

Hierarchies 2 Lasserre / SOS

Given a problem, and an LP relaxation, if the relaxation is weak we can add more constraints in an ad hoc 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 $\{0,1\}^n$), s.t. we have an LP relaxation. And suppose this relaxation has the property that the $K_{LP} \cap \{0,1\}^n =$ ~~Polytope~~ ~~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 $\models \alpha x \geq \beta$ 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's?

- Chvatal: $\exists$ a finite bound on the "rank"
- Eisenbrand & Schulz: $O(n^2 \log n)$. for 0-1 polyhedra.
- Problem: don't know how to take lifts in polytime

[BTW, Matching polytope = lift of base polytope¹ times]
for non-bip

② Lovasz Schrijver / Sherali Adams / Lasserre.

- All known to have rank n . Can be implemented into $2^{O(n^k)}$ time for k lifts.
- We'll talk about Sherali Adams (LP) & Lasserre / SOS (SDP).

Sherali Adams

~~Given constraint $\sum_{i \in S} x_i \geq \beta$, multiplying this by $\prod_{i \in T} x_i$ gives~~

~~still gives $\sum_{i \in S} x_i \geq \beta$~~

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

Sherali-Adams:

What if we throw in constraints that are valid, ~~into~~ generated by the following steps, -

Adding constraints obtained by products of variables + linearizing + consistency on marginals.
non-linear

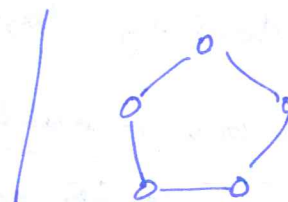
e.g. we introduce variables $X_{ij} = x_i \cdot x_j$

$X_{ijk} = x_i \cdot x_j \cdot x_k$ etc.

~~what~~ what constraints?

Obtained by taking $a^T x \geq b$ and multiply by $\prod_{i \in S} x_i \prod_{j \in T} (1-x_j)$
and setting $x_i^2 = x_i$ (valid for 0-1 solutions) and linearizing.

Example: Stable set $\max \sum x_i$
st $x_i + x_j \leq 1 \quad \forall ij \in E$
 $0 \leq x_i \leq 1 \quad \forall i \in V$



$x_i = 1/2$ solⁿ for all graphs.

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

so get ~~$(x_i + x_j) x_i$~~ $(x_i + x_j) x_i \leq x_i$
 $\Leftrightarrow x_i^2 + x_i x_j \leq x_i \Leftrightarrow x_i x_j = 0$.
(since $x_i^2 = x_i$).

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

but also $x_i x_j + x_j x_k \leq x_k$

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

(Don't forget to use $0 \leq x_i \leq 1$ to get

$$0 \leq x_i \prod_{a \in S} x_a \prod_{b \in T} (1-x_b) \leq$$

these ~~tip~~ ~~value~~ ~~that~~

Can do this for t -level SA

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

Does this ~~help~~ help?

On the 5-cycle:

$$x_1 + x_2 \leq 1 \quad \cdot x_1$$

$$x_2 + x_3 \leq 1 \quad \cdot (1-x_1)$$

$$x_3 + x_4 \leq 1 \quad \cdot x_1$$

$$x_4 + x_5 \leq 1 \quad \cdot (1-x_1)$$

$$x_5 + x_1 \leq 1 \quad \cdot x_1$$

$$\Rightarrow (x_1^2) + x_1 x_2 \leq (x_1)$$

$$x_2 + x_3 - x_1 x_2 - x_1 x_3 \leq 1 - x_1$$

$$x_1 x_3 + x_1 x_4 \leq x_1$$

$$x_4 + x_5 - x_1 x_4 - x_1 x_5 \leq 1 - x_1$$

$$x_1 x_5 + (x_1^2) \leq (x_1)$$

$$\frac{2}{5}x_1 + x_2 + x_3 + x_4 + x_5 \leq 2 \quad !!$$

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

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

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

~~For~~ For a distribution ρ_S on set S of size $|S| \leq k$.

$$\text{st. } \Pr_{\rho(S)} \left[\bigwedge_{i \in S'} (x_i = 1) \right] = x_{S'} \quad \forall S' \subseteq S.$$

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

~~So in the end we get joint distributions over all n vars. i.e. sol's~~

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

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

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

for MaxCut, integrality gap remains $2-\epsilon$ even after $O_\epsilon(n)$ rounds. ☹

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

Looserre / SOS: Similar idea, but also SDP constraints.

