ISS Seminar, Princeton University

16 April 2009

Fundamental Limits of Distributed Estimation and Multiuser Information-Theoretic Games

Mokshay Madiman

Yale University

Includes joint work with: Andrew Barron (Yale),
Abram Kagan (Univ. of Maryland)
and Tinghui Yu (Univ. of Maryland)

Outline

- Distributed estimation of background: motivation and results
- Fundamental limits for more general distributed estimation settings
- Fisher information inequalities and the CLT
- Cooperative Game Theory: background
- The Slepian-Wolf and Gaussian MAC Games
- Resource allocation issues for distributed inference
- Concluding remarks

Motivating Example: Cosmic Microwave Background Radiation

What and Why?

- Radiation left over from the very hot plasma that pervaded the universe soon after the Big Bang– its precise measurement can help to test various key hypotheses about cosmological constants and origins
- Very important problem in astronomy and cosmology
- Has a thermal black body spectrum at a temperature of approx. 2.725 K

How to measure?

- A collection of detectors or sensors focused on different, possibly overlapping regions of the night sky
- Each sensor picks up not only the CMBR (and measurement noise), but also the effects of other sources of similar radiation in the universe, such as intergalactic clouds, whose distribution in different directions is uneven
- May be thought of as "distributed estimation of a background parameter" in the presence of a field of "noise" sources

The typical picture

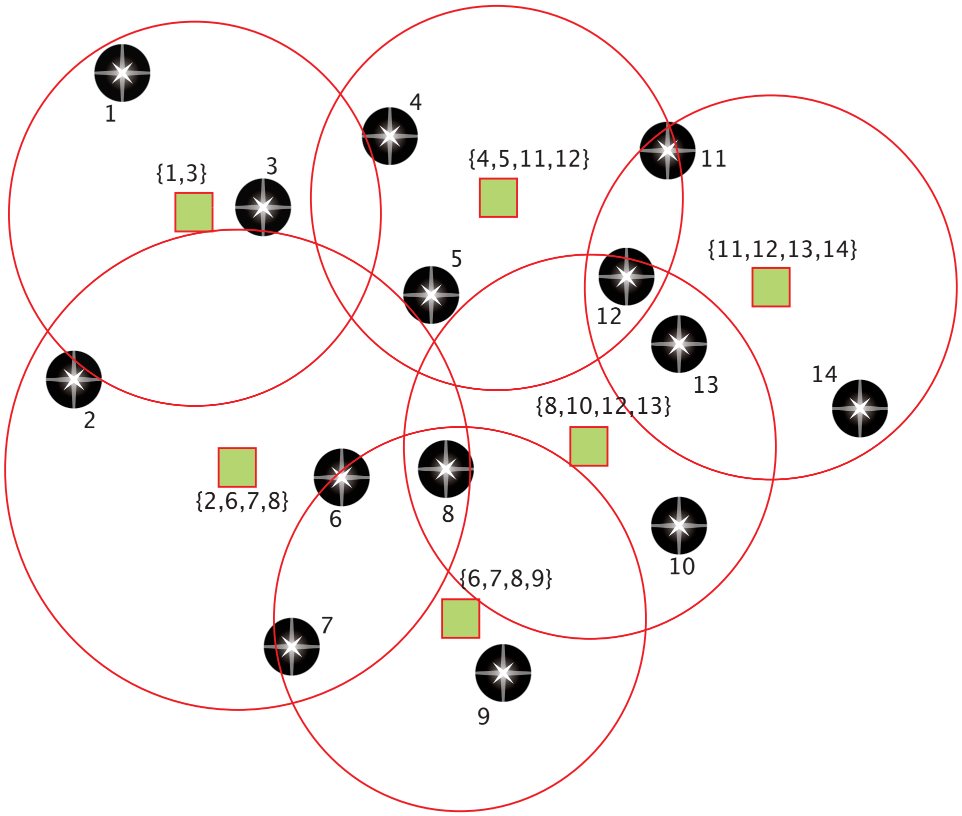

Figure 1: Sources are labelled $1, 2, \ldots, n$, and sensors are labelled by subsets of $[n]$

