①

Hierarchies & Lasserre/SOS

Given a problem, and an LP relaxation, if the relaxation is weak we can add more constraints in an adhoc way (or inspired by the bad examples, and the structure of the problem). But there are also ways to add constraints automatically. There are two ways of looking at it.

Given a combinatorial optimization problem (over the cube $\{0,1\}^n$), sps we have an LP relaxation. And suppose this relaxation has the property that the $K_{LP} \cap \{0,1\}^n =$ ~~Polytope of convex hull of solutions~~ set of feasible solutions. Then can generate feasible inequalities mechanically. Many ways. here are some

1. Gomory-Chvatal (cutting planes): if current constraints $\alpha^T x \geq \beta$ with $\alpha \in \mathbb{Z}^n$ then can infer $\alpha x \geq \lceil \beta \rceil$. So do this for base LP to get a "lift". How many times do we need to lift until get only convex hull of integer sol^{n}s?
   - Chvátal: $\exists$ a finite bound on the "rank"
   - Eisenbrand & Schulz: $O(n^2 \log n)$. for 0-1 polyhedra.
   - Problem: dont know how to take lifts in polytime

   [BTW, Matching polytope for non-bip = lift of base polytope 1 times].

2. Lovasz Schrijver / Sherali Adams / Lasserre.
   - All known to have rank n. Can be implemented with $O(n^{O(k)})$ time for k lifts.

   We'll talk about Sherali Adams (LP) & Lasserre/SOS (SDP).

~~Sherali Adams~~

~~Given constraint $a_i^T x \geq b$, multiplying this by $\prod_{i\in S} x_i \prod_{j\in T}(1-x_j)$~~

~~still gives ...~~

~~the bounds constraints are just per-variable constraints~~

Sherali-Adams:

What if we threw in constraints that are valid, ~~[illegible]~~ generated by the following steps :-

Adding constraints obtained by $\underbrace{\text{products of variables}}_{\text{non-linear}}$ + linearizing + consistency on marginals.

e.g. we introduce variables X_{ij} "=" $x_i \cdot x_j$

X_{ijk} "=" $x_i x_j x_k$ etc.

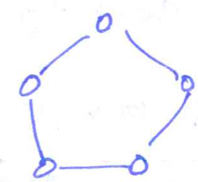
~~[illegible]~~ what constraints?

Obtained by taking $a^T x \geq b$ and multiplying by $\prod_{i \in S} x_i \prod_{j \in S} (1-x_j)$

and setting $x_i^2 = x_i$ (valid for 0-1 solutions) and linearizing.

Example: Stable set

$$\max \sum x_i$$
$$\text{st } x_i + x_j \leq 1 \quad \forall\, ij \in E$$
$$0 \leq x_i \leq 1 \quad \forall\, i \in V$$



$x_i = 1/2$ sol^n for all graphs.

Now multiply each constraint by x_i, and by $(1-x_i)$ etc.

so get ~~$(x_i+x_j)x_i$~~ $(x_i + x_j) x_i \leq x_i$

$\Leftrightarrow \cancel{x_i^2} + x_i x_j \leq \cancel{x_i} \Leftrightarrow x_i x_j = 0.$

(since $x_i^2 = x_i$).

Define vars X_{ij} and this says $X_{ij} = 0 \;\; \forall\, ij \in E$

but also $X_{ik} + X_{jk} \leq X_k$

and so have a variable X_S for each set S of size ≤ 2.

(Don't forget to use $0 \leq x_i \leq 1$ to get $\underbrace{0 \leq x_i \prod_{a \in S} x_a \prod_{b \in T} (1-x_b) \leq}_{}$)

~~[illegible]~~

Can do this for t-level SA

by taking $|S| + |T| \leq t$ and doing this operation

Does this ~~tally~~ help?

On the 5 ~~odd~~ cycle:

$$
\begin{array}{lll}
x_1+x_2 \le 1 & \cdot\, x_1 & \Rightarrow\ x_1^2 + \cancel{x_1x_2} \le x_1 \\
x_2+x_3 \le 1 & \cdot\, (1-x_1) & \quad x_2+x_3-\cancel{x_1x_2}-\cancel{x_1x_3} \le 1-x_1 \\
x_3+x_4 \le 1 & \cdot\, x_1 & \quad \cancel{x_1x_3} + \cancel{x_1x_4} \le x_1 \\
x_4+x_5 \le 1 & \cdot\, (1-x_1) & \quad x_4+x_5-\cancel{x_4x_1}-\cancel{x_5x_1} \le 1-x_1 \\
x_5+x_1 \le 1 & \cdot\, x_1 & \quad \cancel{x_1x_5} + x_1^2 \le x_1 \\
\hline
 & & x_1+x_2+x_3+x_4+x_5 \le 2 \ !!
\end{array}
$$

So the multiplying by $\prod_{i\in S} x_i \prod_{j\in T}(1-x_j)$ is not the useful bit, it's where you use $x_i^2 = x_i$ (only for 0-1 sol^{n}s).

