K arms or probability distributions, an algorithm specifies
  - an adaptive sampling strategy
  - a stopping time τ, and finally
  - a recommendation (a subset from the partition)
- *An algorithm is said to be δ-correct,*
  - if for any set of distributions $\mu = (\mu_1, \mu_2, \ldots, \mu_K)$,
  - it announces in finite time τ, that μ belongs to some set A_j with the probability of error bounded above by δ, for all $\delta > 0$

Lower bound relies on a key Inequality

- Relies on change of measure arguments that go back at least to Lai and Robbins 1985.

Lower bound relies on a key Inequality

- Relies on change of measure arguments that go back at least to Lai and Robbins 1985.
- Under δ-correct algorithm (Kauffman, Cappe, Garivier 2016), for

$$\mu = (\mu_1, \mu_2, \ldots, \mu_K) \in A_i$$

and

$$\nu = (\nu_1, \nu_2, \ldots, \nu_K) \in A_i^c$$

where each arm i is pulled N_i times,

Lower bound relies on a key Inequality

- Relies on change of measure arguments that go back at least to Lai and Robbins 1985.
- Under δ-correct algorithm (Kauffman, Cappe, Garivier 2016), for

$$\mu = (\mu_1, \mu_2, \ldots, \mu_K) \in A_i$$

and

$$\nu = (\nu_1, \nu_2, \ldots, \nu_K) \in A_i^c$$

where each arm i is pulled N_i times,

- we have the 'distribution separation' inequality

$$\boxed{\sum_{i=1}^{K} E_\mu N_i \times KL(\mu_i || \nu_i) \geq \log\left(\frac{1}{\delta}\right)}$$

Rationale for the inequality

- If $\mathbf{X} = (X_{i,j} : i \leq K, j \leq N_j)$ denotes the adaptively generated samples by δ-correct algorithm,

Rationale for the inequality

- If $\mathbf{X} = (X_{i,j} : i \leq K, j \leq N_j)$ denotes the adaptively generated samples by δ-correct algorithm,

$$P_\mu(\mathbf{X} \to A_i) \geq 1 - \delta \quad \text{and, for } \nu \in A_i^c$$

$$P_\nu(\mathbf{X} \to A_i) = E_\mu \exp\left(- \sum_{a=1}^{K} \sum_{j=1}^{N_a} \log \frac{d\mu_a}{d\nu_a}(X_{a,j}) \right) I(\mathbf{X} \to A_i) \leq \delta$$

- This leads to the inequality

$$\boxed{\sum_{i=1}^{K} E_\mu N_i \times KL(\mu_i || \nu_i) \geq \log\left(\frac{1}{\delta}\right).}$$

Max-Min problem for lower bounds

- Lower bound $L(\mu) \times \log(1/\delta)$ on such algorithms, for $\mu \in A_i$,

$$\boxed{\min \sum_{i=1}^{K} t_i \quad \text{s.t.} \quad \inf_{\nu \in A_i^c} \sum_{i=1}^{K} t_i \times KL(\mu_i || \nu_i) \geq 1.}$$

Max-Min problem for lower bounds

- Lower bound $L(\mu) \times \log(1/\delta)$ on such algorithms, for $\mu \in A_i$,

$$\min \sum_{i=1}^{K} t_i \quad \text{s.t.} \quad \inf_{\nu \in A_i^c} \sum_{i=1}^{K} t_i \times KL(\mu_i || \nu_i) \geq 1.$$

Re-express constraint as $\displaystyle \sum_{j=1}^{K} t_j \times \inf_{\nu \in A_i^c} \sum_{i=1}^{K} \frac{t_i}{\sum_{j=1}^{K} t_j} \times KL(\mu_i || \nu_i) \geq 1$

Max-Min problem for lower bounds

- Lower bound $L(\mu) \times \log(1/\delta)$ on such algorithms, for $\mu \in A_i$,

$$\min \sum_{i=1}^{K} t_i \quad \text{s.t.} \quad \inf_{\nu \in A_i^c} \sum_{i=1}^{K} t_i \times KL(\mu_i || \nu_i) \geq 1.$$

Re-express constraint as $\displaystyle \sum_{j=1}^{K} t_j \times \inf_{\nu \in A_i^c} \sum_{i=1}^{K} \frac{t_i}{\sum_{j=1}^{K} t_j} \times KL(\mu_i || \nu_i) \geq 1$

