MATH 567: Lecture 21 (03/27/2025)

* facial disjunctions
* practical algorithm
* rank of cuts
* semidefinite relaxation

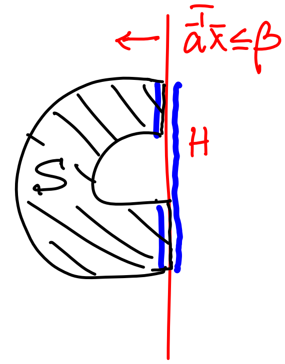
Def D_i is a **facial disjunction** w.r.t. K if D_{ij} are faces of K for $j=1,\dots,k_i$, i.e., $D_{ij} = \{\bar{x} \mid (\bar{a}^{ij})^T\bar{x} = \beta_{ij}\}$ where $(\bar{a}^{ij})^T\bar{x} \le \beta$ is a supporting hyperplane of K.

Theorem 16 If $D_1,\dots,D_p$ are facial disjunctions, then

$$K_p = K \cap_c (D_1 \cap \cdots \cap D_p)$$

in the theoretical algorithm.

Proposition 17 Let S be any set (possibly nonconvex), and $H = \{\bar{x} \mid \bar{a}^T\bar{x} = \beta\}$ is a hyperplane such that $\bar{a}^T\bar{x} \le \beta \;\; \forall \bar{x} \in S$. Then $H \cap \text{conv}(S) = \text{conv}(H \cap S)$.



Proof (Theorem 16)

We show the result for $p=2$, $k_i = 2$, $i=1,2$.

We need to show

$$(K \cap_c D_i) \cap_c D_j = K \cap_c (D_i \cap D_j)$$

$$(K \cap_c D_i) \cap_c D_j = \text{conv}\big[\text{conv}(K \cap D_i) \cap D_j\big] \qquad D_j = D_{j_1} \cup D_{j_2}$$

$$= \text{conv}\Big[\big(\text{conv}(K \cap D_i) \cap D_{j_1}\big) \cup \big(\text{conv}(K \cap D_i) \cap D_{j_2}\big)\Big]$$

$$= \text{conv}\Big[\text{conv}(K \cap D_i \cap D_{j_1}) \cup \text{conv}(K \cap D_i \cap D_{j_2})\Big]$$

by Proposition 17.

$$= \text{conv}\left[(K \cap D_i \cap D_{j_1}) \cup (K \cap D_i \cap D_{j_2})\right]$$

$$= \text{conv}\left[K \cap D_i \cap D_j\right] = K \cap_c (D_i \cap D_j). \qquad \square$$

The theoretical algorithm might not work well in practice. Getting efficient descriptions of the convex hulls in each step might be difficult.

A practical Algorithm

Step 0: $K_0 = K = \{\bar{x} \mid A\bar{x} \le \bar{b}\}$

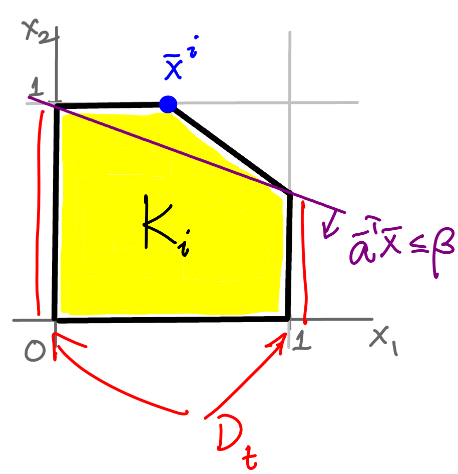
Step i $(i \geq 1)$: K_i : current relaxation, and
$\bar{x}^i$: optimal solution over K_i

Find $\bar{a}^T\bar{x} \le \beta$ such that

① $\bar{a}^T\bar{x} \le \beta$ is valid for $K_i \cap D_t$ for some $t > i$.

② $\bar{a}^T\bar{x} \le \beta$ is violated by $\bar{x}^i$ (i.e., $\bar{a}^T\bar{x}^i > \beta$).