Model for Distributed Estimation of Background

The Model

- **Source**: a star, star cluster or intergalactic cloud that produces radiation (data)
- **A field of n sources**: the i-th source produces the data stream $(X_{i,1}, \ldots, X_{i,T})$, of size T, independently of all other sources
- **A network of sensors**: each sensor corresponds to a subset $s \in \mathcal{C}$, where $\mathcal{C}$ is a collection of subsets of $[n] = \{1, 2, \ldots, n\}$
- **Observations by sensors**: the s-sensor sees the superposition of the background and the data from all the sources in s

The Issues

- Statistical issues– note *finite* samples (focus)
- Communication and computation issues (ignore)
- Local and Global objectives (partially ignore the latter)

A Fundamental Question

A Simplified Model, more precisely

- For each i, $(X_{i,1}, \ldots, X_{i,T})$ comes from a known distribution
- At time t, the s-sensor observes $Y_{s,t} = \theta + \sum_{i \in s} X_{i,t}$, where θ is an unknown number
- From these T observations, the s-sensor constructs an estimate $\hat{\theta}_s(Y_{s,1}, \ldots, Y_{s,T})$ of the location parameter θ

The Question

Consider the minimax mean square risk associated with the s-sensor

$$R_T(s) = \min_{\text{all estimators } \tilde{\theta}_s} \max_{\theta} E[(\tilde{\theta}_s - \theta)^2]$$

How are the minimax risks achievable by the s-sensors (for different s) related to each other?

An inequality for minimax risks

Theorem 1: [Relating minimax risks]

Let $\mathcal{C}$ be a collection of subsets of $[n]$, such that each source i appears in exactly r subsets. Then, for any sample size $T \geq 1$,

$$R_T([n]) \geq \frac{1}{r} \sum_{s \in \mathcal{C}} R_T(s)$$

Remarks

- Compares two sensor configurations: (i) a single $[n]$-sensor, who only sees observations from the total sum, and (ii) a collection of sensors corresponding to $s \in \mathcal{C}$
- Since "data aggregation reduces information", $R_T([n]) \geq R_T(s)$ is obvious, but Theorem 1 goes much deeper
- For the collection $\mathcal{C}_1$ of singleton sets,

$$R_T([n]) \geq \sum_{i \in [n]} R_T(\{i\}) \quad \text{[Kagan '02]}$$

- For the collection $\mathcal{C}_{n-1}$ of leave-one-out sets, $r = n - 1$

Hierarchy of average minimax risks

Theorem 2: [Hierarchy for symmetric collections]

For the collection $\mathcal{C}_k$ of all subsets of size k, let

$$A_k = \frac{1}{\binom{n}{k}} \sum_{s \in \mathcal{C}_k} \frac{R_T(s)}{k}$$

be the average minimax risk per element observed. Then

$$A_1 \leq A_2 \leq \ldots \leq A_{n-1} \leq A_n$$

Remarks

- Using n sensitive sensors picking up the individual sources is better than using $\binom{n}{2}$ rough sensors picking up all pairwise sums
- To our knowledge, this is the first such rigorous result in sensor network theory based on fundamental limits for estimation from *finite* samples

Where we are

- Distributed estimation of background: motivation and results
- Fundamental limits for more general distributed estimation settings
- Fisher information inequalities and the CLT
- Cooperative Game Theory: background
- The Slepian-Wolf and Gaussian MAC Games
- Resource allocation issues for distributed inference
- Concluding remarks

General Setting for Inference from Sensor Network

Why?

