Chapter V Relaxation of quadratic programs

§ 1 Generic scheme

<u>Def 1</u> A <u>quadratic program</u> (QP) is of the following form

$$qp = \sup \; x^T Q_0 x + b_0^T x + \alpha_0$$
$$\text{such that } x \in \mathbb{R}^n$$
$$x^T Q_j x + b_j^T x + \alpha_j = 0 \quad \text{for } j \in [m],$$

where $Q_0, Q_1, \dots, Q_m \in S^n$, $b_0, \dots, b_m \in \mathbb{R}^n$, $\alpha_0, \dots, \alpha_m \in \mathbb{R}$.

Can model many optimization problems as QPs, e.g. problems with 0/1-variables

$$x_j^2 - x_j = 0$$

or with $-1/+1$-variables

$$x_j^2 - 1 = 0.$$

QPs are generally not convex. It is also often NP-hard to find optimal solutions of QPs ($\rightarrow$ V.2).

Goal: Compute upper bounds for qp efficiently (using SDP).

Theorem 2

$$qp \leq \inf y_0$$
$$\text{s.t. } y_0, y_1, \dots, y_m \in \mathbb{R}$$
$$y_0 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \sum_{j=1}^{m} y_j \begin{bmatrix} \alpha_j & \frac{1}{2} b_j^T \\ \frac{1}{2} b_j & Q_j \end{bmatrix} - \begin{bmatrix} \alpha_0 & \frac{1}{2} b_0^T \\ \frac{1}{2} b_0 & Q_0 \end{bmatrix} \in S_+^{n+1}$$

Proof We have

$$x^T Q_j x + b_j^T x + \alpha_j = \left\langle \begin{bmatrix} \alpha_j & \frac{1}{2} b_j^T \\ \frac{1}{2} b_j & Q_j \end{bmatrix}, \underbrace{\begin{bmatrix} 1 & x^T \\ x & xx^T \end{bmatrix}}_{= \begin{bmatrix} 1 \\ x \end{bmatrix} [1 \; x^T] \in S_+^{n+1}} \right\rangle$$

So:

$$qp \leq \sup \left\langle \begin{bmatrix} \alpha_0 & \frac{1}{2} b_0^T \\ \frac{1}{2} b_0 & Q_0 \end{bmatrix}, X \right\rangle$$
$$\text{s.t. } X \in S_+^{n+1}, \quad X_{00} = 1,$$
$$\left\langle \begin{bmatrix} \alpha_j & \frac{1}{2} b_j^T \\ \frac{1}{2} b_j & Q_j \end{bmatrix}, X \right\rangle = 0 \quad \text{for } j \in [m].$$

Dualising this SDP and applying weak duality yields the statement of the theorem. $\boxtimes$

<u>Remarks</u> • If we would impose the condition "rank $X = 1$", then $qp = \sup \ldots$ One <u>relaxes</u> the rank condition in Thm. 2.

• Advantages of SDP: – efficiently computable (by ellipsoid method, IPM)

– any feasible solution provides an upper bound for qp.

[lower bounds can come from feasible solutions which one can find heuristically].

§2 The Goemans-Williamson approximation algorithm for MAXCUT

<u>Def. 1</u> Let $G=(V,E)$ be an undirected graph with weight function $w=(w_{ij})\in\mathbb{R}_+^E$ on the edges. A subset $S\subseteq V$ defines a <u>cut</u> $\delta(S)\subseteq E$ by

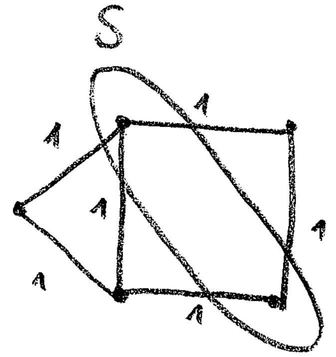
$$\delta(S)=\{\{i,j\}\in E : |\{i,j\}\cap S|=1\}.$$

The <u>weight</u> of $\delta(S)$ is

$$w(\delta(S))=\sum_{\{i,j\}\in\delta(S)} w_{ij}.$$

The <u>MAXCUT problem</u> is

$$mc(G,w)=\max_{S\subseteq V} w(\delta(S))$$



$mc(G,w)=5.$

<u>Facts</u>: • Unlike the MINCUT problem, the MAXCUT problem is NP-hard (Karp, 1972)

• It is even APX-hard: If $P\neq NP$, then one cannot approximate MAXCUT with a factor better than $\frac{16}{17}\approx 0{,}941$ in polynomial time (Håstad, 2001)

Formulation of MAXCUT as QP

$$mc(G,w) = \max \frac{1}{2} \sum_{\{i,j\} \in E} w_{ij} (1 - x_i x_j)$$
$$x \in \mathbb{R}^V$$
$$x_i^2 - 1 = 0 \quad \text{for } i \in V.$$

<u>Now</u>: Consider SDP relaxation and analyze its quality.

