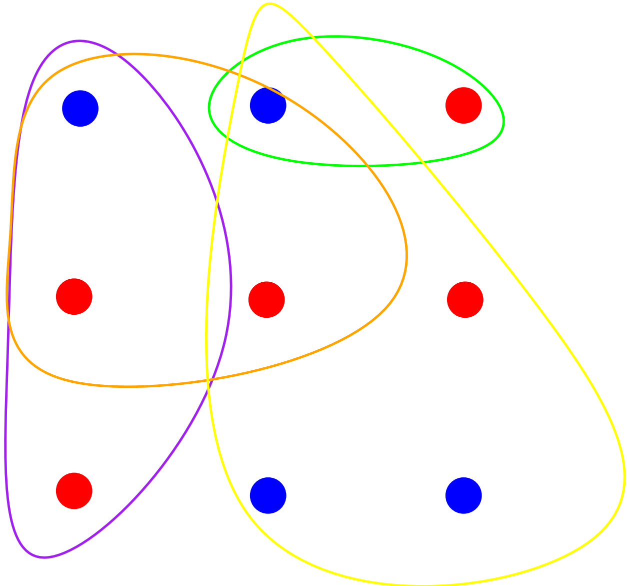
Discrepancy and Approximation Algorithms



Sasho Nikolov

University of Toronto

Outline

1. Basics
2. Discrepancy and Bin Packing
3. Bounds and Algorithms
4. Approximating Discrepancy

Relax – Solve - Round

Powerful paradigm in approximation algorithms

Solve

$$\max c^\intercal x \quad Ax \le b \quad x \in \{0,1\}^n$$

Relax

$$\max c^\intercal x \quad Ax \le b \quad x \in [0,1]^n$$

Round

Rounding

- What do we want from a rounding?
  - $z = R(x)$ is feasible
  - $c^{\mathsf{T}}z \geq \alpha\, c^{\mathsf{T}}x$ → approximation factor α
- Two step approach:
  1. approximately preserve constraints: $(Az)_i \leq b_i + D$
  2. “fix” violated constraints without changing objective value too much
- Step 1: discrepancy theory
- Step 2: problem dependent.

Linear Discrepancy

- Round *x* so that ... approximately sat...

Upper bound on how much the *i*-th constraint can be violated.

$$\min_{z \in \{0,1\}^n} \max_{i=1}^{m} |(Az - Ax)_i| = \min_{z \in \{0,1\}^n} \|A(z-x)\|_\infty$$

  – we can include *c* as one of the rows of *A* to preserve objective value

- Worst case over *x*:

$$\mathrm{lindisc}(A) = \max_{x \in [0,1]^n} \min_{z \in \{0,1\}^n} \|A(z-x)\|_\infty$$

Matrix Discrepancy

- Discrepancy: $\mathrm{disc}(A) = \min_{x \in \{-1,1\}^n} \|Ax\|_\infty$
- Hereditary Discrepancy:

$$\mathrm{herdisc}(A) = \max_{J \subseteq [n]} \mathrm{disc}(A_J)$$

  – A_J is the submatrix of columns indexed by J

Theorem. [G62] A with entries -1, 0, 1 has herdisc(A) = 1 iff A is totally unimodular.

