

An Introduction to Intersection Cuts and Their Applications

Akang Wang

*Dept. of Chemical Engineering
Carnegie Mellon University*

*PSE Seminar
November 30, 2018*

Outline

☐ Intersection Cuts

- Problem Definition
- Derivation
- Geometric Interpretation

☐ Applications

- Mixed Integer Linear Programming
- Reverse Convex Programming
- Polynomial Programming

☐ Comments

Problem Definition

□ Optimization problem:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & c^T x \\ \text{subject to} & x \in P \cap Q \end{array}$$

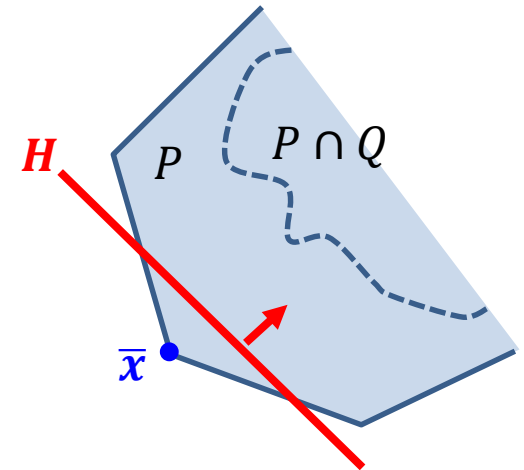
$P := \{x \in \mathbb{R}^n, Ax \leq b, x \geq 0\}$ is a polyhedral set, where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$
 $Q \subseteq \mathbb{R}^n$ represents a non-convex, “complicated” set, such as integrality, reverse convex, etc.

□ A polyhedral relaxation:

$$\begin{array}{ll} \underset{x}{\text{minimize}} & c^T x \\ \text{subject to} & Ax \leq b \\ & x \geq 0 \end{array}$$

Assume the LP optimality is achieved at \bar{x}

Q: How to generate a valid cut H such that $P \cap Q \subseteq H$ and $\bar{x} \notin H$?



Standard Form of an LP

- Introduce **slack** variables **s** and let $t := (x, s)$ represent the variables in LP for convenience

$$\begin{array}{ll} \underset{x, s}{\text{minimize}} & c^T x \\ \text{subject to} & Ax + s = b \\ & x, s \geq 0 \end{array} \longrightarrow \begin{array}{ll} \underset{t}{\text{minimize}} & \tilde{c}^T t \\ \text{subject to} & \tilde{A}t = b \\ & t \geq 0 \end{array}$$

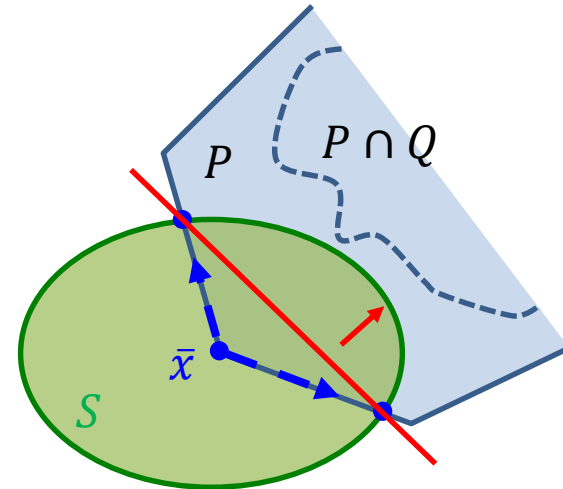
- Notation

- N index set of **structural** variables **x** , $|N| = n$
- I index set of **basic** variables, $|I| = m$
- J index set of **non-basic** variables, $|J| = n$

Intersection Cuts

✓ A **convex** set S contains \bar{x} but **no any feasible point** within its **interior**

- S is a convex set
- $\bar{x} \in \text{int}(S)$
- $\text{int}(S) \cap (P \cap Q) = \emptyset$