- Measuring temperature, chemical concentration, or other variables that are geographically distributed
- Cheap, small sensors can be produced en masse today, but constrained by power, communication & computation ability, sensitivity etc.
- Numerous important applications

Modelling sensor networks

- A field of n "sources", that independently produce data streams of size T (lets say the i-th source produces $(X_{i,1}, \ldots, X_{i,T})$)
- A network of sensors, each corresponding to a subset $s \in \mathcal{C}$, where $\mathcal{C}$ is a collection of subsets of $[n] = \{1, 2, \ldots, n\}$
- The s-sensor has access to a *combination* of the data coming from the sources in s

Three observation models for distributed estimation

Model I: Distributed estimation of background

- For each i, $(X_{i,1}, \ldots, X_{i,T})$ comes from a known distribution
- At time t, the s-sensor observes $Y_{s,t} = \theta + \sum_{i \in s} X_{i,t}$, where θ is an unknown number

Model II: Distributed estimation from sums

- For each i, $(X_{i,1}, \ldots, X_{i,T})$ comes from $f_i(\cdot - \theta 1_T)$, where f_i is known
- At time t, the s-sensor observes $Y_{s,t} = \sum_{i \in s} X_{i,t}$

Model III: Distributed estimation using discriminating sensors

- For each i, $(X_{i,1}, \ldots, X_{i,T_i})$ comes from $f_i(\cdot - \theta 1_{T_i})$, where f_i is known (asynchronous signals allowed)
- The s-sensor observes $(X_{i,t} : t \in T_i, i \in s)$

In all the models, the parameter θ to be estimated is common to all sources, and the sources are independent of each other

Distributed estimation using discriminating sensors

- Independent sources, with i-th source producing $(X_{i,1}, \ldots, X_{i,T_i})$ from $f_i(\cdot - \theta 1_{T_i})$, where f_i is known (asynchronous signals allowed)
- The s-sensor observes the "concatenation" $(X_{i,t} : t \in T_i, i \in s)$
- From these observations, the s-sensor constructs an estimate $\hat{\theta}_s$ of the location parameter θ (which is common to all sources)

The Question

Consider the minimax mean square risk associated with the s-sensor

$$\bar{R}(s) = \min_{\text{all estimators } \tilde{\theta}_s} \max_{\theta} E[(\tilde{\theta}_s - \theta)^2]$$

How are the minimax risks achievable by the s-sensors related?

Theorem 3: [Relating minimax risks]

If the minimax mean square risk associated with the s-sensor is $\bar{R}(s)$, and if each source i appears in exactly r subsets in $\mathcal{C}$, then

$$\frac{1}{\bar{R}([n])} \geq \frac{1}{r} \sum_{s \in \mathcal{C}} \frac{1}{\bar{R}(s)}$$

Remarks on Proofs

Proof ingredients

- The Pitman estimator is defined as the minimum risk estimator (with respect to squared error loss) among all location-equivariant estimators
- It can be written as

$$\hat{\theta}_s(\mathbf{Y}_s) = \bar{Y}_s - E_{\theta=0}(\bar{Y}_s \,|Y_{s,1} - \bar{Y}_s \,, \ldots, \, Y_{s,T} - \bar{Y}_s)$$

- The Pitman estimator is minimax
- One needs the "variance drop lemma" of [M.M.–Barron '07], which relates variance of a sum of functions of subsets of variables to the variances of the functions themselves

Where we are

- Distributed estimation of background: motivation and results
- Fundamental limits for more general distributed estimation settings
- Fisher information inequalities and the CLT
- Cooperative Game Theory: background
- The Slepian-Wolf and Gaussian MAC Games
- Resource allocation issues for distributed inference
- Concluding remarks

The Entropic Central Limit Theorem

If X has density $f(x)$, the Shannon entropy is $H(X) = E[-\log f(X)]$

Entropic CLT

- Gaussian is MaxEnt: $N(0, \sigma^2)$ has maximum entropy among all densities on $\mathbb{R}$ with variance $\leq \sigma^2$
- Entropic Convergence: Let X_i, drawn i.i.d. from density f, have zero mean and variance σ^2. If $S_T := \frac{1}{\sqrt{T}} \sum_{i=1}^{T} X_i$, then

$$D(S_T \| N(0, \sigma^2)) \downarrow 0 \qquad \text{or equivalently,} \qquad H(S_T) \uparrow H(N(0, \sigma^2))$$

Remarks

- Monotonicity in T indicates that the entropy is a *natural measure* for CLT convergence (cf. second law of thermodynamics)
- This powerful intuition comes with powerful techniques; key insight is to use Fisher information or MMSE as an intermediary to entropy [Barron '86, Artstein-Ball-Barthe-Naor '04, Johnson-Barron '04, Guo-Shamai-Verdú '05-'07 . . .]

