Constructive Algorithms for Discrepancy Minimization (Nikhil Bansal, FOCS 2010)

Pravesh Kothari

The University of Texas at Austin
kothari@cs.utexas.edu

January 30, 2012

Discrepancy

- *" Approximating a probability space by a smaller probability space"*

Discrepancy

- *" Approximating a probability space by a smaller probability space"*
- Example: Approximating a Continuous Probability Space by A Discrete Space
- Question: How uniformly can you distribute n-points in a unit square?
- What does UNIFORM mean?

Axis Parallel Rectangles in a Grid

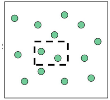
- Question: How uniformly can you distribute n-points in a unit square?
- What does UNIFORM mean?
- 

Figure: Distributing Points in a Unit Square

For any axis parallel rectangle R inside the unit square,

$$|Prob_{x \in [0,1]^2}[x \in R] - Prob_{y \in S}[y \in R]|$$

is small.

- $|(\#\text{of points from S in R}) - n \cdot Area(R)|$ is small.

Discrepancy: Example 2

- "Approximating a large discrete space by a smaller discrete space."
- **Given:** A set X of n points in a unit square.
- **Problem:** Color each point in X red/blue such that each *axis parallel rectangle* has "roughly" the same number of red or blue points.
- Discrepancy: max over Rectangle R

$$((\# \text{ of red points in } R) - (\# \text{ of blue points in} R))$$

Discrepancy: Example 2

- "Approximating a large discrete space by a smaller discrete space."
- **Given:** A set X of n points in a unit square.
- **Problem:** Color each point in X red/blue such that each *axis parallel rectangle* has "roughly" the same number of red or blue points.
- Discrepancy: max over Rectangle R

$$((\# \text{ of red points in } R) - (\# \text{ of blue points in} R))$$

- Non Trivial:Random Coloring: $O(\sqrt{n \log n})$
  Best:

Discrepancy: Example 2

- "Approximating a large discrete space by a smaller discrete space."
- **Given:** A set X of n points in a unit square.
- **Problem:** Color each point in X red/blue such that each *axis parallel rectangle* has "roughly" the same number of red or blue points.
- Discrepancy: max over Rectangle R

$$((\# \text{ of red points in } R) - (\# \text{ of blue points in} R))$$

- Non Trivial:Random Coloring: $O(\sqrt{n \log n})$
  Best: $O(\log^{2.5} n)$

Combinatorial Discrepancy

- **Given :** Set System $(X, \mathbb{S})$. $X = \{1, 2, ..., n\} = [n]$ and $\mathbb{S} = \{S_1, S_2, ..., S_m\}$. Each $S_i \subseteq X$.
- **Problem:** Find a coloring in red and blue such that for every $S_i \in \mathbb{S}$,

$$|(\# \text{ of } x\text{colored } \textit{red} \text{ in } S_i) - (\# \text{ of } x\text{colored } \textit{blue} \text{ in } S_i)|$$

  is small.
- A huge number of Applications!
- **CS:** Computational Geometry, Comb. Opt. , Monte Carlo Simulation, Machine Learning, Complexity, Communication Complexity, Pseudorandomness...
  Math (but not so CS) : Dynamical Systems, Combinatorics, Math. Finance, Number Thy, Ramsey Thy, Algebra, Measure Theory, Numerical Calculus...

Get to the Business: Discrepancy Minimization

- **Given:** A set system $(X, \mathbb{S})$ Universe X:$[n]$
  Set of Subsets of X, $\mathbb{S} = \{S_1, S_2, ..., S_m\}$
- **Problem:** Find $\chi : X \rightarrow \{-1, 1\}^n$ to
  minimize $|\chi(S)|_\infty = \max S_i \in \mathbb{S}, |\sum_{x \in S_i} \chi(x)|$

Best Known *Algorithm*: Random Coloring

- $T = (X, \mathbb{S})$: A set system with $|X| = n$ and $|\mathbb{S}| = m$.
- $\chi : X \to \{-1, 1\}^n$ a random coloring.
- Chernoff Bound: $Prob[|\chi(S)| \geq c\sqrt{n}] \leq e^{-\frac{c^2}{4}}$