Set $K_{i+1} \leftarrow$ LP relaxation of $K_i \cap D_t \cap \{\bar{x} \mid \bar{a}^T\bar{x} \le \beta\}$.



How do we find $(\bar{a}, \beta)$?

Solve an LP with $\bar{a}, \beta$ as variables.

$$\max \quad \beta - \bar{a}^T \bar{x}^i \quad \text{is given}$$

$$\text{s.t.} \left(\bar{a}^T \bar{x} \le \beta \text{ is derivable from } K_i \cap D_t \text{ by Farkas' Lemma} \right) \quad \text{———} (1)$$

normalization constraint : e.g.,

$$\sum_{i=1}^{n} |a_i| + |\beta| \le 1. \longrightarrow \text{we can linearize this constraint}$$

Without the normalization, the LP could be unbounded. If $(\beta, \bar{a}^T)$ works, then $100\beta - 100\,\bar{a}^T\bar{x}^i$ gives a bigger separation. And so does $10000\beta - 10000\,\bar{a}^T\bar{x}^i$.

For (1), we will use variables representing the multipliers for deriving the constraint $\bar{a}^T\bar{x}^i \le \beta$ from $K_i \cap D_t$.

Q. How good is any such cutting plane algorithm? First, we define rank of cuts.

Rank of cuts

$K = \{\bar{x} \mid A\bar{x} \le \bar{b}\}$: All inequalities in $A\bar{x} \le \bar{b}$, or derivable from $A\bar{x} \le \bar{b}$ are **rank 0 cuts** (or inequalities).

Let $\bar{a}^T\bar{x} \le \beta$ be valid for K. Then the CG cut $\lfloor \bar{a}^T \rfloor \bar{x} \le \lfloor \beta \rfloor$ is a **rank 1 CG cut**.

Combining some (or all) rank-1 CG cuts and applying the CG procedure again gives me a **rank 2 CG cut**.

Similar notion of rank can be defined for the LS-procedure, MIG cuts, etc.

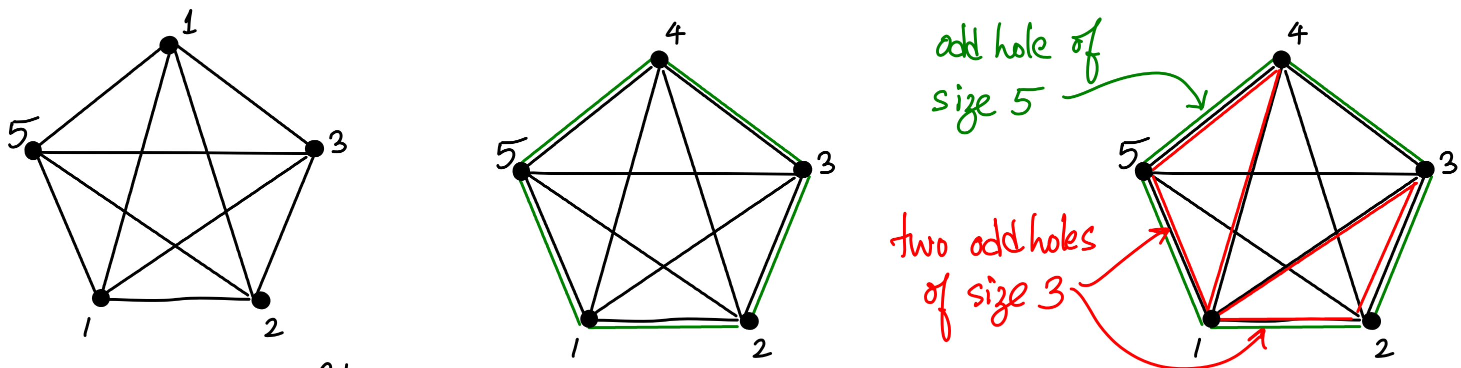
Ideally, we want to derive inequalities with small rank, so that we do not have to apply the cut-procedure too many times.

Back to LS Procedure

Q: How many steps to generate a good inequality?