Theorem. [LSV86] $\mathrm{lindisc}(A) \leq \mathrm{herdisc}(A)$

Proof

- Observation: $\text{disc}(A) = \min_{x\in\{-1,1\}^n} \|Ax\|_\infty$

Algorithmically, we need to compute low-discrepancy *x* for any A_J

$\min_{z\in\{0,1\}^n} \|A(z - \frac{1}{2}e)\|_\infty$

  – i.e. we can round 1/2 … most 0.5 * herdisc(…

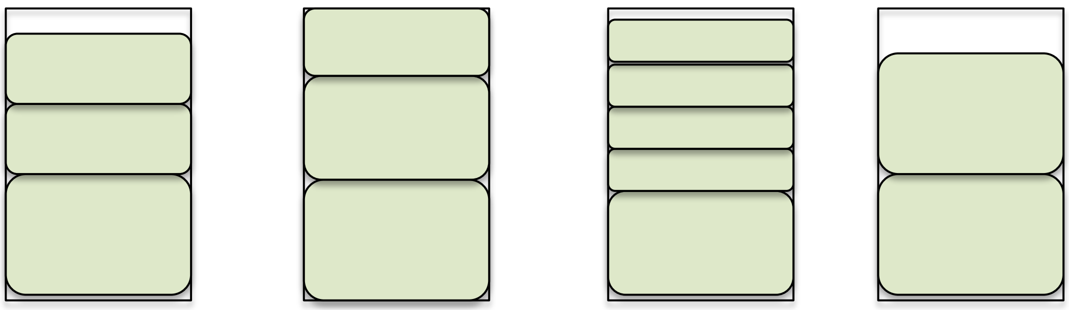
Cost of rounding:
≤ (1/2 + 1/4 + 1/8 + …) * herdisc(*A*)
≤ herdisc(*A*)

- In general: write in binary … bit

x_1 = 0. 1 0 1 0	x_1 = 0. 1 0 1 0	x_1 = 0. **1 0** 0 0	x_1 = 0. 1 0 0 0
x_2 = 0. 1 1 0 0	x_2 = 0. 1 1 0 0	x_2 = 0. 1 1 0 0	x_2 = **1. 0** 0 0 0
x_3 = 0. 0 0 1 1	x_3 = 0. 0 **1 0** 0	x_3 = 0. 0 1 0 0	x_3 = **0. 1** 0 0 0

Bin Packing

Problem: Pack items of sizes $1 \geq s_1 \geq s_2 \geq \ldots \geq s_n$ into the *fewest* bins of size 1



[KK82] $\mathrm{OPT} + O(\log^2 \mathrm{OPT})$. If $s_n = \Omega(1)$, $\mathrm{OPT} + O(\log \mathrm{OPT})$

[HR17] $\mathrm{OPT} + O(\log \mathrm{OPT})$ for all sizes.

Conjecture. $\mathrm{OPT} + O(1)$.

LP Relaxation

Configuration LP: p in $\{0,1\}^n$ is *feasible* if $\Sigma_i\, p_i\, s_i \leq 1$

(how to pack a single bin)

Solve using multiplicative weights and PST framework.

$$\min e^{\intercal} x$$
$$Px \geq e$$
$$x \geq 0$$

The rows of P are all feasible patterns.

“Smallest number of feasible patterns that cover all items”

Karmakar-Karp via Discrepancy [EPR 11]

	A							x			
s_1	1	1	0	0	0	0	0	0.5		1	1 - D
s_2	0	0	1	0	0	0	0	0.5		2	2 - D
s_3	0	0	0	1	0	0	0	1		3	3 - D
s_4	0	0	0	0	1	1	0	1	=	3	4 - D
s_5	0	0	0	0				0.6		4	5 - D
s_6	0	0	0					0		6	6 - D
s_7	0	0	0	2				1		8	7 - D
s_8	2	0	0	2	3	2	3			8 10 11	8 - D
s_9	2	2	2	2	3	3	4			20 19	9 - D
C	1	1	1	1	1	1	1			5	

Sum of first i rows

Number of available slots of size **at least** s_i

D = herdisc(B)

B: modified matrix

Exercise: after adding D more bins we can pack all items.

Karmakar-Karp, contd.

- Assume at most k items fit per bin:
  - the matrices are monotone down each column and have entries bounded by k
- The discrepancy of such matrices is $O(k \log n)$
- Implies + O(log OPT) approximation if all item sizes are *constant*.
  - [HR17] reduce the general case to this case without further loss.

- **[NNN12]** No rounding which only uses the support of an optimal LP solution can do better for this case.

(Efficient) Partial Coloring

Theorem. [LM12] Let x in $[-1,1]^n$, and $\lambda_1, \lambda_2, \ldots, \lambda_m$ s.t.

$$\sum_i \exp(-\lambda_i^2/16) \le \frac{n}{2}$$

Then we can *efficien*

Earlier non-algorithmic versions by Beck and Spencer. First algorithmic work by Bansal.