- Equivalent Max-Min problem

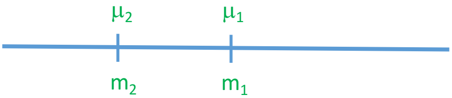
$$L(\mu)^{-1} = \max_{\sum_{i=1}^{K} w_i = 1, w_i \geq 0} \inf_{\nu \in A_i^c} \sum_{i=1}^{K} w_i \, KL(\mu_i || \nu_i)$$

Some restrictions necessary on distributions of underlying arms

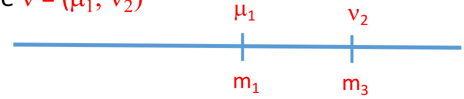
Restriction on underlying distributions for fast algorithms.

Selecting the best arm Glynn and J 2015

Consider instance $\mu = (\mu_1, \mu_2)$ with means (m_1, m_2) $m_1 > m_2$



and instance $\nu = (\mu_1, \nu_2)$



- Under δ-correct algorithm lower bound on expected number of samples given to arm 2 under P

$$E_\mu N_2 \times KL(\mu_2 || \nu_2) \geq \log\left(\frac{1}{\delta}\right).$$

- Under δ-correct algorithm lower bound on expected number of samples given to arm 2 under P

$$E_\mu N_2 \times KL(\mu_2 || \nu_2) \geq \log\left(\frac{1}{\delta}\right).$$

- Then,

$$\boxed{E_\mu N_2 \geq \frac{1}{\inf_{\nu_2 : m_3 > m_1} KL(\mu_2 || \nu_2)} \log\left(\frac{1}{\delta}\right)}$$

- Under δ-correct algorithm lower bound on expected number of samples given to arm 2 under P

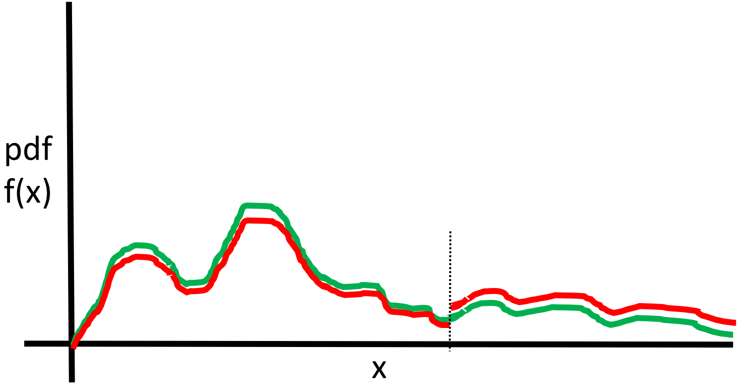
$$E_\mu N_2 \times KL(\mu_2 || \nu_2) \geq \log\left(\frac{1}{\delta}\right).$$

- Then,

$$\boxed{E_\mu N_2 \geq \frac{1}{\inf_{\nu_2 : m_3 > m_1} KL(\mu_2 || \nu_2)} \log\left(\frac{1}{\delta}\right)}$$

- Glynn and J. show that if distributions are unbounded, $KL(\mu_2 || \nu_2)$ can be made arbitrarily small, hence finite expected time algorithms not feasible without further restrictions

Two dist. - Mean arbitrarily far, KL arbitrarily close



We first restrict to one parameter exponential families

- Examples include Binomial, Poisson, Gaussian with known variance, Gamma distribution with known shape parameter.

We first restrict to one parameter exponential families

- Examples include Binomial, Poisson, Gaussian with known variance, Gamma distribution with known shape parameter.
- This allows us to think of Kullbach Leibler divergence as a function of the means of the distributions.

We first restrict to one parameter exponential families

- Examples include Binomial, Poisson, Gaussian with known variance, Gamma distribution with known shape parameter.
- This allows us to think of Kullbach Leibler divergence as a function of the means of the distributions.
- In the remaining talk, $\mathcal{Q}$ is a collection of vector of parameters in $\Re^K$.

