Scaling and Optimization over Counting

Mohit Singh
Georgia Tech

Permanent

- Given a non-negative $n \times n$ matrix A

$$Per(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} A_{i\sigma_i} = \sum_{M \in \mathcal{M}} \prod_{(i,j) \in M} A_{ij} = \sum_{M \in \mathcal{M}} A^M$$

where $\mathcal{M}$ is set of perfect matchings between rows and columns.

How to compute Permanent (for simplicity, assume A positive).

$$Per(A) \leq \prod_{i=1}^{n} \left(\sum_{j=1}^{n} A_{ij} \right) = \sum_{M \in \mathcal{M}} A^M + \cdots .$$

$$Per(A) \leq min \left\{ \prod_{i=1}^{n} \left(\sum_{j=1}^{n} A_{ij}\, y_{ij} \right) : y^M = 1\ \forall M \in \mathcal{M}, y > 0 \right\}$$

Scaling 1: Sinkhorn scaling

Lemma(Folklore+Sinkhorn Scaling):

$$Per(A) \leq \min \left\{ \prod_{i=1}^{n} \left(\sum_{j=1}^{n} A_{ij}\, y_{ij} \right) : y^M = 1\ \forall\, M \in \mathcal{M}\,, y > 0 \right\}$$

$$= \min \left\{ \prod_{i=1}^{n} \left(\sum_{j=1}^{n} A_{ij}\, z_j \right) : \prod_{j=1}^{n} z_j = 1\,, z > 0 \right\}$$

$$= \exp\left(\max\left\{ b_{ij} \log \frac{a_{ij}}{b_{ij}} : b \in P(\mathcal{M}) \right\} \right)$$

Here $P(\mathcal{M}) = \left\{ b \in R_+^{n\times n} : \sum_i b_{ij} = 1 \quad \forall i, \sum_j b_{ij} = 1\ \forall j \right\}$ is the matching polytope.

Proof: First equality. $y_{ij} = z_i z_j$ from complementary slackness in matching polytope.

Second equality. Convex duality. LHS/RHS is not convex program as such.

The bound is symmetric in rows and columns.

Gurvits' Result

- Theorem (Egorychev, Falikman'71):

$$Per(A) \leq min\,\{\prod_{i=1}^{n}(\sum_{j=1}^{n} A_{ij}\, z_j) : \prod_{j=1}^{n} z_j = 1, z > 0\,\} \leq \frac{n^n}{n!}\, Per(A).$$

Theorem (Gurvits): If $p(z_1, \ldots, z_n)$ is a n variate degree n homogenous real stable polynomial with non-negative coefficients

$$\frac{\mathrm{n}^n}{n!} \times coefficient\ of\ z_1 z_2 \ldots z_n \geq min\,\left\{p(z): \prod_{j=1}^{n} z_j = 1, z > 0\right\}$$

Pf: Wait till Monday.

Moral: Scaling helps to improve the bound.

Duality

Lemma:

$$Per(A) \leq \ min \left\{ \prod_{i=1}^{n} \left(\sum_{j=1}^{n} A_{ij}\, y_{ij} \right) : y^M \geq 1 \; \forall \; M \in \mathcal{M} \; , y > 0 \right\}$$

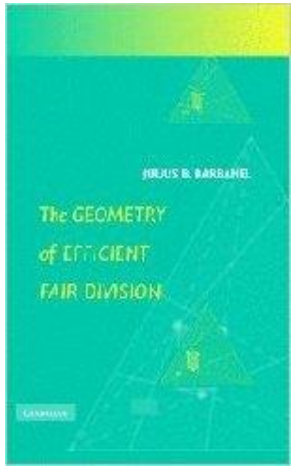
$$= sup_{\theta \in P(\mathcal{M})} inf_{y>0} \frac{\prod_{i=1}^{n} \left(\sum_{j=1}^{n} \; A_{ij} y_{ij} \right)}{y^{\theta}}$$