And then linearize (which is lossy but necessary for poly time).

Intuition: get "local distributions" over sets of size $\le k$ (if we take k-levels) i.e.

$\exists$ a distribution $\mathcal{D}_S$ on set S of size $|S| \le k$.

s.t. $\Pr_{\mathcal{D}(S)}\left[\bigwedge_{i\in S'}(x_i=1)\right] = x_{S'}$ $\forall S' \subseteq S$.

And consistency also (i.e. if $S' \subseteq S\cap T$ then $\Pr_{\mathcal{D}(S)}\left[\bigwedge_{i\in S'}(x_i=1)\right] = \Pr_{\mathcal{D}(T)}\left[\bigwedge_{i\in S'}(x_i=1)\right]$.)

~~So in a perfect world get joint distributions over all n vars, i.e. sol^n sets~~

Imp: These distributions satisfy all constraints on $\le k$ vars. (and more) so locally "look like" solutions.

Of course, there may not be a global sol^n that agrees with this local solutions.
But maybe the integrality gap has improved.

Bad news: for Ind Set, does not capture clique constraints (which even the basic SDP, i.e. Lovász ϑ-function, captures) ☹

for Max Cut, integrality gap remains $2-\varepsilon$ even after $O_\varepsilon(n)$ rounds. ☹

Good news: Helps for bounded TW problems, (Ind set, Sparsest Cut etc).
and for other problems where "remembering small state" works.

Lasserre / SoS: Similar idea, but also SDP constraints.

e.g. for independent set max $\sum_i' \|X_i\|^2$

$$\text{st} \quad \langle X_i, X_j \rangle = 0 \quad \forall\, i \sim j$$

$$\langle X_S, X_T \rangle = \langle X_{S'}, X_{T'} \rangle \text{ if } S \cup T = S' \cup T'$$

$$\|X_\phi\|^2 = 1.$$

Says things like $\langle X_i, X_\phi \rangle = \langle X_i, X_i \rangle$

to scale down to 1.

and

$\langle X_i, X_{\{i \cup j\}} \rangle = \langle X_i, X_j \rangle$

all use that intended solutions satisfy $\boxed{x_i^2 = x_i}$

Gives a lot more power. Don't know bad examples any more. for lots of problems.

. In fact, all "standard" bad examples fail — for many problems.

. This is the optimal LP for constraint satisfaction problems (unless UGC fails).

Bad examples : k-XOR — linear rounds still cannot detect sat'able instances from unsat.

dynamic programming problems like knapsack not exactly solved.

But still — don't know how to use it, don't know bad examples, for many problems.

Here's a different perspective coming from polynomial optimization.

Want to solve $\max P(x)$

st $f(x) \geq 0 \quad \forall f \in F$ } semialgebraic set

$g(x) = 0 \quad \forall g \in G.$]

[This set denoted by K_{FG} is not convex in general, may not even be connected.

e.g. $\sum_i x_i^2 = 1$ (surface of unit ball) or $x_i^2 = x_i \ \forall i$ (Cube).]

Let's make it a convex problem by considering a measure μ on K_{FG}, now we want to optimize

$$\max \mathop{\mathbb{E}}_{x \sim \mu} [P(x)]$$

st μ is supported on $K_{F,G}$

Too many degrees of freedom! Suppose $P(x)$ has degree $\leq d$. Then suppose we say that $\mathop{\mathbb{E}}_{x \sim \mu} [x^\alpha] = L(x^\alpha)$ $\forall \alpha$ with $\|\alpha\|_1 \leq d$, $\in \mathbb{Z}^n_{\geq 0}$ and L is a linear map (so extends to all degree-d polynomials by $L(q(x)) = \sum_\beta q_\beta \cdot L(x^\beta)$.)

Then we want to maximize

(L is often called the "pseudoexpectation" in the jargon of the area.)

$\max L(p(x))$

st $L(1) = 1.$

L is a linear map

$\exists$ measure μ on K_{FG} that has moments $L(x^\alpha)$.

this is a hard problem. So let's relax the problem finally.

forall $f \in F$, add constraint $L(fq^2) \geq 0 \quad \forall q$ of degree $\leq \frac{d - \deg(f)}{2}$

forall $g \in G$, add constraint $L(gq) = 0 \quad \forall q$ of degree $\leq d - \deg(g)$.

"degree-d Lasserre"

so finally $\max L(p(x))$

st $L(1) = 1$ // think of this as $1 \geq 0$.

L linear $\quad L(q^2) \geq 0.$

$L(fq^2) \geq 0 \quad \forall q$ of degree $\leq \frac{d - \deg(f)}{2} \ \forall f \in F$

$L(gq) = 0 \quad \forall q$ —— $\leq d - \deg(g) \ \forall g \in G.$

- Claim: this is an SDP with $n^{O(d)}$ size.