Characterizing the solution to lower bound

Sets A and A^c are half spaces

A, a half-space

- Given

$$\mu \in A \triangleq \{\nu \in \Omega : \sum_{i=1}^{K} a_i \nu_i < b\}$$

what restrictions do $\nu \in A^c$ impose on $E_\mu N_a$ for each arm a

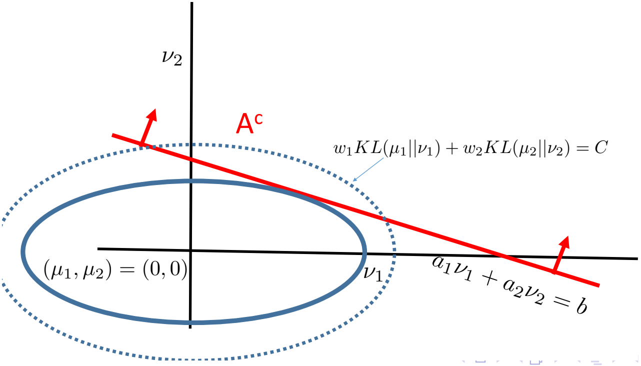
A geometric view when A is a half-space

Recall $\max_{w_1+w_2=1, w_i \geq 0} \inf_{\nu \in A^c} \left(w_1\, KL(\mu_1 || \nu_1) + w_2\, KL(\mu_2 || \nu_2) \right)$

A geometric view when A is a half-space

$$\text{Recall} \max_{w_1+w_2=1, w_i\geq 0} \inf_{\nu\in A^c} \left(w_1\, KL(\mu_1||\nu_1) + w_2\, KL(\mu_2||\nu_2)\right)$$

For inner optimization problem, look for smallest level set that intersects with A^c.



Solving the lower bound optimization problem

- Recall the lower bound problem

$$\boxed{\max_{\sum_{a=1}^{K} w_a=1, w_a\geq 0} \inf_{\nu \in A^c} \sum_{a=1}^{K} w_a \, KL(\mu_a || \nu_a)}$$

Solving the lower bound optimization problem

- Recall the lower bound problem

$$\boxed{\max_{\sum_{a=1}^K w_a=1, w_a\geq 0} \inf_{\nu\in A^c} \sum_{a=1}^K w_a\, KL(\mu_a||\nu_a)}$$

- Using minimax theorem

$$\inf_{\nu\in A^c} \max_{\sum_{a=1}^K w_a=1, w_a\geq 0} \sum_{a=1}^K w_a\, KL(\mu_a||\nu_a)$$

Solving the lower bound optimization problem

- Recall the lower bound problem

$$\boxed{\max_{\sum_{a=1}^K w_a=1, w_a\geq 0} \inf_{\nu\in A^c} \sum_{a=1}^K w_a \, KL(\mu_a||\nu_a)}$$

- Using minimax theorem

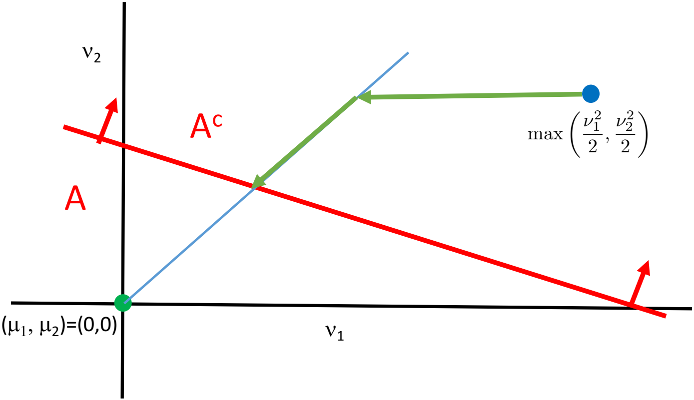
$$\inf_{\nu\in A^c} \max_{\sum_{a=1}^K w_a=1, w_a\geq 0} \sum_{a=1}^K w_a \, KL(\mu_a||\nu_a)$$

- This equals

$$\inf_{\nu\in A^c} \max_{a} KL(\mu_a||\nu_a).$$

Solving $\inf_{\nu \in A^c} \max_a KL(\mu_a || \nu_a)$

- Set $(\mu_1, \mu_2) = (0, 0)$.
- Gaussian distribution with variance 1, so $KL(\mu_i || \nu_i) = \nu_i^2/2$.



