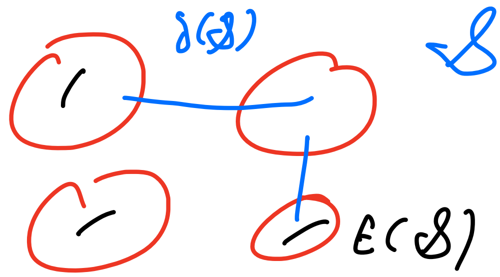
More SDPs:

Today: examples where QP formulation isn't obvious or doesn't exist



Correlation Clustering:

Input: - $G=(V,E)$
- $w^-: E \to \mathbb{R}^+$
- $w^+: E \to \mathbb{R}^+$

Feasible solution: Partition $\mathcal{S}$ of V

$\delta(\mathcal{S})$ = edges go between parts of $\mathcal{S}$

$E(\mathcal{S})$ = edges where both endpoints in same part of $\mathcal{S}$

Objective: $\max W(\mathcal{S}) = \max \sum_{e \in E(\mathcal{S})} w^+(e) + \sum_{e \in \delta(\mathcal{S})} w^-(e)$

Simple 1/2-approximation: return better of $\mathcal{S}_1 = \{V\}$, $\mathcal{S}_2 = \{\{i\} : i \in V\}$

value of $\mathcal{S}_1 = W(\mathcal{S}_1) = \sum_{e \in E} w^+(e)$

value of $\mathcal{S}_2 = W(\mathcal{S}_2) = \sum_{e \in E} w^-(e)$

Every partition $\mathcal{S}$ has $W(\mathcal{S}) \leq \sum_{e \in E} w^+(e) + \sum_{e \in E} w^-(e)$

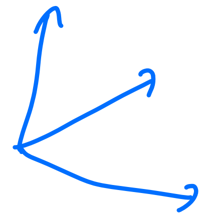
$\Rightarrow \max(W(\mathcal{S}_1), W(\mathcal{S}_2)) \geq \frac{1}{2} \cdot OPT$

Vector Programming Formulation:

$$\max \sum_{\{i,j\}\in E} \left(w^+(i,j)\langle v_i, v_j\rangle + w^-(i,j)(1-\langle v_i,v_j\rangle)\right)$$

s.t. $v_i \in \{e_1, e_2, \dots, e_n\}$ $\quad \forall i\in[n]$

standard basis vectors



Thm: This is an exact formulation

Pf: Let $\mathcal{S} = (S_1, S_2, \dots, S_k)$ partition of V

For $j\in[k]$ and $i \in S_j$, set $v_i = e_j$

$$\Rightarrow \sum_{\{i,j\}\in E} \left(w^+(i,j)\langle v_i, v_j\rangle + w^-(i,j)(1-\langle v_i,v_j\rangle)\right)$$

$$= \sum_{\{i,j\}\in E(\mathcal{S})} w^+(i,j) + \sum_{\{i,j\}\in\delta(\mathcal{S})} w^-(i,j) = w(\mathcal{S})$$

Other direction: Let $\{v_i\}$ solution to vector program.

$\forall j\in[n]$, let $S_j = \{i\in V: v_i = e_j\}$

$\Rightarrow \{S_1,S_2,\dots,S_n\}$ form partition of value = vector program objective

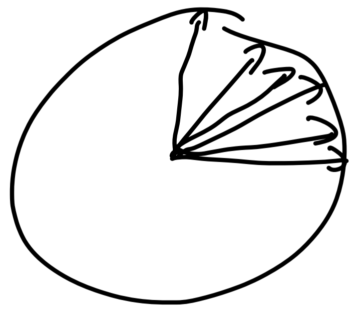
Relax to SDP:

$$\max \sum_{\{i,j\} \in E} \left(w^+(i,j) \langle v_i, v_j \rangle + w^-(i,j)(1 - \langle v_i, v_j \rangle) \right)$$

$$\text{s.t.} \quad \langle v_i, v_i \rangle = 1 \qquad \forall i \in V \qquad \text{(unit vectors)}$$

$$\langle v_i, v_j \rangle \geq 0 \qquad \forall i,j \in V$$

$$v_i \in \mathbb{R}^n \qquad \forall i \in V$$



Rounding: two random hyperplanes

- Let r_1, r_2 unit vectors drawn independently, uniformly at random from set of all unit vectors

- Let $R_1 = \{ i \in V : \langle r_1, v_i \rangle \geq 0, \ \langle r_2, v_i \rangle \geq 0 \}$

$R_2 = \{ i \in V : \langle r_1, v_i \rangle \geq 0, \ \langle r_2, v_i \rangle < 0 \}$

$R_3 = \{ i \in V : \langle r_1, v_i \rangle < 0, \ \langle r_2, v_i \rangle \geq 0 \}$

$R_4 = \{ i \in V : \langle r_1, v_i \rangle < 0, \ \langle r_2, v_i \rangle < 0 \}$