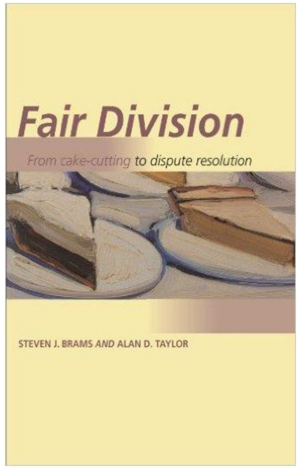
Fair Allocations

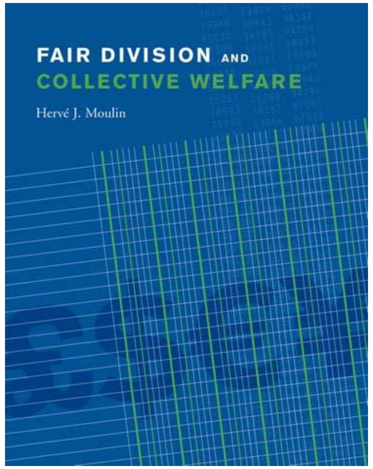
- How to divide given set of items among players?

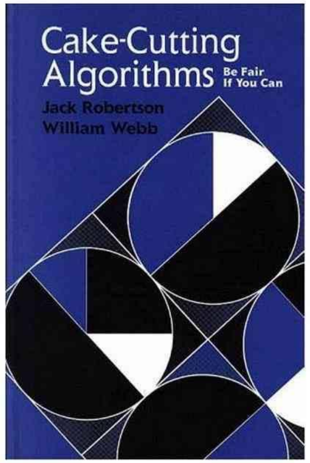
- Utilitarian Approach: Assign to maximize the sum total utility.
  - Disregards fairness

- Egalitarian Approach: Assign to maximize the minimum utility of players.
  - Disregards average welfare.









Nash Social Welfare

- Takes an intermediate approach.
  - Maximize the product (or geometric mean) of the utilities of agents.

- Axiomatic approach: Under a set of natural axioms, Nash social welfare is the objective to optimize. [Kaneko-Nakamura'79]
  - Develops on Nash Bargaining Game. [Nash'50]

Computability

- I: set of n players, J: set of m items. Additive Valuations.
  - v_{ij} valuation of player i for good j.
  - If player i assigned a set S of goods, then his utility is $\sum_{j \in S} v_{ij}$.
- Divisible Items.
- Convex Program

$$
\begin{aligned}
\max \quad & \left(\prod_{i \in I}\left(\sum_{j \in J} v_{ij} x_{ij}\right)\right) \\
\text{s.t.} \quad & \\
& \sum_{i \in I} x_{ij} = 1 \quad \forall j \in J \\
& x \geq 0
\end{aligned}
$$

Indivisible items

- Items have to be assigned to agents.
- Formally find an assignment $\sigma : J \to I$ such that

$$\max\left(\prod_{i \in I}\left(\sum_{j:\sigma(j)=i} v_{ij}\right)\right)$$

Equivalent to selecting a weighted subgraph that maximizes matchings.

Judgement of Solomon by Giorgono

Gurvits-style result

Theorem[Anari, Oveis-Gharan, Saberi, S'17]: There is a e^n-approximation for the Nash Social Welfare problem.

Convex Program

$$\max \quad \left(\prod_{i\in I}\left(\sum_{j\in J} v_{ij}x_{ij}\right)\right)$$

s.t.

$$\sum_{i\in I} x_{ij} = 1 \quad \forall j \in J$$

$$x \geq 0$$

Fractional OPT: $\frac{M}{n}$

Integral OPT: $M^{\frac{1}{n}}$

New Convex Program

- $\max_{x \in P} \min_{y \in Q} f(x,z) = \left(\prod_{i \in I}\left(\sum_{j \in J} v_{ij} x_{ij} z_{ij}\right)\right)$
- P={x: $\sum_{i \in I} x_{ij} = 1 \quad \forall j \in J, x \geq 0$ }
- $Q = \{z: z^M \geq 1 \, \forall \, matchings \; M \in \mathcal{M}_n\}$
- $\mathcal{M}_n$:= matchings of size n.