- More generally, the optimal (w^*, ν^*) corresponds to

$$KL(\mu_i || \nu_i^*) = KL(\mu_1 || \nu_1^*) \quad \forall i,$$

$$\sum_{i=1}^{K} a_i \nu_i^* = b.$$

- More generally, the optimal (w^*, ν^*) corresponds to

$$KL(\mu_i||\nu_i^*) = KL(\mu_1||\nu_1^*) \quad \forall i,$$

$$\sum_{i=1}^{K} a_i \nu_i^* = b.$$

- The slope matching condition

$$\frac{w_i^*}{a_i} KL'(\mu_i||\nu_i^*) = \frac{w_1^*}{a_1} KL'(\mu_1||\nu_1^*).$$

- More generally, the optimal (w^*, ν^*) corresponds to

$$KL(\mu_i||\nu_i^*) = KL(\mu_1||\nu_1^*) \quad \forall i,$$

$$\sum_{i=1}^{K} a_i \nu_i^* = b.$$

- The slope matching condition

$$\frac{w_i^*}{a_i} KL'(\mu_i||\nu_i^*) = \frac{w_1^*}{a_1} KL'(\mu_1||\nu_1^*).$$

- And lower bound on expected generated samples

$$KL(\mu_1||\nu_1^*)^{-1} \times \log(\frac{1}{\delta}).$$

δ-correct algorithm that matches lower bounds

The tracking algorithm

- Decisions: Which arm to sample and when to stop

The tracking algorithm

- Decisions: Which arm to sample and when to stop
- Closely follows Garivier and Kaufmann (2016) that was proposed in the best arm setting

The tracking algorithm

- Decisions: Which arm to sample and when to stop
- Closely follows Garivier and Kaufmann (2016) that was proposed in the best arm setting
- At sampling step n, ensure that at least about $\sqrt{n}$ samples allocated to each arm.

The tracking algorithm

- Decisions: Which arm to sample and when to stop
- Closely follows Garivier and Kaufmann (2016) that was proposed in the best arm setting
- At sampling step n, ensure that at least about $\sqrt{n}$ samples allocated to each arm.
- This ensures that with high probability $\hat{\mu}_n$ approximates μ and thus $w^*(\hat{\mu}_n)$ approximates $w^*(\mu)$.

The tracking algorithm

- Decisions: Which arm to sample and when to stop
- Closely follows Garivier and Kaufmann (2016) that was proposed in the best arm setting
- At sampling step n, ensure that at least about $\sqrt{n}$ samples allocated to each arm.
- This ensures that with high probability $\hat{\mu}_n$ approximates μ and thus $w^*(\hat{\mu}_n)$ approximates $w^*(\mu)$.
- Choose an arm to match proportions $w^*(\hat{\mu}_n)$.

Stopping rule motivated by Generalized Likelihood Ratio Method (Chernoff)

- After iteration n, suppose $\hat{\mu}(n) \in \tilde{A}$ (either A or A^c)
- Compute logarithm of

$$\frac{\max_{\mu \in \tilde{A}} \text{ Likelihood value } (\mu)}{\max_{\nu \in \tilde{A}^c} \text{ Likelihood value } (\nu)}.$$

- This equals

$$\inf_{\nu \in \tilde{A}^c} \sum_i \frac{N_i(n)}{n} \times KL(\hat{\mu}_n || \nu_i)$$

Stopping rule

- Let the separation function

$$\beta(n,\delta) = \log\left(\frac{cn}{\delta}\right)$$

for well chosen c.

Stopping rule

- Let the separation function

$$\beta(n,\delta) = \log\left(\frac{cn}{\delta}\right)$$

  for well chosen c.

- After iteration n, suppose $\hat{\mu}(n) \in \tilde{A}$ (either A or A^c)

- **If**

$$\inf_{\nu \in \tilde{A}^c} \sum_i \frac{N_i(n)}{n} \times KL(\hat{\mu}_n || \nu_i) \geq \frac{1}{n}\beta(n,\delta)$$

- **then** declare $\mu \in \tilde{A}$

Stopping rule

- Let the separation function

$$\beta(n,\delta) = \log\left(\frac{cn}{\delta}\right)$$

for well chosen c.