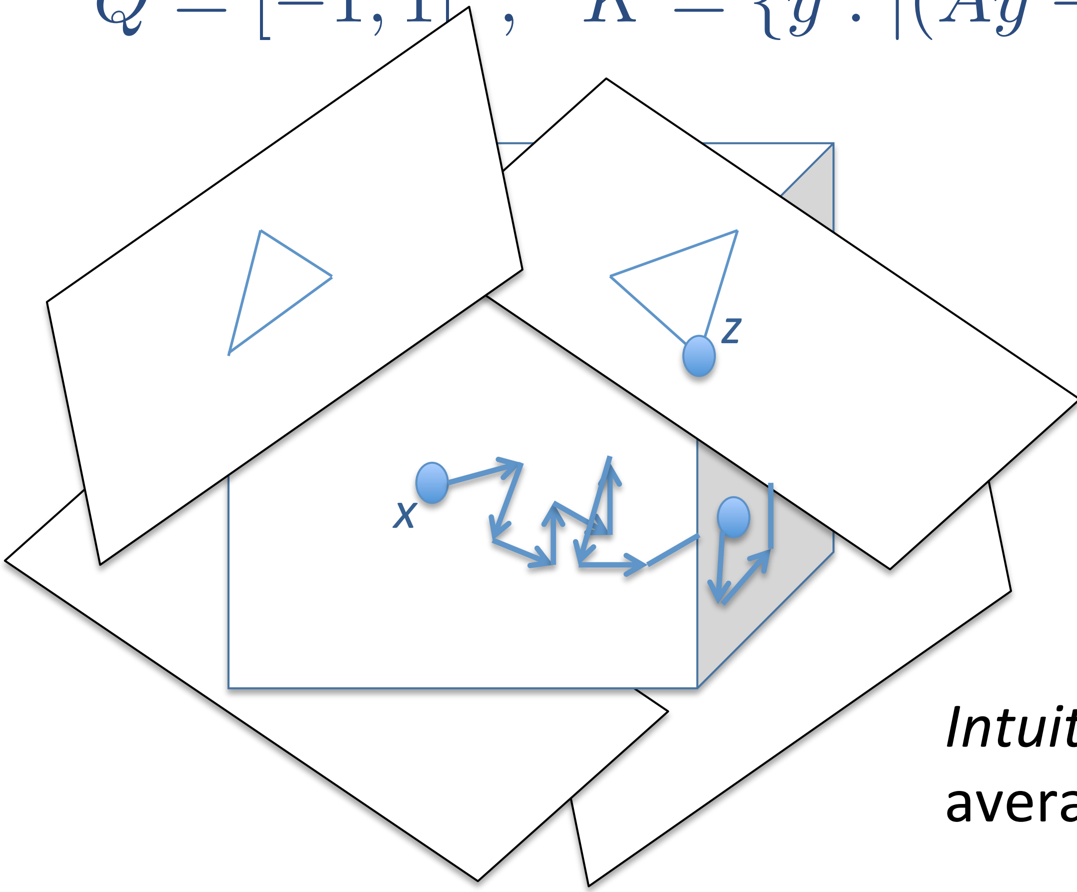
$$\forall i: \quad |(Az - Ax)_i| \le \lambda_i \|a_{i,*}\|_2$$

and at least $n/10$ coordinates of z are -1 or 1.