Fisher Information and CLT

Fisher Information

- The Fisher information of a r.v. X with differentiable density $f(x)$ is

$$I(X) = E\left[\frac{d}{dx}\log f(X)\right]^2$$

- Under the variance constraint $\mathrm{var}(X) = \sigma^2$, the normal has minimum Fisher information $\frac{1}{\sigma^2}$

CLT via Fisher information

For i.i.d. X_i with zero mean and finite variance σ^2, let $S_T = \frac{1}{\sqrt{T}}\sum_{i=1}^{T} X_i$

- Monotonicity: If X_i have differentiable densities,

  $I(S_T) \leq I(S_{T-1})$ [Artstein et al. '04, Tulino–Verdú '06, M.M.–Barron '06]

- Convergence: $I(S_T) \to I(N(0,\sigma^2))$ [Johnson & Barron '04]

A Fisher Information Inequality

Theorem 4: [Fisher Information Inequality]

For independent X_1, $X_2, \ldots, X_n$ with differentiable densities,

$$\frac{1}{I(\text{sum}_{[n]})} \geq \frac{1}{r} \sum_{s \in \mathcal{C}} \frac{1}{I(\text{sum}_s)}$$ [M.M. & Barron '07]

Remarks

- Special case of $\mathcal{C}_1$:

  $$\frac{1}{I(X_1 + X_2)} \geq \frac{1}{I(X_1)} + \frac{1}{I(X_2)}$$ [Stam '59]

  Implies CLT monotonicity for doubling of sample size, $I(S_{2n}) \leq I(S_n)$

- Special case of $\mathcal{C}_{n-1}$:

  $$\frac{1}{I(\text{sum}_{[n]})} \geq \frac{1}{n-1} \sum_{s \in \mathcal{C}_{n-1}} \frac{1}{I(\text{sum}_s)}$$ [Artstein et al. '04]

  Implies CLT monotonicity at every step, $I(S_n) \leq I(S_{n-1})$

- We give a new proof of Theorem 4 as an application of Theorem 1

Proof of Theorem 4

Theorem 4: [FISHER INFORMATION INEQUALITY]

For independent $X_1, X_2, \ldots, X_n$ with differentiable densities,

$$\frac{1}{I(\mathsf{sum}_{[n]})} \geq \frac{1}{r} \sum_{s \in \mathcal{C}} \frac{1}{I(\mathsf{sum}_s)}$$

[M.M. & Barron '07]

New Proof [M.M., Barron, Kagan, Yu '08]

By standard asymptotic theory from statistics,

$$\lim_{T \to \infty} T\, R_T(s) = \frac{1}{I(\sum_{i \in s} X_i)} = \frac{1}{I(\mathsf{sum}_s)}$$

for each $s \subset [n]$

Thus letting $T \to \infty$ in

$$R_T([n]) \geq \frac{1}{r} \sum_{s \in \mathcal{C}} R_T(s)$$

proves Theorem 4

Where we are

- Distributed estimation of background: motivation and results
- Fundamental limits for more general distributed estimation settings
- Fisher information inequalities and the CLT
- Cooperative Game Theory: background
- The Slepian-Wolf and Gaussian MAC Games
- Resource allocation issues for distributed inference
- Concluding remarks

The Slepian-Wolf rate region

The distributed data compression problem

Scenario: n sources produce data $(X_1, X_2, \ldots, X_n)$ that have some known joint distribution (say, i.i.d. copies at each time)
Goal: Want to optimally compress data, i.e., if R_i is the number of bits per unit time used to encode X_i, want to minimize the sum rate $\sum_{i\in[n]} R_i$
Constraint: Source i can only use its own data in its encoding procedure, without access to the data produced by the other sources
First Question: Can one characterize the region of achievable rate vectors $(R_1, \ldots, R_n)$?