Best Known *Algorithm*: Random Coloring

- $T = (X, \mathbb{S})$: A set system with $|X| = n$ and $|\mathbb{S}| = m$.
- $\chi : X \to \{-1, 1\}^n$ a random coloring.
- Chernoff Bound: $Prob[|\chi(S)| \geq c\sqrt{n}] \leq e^{-\frac{c^2}{4}}$
- Choose $c = \sqrt{4 \log 4m}$ to obtain

$$Prob[|\chi(S)| \geq \sqrt{4n \log 4m}] \leq \frac{1}{2m}$$

Best Known *Algorithm*: Random Coloring

- $T = (X, \mathbb{S})$: A set system with $|X| = n$ and $|\mathbb{S}| = m$.
- $\chi : X \to \{-1, 1\}^n$ a random coloring.
- Chernoff Bound: $Prob[|\chi(S)| \geq c\sqrt{n}] \leq e^{-\frac{c^2}{4}}$
- Choose $c = \sqrt{4 \log 4m}$ to obtain

$$Prob[|\chi(S)| \geq \sqrt{4n \log 4m}] \leq \frac{1}{2m}$$

- Union Bound over all sets:

$$Prob[\exists S \in \mathbb{S} : |\chi(S)| \geq \sqrt{4n \log 4m}] \leq \frac{1}{2}$$

Best Known *Algorithm*: Random Coloring

- $T = (X, \mathbb{S})$: A set system with $|X| = n$ and $|\mathbb{S}| = m$.
- $\chi : X \to \{-1, 1\}^n$ a random coloring.
- Chernoff Bound: $Prob[|\chi(S)| \geq c\sqrt{n}] \leq e^{-\frac{c^2}{4}}$
- Choose $c = \sqrt{4 \log 4m}$ to obtain

$$Prob[|\chi(S)| \geq \sqrt{4n \log 4m}] \leq \frac{1}{2m}$$

- Union Bound over all sets:

$$Prob[\exists S \in \mathbb{S} : |\chi(S)| \geq \sqrt{4n \log 4m}] \leq \frac{1}{2}$$

- Thus with probability at least $\frac{1}{2}$, χ achieves a discrepancy of $\mathbf{O}(\sqrt{n \log m})$.

Best Known *Algorithm*: Random Coloring

- $T = (X, \mathbb{S})$: A set system with $|X| = n$ and $|\mathbb{S}| = m$.
- $\chi : X \to \{-1, 1\}^n$ a random coloring.
- Chernoff Bound: $Prob[|\chi(S)| \geq c\sqrt{n}] \leq e^{-\frac{c^2}{4}}$
- Choose $c = \sqrt{4 \log 4m}$ to obtain

$$Prob[|\chi(S)| \geq \sqrt{4n \log 4m}] \leq \frac{1}{2m}$$

- Union Bound over all sets:

$$Prob[\exists S \in \mathbb{S} : |\chi(S)| \geq \sqrt{4n \log 4m}] \leq \frac{1}{2}$$

- Thus with probability at least $\frac{1}{2}$, χ achieves a discrepancy of $\mathbf{O}(\sqrt{n \log m})$.
- Analysis is **tight** !

Surprise! Better Colorings *Exist*!

- Spencer (1985) : **"Six Standard Deviations Suffice"**
  There always exists a coloring with discrepancy at most $6\sqrt{n}$.
  (Here, $m = n$)

Surprise! Better Colorings *Exist*!

- Spencer (1985) : **"Six Standard Deviations Suffice"**
  There always exists a coloring with discrepancy at most $6\sqrt{n}$.
  (Here, $m = n$)
- For arbitrary m, discrepancy $= O\left(\sqrt{n \log \frac{m}{n}}\right)$

Surprise! Better Colorings *Exist*!

- Spencer (1985) : **"Six Standard Deviations Suffice"**
  There always exists a coloring with discrepancy at most $6\sqrt{n}$.
  (Here, $m = n$)
- For arbitrary m, discrepancy $= O\left(\sqrt{n \log \frac{m}{n}}\right)$
- **Problem :** "Highly" non constructive Proof. (PHP on exponentially large universe)
- **Challenge: Can we find such a coloring algorithmically ?** [Spencer 1985]

Surprise! Better Colorings *Exist*!