- compare with a random rounding: $\sum_i \exp(-\lambda_i^2/4) \le \frac{1}{2}$
- if $m \le n/2$, can set $\lambda_1 = \lambda_2 = \ldots = \lambda_m = 0$: basic feasible solution
- interpolates between randomized and iterative rounding

Lovett-Meka Algorithm

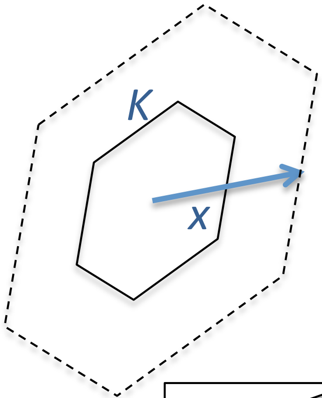
$$Q = [-1,1]^n, \quad K = \{y : |(Ay - Ax)_i| \le \lambda_i \|a_{i,*}\|_2\}$$



Run Brownian motion in $K \cap Q$. Once you hit a face, stick to it.

Success: $n/10$ tight constraints at z from Q.

Intuition: the facets of K are on average further than the facets of Q.



Banaszczyk's Theorem

$$\|x\|_K = \min\{t : x \in tK\}$$

Th columns have Euclidean is a symmetric convex body ($K = -K$), then here exists an x in $\{-1, 1\}^n$ s.t. $\|Ax\|_K \leq 10 \cdot \mathbb{E}\|G\|_K$, where G is a standard Gaussian random vector.

constant if K has Gaussian measure $\geq 1/2$.

Theorem. [DGLN16] If A is Euclidean norm at most 1, then μ over $\{-1, 1\}^n$ s.t. AX is O(1)-Subgaussian, for $X \sim \mu$.

Constant variance and sub-Gaussian tail bounds in every direction: $\mathbb{P}(\langle \theta, AX \rangle \geq t) \leq e^{-t^2/C}$

Algorithmic Banaszcyk Thm

- If A is orthonormal: uniformly random X from $\{-1, 1\}^n$
- If all columns of A the same: $X = \pm(+1,-1, +1, \ldots)$

- **[BDGL17]** Can efficiently sample μ.
  - Random walk in Q.
  - Intuitively: combine the two cases above.

Komlos Problem

- **Komlos conjecture:** for any A with columns of Euclidean norm at most 1, there exists an x in $\{-1, 1\}^n$ s.t. $\|Ax\|_\infty = O(1)$.

- **[B98]** $\|Ax\|_\infty = O(\sqrt{\log m})$
  – Proof: $\mathbb{E}\|G\|_\infty = O(\sqrt{\log m})$, and apply theorem.

Complexity of Discrepancy

- **[CNN11]** NP-hard to d

  Largest it can be **[S85]**.

  disc(A) = 0 and disc(A) = $\Theta(n^{1/2})$ for binary O(n)×n matrix A.

- **[NT14]** Can approximate herdisc(A) up to $O((\log m)^{3/2})$.
  - **[DNTT17]** … up to polylog(rank A)
  - **[AHG14]** NP-hard to apx better than factor 2.

Approximating HerDisc

maximum Euclidean norm of a column of *A*

- First upper bound: by Banaszczyk

$$\mathrm{herdisc}(A) \leq 10(\mathbb{E}\|G\|_\infty) \cdot (\mathrm{col}(A))$$

  - we use that $\mathrm{col}(A_J) \leq \mathrm{col}(A)$
- Observe: $\|Ax\|_\infty = \|Ax\|_Q = \|TAx\|_{TQ}$, for any invertible T.
- Better upper bound:

$$\mathrm{herdisc}(A) \leq 10 \cdot \inf_T (\mathbb{E}\|G\|_{TQ}) \cdot (\mathrm{col}(TA))$$

$$=: \lambda(A)$$

Approximating HerDisc

$$\mathrm{herdisc}(A) \le \lambda(A) \le O((\log m)^{3/2}) \cdot \mathrm{herdisc}(A)$$

- Proof sketch:
  - formulate $\lambda(A)$ as the value of a convex program *P*
  - the dual *D* of *P* is a maximizati

  Volumetric argument
  - feasible solution to *D* → lower bound on herdisc(*A*)

- The program *P* can be solved efficiently
- $\lambda(A)$ can be relaxed to an SDP.

Open Problems

- Approximate lindisc(A)
- Use discrepancy rounding for other approximation problems
- Get + o(log OPT) approximation for Bin Packing
- Solve Komlos problem