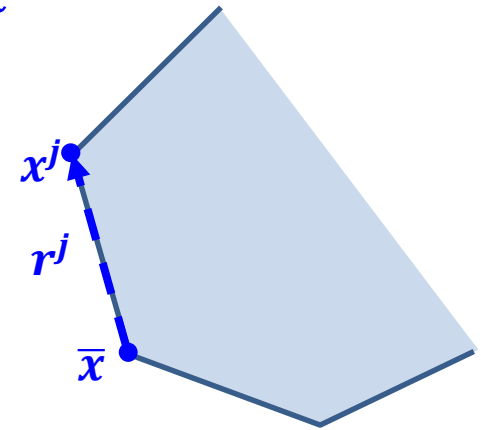


✓ Follow the **extreme rays** at \bar{x} and find the **intersection points**

✓ Obtain the **intersection cut** that goes through all intersection points

Extreme Rays

- Find its neighboring extreme point x^j , then $r^j := x^j - \bar{x}$
- Move from one extreme point \bar{x} to its neighboring extreme point x^j when a **non-basic** variable **enters** the basis and a **basic** variable **leaves** the basis
- Simple tableau



Focus on structural variables x rows

$$\begin{array}{cc}
 \text{Basic} & \text{Non-basic} \\
 \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{bmatrix} & \begin{bmatrix} * & * \\ * & \bar{a}_{ij} & * \\ * & * & * \end{bmatrix}
 \end{array}
 \begin{bmatrix} t_I \\ t_J \end{bmatrix} = \begin{bmatrix} \bar{t}_I \\ \bar{t}_J \end{bmatrix}$$

$$\begin{array}{l}
 x_i = \bar{x}_i - \sum_{j \in J} \bar{a}_{ij} t_j \quad \forall i \in I \cap N \quad \text{Basic} \\
 x_i = 0 \quad \forall i \in J \cap N \quad \text{Non-basic}
 \end{array}$$

\bar{x}

Extreme Rays

- Choose a **non-basic** variable (structural or slack) t_j for some $j \in J$ and let t_j **enter the basis** (assume non-degeneracy)

Other non-basic variables will stay unchanged (still at 0)

$$\begin{aligned} x_i &= \bar{x}_i - \sum_{j \in J} \bar{a}_{ij} t_j & \forall i \in I \cap N \\ x_i &= 0 & \forall i \in J \cap N \end{aligned}$$

\bar{x}

pivot

$$\begin{aligned} x_i &= \bar{x}_i - \bar{a}_{ij} \xi & \forall i \in I \cap N \\ x_i &= 0 & \forall i \in J \cap N \setminus \{j\} \\ x_j &= \xi & \text{if } j \in N \end{aligned}$$

x^j

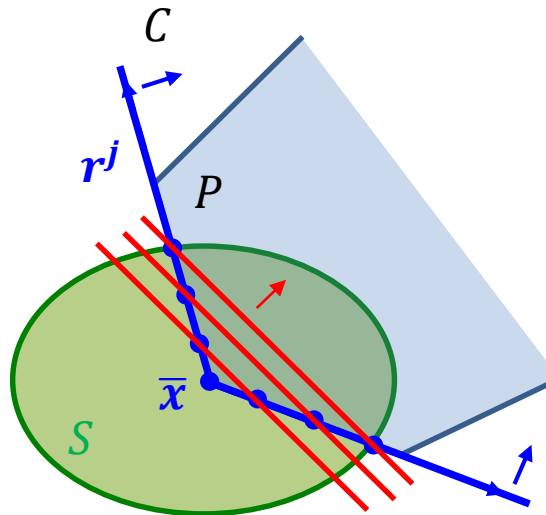
No need to track slack variables s rows

- An extreme ray $r^j = x^j - \bar{x}$

$$\begin{aligned} r_i^j &= -\bar{a}_{ij} \xi & \forall i \in I \cap N \\ r_i^j &= 0 & \forall i \in J \cap N \setminus \{j\} \\ r_j^j &= \xi & \text{if } j \in N \end{aligned} \quad \xrightarrow{\xi > 0} \quad \begin{aligned} r_i^j &= -\bar{a}_{ij} \\ r_i^j &= 0 \\ r_j^j &= 1 \end{aligned}$$