- Spencer (1985) : **"Six Standard Deviations Suffice"**
  There always exists a coloring with discrepancy at most $6\sqrt{n}$.
  (Here, $m = n$)
- For arbitrary m, discrepancy $= O\left(\sqrt{n \log \frac{m}{n}}\right)$
- **Problem :** "Highly" non constructive Proof. (PHP on exponentially large universe)
- **Challenge: Can we find such a coloring algorithmically ?** [Spencer 1985]
- **Conjecture:** Not Possible! [Alon and Spencer]

Approximating Discrepancy

- **Question :** Suppose a set system has low discrepancy (much lower than $\sqrt{n}$). Is it possible to find it?
- **Answer:**

Approximating Discrepancy

- **Question :** Suppose a set system has low discrepancy (much lower than $\sqrt{n}$). Is it possible to find it?
- **Answer:** NO! [Charikar-Newman-Nikolov 2011] : Distinguishing between discrepancy of 0 and $\sqrt{n}$ is NP-Complete

Approximating Discrepancy

- **Question :** Suppose a set system has low discrepancy (much lower than $\sqrt{n}$). Is it possible to find it?
- **Answer:** NO! [Charikar-Newman-Nikolov 2011] : Distinguishing between discrepancy of 0 and $\sqrt{n}$ is NP-Complete
- What if a set system has low *hereditary discrepancy* ? [Matousek]
- $$Herdisc(X, \mathbb{S}) = max_{Y \subseteq X} disc(Y, \mathbb{S}_Y)$$
- Why Hereditary Discrepancy ?
  Well studied, more robust notion of discrepancy, much used in Computational Geometry PLUS
  fits naturally into the technique!

The Results

Theorem (Constructive Spencer)

Given a set system $(X, \mathbb{S})$ with $|X| = |\mathbb{S}| = n$, there exists a poly(m, n) algorithm to construct a coloring χ with discrepancy $O(\sqrt{n})$.

Theorem (Constructive Low Hereditary Discrepancy)

Given a set system $(X, \mathbb{S})$ with $|X| = n, |\mathbb{S}| = m$, and a guarantee that herdisc$(X, \mathbb{S}) \leq \lambda$, there exists an algorithm running in time poly(m, n) to construct a coloring with discrepancy $O(\lambda \log(mn))$.

- A neat technique! May have applications waiting to be discovered in making several other existential results algorithmic.

A first Attempt: Convex Relaxations

Theorem (Constructive Low Hereditary Discrepancy)

Given a set system $(X, \mathbb{S})$ with $|X| = n, |\mathbb{S}| = m$, and a guarantee that $herdisc(X, \mathbb{S}) \leq \lambda$, there exists an algorithm running in time $poly(m, n)$ to construct a coloring with discrepancy $O(\lambda \log(mn))$.

A first Attempt: Convex Relaxations

Theorem (Constructive Low Hereditary Discrepancy)

Given a set system $(X, \mathbb{S})$ with $|X| = n, |\mathbb{S}| = m$, and a guarantee that $herdisc(X, \mathbb{S}) \leq \lambda$, there exists an algorithm running in time $poly(m, n)$ to construct a coloring with discrepancy $O(\lambda \log(mn))$.

- **A first attempt:** LP

$$-\sqrt{n} \leq \sum_{x \in S_i} u_x^+ - \sum_{x \in S_i} u_x^- \leq \sqrt{n} \; \forall S_i \in \mathbb{S}$$

$$u_x^+ + u_x^- = 1 \; \forall u \in X$$

$$u_x^+, u_x^- \geq 0 \; \forall u \in X$$

A first Attempt: Convex Relaxations

Theorem (Constructive Low Hereditary Discrepancy)

Given a set system $(X, \mathbb{S})$ with $|X| = n, |\mathbb{S}| = m$, and a guarantee that herdisc$(X, \mathbb{S}) \leq \lambda$, there exists an algorithm running in time poly(m, n) to construct a coloring with discrepancy $O(\lambda \log(mn))$.