The Slepian-Wolf solution

A vector of rates $(R_1, \ldots, R_n)$ is achievable in a distributed fashion if and only if

$$\sum_{i\in s} R_i \geq H(X_s|X_{s^c}) \quad \text{for all } s \subset \{1, \ldots, n\},$$

where X_s denotes $(X_i : i \in s)$

Why is there no loss of sum rate?

The natural question

The optimal compression rate for a centralized compressor is $H(X_1, \ldots, X_n)$ bits per symbol. For the distributed compression scenario, can we achieve a sum rate equal to this, or do we need to take a hit because of the distributed nature of the compression?

The usual approach

- from the polymatroidal property of entropy, or
- by constructing an explicit rate point with optimal sum rate and showing that it satisfies all necessary constraints

The "right" answer: dualities

Information inequalities $<-->$ Properties of rate regions

Illuminated by cooperative game theory

Cooperative Games

A cooperative game is specified by $([n], v)$, where:

- Set of n players: $[n] = \{1, 2, \dots, n\}$
- Players can form arbitrary coalitions $s \subset [n]$
- Values $v(s)$ of coalitions s are defined by $v : 2^{[n]} \to \mathbb{R}$, where $v(\phi) = 0$

Cost allocation interpretation

– If player i contributes t_i, and value is transferable, the cumulative cost for the coalition s is $\sum_{i \in s} t_i$

– Since each coalition must pay its due, the set of *aspirations* of the game (cost allocations that the players can aspire to) is

$$A(v) = \left\{ t \in \mathbb{R}^n : \sum_{i \in s} t_i \geq v(s) \text{ for each } s \subset [n] \right\}$$

Goal of the game: Minimize social cost, i.e., the total sum $\sum_{i \in [n]} t_i$.

Core

The *core* of a game v is the set of aspiration vectors $t \in \mathbb{R}^n$ such that $\sum_{i \in [n]} t_i = v([n])$.

Game Theory Review - I

Balanced Games

Given a collection $\mathcal{C}$ of subsets of $[n]$, a function $\alpha : \mathcal{C} \rightarrow \mathbb{R}_+$ is a *fractional partition using* $\mathcal{C}$ if for each $i \in [n]$, we have $\sum_{s \in \mathcal{C} : i \in S} \alpha(s) = 1$.

A game is $balanced$ if

$$v([n]) \geq \sum_{s \in \mathcal{C}} \alpha(s) v(s) \tag{1}$$

for any fractional partition α using any collection $\mathcal{C}$

The Bondareva-Shapley theorem

The core of a game is non-empty iff the game is balanced

Idea of proof: LP duality

Game Theory Review - II

Convex Games

A game is $convex$ if

$$v(s \cup t) + v(s \cap t) \geq v(s) + v(t)$$

for any sets s and t (equivalently: v is "supermodular")

Convex games are nice

- A convex game has a non-empty core
- A convex game has a "large" core: If $\sum_{i \in s} y_i \geq v(s)$ for each s, there exists x in the core such that $x \leq y$ (component-wise)
- There exists a unique point $\phi[v]$ (the Shapley value) in the core of any convex game v satisfying the following axioms:
  1. it is invariant under permutation of players,
  2. if u and v are two games, then $\phi[u + v] = \phi[u] + \phi[v]$

  This formalizes the notion of a "fair allocation" of cost to the players
- Convex games are closely related to contra-polymatroids

Where we are

- Distributed estimation of background: motivation and results
- Fundamental limits for more general distributed estimation settings
- Fisher information inequalities and the CLT
- Cooperative Game Theory: background
- The Slepian-Wolf and Gaussian MAC Games
- Resource allocation issues for distributed inference
- Concluding remarks

The Slepian-Wolf Game - I

The S-W rate region is the set of aspirations of the cooperative game

$$v_{SW}(s) = H(X_s | X_{s^c})$$

Remarks

- The S-W game is balanced:

  $$\sum_{S \in \mathcal{C}} \beta(s) H(X_s | X_{s^c}) \leq H(X_{[n]}) \quad \text{[M.M. \& Tetali '07]}$$

  Thus the core is non-empty and a sum rate of $H(X_{[n]})$ is achievable
- In fact, the S-W game is a convex game. Supermodularity of v_{SW} pointed out in [Fujishige '78]

The Slepian-Wolf Game - II

Robustness to certain failures

Suppose the users can only drop out in a certain order (w.l.o.g., assume that the first user to potentially drop out would be user n, followed by user $n-1$, etc.). Then there exists a rate point for S-W coding which is feasible and optimal irrespective of the number of users that have dropped out.