Example: node packing on complete graph with 5 nodes.

$$K = \left\{ \begin{array}{l} x_i + x_j \le 1 \ \forall (i,j) \\ 0 \le x_i \le 1 \ \forall i \end{array} \right\}$$



complete graph!

Node packing problem:

A good inequality is $x_1 + x_2 + x_3 + x_4 + x_5 \le 1$ ——— $\circledast$

We can pick at most one node.

Recall the definitions of $M_j^{NL}(K)$, $M_j(K)$, and $P_j(K)$. We have

$M_1^{NL}(K)$ specified as
$$\left\{ \begin{array}{c} (A\bar{x} - \bar{b})x_1 \le 0 \\ (A\bar{x} - \bar{b})(1 - x_1) \le 0 \\ x_1(1 - x_1) = 0 \end{array} \right\}.$$

we get odd hole inequalities of size 3:

$$x_1 + x_2 + x_3 \le 1$$
$$x_1 + x_2 + x_4 \le 1$$
$$\vdots$$

and $x_1 + x_2 + x_3 + x_4 + x_5 \le 2$, odd-hole inequality of size 5.

We do not get ⊛.

In $P_2(P_1(K))$, we get $x_1 + x_2 + x_3 + x_4 \le 1$ —— (2)

To see (2) is valid for $P_2(P_1(K))$, we verify that (2) is valid for $P_1(K) \cap (x_2 = 0 \vee x_2 = 1)$.

$P_1(K) \cap (x_2 = 0)$ gives $x_1 + x_3 + x_4 \le 1$, which is there in $P_1(K)$.

$P_1(K) \cap (x_2 = 1)$ gives $x_1 + x_3 + x_4 \le 0$, which is also valid for $P_1(K)$, since we originally (in K itself) have $x_i + x_j \le 1 \ \forall (i,j)$. Hence $x_2 = 1$ immediately forces $x_j = 0 \ \forall j \neq 2$.

But we still do not get ⊛ in $P_2(P_1(K))$.

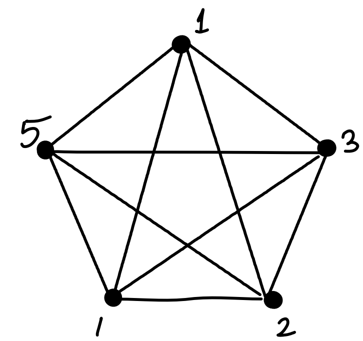
$P_3(P_2(P_1(K)))$ gives $x_1 + \cdots + x_5 \le 1$ —— ⊛ !

$\Rightarrow$ LS-rank of ⊛ is 3.

In general, LS rank of $x_1 + \cdots + x_k \le 1$ is $\le k-2$, and is often $= k-2$.

Q. Could we derive $x_1+x_2+x_3+x_4+x_5 \le 1$ — (*)
in one step? As a rank-1 cut.

YES!



Semidefinite Relaxation

Recall:

$$M_j^{NL}(K) \equiv \left\{ \begin{array}{c} (A\bar{x}-\bar{b})x_j \le \bar{0} \\ (A\bar{x}-\bar{b})(1-x_j) \le \bar{0} \\ x_j(1-x_j) = 0 \end{array} \right\}, \quad j = 1, \ldots, n.$$

We linearize $M_j^{NL}(K)$ $(x_j^2 \leftarrow x_j,\ x_i x_j \leftarrow y_i)$ to $M_j(K)$.

Define $M^{NL}(K) = \bigcap_{j=1}^{n} M_j^{NL}(K)$, and $M(K) = \bigcap_{j=1}^{n} M_j(K)$

We could linearize all systems, and then take their intersection. Equivalently, we could take the intersection of the nonlinear systems, and then linearize.