- **A first attempt:** LP

$$-\sqrt{n} \leq \sum_{x \in S_i} u_x^+ - \sum_{x \in S_i} u_x^- \leq \sqrt{n} \; \forall S_i \in \mathbb{S}$$

$$u_x^+ + u_x^- = 1 \; \forall u \in X$$

$$u_x^+, u_x^- \geq 0 \; \forall u \in X$$

- **Problem:** Absolutely No Information! $u_x^+ = u_x^- = \frac{1}{2}$

A modified first attempt: SDPs

- SDP :

$$||\sum_{x \in S_i} v_x||^2 \leq n \ \forall S_i \in \mathbb{S}$$

$$||v_x||^2 = 1 \ \forall x \in X$$

$$v_x \in \mathbb{R}^n \ \forall x \in X$$

A modified first attempt: SDPs

- SDP :

$$||\sum_{x\in S_i} v_x||^2 \leq n \;\forall S_i \in \mathbb{S}$$

$$||v_x||^2 = 1 \;\forall x \in X$$

$$v_x \in \mathbb{R}^n \;\forall x \in X$$

- **Problem :** $v_x = e_x \;\forall x \in X$ always a solution!

A modified first attempt: SDPs

- SDP :

$$||\sum_{x\in S_i} v_x||^2 \leq n \ \forall S_i \in \mathbb{S}$$

$$||v_x||^2 = 1 \ \forall x \in X$$

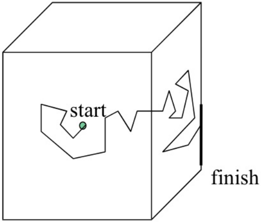
$$v_x \in \mathbb{R}^n \ \forall x \in X$$

- **Problem :** $v_x = e_x \ \forall x \in X$ always a solution!
- **Observation:** Tighten Some Constraints a Little Bit for some sets, and the trivial solution disappears...
- Example :

$$||\sum_{x\in S_i} v_x||^2 \leq \frac{n}{\log n}$$

The Algorithm

- A random walk over *fractional* colorings of X



- Each move adds an increment generated by a suitable SDP δ_t to the current coloring.
- Whenever a label hits $\{-1, 1\}^n$, we freeze the label.
- Random Walk over different *directions* is (nicely!) correlated.

The Algorithm: Analysis

- Random Walk over each direction has a *high* variance. $\Rightarrow$ converge to $\{-1, 1\}^n$ fast.
- Walk over Discrepancies of each set has *low* variance $\Rightarrow$ Discrepancy is bounded!

The Algorithm: Generating a Step

- Let $U \subseteq X$ be the set of vertices which haven't hit $\{-1,1\}^n$ at the start of a fixed iteration.
- The following SDP is feasible:

$$||\sum_{x \in S_i \cap U} v_x||^2 \leq \lambda \ \forall S_i \in \mathbb{S}$$

$$||v_x||^2 = 1 \ \forall x \in U$$

- Rounding: Choose a random vector of independent $\mathcal{N}(0,1)$ variables $(g_1, g_2, .., g_n)$
- For each x, let the update be: $\eta_x = \langle v_x, g \rangle$

Why the rounding works

Fact (Additivity of Gaussians)

If $(g_1, g_2, ..., g_n)$ is a vector of independent gaussians of variance 1, *then* $\langle v_x, g \rangle \sim \mathcal{N}(0, ||v_x||^2)$

- **High Variance Walk on coordinates** :
  $\eta_x = \langle v_x, g \rangle \sim \mathcal{N}(0, 1)$
- **Low Variance Walk on Discrepancies** For each set $S_i \in \mathbb{S}$, $\sum_{x \in S_i} \eta_x \sim \mathcal{N}(0, \sigma^2)$ with $\sigma \leq \lambda$.

Algorithm

- Start with the coloring $\chi_0 = (0, 0, ..., 0)$
- Update each coloring: $\chi_t = \chi_{t-1} + \gamma \cdot \eta$
  (γ is tiny $\sim \frac{1}{n^2}$ step size, chosen to control error due to rounding)
- Whenever $\chi(x) \geq (1 - 1/n)$ or $\chi(x) \leq (-1 + 1/n)$ round $\chi(x)$ and fix.
- Thus, color of an element does a random walk with each step $\gamma\mathcal{N}(0, 1)$
  discrepancy of each set does a random walk with each step $\gamma\mathcal{N}(0, \lambda^2)$

Analysis

Claim (Half the vertices are colored in each iteration)

With probability at least $\frac{1}{2}$, $\frac{n}{2}$ *coordinates are colored in the first iteration.*

Analysis

Claim (Half the vertices are colored in each iteration)

With probability at least $\frac{1}{2}$, $\frac{n}{2}$coordinates are colored in the first iteration.