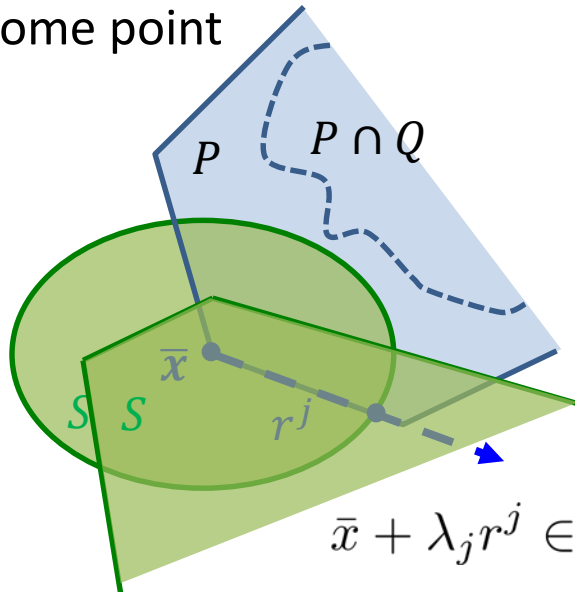
Simplicial Conic Relaxation

- # of extreme rays = # of non-basic variables = $|J| = n$
- These extreme rays are **linearly independent**
- Define a set $C := \{x | x = \bar{x} + \sum_{j \in J} \lambda_j r^j, \lambda_j \geq 0 \forall j \in J\}$, then $P \subseteq C$



Intersection Points

- The convex set S is intersected by a halfline $\eta^j = \bar{x} + \lambda_j r^j$, where $\lambda_j \geq 0$ at some point



$$\begin{array}{ll} \text{maximize} & \lambda_j \\ \text{subject to} & \lambda_j \geq 0 \end{array} \quad (*)$$

$$\text{subject to } \bar{x} + \lambda_j r^j \in S$$

$$\bar{x} + \lambda_j r^j \in S \quad \forall \lambda_j \geq 0$$

- This problem (*) can be solved in **polynomial time** (e.g. line search) and two cases will arise:

- (*) has a **unique solution** $\bar{\lambda}_j > 0$ J_1
- The obj. is **unbounded** ($r^j \in \text{Rec}(S)$, set $\bar{\lambda}_j = +\infty$) J_2

Intersection Cuts

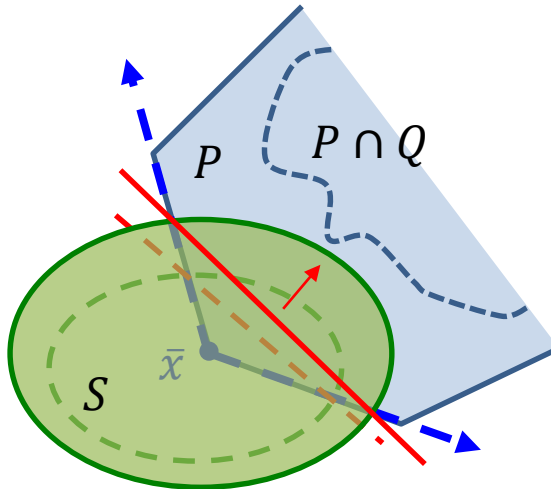
- The intersection cut $\beta^T x \leq \beta_0$ is the halfspace whose boundary contains each **intersection point** ($j \in J_1$) and that is parallel to each **extreme ray** ($j \in J_2$) in $\text{Rec}(S)$

$$\begin{aligned}\beta^T(\bar{x} + \bar{\lambda}_j r^j) &= \beta_0 & \forall j \in J_1 \\ \beta^T r^j &= 0 & \forall j \in J_2\end{aligned}$$

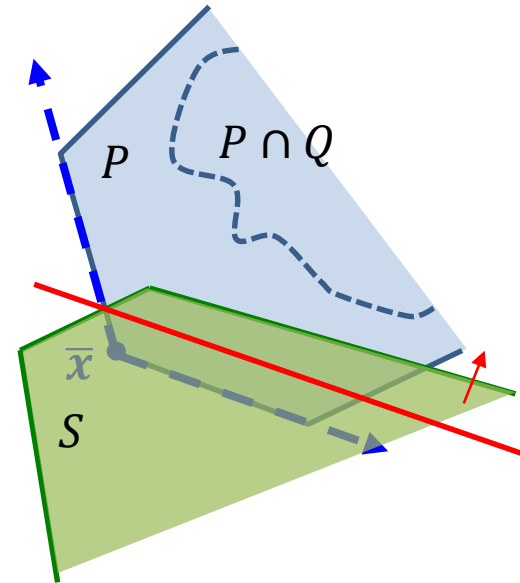
- A system of linear equalities
 - $|J_1| + |J_2| = n$ equations and $n + 1$ variables ($\beta \in \mathbb{R}^n, \beta_0 \in \mathbb{R}$)
 - a **unique solution** (except for a constant factor) since $\{r^j | j \in J\}$ are linearly independent
 - analytical solution: $\beta_0 = \sum_{i \in J} \frac{1}{\bar{\lambda}_i} \mathbf{b}_i - 1$ $\beta_j = \sum_{i \in J} \frac{1}{\bar{\lambda}_i} \mathbf{a}_{ij} \quad \forall j \in N$
- An equivalent but more popular version in the literature