Fractional OPT: $M^{\frac{1}{n}}$

Integral OPT: $M^{\frac{1}{n}}$

New Convex Program

- $\max_{x \in P} \min_{y \in Q} f(x,z) = \left(\prod_{i \in I} \left(\sum_{j \in J} v_{ij} x_{ij} z_{ij}\right)\right)$
- P={x: $\sum_{i \in I} x_{ij} = 1 \quad \forall j \in J, x \geq 0$ }
- $Q = \{z : z^M \geq 1 \ \forall \ matchings \ M \in \mathcal{M}_n\}$
- $\mathcal{M}_n$:= matchings of size n.

Is the above optimization problem efficiently solvable?

- The objective is log-concave in x.
- The objective is log-convex in log z (coordinate wise).
- Q is a polyhedron in log z.

Is it a relaxation to our problem?

Fractional OPT: $M^{\frac{1}{n}}$

Integral OPT: $M^{\frac{1}{n}}$

Similar polynomial optimization relaxation: Maximum sub-determinant problem [Nikolov, S'16]

Relaxation

- $\max_{x \in P} \min_{y \in Q} f(x, y) = \left(\prod_{i \in I}\left(\sum_{j \in J} v_{ij} x_{ij} z_{ij}\right)\right)$
- P={x: $\sum_{i \in I} x_{ij} = 1 \;\; \forall j \in J, x \geq 0$ }
- $Q = \{z: z^M \geq 1 \; \forall \; matchings \; M \in \mathcal{M}_n\}$
- $\mathcal{M}_n$:= matchings of size n.

$z_{ij} = y_j$ for each i and j.

Lemma: $OPT <= \max_{x \in P} \min_{z \in Q} f(x, z)$

Proof: Let $\bar{x}$ denote the indicator of optimal integral assignment $\sigma: J \rightarrow I$. For any $y \in Q$, we have

$$f(\bar{x}, z) = \left(\prod_{i \in I}\left(\sum_{j \in J} v_{ij} \bar{x}_{ij} \; z_{ij}\right)\right)$$
$$= \sum_{M \in \mathcal{M}_n : \bar{x}^M = 1} v^M z^M$$
$$= \sum_{M \in \mathcal{M}_n : \bar{x}^M = 1} v^M z^M$$
$$= \left(\prod_{i \in I}\left(\sum_{j : \sigma(j) = i} v_{ij}\right)\right) = OPT$$

Randomized Algorithm

- Solve the program to obtain x^*.
- Independently, for each item j, assign it to player i with probability x_{ij}^*.

Let $p(y) = f(x^*, y)$ be homogenous degree n polynomial in m variables.

$$p(y) = \left(\prod_{i \in I}\left(\sum_{j \in J} v_{ij} x_{ij}^* \, y_j\right)\right)$$

We will let c_S denote the coefficient of square-free term $y^S := \prod_{j \in S} y_j$ for any $S \subseteq J, |S| = n$

$$c_S = \sum_{\sigma : I \to S, \sigma \ onto} \prod_{i \in I} v_{i\sigma(i)} x_{i\sigma(i)}^*$$

Analysis

- Lemma: The expected objective of the solution returned by the algorithm is the sum of square free coefficients in $p(y)$.
- Proof Sketch: Let X_{ij} denote the random variable that player i gets item j.

$$E[Alg] = E\left[\prod_{i \in I}\left(\sum_{j \in J} X_{ij} v_{ij}\right)\right] = \sum_{\sigma: I \rightarrow J} E\left[\prod_{i \in I} X_{i\sigma(i)} v_{i\sigma(i)}\right]$$

$$= \sum_{\sigma: I \rightarrow J} \Pr\left[X_{i\sigma(i)} = 1\ \forall i \in I\right] \left(\prod_i v_{i\sigma(i)}\right)$$

For any $\sigma: I \rightarrow J$, $\Pr\left[X_{i\sigma(i)} = 1\ \forall i \in I\right] = 0$ if σ is not 1-1.