Since the MAXCUT QP does not have linear terms, we can use an SDP relaxation of size $|V|$, instead of $|V|+1$.

$$sdp(G,w) = \max \frac{1}{2} \sum_{\{i,j\} \in E} w_{ij} (1 - X_{ij})$$
$$X \in S_+^V$$
$$X_{ii} - 1 = 0 \quad \text{for } i \in V.$$

<u>Clear</u>: $mc(G,w) \leq sdp(G,w)$

<u>Proof</u>: Every $x \in \{-1,+1\}^V$ is a feasible solution for the MAXCUT QP. Define $X = xx^T$. This is a feasible solution for the SDP relaxation with $\frac{1}{2} \sum_{ij \in E} w_{ij} (1 - x_i x_j) = \frac{1}{2} \sum_{ij \in E} w_{ij} (1 - X_{ij})$. $\boxtimes$

<u>Remark</u> If X feasible for SDP relaxation and rank $X = 1$, then X is of the form $X = xx^T$ with $x \in \{-1, +1\}^V$ ($\to$ Exercises).

How good is the SDP relaxation?

<u>Theorem 2</u> (Goemans, Williamson, 1995)
For $G = (V, E)$, $w \in \mathbb{R}_+^E$ we have

$$\mathrm{sdp}(G, w) \geq \mathrm{mc}(G, w) \geq 0{,}878\ldots \, \mathrm{sdp}(G, w).$$

<u>Proof</u> We prove the inequality with the help of an algorithm. It yields a randomized polynomial time algo. to approximate MAXCUT within a factor of $0{,}878\ldots$

<u>Algorithm</u>

1. Solve $\mathrm{sdp}(G, w)$. Let X be an optimal solution.
2. Perform Cholesky decomposition of X. This gives normal vectors $v_i \in \mathbb{R}^V$, $i \in V$, $\|v_i\| = 1$ with

$$X_{ij} = v_i^T v_j \quad \text{for } i, j \in V.$$

Randomized rounding:

3. Choose $r \in \mathbb{R}^V$, $\|r\| = 1$, randomly from the probability distribution on the unit sphere which is invariant under orthogonal transformations ($\rightarrow$ Haar measure)

4. Define a cut $\delta(S)$ by
$$i \in S \iff \operatorname{sign}(v_i^T r) \geq 0.$$

That means
$$x_i = \begin{cases} \operatorname{sign}(v_i^T r) & , \text{ if } \operatorname{sign}(v_i^T r) \neq 0 \\ 1 & , \text{ otherwise.} \end{cases}$$

Claim The expected value
$$\mathbb{E}[w(\delta(S))] \geq 0{,}878\ldots \operatorname{sdp}(G, w).$$

Lemma 3 (Grothendieck's identity)

Let $u, v \in \mathbb{R}^d$ be vectors with $\|u\| = \|v\| = 1$.
Let $r \in \mathbb{R}^d$ be a random vector with $\|r\| = 1$ chosen from the $O(d)$-invariant probability distribution on the unit sphere S^{d-1}. Then,

(i) $\mathbb{P}[\mathrm{sign}(u^T r) \neq \mathrm{sign}(v^T r)] = \dfrac{\arccos(u^T v)}{\pi}$

(ii) $\mathbb{E}[\mathrm{sign}(u^T r)\,\mathrm{sign}(v^T r)] = \frac{2}{\pi} \arcsin(u^T v)$.