Remarks

- Solution related to a modified S-W game
$$\bar{v}_{SW}(s) = H(X_s | X_{s^c \setminus >s})$$
where $>s = \{i \in [n] : i > j \text{ for every } j \in s\}$, since core of this game $=$ set of rate points that are simultaneously in the S-W rate region of every $[k]$, for $k \in [n]$.
- [M.M. & Tetali '07] showed that the game $\bar{v}_{SW}$ is balanced
- The robust rate point is an extreme point of the S-W rate region; the importance of these is well known (see, e.g., [Rimoldi-Urbanke '96], [Coleman et. al. '06] and [Cristescu et al. '05])

The Basic Gaussian MAC Game

The Gaussian multiple access channel (g-MAC) produces

$$Y = \sum_{i \in [n]} X_i + Z$$

at the receiver, where X_i are independent sources (whose powers are constrained to P_i), and $Z \sim N(0, N)$ is independent of the sources

Fact: The capacity region of the n-user g-MAC is the aspiration set of the "g-MAC" game

$$v_g(s) = C\left(\frac{\sum_{i \in s} P_i}{N}\right)$$

where $C(x) = \frac{1}{2}\log(1 + x)$

Remarks

- [Han '78] showed that the g-MAC game is a concave game. In particular, its core is non-empty, and a sum capacity of $C\left(\frac{\sum_{i \in [n]} P_i}{N}\right)$ is achievable
- Can interpret as for S-W game: robustness to senders dropping out; Shapley value; large core

The La-Anantharam Gaussian MAC Game

Consider an arbitrarily varying Gaussian multiple access channel, where the users are potentially hostile and aware of each others' codebooks

Goal: Quantify the capacities achievable by a coalition s even when the users in s^c coherently combine to jam the channel

Solution: The capacity region of this channel is the aspiration set of the "La-Anantharam game":

$$v_{LA}(s) = C\left(\frac{P_{\hat{s}}}{\Lambda_{s^c} + N}\right)$$

where $P_s = \sum_{i \in s} P_i$, $\Lambda_s = [\sum_{i \in s} \sqrt{P_i}]^2$, and $\hat{s} = \{i \in s : P_i \geq \Lambda_{s^c}\}$

Remarks

The La-Anantharam game is not a concave game, but it has a non-empty core. In particular, a sum capacity of $C\left(\frac{\sum_{i \in [n]} P_i}{N}\right)$ is achievable. [La & Anantharam '02]

Examples of information-theoretic games

- Discussed: Slepian-Wolf game, Gaussian MAC games
- Other classes of memoryless multiple access channels (Poisson, discrete, fading, etc.)
- Potentially more delicate: multiuser channels with memory, feedback, etc.
- Robust hypothesis testing (a la [Huber & Strassen '73]) and variants, including work on robust rate distortion and channel coding by [Poor '82] and [Geraniotis '85-'86]
- A new example: distributed estimation games

See [M.M., "Cores of Cooperative Games in Information Theory", EURASIP J. on Wireless Comm. and Networking (2008)]

Distributed estimation of a background (recall)

Model I

- For each i, $(X_{i,1}, \ldots, X_{i,T})$ comes from a known distribution
- At time t, the s-sensor observes $Y_{s,t} = \theta + \sum_{i \in s} X_{i,t}$, where θ is an unknown number
- From these T observations, the s-sensor constructs an estimate $\hat{\theta}_s(Y_{s,1}, \ldots, Y_{s,T})$ of the location parameter θ

Theorem 1': [Relating minimax risks]

Consider the minimax mean square risk associated with the s-sensor

$$R_T(s) = \min_{\text{all estimators } \tilde{\theta}_s} \max_{\theta} E[(\tilde{\theta}_s - \theta)^2]$$