- After iteration n, suppose $\hat{\mu}(n) \in \tilde{A}$ (either A or A^c)

- **If**

$$\inf_{\nu \in \tilde{A}^c} \sum_i \frac{N_i(n)}{n} \times KL(\hat{\mu}_n || \nu_i) \geq \frac{1}{n}\beta(n,\delta)$$

- **then** declare $\mu \in \tilde{A}$

- Else, sample again

Result

Theorem

The algorithm is δ-correct. If $\tau(\delta)$ denotes the stopping time, then

$$\limsup_{\delta \to 0} \frac{E_\mu \tau(\delta)}{\log(1/\delta)} = KL(\mu_1 || \nu_1^*)^{-1}.$$

Characterizing the solution to lower bound when A or A^c is convex

When A^c is convex

- Recall the min-max lower bound problem

$$\max_{\sum_{a=1}^K w_a=1, w_a\geq 0} \inf_{\nu\in A^c} \sum_{a=1}^{K} w_a \, KL(\mu_a||\nu_a)$$

When A^c is convex

- Recall the min-max lower bound problem

$$\max_{\sum_{a=1}^{K} w_a=1, w_a \geq 0} \inf_{\nu \in A^c} \sum_{a=1}^{K} w_a \, KL(\mu_a || \nu_a)$$

- **Theorem:** Let (w^*, ν^*) denote an optimal solution.

When A^c is convex

- Recall the min-max lower bound problem

$$\max_{\sum_{a=1}^{K} w_a = 1, w_a \geq 0} \inf_{\nu \in A^c} \sum_{a=1}^{K} w_a \, KL(\mu_a || \nu_a)$$

- **Theorem:** Let (w^*, ν^*) denote an optimal solution.
  - ν^* is unique. It solves: $\min_{\nu \in A^c} \max_i K_i(\mu_i | \nu_i)$

When A^c is convex

- Recall the min-max lower bound problem

$$\max_{\sum_{a=1}^{K} w_a=1, w_a \geq 0} \inf_{\nu \in A^c} \sum_{a=1}^{K} w_a \, KL(\mu_a || \nu_a)$$

- **Theorem:** Let (w^*, ν^*) denote an optimal solution.
  - ν^* is unique. It solves: $\min_{\nu \in A^c} \max_i K_i(\mu_i | \nu_i)$
  - There exists a maximal $\mathcal{I} \subset \{1, 2, \ldots, K\}$ such that $w_i^* > 0$ for $i \in \mathcal{I}$, $w_i^* = 0$ for rest of i,

When A^c is convex

- Recall the min-max lower bound problem

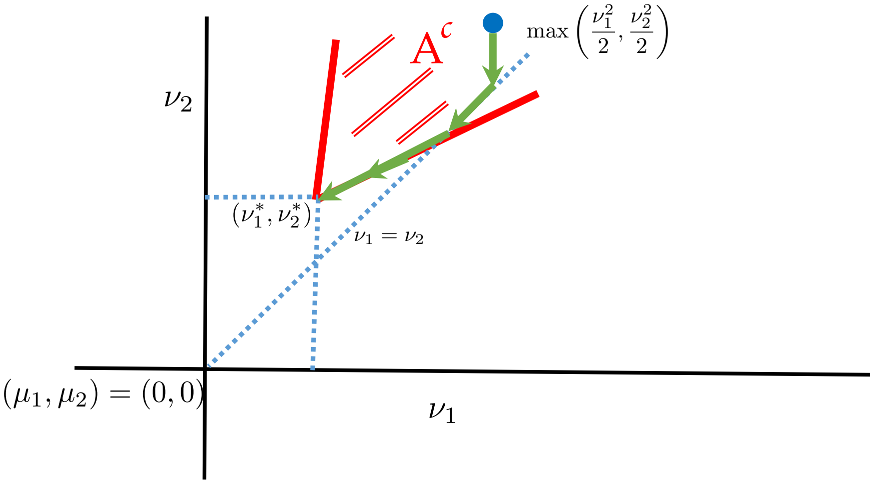
$$\max_{\sum_{a=1}^K w_a = 1, w_a \geq 0} \inf_{\nu \in A^c} \sum_{a=1}^{K} w_a \, KL(\mu_a || \nu_a)$$

- **Theorem:** Let (w^*, ν^*) denote an optimal solution.
  - ν^* is unique. It solves: $\min_{\nu \in A^c} \max_i K_i(\mu_i | \nu_i)$
  - There exists a maximal $\mathcal{I} \subset \{1, 2, \ldots, K\}$ such that $w_i^* > 0$ for $i \in \mathcal{I}$, $w_i^* = 0$ for rest of i,