$$\sum_{j \in J} \frac{1}{\bar{\lambda}_j} t_j \geq 1$$

Geometric Interpretation

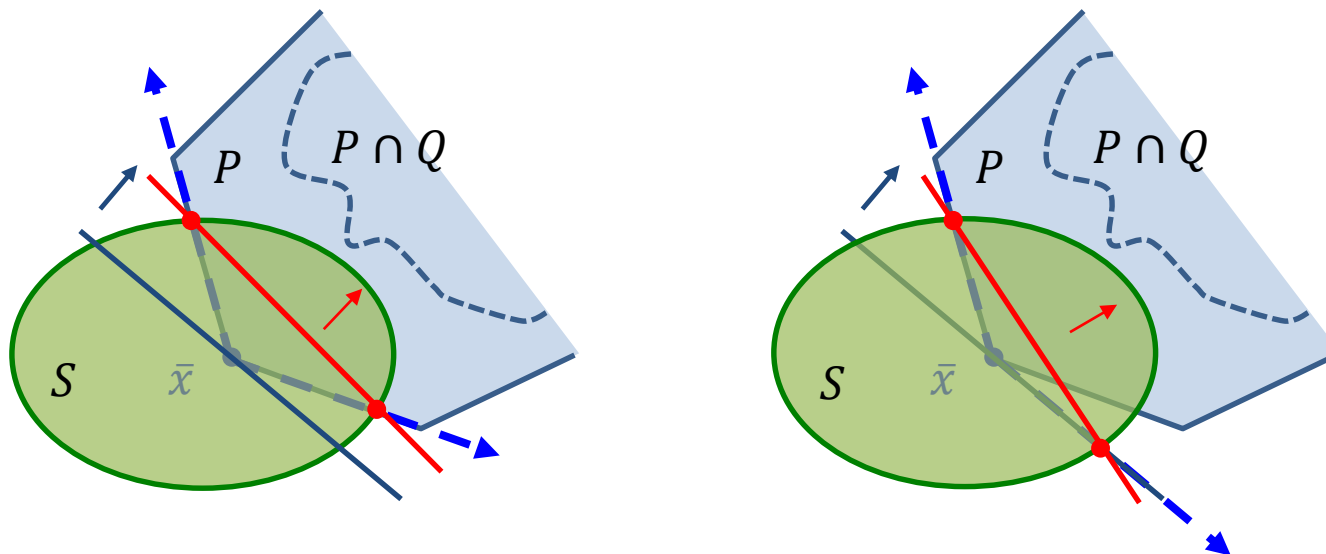


The larger $S \rightarrow$ the deeper cut



The intersection cut is parallel to an extreme ray in $\text{Rec}(S)$

Degeneracy



- ❑ The **degeneracy** will not affect the correctness of the intersection cut formula
- ❑ The **choice of a basis** will lead to different (and valid) intersection cuts
- ❑ In general, **no dominance** relationship among these cuts is guaranteed

Implementation Details

- ✓ $\bar{\lambda}_i$ should be approximated below for **numerical validity**
 - a valid approximation to the intersection cut
- ✓ Scale a cut and perform reduction on small coefficients if necessary for **numerical stability**
- ✓ For a more generic LP as follows, the derivation for intersection cut has to be updated

- **extreme rays** r^j
- **intersection cut formula**

$$\sum_{j \in J^L} \frac{1}{\bar{\lambda}_j} t_j + \sum_{j \in J^U} \frac{-1}{\bar{\lambda}_j} t_j \geq 1$$