$= \prod_{i \in \mathrm{I}} x^*_{i\sigma(i)}$ if σ is 1-1.

$$E[Alg] = \sum_{\sigma: I \rightarrow J, \sigma\ 1-1} \prod_{i \in \mathrm{I}} x^*_{i\sigma(i)} v_{i\sigma(i)} = \sum_{S \subseteq J: |S| = n} \sum_{\sigma: I \rightarrow S, \sigma\ onto} \prod_{i \in I} v_{i\sigma(i)} x^*_{i\sigma(i)} = \sum_{S \subseteq J: |S| = n} c_S$$

Recap

- New Polynomial optimization problem.
- Simple randomized algorithm.

$$OPT^* = \min_{y \in Q} p(y) = \left(\prod_{i \in I} \left(\sum_{j \in J} v_{ij} x^*_{ij} \, y_j \right) \right)$$

$E[Alg]$= sum of square free coefficients of $p(y)$.

Theorem: sum of square free coefficients of $p(y) \geq e^{-n} \min_{y \in Q} p(y)$.

Analysis

- Theorem[Anari,Oveis-Gharan, Saberi, S'17]: Let p be a real stable degree n homogenous polynomial in variables $y_1, \dots, y_m$ where the coefficient of y^S is c_S for any $|S| = n$. Then

$$\sum_{S:|S|=n} c_S \geq e^{-n} \inf_{y \succ 0: y^S \geq 1\ \forall S, |S|=n} p(y)$$

Proof: Apply Gurvits' Theorem to $p(y) \cdot E_{m-n}(y_1, \dots y_m)$.

$$\sum_{S:|S|=n} c_S \geq e^{-m}\, inf_{y>0} \frac{p(y) \cdot E_{m-n}(y_1, \dots y_m)}{y_1 . y_2 \dots y_m}$$

$$\geq e^{-m} \sup_{\theta:\theta^T 1 = \mathrm{n}} \inf_{y>0} \frac{p(y)}{y^\theta} \inf_{\alpha \in [0,1]^m : \alpha^T 1 = m-n} \inf_{y>0} \frac{E_{m-n}(y_1, \dots, y_m)}{y^\alpha}$$

$$\geq e^{-n} \sup_{\theta:\theta^T 1 = \mathrm{n}} \inf_{y>0} \frac{p(y)}{y^\theta} = e^{-n} \inf_{y \succ 0: y^S \geq 1\ \forall S, |S|=n} p(y)$$

Efficient Algorithm

- Solve the program to obtain x^*.
- Independently, for each item j, assign it to player i with probability x^*_{ij}.

Theorem: The expected value of the algorithm is a e^{-n}-approximation.

Method of conditional expectations: Assign items one by one while ensuring the conditional expectation does not decrease.
Evaluating the conditional expectation = problem of counting (weighted) matchings in a bipartite graph. FPRAS [Jerrum, Sinclair, Vigoda'04].

Theorem: There is polynomial time algorithm that returns an e^{-n}-approximation with high probability.

Back to Permanents

Exposition based on [Anari, Oveis-Gharan'17, Starzak-Vishnoi'17]

- $Per(A) \leq min\left\{\prod_{i=1}^{n}\left(\sum_{j=1}^{n} A_{ij}\, y_{ij}\right) : y^M \geq 1\ \forall M \in \mathcal{M}, y > 0\right\}$

$$= \sup_{\theta \in P(\mathcal{M})} inf_{y>0} \frac{p(y)}{y^\theta}$$

Where $p(y) = \prod_{i=1}^{n}\left(\sum_{j=1}^{n} A_{ij}\, y_{ij}\right)$.

$$Per(A) \leq \sup_{\theta \in P(\mathcal{M})} inf_{y>0,x>0} \frac{p(y) \cdot \prod_{j=1}^{n}\left(\sum_{i=1}^{n} x_{ij}\right)}{y^\theta x^\theta}$$

Theorem [Anari, Oveis-Gharan'17]:

$$Per(A) \leq \sup_{\theta \in P(\mathcal{M})} inf_{y>0,x>0} \frac{p(y) \cdot \prod_{j=1}^{n}\left(\sum_{i=1}^{n} x_{ij}\right) \theta^\theta}{y^\theta x^\theta}$$