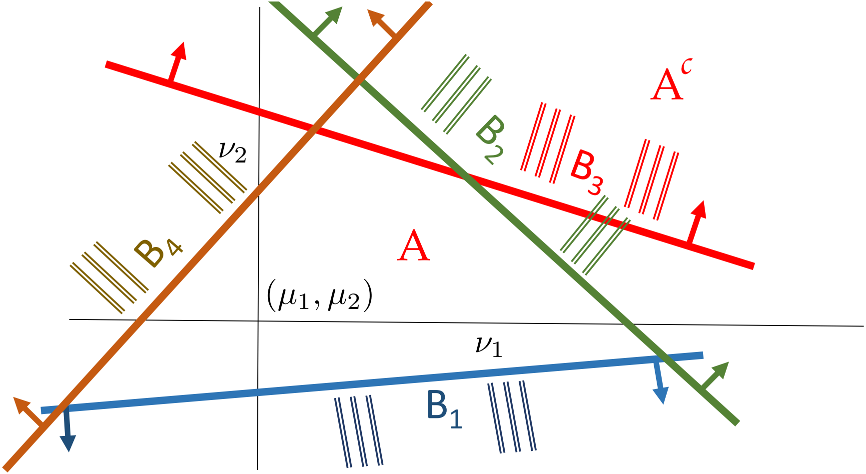
$$KL(\mu_i || \nu_i^*) = Const. \text{ for } i \in \mathcal{I},$$

$$KL(\mu_i | \nu_i^*) < Const \text{ for } i \in \mathcal{I}^c.$$

The algorithm and the optimal point



Algorithm when A^c is a union of half-spaces



A^c
B_2
B_3
ν_2
B_4
A
(μ_1, μ_2)
ν_1
B_1

Algorithm

- Recall the min-max lower bound problem

$$\max_{\sum_{a=1}^{K} w_a = 1, w_a \geq 0} \inf_{\nu \in A^c} \sum_{a=1}^{K} w_a \, KL(\mu_a || \nu_a)$$

Algorithm

- Recall the min-max lower bound problem

$$\max_{\sum_{a=1}^{K} w_a=1, w_a\geq 0} \inf_{\nu\in A^c} \sum_{a=1}^{K} w_a \, KL(\mu_a||\nu_a)$$

- For given $(w_a : a \leq K)$ solve

$$\inf_{\nu\in A^c} \sum_{a=1}^{K} w_a \, KL(\mu_a||\nu_a)$$

by finding the nearest hyperplane.

Algorithm

- Recall the min-max lower bound problem

$$\max_{\sum_{a=1}^{K} w_a=1, w_a \geq 0} \inf_{\nu \in A^c} \sum_{a=1}^{K} w_a \, KL(\mu_a || \nu_a)$$

- For given $(w_a : a \leq K)$ solve

$$\inf_{\nu \in A^c} \sum_{a=1}^{K} w_a \, KL(\mu_a || \nu_a)$$

  by finding the nearest hyperplane.

- Update $(w_a : a \leq K)$ using steepest descent. Repeat

General Distributions

$$\mathcal{L} = \{\eta \in \mathcal{P}(\Re) : \mathbb{E}_{X\sim\eta}(f(|X|)) \leq B\}$$

Recall lower bound Prob $\max_{w\in\Sigma_K} \inf_{\nu\in A^c} \sum_{a=1}^K w_a\ KL(\mu_a||\nu_a)$

- We consider sets A of the form

$$\{\nu = (\nu_1,\ldots,\nu_K) : \nu_i \in \mathcal{L}, (m(\nu_1), m(\nu_2),\ldots, m(\nu_K)) \in B \subset \Re^K\}$$

where $m(\nu_i)$ denotes the mean under ν_i.