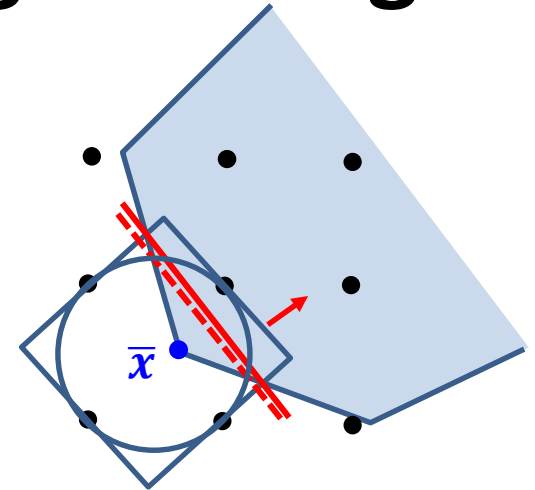
J^L : index set of non-basic variables at lower bounds

J^U : index set of non-basic variables at upper bounds

$$\begin{array}{ll} \underset{x}{\text{minimize}} & c^T x \\ \text{subject to} & Ax \leq b \\ & x^L \leq x \leq x^U \end{array}$$

Mixed Integer Linear Programming

$$\begin{array}{ll}\underset{x}{\text{minimize}} & c^T x \\ \text{subject to} & Ax = b \\ & x \geq 0 \\ & x_i \in \mathbb{Z} \quad i \in N\end{array}$$



❑ The hypersphere can be selected as a valid convex set S

- S is a convex set



- $\bar{x} \in \text{int}(S)$



- $\text{int}(S) \cap (P \cap Q) = \emptyset$



❑ Hard to find the “optimal” set S

❑ $\bar{\lambda}_j$ can be identified analytically

Reverse Convex Programming

A constraint $g(x) \geq 0$ is called **reverse convex** if g is convex

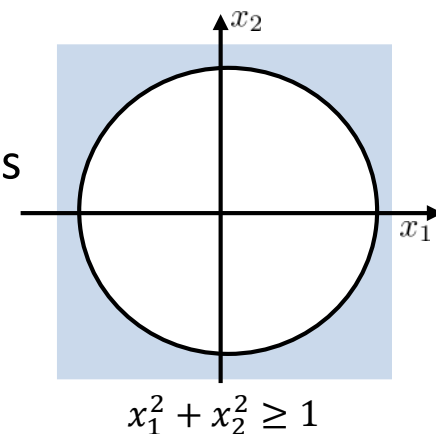
$$\begin{array}{ll} \underset{x \in \mathbb{R}^n}{\text{minimize}} & c^T x \\ \text{subject to} & f_k(x) \leq 0 \quad \forall k = 1, 2, \dots, p \\ & g_l(x) \geq 0 \quad \forall l = 1, 2, \dots, q \end{array}$$

Convex

Reverse Convex

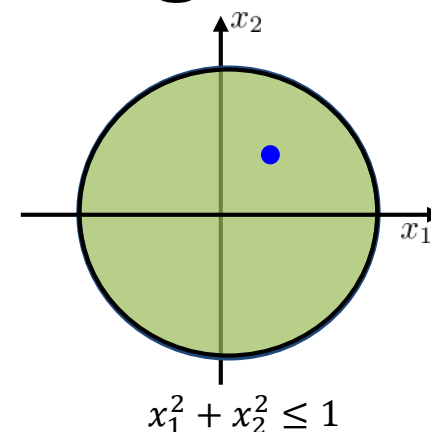
where $f_k(x)$ and $g_l(x)$ are both convex on \mathbb{R}^n

- $f_k(x) \leq 0$ can be outer-approximated by linear inequalities
- $g_l(x) \geq 0$ represent the “complicated” constraints



Reverse Convex Programming

$$\begin{array}{ll}\underset{x}{\text{minimize}} & c^T x \\ \text{subject to} & Ax \leq b \\ & x \geq 0\end{array}$$



□ Define $S = \{x \in \mathbb{R}^n : g_{\bar{l}}(x) \leq 0\}$ for some \bar{l} such that $g_{\bar{l}}(\bar{x}) < 0$

- S is a convex set



- $\bar{x} \in \text{int}(S)$



- $\text{int}(S) \cap (P \cap Q) = \emptyset$



□ $\bar{\lambda}_j$ can be identified via solving $g_{\bar{l}}(\bar{x} + \lambda_j r^j) = 0$ with $\lambda_j \geq 0$