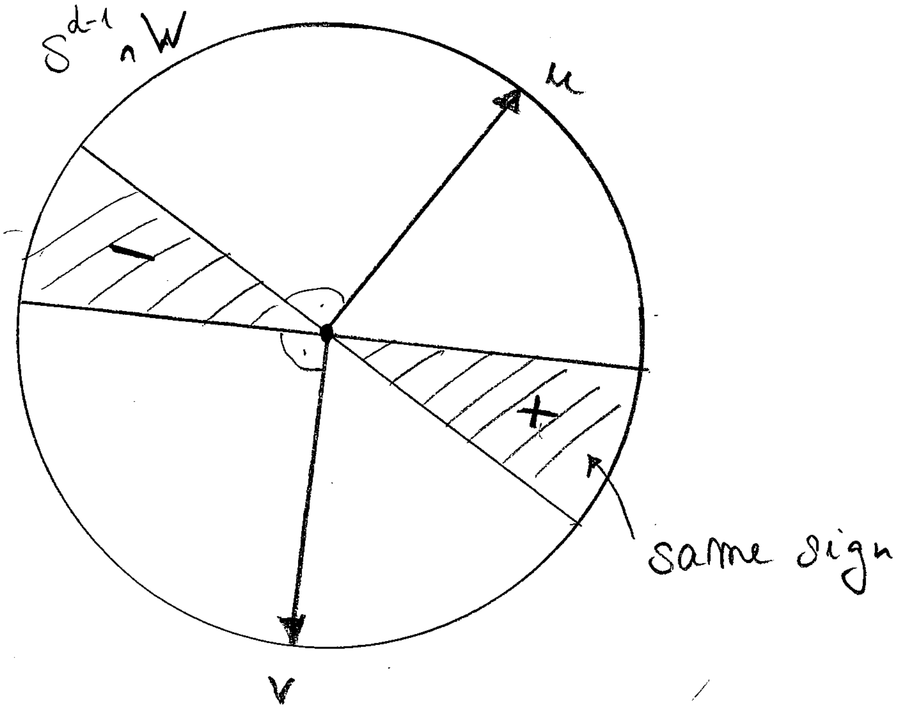
Proof (i) If $u = v$, then $\arccos(u^T v) = 0$,
if $u = -v$, then $\arccos(u^T v) = \pi$.

In both cases, (i) is obvious.

Now assume $W = \mathrm{span}\{u, v\}$ has dimension 2.
Project r orthogonally onto W. This gives $s \in W$ such that $u^T r = u^T s$, and $v^T r = v^T s$, and $\frac{s}{\|s\|}$ is uniformly distributed on the unit circle $S^{d-1} \cap W$ by the $O(d)$-invariance of the prob. distribution.

Hence,

$$\mathbb{P}[\operatorname{sign}(u^T r) \neq \operatorname{sign}(v^T r)] = 2 \cdot \frac{1}{2\pi} \arccos(u^T v)$$



(ii) $\mathbb{E}[\operatorname{sign}(u^T r)\operatorname{sign}(v^T r)]$

$$= (+1) \cdot \mathbb{P}[\operatorname{sign}(u^T r) = \operatorname{sign}(v^T r)]$$
$$+ (-1) \cdot \mathbb{P}[\operatorname{sign}(u^T r) \neq \operatorname{sign}(v^T r)]$$
$$= 1 - 2\,\mathbb{P}[\operatorname{sign}(u^T r) \neq \operatorname{sign}(v^T r)]$$
$$= 1 - \frac{2}{\pi} \arccos(u^T v)$$
$$= \frac{2}{\pi} \arcsin(u^T v)$$ because $\arcsin t + \arccos t = \frac{\pi}{2}$.

$\boxtimes$

Back to the proof of Thm. 2:

Let X be an optimal solution of $sdp(G,w)$, and $v_i \in \mathbb{R}^V$ so that $X_{ij} = v_i^T v_j$ as constructed in Step 2. Consider the cut $\delta(S)$ defined in Step 4.

Then:

$$\mathbb{E}[w(\delta(S))] = \sum_{\{i,j\}\in E} w_{ij}\, \mathbb{P}[\{i,j\} \in \delta(S)]$$

$$= \sum_{\{i,j\}\in E} w_{ij}\, \mathbb{P}[\operatorname{sign}(v_i^T r) \neq \operatorname{sign}(v_j^T r)]$$

$$= \sum_{\{i,j\}\in E} w_{ij} \frac{\arccos(v_i^T v_j)}{\pi}$$

$$= \sum_{\{i,j\}\in E} w_{ij} \left(\frac{1-v_i^T v_j}{2}\right) \cdot \underbrace{\left(\frac{2}{\pi} \frac{\arccos(v_i^T v_j)}{1-v_i^T v_j}\right)}_{\geq 0{,}878\ldots}$$

$$\geq 0{,}878\ldots\, sdp(G,w).$$

$$\left[\text{because } \min_{t\in[-1,1]} \frac{2}{\pi}\frac{\arccos t}{1-t} = 0{,}878\ldots\right]$$

Since $mc(G,w) \geq \mathbb{E}[w(\delta(S))]$, we proved Thm. 2. $\boxtimes$

Remarks :

- probability distribution on unit sphere S^{d-1} invariant under $O(d-1)$ is uniquely determined.

Realisation: Choose $x_1, \dots, x_d \in \mathbb{R}$, $x_i \sim N(0,1)$ from normal distribution (density function $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$), then $v = \frac{1}{\|x\|} \begin{pmatrix} x_1 \\ \vdots \\ x_d \end{pmatrix} \in S^{d-1}$

is chosen randomly according to the $O(d)$-invariant prob. distribution.

- one can derandomize the algorithm (but then it becomes ugly).

- approximation factor might be optimal ($\rightarrow$ Khot's unique games conjecture).