Let $M_+(K) \equiv M(K) \cap \left\{ \bar{x} \;\middle|\; \underline{(a_0 + \bar{a}^T \bar{x})^2 \ge 0} \;\; \forall \begin{bmatrix} a_0 \\ \bar{a} \end{bmatrix} \in \mathbb{R}^{n+1} \right\}.$

$\begin{bmatrix} 1 & \bar{x}^T \\ \bar{x} & \bar{x}\bar{x}^T \end{bmatrix}$ is positive semidefinite

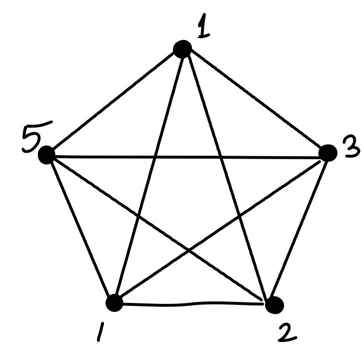
You would think this restriction is always satisfied! But imposing it explicitly makes the difference — see Notes to follow...

$M_+(K)$ is the **semidefinite relaxation** of the problem.

Def $A \in \mathbb{R}^{n \times n}$ is **positive semidefinite** (PSD) if $\bar{x}^T A \bar{x} \geq 0 \ \forall \bar{x} \in \mathbb{R}^n$. We write $A \succeq 0$

If A is PSD, all its eigenvalues are nonnegative.

Back to example on vertex packing



$M_+(K)$ has $(1 - x_1 - x_2 \cdots x_5)^2 \geq 0$

$$\Rightarrow 1 - 2\sum_{i=1}^{5} x_i + \sum_{i=1}^{5} \underbrace{x_i^2}_{x_i} + 2\sum_{i \neq j} \underbrace{x_i x_j}_{=0} \geq 0$$

$$\Rightarrow \quad 1 - \sum_{i=1}^{5} x_i \geq 0,$$ which is (*)!

$M(K)$ has $(x_i + x_j \leq 1)x_i \implies x_i x_j \leq 0$ as $x_i^2 = x_i$

$M(K)$ also has $(-x_j \leq 0)x_i \implies x_i x_j \geq 0$

$$\implies x_i x_j = 0.$$

Hence the semidefinite relaxation rank of (*) is 1!

Notes

① $M^{NL}(K)$ contains inequalities for

$$X = \begin{pmatrix} 1 & \bar{x}^T \\ \bar{x} & \bar{x}\bar{x}^T \end{pmatrix} = \left[\begin{array}{c|cccc} 1 & x_1 & x_2 & \cdots & x_n \\ \hline x_1 & x_1^2 & & & \\ x_2 & & x_2^2 & & x_i x_j \\ \vdots & & & \ddots & \\ x_n & & x_j x_i & & x_n^2 \end{array}\right].$$

Inequalities for X can be written as

$$B \bullet X \geq 0, \text{ where } A \bullet B = \text{trace}(A^T B)$$

② To get $M(K)$ from $M^{NL}(K)$, we replace x_i^2 by x_i and $x_i x_j$ by y_{ij}. Hence inequalities of $M(K)$ can be written as

$$C \bullet \begin{pmatrix} 1 & \bar{x}^T \\ \bar{x} & Y \end{pmatrix} \geq 0, \text{ where } \text{diag}(Y) = \bar{x}.$$

③ $(a_0 + \bar{a}^T \bar{x})^2 \geq 0$ can be written equivalently as

$$[a_0 \ \bar{a}^T] \begin{bmatrix} 1 & \bar{x}^T \\ \bar{x} & \bar{x}\bar{x}^T \end{bmatrix} \begin{bmatrix} a_0 \\ \bar{a} \end{bmatrix} \geq 0 \Rightarrow [a_0 \ \bar{a}^T] \begin{bmatrix} a_0 + \bar{a}^T \bar{x} \\ a_0 \bar{x} + (\bar{x}\bar{x}^T)\bar{a} \end{bmatrix} = [a_0 \ \bar{a}^T] \begin{bmatrix} (a_0 + \bar{a}^T \bar{x}) \\ (a_0 + \bar{a}^T \bar{x})\bar{x} \end{bmatrix}$$

$$= (a_0 + \bar{a}^T \bar{x})(a_0 + \bar{a}^T \bar{x}) \geq 0.$$

In other words, $M_+(K)$ has $\begin{bmatrix} 1 & \bar{x}^T \\ \bar{x} & Y \end{bmatrix} \succeq 0$ as added constraints.