- Max-min problem may be re-expressed as

$$\max_{w\in\Sigma_K} \inf_{x\in B^c} \sum_{i=1}^{K} w_i KL_{inf}(\mu_i, x_i).$$

- where, for $x \in \Re$, $KL_{inf}(\eta, x)$ is defined as the solution to

$$\min_{\kappa\in\mathcal{L}} KL(\eta,\kappa); \quad \text{s.t.} \int_{y\in\Re} y d\kappa(y) = x$$

$KL_{inf}(\eta, x)$

- Similar definition $KL_{inf}(\eta, x)$ for bounded random variables was used by Burnetas and Katehakis 1996, and Honda and Takemura 2015 in regret minimization setting

$KL_{inf}(\eta, x)$

- Similar definition $KL_{inf}(\eta, x)$ for bounded random variables was used by Burnetas and Katehakis 1996, and Honda and Takemura 2015 in regret minimization setting
- *Result:* $KL_{inf}(\eta, x)$ is convex and twice differentiable in x and continuous in η in the *Wasserstein distance*

$KL_{inf}(\eta, x)$

- Similar definition $KL_{inf}(\eta, x)$ for bounded random variables was used by Burnetas and Katehakis 1996, and Honda and Takemura 2015 in regret minimization setting

- *Result:* $KL_{inf}(\eta, x)$ is convex and twice differentiable in x and continuous in η in the *Wasserstein distance*

- **Our analysis of lower bounds is identical to earlier one with $KL_{inf}(\mu_i, x)$ replacing the earlier $KL(\mu_i|\nu_i)$ in SPEF settings.**

Dual characterization of $\mathbf{KL_{inf}}(\eta, \mathbf{x})$

$$\boxed{KL_{inf}(\eta, x) = \max_{\lambda_1, \lambda_2 \in \mathcal{S}} \mathbb{E}_\eta \left(\log\left(1 - (X - x)\lambda_1 - (B - f(X))\lambda_2\right)\right),}$$

where $\mathcal{S} = \left[0, \left(f^{-1}(B) - x\right)^{-1}\right] \times \left[0, (B - f(x))^{-1}\right]$

New concentration inequalities needed in our analysis

- $\hat{\kappa}(n)$ denotes the empirical distribution corresponding to n samples from distribution κ

New concentration inequalities needed in our analysis

- $\hat{\kappa}(n)$ denotes the empirical distribution corresponding to n samples from distribution κ
- $m(\kappa)$ denotes the mean under κ, $u \geq 0$, and a constant D depends on κ

New concentration inequalities needed in our analysis

- $\hat{\kappa}(n)$ denotes the empirical distribution corresponding to n samples from distribution κ
- $m(\kappa)$ denotes the mean under κ, $u \geq 0$, and a constant D depends on κ
- Then,

$$\mathbb{P}\left(KL_{inf}(\hat{\kappa}(n), m(\kappa)) \geq u\right) \leq 4D(1+n)^2 e^{-nu}$$

New concentration inequalities needed in our analysis

- $\hat{\kappa}(n)$ denotes the empirical distribution corresponding to n samples from distribution κ
- $m(\kappa)$ denotes the mean under κ, $u \geq 0$, and a constant D depends on κ
- Then,

$$\boxed{\mathbb{P}\left(KL_{inf}(\hat{\kappa}(n), m(\kappa)) \geq u\right) \leq 4D(1+n)^2 e^{-nu}}$$

- Furthermore, for $n \in \mathbb{N}$, and $\Gamma > K+1$, and a constant $\hat{D}$,

$$\mathbb{P}\left(\sum_a N_a(n)\, KL_{inf}(\hat{\mu}_a(N_a(n)), m(\mu_a)) \geq \Gamma\right)$$
$$\leq \quad e^{K+1}\hat{D}\left(\frac{4(1+n)^2\Gamma^2 \log n)}{K}\right)^K e^{-\Gamma}$$

Conclusions

- We consider pure exploration multi-armed bandit problems in a general unifying framework. This involves finding the correct partition for a vector of distributions from a superset of such collections.

Conclusions

- We consider pure exploration multi-armed bandit problems in a general unifying framework. This involves finding the correct partition for a vector of distributions from a superset of such collections.

- We developed geometric and duality based methodology for computing lower bounds on computational effort for δ - correct algorithms.

Conclusions

- We consider pure exploration multi-armed bandit problems in a general unifying framework. This involves finding the correct partition for a vector of distributions from a superset of such collections.

- We developed geometric and duality based methodology for computing lower bounds on computational effort for δ - correct algorithms.

- Using these, we develop algorithms with matching computational bounds.

Conclusions

- We consider pure exploration multi-armed bandit problems in a general unifying framework. This involves finding the correct partition for a vector of distributions from a superset of such collections.
- We developed geometric and duality based methodology for computing lower bounds on computational effort for δ - correct algorithms.
- Using these, we develop algorithms with matching computational bounds.
- Extended to general distributions