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

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

$$\langle x_S, x_T \rangle = \langle x_{S'}, x_{T'} \rangle \text{ if } S \cup T = S' \cup T'$$

$$\|x_\emptyset\|^2 = 1.$$

Says things like $\langle x_i, x_\emptyset \rangle = \langle x_i, x_i \rangle$

to scale down to 1.

and

$$\langle x_i, x_{\{i,j\}} \rangle = \langle x_i, x_j \rangle$$

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 $f(x) \geq 0 \quad \forall f \in F$
 $g(x) = 0 \quad \forall g \in G.$ } semialgebraic set

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

eg. $\sum x_i^2 = 1$ (surface of unit ball) or $x_i^2 = x_i \quad \forall i$ (Cube).]

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

optimize $\max_{\mu} \mathbb{E}[P(x)]$

st μ is supported on K_{FG}

Too many degrees of freedom! Suppose $P(x)$ has degree $\leq d$. Then suppose we say

that $\mathbb{E}[x^\alpha] = L(x^\alpha)$ and L is a linear map (so extends to all degree d polynomials by
 $L(q(x)) = \sum_B q_B \cdot L(x^B)$)
 $\forall \alpha$ with $\|\alpha\|_1 \leq d$
 $\in \mathbb{Z}_{\geq 0}^n$

Then we want to maximize

$\max L(P(x))$

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.

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

"degree- d
Lasserre"

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

st $L(1) = 1$

L linear

$L(fq^2) \geq 0$

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

// think of this as $1 \geq 0$.

$L(gq) = 0$

$\forall q$ —

$\leq \frac{d - \deg(f)}{2} \quad \forall f \in F$

$\leq d - \deg(g) \quad \forall g \in G.$

• Claim: this is an SDP with n^{ord} size.

Pf sketch: ~~the SOS proof~~

• For the g 's, suffices to check $L(g(x)^p) = 0. = \sum_{\alpha} g_{\alpha} \cdot L(x^{\alpha+p}) = 0.$
 $\uparrow$ variables in an LP!

• For the f 's, $L(f(x)q(x)^2) = \sum_{\beta, \gamma} f_{\alpha} q_{\beta} q_{\gamma} \cdot L(x^{\alpha+\beta+\gamma}) \geq 0 \quad \forall q \in \mathbb{R}^{\binom{2d}{2}}$
 $\Leftrightarrow \left(\sum_{\alpha} f_{\alpha} L(x^{\alpha+p+\delta}) \right)_{\beta, \gamma} \succeq 0.$ PSDness.

Dual SOS
degree-d

inf λ
s.t. $p = \lambda - \vartheta + \omega$
 $\lambda \in \mathbb{R}$
 $\vartheta \in Q_d(FU\mathbb{B})$
 $\omega \in I_d(G)$

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

Weak duality: given L , and $(\lambda, \vartheta, \omega)$

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

Note: ~~we~~ in general, we could throw in more conditions, eg

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

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

Before we give the convergence result of Lasserre (which says that as $d \rightarrow \infty$, 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. $\nexists$ no solⁿ to $Ax \geq b$,
 then $\exists$ a linear combo of constraints that proves $-1 \geq 0$.

What about poly systems in general?

Positivstellensatz [Schnüdelgen]

• $K_{F,G}$ be compact. then if $f(x)$ is positive on $K_{F,G}$ then

$$f(x) = \underbrace{\sum h_i^2(x)}_{\text{SOS}} + \underbrace{\sum g(x) q(x)}_{I(G) \text{ ideal members}} + \sum_{I \in F} \left(\prod_{f \in I} f(x) \right) \underbrace{h_I(x)}_{\text{SOS}}$$

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

$$-1 = \text{SOS} + \underbrace{\sum_{g \in I(G)} w_g(x)}_{I(G)} + c'(x) \quad \text{of this form.}$$

Positivstellensatz [Putinar]

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

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

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

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 $d \rightarrow \infty$, the SOS/Lasserre systems converge to OPT.

~~What is Archimedean?~~
So direction in TCS: ~~can we~~ 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

MAXCut - 0.878 SDP $\in$ SOS. of small degree

[See David Steurer's notes]

$\Rightarrow$ SDP $\geq \frac{\text{MaxCut}}{0.878}$ 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?