Polynomial Programming

$$\begin{aligned} & \underset{x}{\text{minimize}} && p_0(x) \\ & \text{subject to} && p_i(x) \leq 0 \quad \forall i = 1, 2, \dots, m \end{aligned}$$

where $p_i(x)$ is a **polynomial** function with respect to $x \in \mathbb{R}^n$

– e.g. $p_i(x) = 2 + 3x_1 - 3.2x_1x_2^2 + 4x_2^4, d = 4$

Define $m_r(x) := [1, x_1, x_2, \dots, x_n, x_1^2, x_1x_2, \dots, x_n^2, \dots, x_n^r]^T$, where $r = \lceil d_{\max}/2 \rceil$

$$p_i(x) = m_r^T(x) A_i m_r(x) \leq 0 \iff \langle A_i, m_r(x) \cdot m_r^T(x) \rangle \leq 0$$

where A_i is an appropriately defined symmetric matrix

Moment-based Reformulation (lifted space)

$$X := \begin{bmatrix} X_{11} & X_{12} & X_{13} & X_{14} & X_{15} & X_{16} \\ X_{21} & X_{22} & X_{23} & X_{24} & X_{25} & X_{26} \\ X_{31} & X_{32} & X_{33} & X_{34} & X_{35} & X_{36} \\ X_{41} & X_{42} & X_{43} & X_{44} & X_{45} & X_{46} \\ X_{51} & X_{52} & X_{53} & X_{54} & X_{55} & X_{56} \\ X_{61} & X_{62} & X_{63} & X_{64} & X_{65} & X_{66} \end{bmatrix} = \begin{bmatrix} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 \\ x_1 & x_1^2 & x_1x_2 & x_1^3 & x_1^2x_2 & x_1x_2^2 \\ x_2 & x_1x_2 & x_2^2 & x_1^2x_2 & x_1x_2^2 & x_2^3 \\ x_1^2 & x_1^3 & x_1^2x_2 & x_1^4 & x_1^3x_2 & x_1^2x_2^2 \\ x_1x_2 & x_1^2x_2 & x_1x_2^2 & x_1^3x_2 & x_1^2x_2^2 & x_1x_2^3 \\ x_2^2 & x_1x_2^2 & x_2^3 & x_1^2x_2^2 & x_1x_2^3 & x_2^4 \end{bmatrix} = m_r(x) m_r^T(x)$$

$\langle A, B \rangle = \sum_i \sum_j a_{ij} b_{ij}$

$\langle A_i, X \rangle \leq 0 \quad \forall i = 1, 2, \dots, m$

$m_r^T(x) X = m_r^T(x) \cdot m_{A_i}(x)$

Polynomial Programming

Linear

Convex

Non-convex

$$X = m_r(x) \cdot m_r^T(x) \iff \text{consistency, } X \succeq 0, \text{rank}(X) \leq 1$$

$$X = \begin{bmatrix} 1 & x_1 & x_2 & x_1^2 & x_1x_2 & x_2^2 \\ x_1 & x_1^2 & x_1x_2 & x_1^3 & x_1^2x_2 & x_1x_2^2 \\ x_2 & x_1x_2 & x_2^2 & x_1^2x_2 & x_1x_2^2 & x_2^3 \\ x_1^2 & x_1^3 & x_1^2x_2 & x_1^2x_2^2 & x_1^3x_2 & x_1^2x_2^2 \\ x_1x_2 & x_1^2x_2 & x_1x_2^2 & x_1^3x_2 & x_1^2x_2^2 & x_1x_2^3 \\ x_2^2 & x_1x_2^2 & x_2^3 & x_1^2x_2^2 & x_1x_2^3 & x_2^4 \end{bmatrix}$$

Diagonal entries $X_{ii} \geq 0$

$$X_{62} = X_{53} = X_{35} = X_{26}$$

$$Q := \{X \in \mathbb{S}^{n \times n} | X \succeq 0, \text{rank}(X) \leq 1\}$$

Polyhedral Relaxation

$$\underset{X}{\text{minimize}} \quad \langle A_0, X \rangle$$

$$\text{subject to} \quad \langle A_i, X \rangle \leq 0 \quad \forall i = 1, 2, \dots, m$$