A round $T = O(\frac{1}{\gamma^2})$ steps.

Claim (Low Discrepancy Rise)

Each set has discrepancy $O(\lambda)$ in expectation in each round.

Analysis

Claim (Half the vertices are colored in each iteration)

With probability at least $\frac{1}{2}$, $\frac{n}{2}$ coordinates are colored in the first iteration.

A round $T = O(\frac{1}{\gamma^2})$ steps.

Claim (Low Discrepancy Rise)

Each set has discrepancy $O(\lambda)$ in expectation in each round.

Claim

Every element is colored in $O(\log n)$ rounds.

Analysis

Claim (Half the vertices are colored in each iteration)

With probability at least $\frac{1}{2}$, $\frac{n}{2}$ coordinates are colored in the first iteration.

A round $T = O(\frac{1}{\gamma^2})$ steps.

Claim (Low Discrepancy Rise)

Each set has discrepancy $O(\lambda)$ in expectation in each round.

Claim

Every element is colored in $O(\log n)$ rounds.

Claim

Expected discrepancy for any set in all rounds is at most $O(\lambda \log n)$

Analysis

Claim (Half the vertices are colored in each iteration)

With probability at least $\frac{1}{2}$, $\frac{n}{2}$ coordinates are colored in the first iteration.

A round $T = O(\frac{1}{\gamma^2})$ steps.

Claim (Low Discrepancy Rise)

Each set has discrepancy $O(\lambda)$ in expectation in each round.

Claim

Every element is colored in $O(\log n)$ rounds.

Claim

Expected discrepancy for any set in all rounds is at most $O(\lambda \log n)$

Summary

- Each Step generated by some **SDP** which introduces correlation between updates on colors of different elements.
- Use Hereditary Discrepancy to guarantee feasibility.
- High Variance of the Walk along each coordinate.
- Low Variance walk on the set of discrepancies.

- Spencer's Setting: $(X, \mathbb{S})$, $|X| = n$ and $|\mathbb{S}| = m$.
- Suppose we use $\lambda = \sqrt{n}$ we obtain : discrepancy of $O(\sqrt{n}\sqrt{\log n \log m})$.

- Spencer's Setting: $(X, \mathbb{S})$, $|X| = n$ and $|\mathbb{S}| = m$.
- Suppose we use $\lambda = \sqrt{n}$ we obtain : discrepancy of $O(\sqrt{n}\sqrt{\log n \log m})$.
- **Idea! :** Tighten bounds on constraints in successive SDPs by using Spencer's bound itself!

- Spencer's Setting: $(X, \mathbb{S})$, $|X| = n$ and $|\mathbb{S}| = m$.
- Suppose we use $\lambda = \sqrt{n}$ we obtain : discrepancy of $O(\sqrt{n}\sqrt{\log n \log m})$.
- **Idea! :** Tighten bounds on constraints in successive SDPs by using Spencer's bound itself!
- Nope! :-(Doesn't work. Gives $O(\sqrt{n \log m})$
- **Issue:** The notorious Union Bound.

Cool Idea: SDP from Spencer's Proof

Want:

1) Tighter Constraints on SDPs to guarantee lower discrepancy.
2) Forbid the nasty union bound.

Open Questions

Problem (Beck Fiala Conjecture)

Given a set system $(X, \mathbb{S})$ such that each $x \in X$ appears in at most t sets in $\mathbb{S}$, construct a coloring of X with minimum discrepancy.

- History: Discrepancy $2t - 1$ [Beck and Fiala, 1981], $O(C\sqrt{t \log m} \log n)$ [Beck, 1981], $2t - 3$ by [Bednarchak and Helm, 1997 and Helm, 1999]
- $O(\sqrt{t \log n})$ [Banaszczyk, 1998], and a constructive version by [Srinivasan,]
- Beck and Fiala conjectured a discrepancy of $O(\sqrt{t})$ – WIDELY OPEN!

Thank You!

QUESTIONS?