- Return $\mathcal{S} = \{ R_1, R_2, R_3, R_4 \}$

$$X_{ij} = \begin{cases} 1 & \text{if } i,j \text{ in same cluster} \\ 0 & \text{otherwise} \end{cases}$$

Last time:

$$\Pr[\text{random hyperplane separates } i,j] = \frac{\theta_{ij}}{\pi}$$

$$\Rightarrow E[X_{ij}] = \left(1 - \frac{\theta_{ij}}{\pi}\right)^2$$

$$\Rightarrow E[w(\mathcal{S})] = E\Big[\sum_{\{i,j\} \in E(\mathcal{S})} w^+(i,j) + \sum_{\{i,j\} \in \delta(\mathcal{S})} w^-(i,j)\Big]$$

$$= E\Big[\sum_{\{i,j\} \in E} \big(w^+(i,j)\, X_{ij} + w^-(i,j)\,(1 - X_{ij})\big)\Big]$$

$$= \sum_{\{i,j\} \in E} \big(w^+(i,j)\, E[X_{ij}] + w^-(i,j)(1 - E[X_{ij}])\big)$$

$$= \sum_{\{i,j\} \in E} \left(w^+(i,j)\left(1 - \frac{\theta_{ij}}{\pi}\right)^2 + w^-(i,j)\left(1 - \left(1 - \frac{\theta_{ij}}{\pi}\right)^2\right)\right)$$

Trig facts:
$$\left.\begin{aligned} \left(1 - \frac{\theta_{ij}}{\pi}\right)^2 &\geq \frac{3}{4} \cos \theta_{ij} \\ 1 - \left(1 - \frac{\theta_{ij}}{\pi}\right)^2 &\geq \frac{3}{4}(1 - \cos \theta_{ij}) \end{aligned}\right\} \text{ if } \theta_{ij} \leq \frac{\pi}{2}$$

SDP constraint $\langle v_i, v_j \rangle \geq 0 \Rightarrow \theta_{ij} \leq \frac{\pi}{2}$

$$\geq \sum_{\{i,j\} \in E} \left(w^+(i,j) \cdot \frac{3}{4} \cos \theta_{ij} + w^-(i,j) \cdot \frac{3}{4}(1 - \cos \theta_{ij})\right)$$

$$= \frac{3}{4} \sum_{\{i,j\} \in E} \left(w^+(i,j) \langle v_i, v_j \rangle + w^-(i,j) \left(1 - \langle v_i, v_j \rangle\right) \right)$$

$$= \frac{3}{4} \cdot OPT(SDP)$$

$\Rightarrow \frac{3}{4}$-approximation

Max-2SAT:

Input: - Variables $x_1, x_2, \dots, x_n$

- CNF clauses $C_1, C_2, \dots, C_m$, each with 2 literals

$(x_1 \vee \bar{x_7})$, $(\bar{x_2} \vee x_5)$, $(x_3 \vee x_1)$, ...

Feasible solution: Assignment $\{x_1, \dots, x_n\} \to \{T, F\}$

Objective: max # satisfied (true) clauses

Note: 2SAT not NP-hard, but Max-2SAT is!

Want to write strict quadratic program, where

$y_i = -1$ corresponds to $x_i = T$

$y_i = 1$ corresponds to $x_i = F$

Problem: Symmetry!

$C = x_i \vee \overline{x_j} \Rightarrow y_i = -1,\ y_j = 1$ should be ok, but

$y_i = 1,\ y_j = -1$ should not.

But strictness means can only look at products!

Solution: add "dummy variable" $y_T \in \{-1, 1\}$

$y_i = y_T$ corresponds to $x_i = T$

$y_i = -y_T$ corresponds to $x_i = F$

Consider clause $x_i \vee x_j$, values y_i, y_j, y_T:

$$\frac{3 + y_i y_T + y_j y_T - y_i y_j}{4} = \begin{cases} 0 & \text{if } y_i = y_j = -y_T \\ 1 & \text{otherwise} \end{cases}$$

clause $x_i \vee \overline{x_j}$, values y_i, y_j, y_T: same thing, negate y_j

$$\frac{3 + y_i y_T - y_j y_T + y_i y_j}{4} = \begin{cases} 0 & \text{if } y_i = -y_T,\ y_j = y_T \\ 1 & \text{otherwise} \end{cases}$$

Clause $\overline{x_i} \vee x_j$, values y_i, y_j, y_T:

$$\frac{3 - y_i y_T + y_j y_T + y_i y_j}{4} = \begin{cases} 0 & \text{if } y_i = y_T,\ y_j = -y_T \\ 1 & \text{otherwise} \end{cases}$$

Clause $\overline{x_i} \vee \overline{x_j}$, values y_i, y_j, y_T:

$$\frac{3 - y_i y_T - y_j y_T - y_i y_j}{4} = \begin{cases} 0 & \text{if } y_i = y_j = y_T \\ 1 & \text{otherwise} \end{cases}$$