$$X_{11} = 1$$

$$X_{ii} \geq 0 \quad \forall i = 2, \dots, n$$

consistency

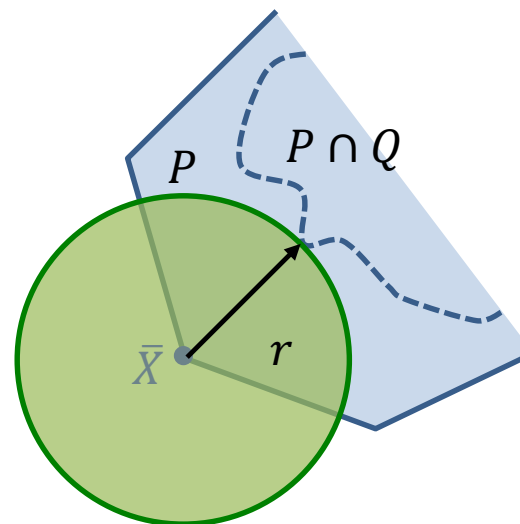
Oracle Ball Cut

□ Define S as a **ball** $B(\bar{X}, r)$ centering at \bar{X} with a radius r

- S is a convex set
- $\bar{X} \in \text{int}(S)$
- $\text{int}(S) \cap (P \cap Q) = \emptyset$

$$\underset{Y}{\text{minimize}} \quad \|\bar{X} - Y\|_F \quad (\#)$$

$$\text{subject to} \quad \begin{array}{l} Y \succeq 0 \\ \text{rank}(Y) \leq 1 \end{array} \quad Q$$



□ Problem (#) : calculate the shortest distance between \bar{X} and a point from Q

– it can be **analytically solved** ($\bar{\lambda}_j = r = (\#) \text{ opt. val.}$)

□ This convex set S can be enlarged (strengthened cut)

2 × 2 Cut

Theorem: $X \succeq 0$ and $\text{rank}(X) = 1$ iff all the 2×2 **principle minors** of X are zero

$$\bar{X} = \begin{matrix} & & \begin{matrix} 3 \\ 5 \end{matrix} & & \\ \begin{matrix} 3 \\ 5 \end{matrix} & \begin{bmatrix} \bar{X}_{11} & \bar{X}_{12} & \bar{X}_{13} & \bar{X}_{14} & \bar{X}_{15} & \bar{X}_{16} \\ \bar{X}_{21} & \bar{X}_{22} & \bar{X}_{23} & \bar{X}_{24} & \bar{X}_{25} & \bar{X}_{26} \\ \bar{X}_{31} & \bar{X}_{32} & \bar{X}_{33} & \bar{X}_{34} & \bar{X}_{35} & \bar{X}_{36} \\ \bar{X}_{41} & \bar{X}_{42} & \bar{X}_{43} & \bar{X}_{44} & \bar{X}_{45} & \bar{X}_{46} \\ \bar{X}_{51} & \bar{X}_{52} & \bar{X}_{53} & \bar{X}_{54} & \bar{X}_{55} & \bar{X}_{56} \\ \bar{X}_{61} & \bar{X}_{62} & \bar{X}_{63} & \bar{X}_{64} & \bar{X}_{65} & \bar{X}_{66} \end{bmatrix} & \begin{matrix} 3 \\ 5 \end{matrix} \end{matrix}$$

$X_{[i,j]}$: submatrix induced by i, j

$$\det(X_{[i,j]}) = 0$$

If $\bar{X}_{[i,j]} > 0$ for some i, j ($1 \leq i < j \leq n$), define $S := \{X \in \mathbb{S}^{n \times n} | X_{[i,j]} \succeq 0\}$

• S is a convex set



• $\bar{x} \in \text{int}(S)$



• $\text{int}(S) \cap (P \cap Q) = \emptyset$



$$\text{int}(S) : X_{[i,j]} \succ 0 \Rightarrow \det(X_{[i,j]}) > 0$$

$$Q := \{X \in \mathbb{S}^{n \times n} | X \succeq 0, \text{rank}(X) \leq 1\}$$

$$\det(X_{[i,j]}) = 0 \quad \forall X \in Q$$

2 × 2 Cut

How to find the **intersection points**?

$$\begin{aligned} \bar{X} + \lambda R \in S &= \{X \in \mathbb{S}^{n \times n} \mid X_{[i,j]} \succeq 0\} \\ &\iff \bar{X}_{[i,j]} + \lambda R_{[i,j]} \succeq 0 \\ &\iff \begin{bmatrix} \bar{X}_{ii} + \lambda R_{ii} & \bar{X}_{ij} + \lambda R_{ij} \\ \bar{X}_{ji} + \lambda R_{ji} & \bar{X}_{jj} + \lambda R_{jj} \end{bmatrix} \succeq 0 \end{aligned}$$