Lower Bound

Theorem [Anari, Oveis-Gharan'17]:

$$Per(A) \leq \sup_{\theta \in P(\mathcal{M})} inf_{y>0,x>0} \frac{p(y) \cdot \prod_{j=1}^{n} \left(\sum_{i=1}^{n} x_{ij}\right) \theta^{\theta}}{y^{\theta} x^{\theta}}$$

- Proof: RHS =

$$\inf_{y>0} p(y) \cdot \sup_{\theta \in P(\mathcal{M})} inf_{y>0,x>0} \frac{\prod_{j=1}^{n} \left(\sum_{i=1}^{n} x_{ij}\right) \theta^{\theta}}{y^{\theta} x^{\theta}}$$

Fix y to be minimizer. We have

$$RHS \geq \sum_{M \in \mathcal{M}} A^M y^M \sup_{\theta \in P(\mathcal{M})} inf_{y>0,x>0} \frac{\prod_{j=1}^{n} \left(\sum_{i=1}^{n} x_{ij}\right) \theta^{\theta}}{y^{\theta} x^{\theta}}$$

For M summand, pick $\theta = 1_M$. $\theta^{\theta} = 1$. Thus the scaling doesn't affect the terms of the permanent! Use.

$$inf_{x>0} \frac{\prod_{j=1}^{n} \left(\sum_{i=1}^{n} x_{ij}\right)}{x^M} \geq 1$$

Stronger bounds

- $Per(A) \leq \sup_{\theta \in P(\mathcal{M})} inf_{y>0,x>0} \frac{p(y) \cdot \prod_{j=1}^{n} (\sum_{i=1}^{n} x_{ij}) \theta^{\theta}}{y^{\theta} x^{\theta}}$

RHS is computable bound. Not straightforward.

Theorem[Anari,Oveis-Gharan'17]:

$$\sup_{\theta \in P(\mathcal{M})} inf_{y>0,x>0} \frac{p(y) \cdot \prod_{j=1}^{n} \left(\sum_{i=1}^{n} x_{ij}\right) \theta^{\theta} (1-\theta)^{1-\theta}}{y^{\theta} x^{\theta}} \leq Per(A)$$

Proof: Relies on stable polynomials. Not clear if LHS is computable.

Bethe Permanent

- Corollary: $\sup_{\theta \in P(\mathcal{M})} inf_{y>0,x>0} \frac{p(y) \cdot \prod_{j=1}^{n} (\sum_{i=1}^{n} x_{ij}) \theta^{\theta} (1-\theta)^{1-\theta}}{y^{\theta} x^{\theta}} \leq Per(A) \leq \sup_{\theta \in P(\mathcal{M})} inf_{y>0,x>0} \frac{p(y) \cdot \prod_{j=1}^{n} (\sum_{i=1}^{n} x_{ij}) \theta^{\theta}}{y^{\theta} x^{\theta}}$

The gap between the bounds is at most $(1-\theta)^{1-\theta} \geq \prod_{ij} e^{-\theta_{ij}} = e^{-n}$

Theorem[Straszak, Vishnoi'17]:

$$\log \sup_{\theta \in P(\mathcal{M})} inf_{y>0,x>0} \frac{p(y) \cdot \prod_{j=1}^{n} (\sum_{i=1}^{n} x_{ij}) \theta^{\theta} (1-\theta)^{1-\theta}}{y^{\theta} x^{\theta}}$$
$$= \sup_{b \in P(\mathcal{M})} \sum_{ij} \mathrm{b_{ij}} \log \frac{a_{ij}}{b_{ij}} + (1-b_{ij}) \log(1-b_{ij})$$

LHS bound is within $2^{-\mathrm{n}}$ of the Per(A). [Gurvits, Samorodnitsky'14].

Both bounds are computable. [Vontobel'12].

More generally

- Nothing special about permanent.
- Pick polynomials p and q stable. Allows to obtain Schrijver's stronger bound for permanent.
- Replace matching by matroid intersection for special matroids.
- Relationship between Bethe Free Energy and Polynomial relaxations. [Straszak-Vishnoi'17].

Open questions:

1. Bounds are not tight. Can this view help.
2. Duality

Thank You!