<u>Strict Quadratic Programming Formulation:</u>

$$\max \sum_{\text{clauses } x_i \vee x_j} \frac{3 + y_i y_T + y_j y_T - y_i y_j}{4} + \sum_{\text{clauses } x_i \vee \overline{x_j}} \frac{3 + y_i y_T - y_j y_T + y_i y_j}{4}$$

$$+ \sum_{\substack{\text{clauses} \\ \overline{x_i} \vee x_j}} \frac{3 - y_i y_T + y_j y_T + y_i y_j}{4} + \sum_{\substack{\text{clauses} \\ \overline{x_i} \vee \overline{x_j}}} \frac{3 - y_i y_T - y_j y_T - y_i y_j}{4}$$

s.t. $\quad y_i^2 = 1 \qquad \forall i \in [n]$

$\quad y_T^2 = 1$

Relax to SDP:

$$\max \sum_{\substack{\text{clauses}\\ x_i \vee x_j}} \frac{3 + \langle v_i, v_T\rangle + \langle v_j, v_T\rangle - \langle v_i, v_j\rangle}{4} + \sum_{\substack{\text{clauses}\\ x_i \vee \overline{x_j}}} \frac{3 + \langle v_i, v_T\rangle - \langle v_j, v_T\rangle + \langle v_i, v_j\rangle}{4}$$

$$+ \sum_{\substack{\text{clauses}\\ \overline{x_i} \vee x_j}} \frac{3 - \langle v_i, v_T\rangle + \langle v_j, v_T\rangle + \langle v_i, v_j\rangle}{4} + \sum_{\substack{\text{clauses}\\ \overline{x_i} \vee \overline{x_j}}} \frac{3 - \langle v_i, v_T\rangle - \langle v_j, v_T\rangle - \langle v_i, v_j\rangle}{4}$$

s.t. $\langle v_i, v_i\rangle = 1 \quad \forall i \in [n]$

$v_i \in \mathbb{R}^n \quad \forall i \in [n]$

$\langle v_T, v_T\rangle = 1$

$v_T \in \mathbb{R}^n$

Relaxation $\Rightarrow$ OPT(SDP) $\geq$ OPT

Rounding: random hyperplane!

- Choose unit vector $r \in \mathbb{R}^n$ u.a.r.

- Set $x_i = \begin{cases} T & \text{if } \operatorname{sign}(\langle r, v_i\rangle) = \operatorname{sign}(\langle r, v_T\rangle) \\ F & \text{if } \operatorname{sign}(\langle r, v_i\rangle) \neq \operatorname{sign}(\langle r, v_T\rangle) \end{cases}$

Consider clause $x_i \vee x_j$ (other types similar)

$$\frac{3 + \langle v_i, v_T\rangle + \langle v_j, v_T\rangle - \langle v_i, v_j\rangle}{4} = \frac{1}{4}\Big((1 + \langle v_i, v_T\rangle) + (1 + \langle v_j, v_T\rangle) + (1 - \langle v_i, v_j\rangle)\Big)$$

$\Rightarrow$ all terms $(1 \pm \langle v_k, v_l\rangle)$ for some k, l (possibly $= T$)

Analyze each term

Recall $\alpha_{GW} = \inf_{0 \le \theta \le \pi} \frac{2\theta}{\pi(1-\cos\theta)}$

Consider $1 - \langle v_k, v_l \rangle$

$\Rightarrow$ contribution to SDP is $1 - \langle v_k, v_l \rangle = 1 - \cos\theta_{kl}$

$\Pr[v_k, v_l \text{ separated by hyperplane}] = \frac{\theta_{kl}}{\pi}$

$\Rightarrow E[1 - y_k y_l] = 2 \cdot \Pr[v_k, v_l \text{ separated}] = 2 \frac{\theta_{kl}}{\pi} \ge \alpha_{GW}(1 - \cos\theta_{kl})$

$\Rightarrow$ in expectation, rounded solution gets $\ge \alpha_{GW} \cdot SDP$

Consider $1 + \langle v_k, v_l \rangle$

$\Rightarrow$ contribution to SDP $= 1 + \cos\theta_{kl}$

Contribution to rounded solution: $E[1 + y_k y_l] = 2\left(1 - \frac{\theta_{kl}}{\pi}\right)$

$\Rightarrow$ ratio b/w contributions $= \dfrac{2\left(1 - \frac{\theta_{kl}}{\pi}\right)}{1 + \cos\theta_{kl}} = \dfrac{2(\pi - \theta_{kl})}{\pi(1 + \cos\theta_{kl})}$

Let $\theta' = \pi - \theta_{kl} \Rightarrow = \dfrac{2\theta'}{\pi(1 - \cos\theta')} \ge \alpha_{GW}$

$\cos(\pi - \theta) = -\cos\theta$

$\Rightarrow$ rounded sol. $\ge \alpha_{GW} \cdot SDP$