Let α be a fractional partition using $\mathcal{C}$. Then, for any sample size $T \geq 1$,

$$R_T([n]) \geq \sum_{s \in \mathcal{C}} \alpha(s) R_T(s)$$

Resource Allocation Game for estimation of a background

Variance permissions

- Suppose we can give *variance permissions* V_i for each source, i.e., the s-sensor is only allowed an estimator with mean squared error $\leq \sum_{i \in s} V_i$
- For an arbitrary sensor configuration to be feasible, need

$$\sum_{i \in s} V_i \geq R_T(s) \qquad \text{for each } s \subset [n]$$

Theorem 5: [Resource Allocation Game for Model I]

It is possible to allot variance permissions to all sources in such a way that there is no wasted total variance, i.e., $\sum_{i \in [n]} V_i = R_T([n])$

Proof:

Theorem 1' showed that the game with value function $R_T(s)$ is balanced; hence its core is non-empty

Distributed estimation using discriminating sensors (recall)

Model III

- Independent sources, with i-th source producing $(X_{i,1}, \ldots, X_{i,T_i})$ from $f_i(\cdot - \theta 1_{T_i})$, where f_i is known (asynchronous signals allowed)
- The s-sensor observes the "concatenation" $(X_{i,t} : t \in T_i, i \in s)$
- From these observations, the s-sensor constructs an estimate $\hat{\theta}_s$ of the location parameter θ (which is common to all sources)

Theorem 3': [Relating minimax risks]

Consider the minimax mean square risk associated with the s-sensor

$$\bar{R}(s) = \min_{\text{all estimators } \tilde{\theta}_s} \max_{\theta} E[(\tilde{\theta}_s - \theta)^2]$$

If α is a fractional partition using $\mathcal{C}$, then

$$\frac{1}{\bar{R}([n])} \geq \sum_{s \in \mathcal{C}} \frac{\alpha(s)}{\bar{R}(s)}$$

The discriminating sensors game

Data pricing

- The value of the data bundle produced by the sources in s is $\frac{1}{\bar{R}(s)}$; thus the easier it is to estimate θ from the bundle, the more valuable it is
- Thus any coalition s of sources can distribute the amount it collects (upto $\frac{1}{\bar{R}(s)}$ if acting separately from the sources in s^c) among its members; if P_i is the payoff for the i-th source, $\sum_{i \in s} P_i$ cannot exceed what the coalition collects

Theorem 6: [Resource Allocation Game for Model III]

It is always advantageous to the data owners (sources) to form a grand coalition and sell the data together. In other words, there always exists payoffs P_i such that $\sum_{i \in [n]} P_i = \frac{1}{\bar{R}([n])}$ and

$$\sum_{i \in s} P_i \geq \frac{1}{\bar{R}(s)} \qquad \text{for each } s \subset [n]$$

Proof: Theorem 3' showed that the game with value function $v(s) = \frac{1}{\bar{R}(s)}$ is balanced; hence its core is non-empty

Summary

- First decision-theoretic analysis of a sensor network model based on finite samples
- Applications include comparing network configurations and solving a risk allocation problem for a toy model
- New statistical proof of Fisher information inequalities via a finite-sample generalization (and thereby of monotonicity in Information-based CLT's)
- Perspectives from cooperative game theory shed light on fundamental aspects of multiuser scenarios (in data compression, communication, and statistics)

Thank you!

○ — ○ — ○

References

- M.M.–Barron–Kagan–Yu, "Minimax risks for distributed estimation of the background in a field of noise sources", WITS 2008
- M.M.–Barron–Kagan–Yu, "A model for pricing data bundles based on minimax risks for estimation of a location parameter", ITW 2009
- M.M.–Barron, "Generalized Entropy Power Inequalities and Monotonicity Properties of Information", IEEE Trans. Inform. Theory (2007)
- M.M.–Tetali, "Information Inequalities for Joint Distributions, with Interpretations and Applications", IEEE Trans. Inform. Theory (2009, to appear)
- M.M., "Cores of Cooperative Games in Information Theory", EURASIP J. on Wireless Comm. and Networking (2008)

○ — ○ — ○