Pf sketch: ~~the SOS prog~~

- For the gs, suffices to check $L(g(x)x^\beta)=0$. $=\sum_\alpha' g_\alpha \cdot L(x^{\alpha+\beta})=0$.

variables in an LP!

- For the fs, $L(f(x)q(x)^2)=\sum_{\beta,\gamma}'\sum_\alpha' f_\alpha\, q_\beta q_\gamma \cdot L(x^{\alpha+\beta+\gamma})\geq 0 \quad \forall q \in \mathbb{R}^{\binom{n}{\leq d}}$

$\iff \left(\sum_\alpha' f_\alpha L(x^{\alpha+\beta+\gamma})\right)_{\beta,\gamma} \succeq 0$. psdness.

Dual SOS degree-d

$$\begin{aligned} &\inf \lambda \\ \text{s.t. } & p=\lambda-\sigma+\omega \\ & \lambda \in \mathbb{R} \\ & \sigma \in Q_d(F\cup\{1\}) \\ & \omega \in I_d(G) \end{aligned}$$

[Notation: $Q_d(F)$ = "quadratic module" of degree $\leq d$ wrt F
$= \text{cone}(fq^2 : q \in \mathbb{R}_{\frac{d-\deg(f)}{2}}[\bar{x}])$
total degree smaller than d.

$I_d(G)$ = "ideal" of degree $\leq d$ wrt G
$= \text{span}(gq : q \in \mathbb{R}_{d-\deg(g)}[x])$]

Weak duality: given L, and (λ,σ,ω)

$$L(p(x)) = L(\lambda-\sigma+\omega) = \lambda - \underbrace{L(\sigma)}_{\geq 0} + \underbrace{L(\omega)}_{=0} \leq \lambda.$$

Note: in general, we could throw in more conditions, e.g.

$$L\Big(\big(\prod_{f\in I} f\big)q^2\Big)\geq 0 \quad \forall I\subseteq F.$$

But don't need this for "well conditioned" ("Archimedean") problems.

Before we give the convergence result of Lasserre (which says that as $d \to \infty$, the degree d relaxation and the degree-d SOS both converge to OPT).

Let's give some perspective

LP duality (Farkas' Lemma) says that

sps. $Ax \geq b$ implies $\alpha x \geq \beta$. (i.e. $\alpha x \geq \beta$ on all points in $K = \{x \mid Ax \geq b\}$

then we can write the inequality $\alpha x \geq \beta$ by positive ~~convex~~ linear combo of $\{a_i x \geq b_i\}_i$. also use $1 \geq 0$.

Equivalently, if $Ax \geq b$ is inconsistent, i.e. $\exists$ no sol^n to $Ax \geq b$.

then $\exists$ a linear combo of constraints that proves $-1 \geq 0$.

What about poly systems in general?

Positivstellensatz [Schmüdgen]

- $K_{F,G}$ be compact. then if $p(x)$ is strictly positive on $K_{F,G}$ then

$$p(x) = \underbrace{h_0(x)}_{\text{SOS}} + \underbrace{\sum g(x)\, q(x)}_{I(G)\text{ ideal members}} + \sum_{I \subseteq F} \Big(\prod_{f \in I} f(x)\Big) \underbrace{h_I(x)}_{\text{SOS}}$$

- or in the contrapositive form. if $K_{F,G}$ is empty then

$$-1 = \text{SOS} + \underset{I(G)}{\omega(x)} + \omega'(x)$$

of this form.

This form is called the Krivine-Stengle form and does not require compactness.

Positiv Stellensatz [Putinar]

- if $K_{F,G}$ is Archimedean, then don't need the products of the fs. i.e.. with the add'l Archimedean assumption, if $p(x) \geq 0$ on $K_{F,G}$

$$p(x) \in \cancel{\text{SOS}}\ I(G) + Q(F \cup \{1\})$$

Note: No bound on the degrees of the polynomials on the right-!! Even if all $p(x)$, $f(x)$, $g(x)$ all have low degrees.

What is Archimedean?

It means that the statement $\|x\|^2 \leq R$ can be proven in the same system.

That is $\boxed{\{R - \|x\|^2\} \in I(G) + Q(F \cup \{1\}).}$

this implies that $K_{F,G}$ is compact.

OK. So we have the theorems saying the system is sound and complete. But the degrees of the proofs may not be controlled. And the convergence theorem of Lasserre says pretty much that as the $d \to \infty$, the SOS / Lasserre systems converge to OPT.

So direction in TCS: ~~can be~~ what statements have low-degree proofs? (and use small numbers – see Ryan's paper) ~~that set can be~~

E.g. the GW max-cut result implies that we prove

[See Barak-Steurer notes]

$$\text{MAXCut} - 0{\cdot}878\ \text{SDP} \in \text{SOS. of small degree}$$

$$\Rightarrow \text{SDP} \geq \frac{\text{MaxCut}}{0{\cdot}878}$$ and so integrality gap is not too large.

- Can we show that gap gets better with higher degree SOS?
- What limits to this general technique?