- If $R_{[i,j]} \succcurlyeq 0$, no intersection point (set $\bar{\lambda} = +\infty$)
- Else, $\bar{\lambda}$ can be analytically computed

Computational Results

- ❑ Implementation: Python 2.7.13 / Gurobi 7.0.1
- ❑ Instances:
 - 26 Quadratically Constrained Quadratic Programs (**QCQP**) from GLOBALlib, $n = 6 \sim 63$
 - 99 **BoxQP** (non-convex quadratic objective, bound constraints), $n = 12 \sim 126$
- ❑ Compare the **root node bound**
 - McCormick estimator and RLT (Reformulation Linearization Technique) relaxation

$OPT = 100$
 $RLT = 80$
 $GLB = 90$
- ❑ Stopping conditions:
 - Time limit 600 sec
 - No improvement in obj. val. (10 iter)
 - No violated cut
 - LP becomes numerically unstable

$$\text{Initial Gap} = \frac{OPT - RLT}{|OPT| + \epsilon} \quad 20/100$$
$$\text{End Gap} = \frac{OPT - GLB}{|OPT| + \epsilon} \quad 10/100$$
$$\text{Gap Closed} = \frac{GLB - RLT}{OPT - RLT} \quad 10/20$$

Computational Results

OB: Oracle Ball Cuts

SO: Strengthened OB

OA: Outer Approximation cuts for $X \geq 0$

2x2: 2×2 cuts

Cut Family	Initial Gap	End Gap	Closed Gap	# Cuts	Iters	Time (s)	LPTime (%)
OB	1387.92%	1387.85%	1.00%	16.48	17.20	2.59	2.06%
SO		1387.83%	8.77%	18.56	19.52	4.14	2.29%
OA		1001.81%	8.61%	353.40	83.76	33.25	7.51%
2x2 + OA		1003.33%	32.61%	284.98	118.08	30.40	15.03%
SO+2x2+OA		1069.59%	31.91%	174.79	107.16	29.55	12.56%

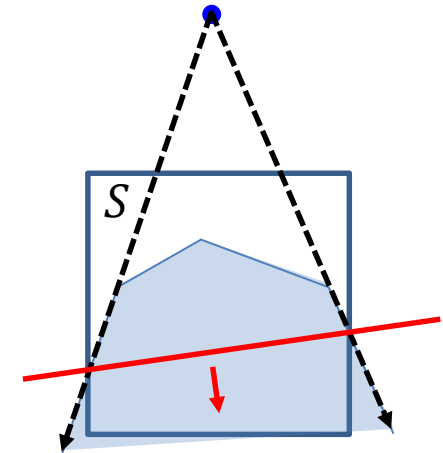
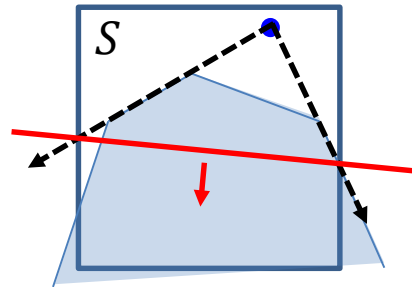
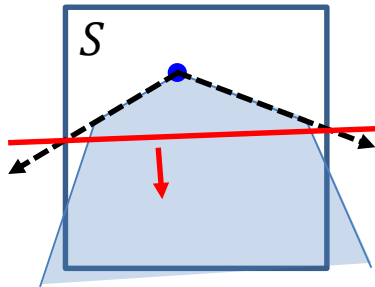
Averages for GLOBALlib instances

Cut Family	Initial Gap	End Gap	Closed Gap	# Cuts	Iters	Time (s)	LPTime (%)
OB	103.59%	103.56%	0.04%	12.84	13.62	127.15	0.40%
SO		103.33%	0.34%	14.34	15.45	132.07	0.49%
OA		30.88%	75.55%	676.90	137.52	459.28	31.80%
2x2 + OA		32.84%	74.52%	349.21	140.40	473.18	28.76%
SO+2x2+OA		33.43%	74.03%	227.39	136.93	475.38	26.59%

Averages for BoxQP instances

Comments

- ❑ The intersection cut is quite **generic** and **computationally cheap** to generate if a set S is given
- ❑ How to find a valid set S for your problem? **NO GENERIC ANSWER**
- ❑ **Research opportunities**
 - Find a valid set S in your application
 - Strengthen the intersection cut



THANK YOU!