Introduction to Optimization, DIKU 2007
Tuesday, March 6
David Pisinger

Program of the day:

- Convex Hull, Totally Unimodular (Wolsey sec. 3.1-3.2)
- Cutting planes — a method to obtain tighter bounds and faster convergence to integer solutions (Wolsey chap. 8)
- Application: branch-and-cut algorithms

Solving IP models

Some IP naturally lead to integer solutions

- Totally unimodular (TU) matrices

Otherwise, reformulation

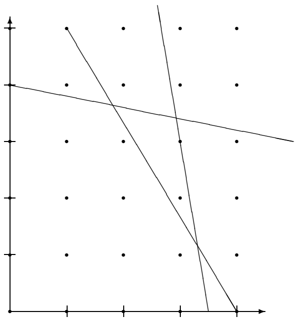
- Reformulation of constraints to TU
- Tightening M, m
- Fixation of variables

Finally, enumrative methods

- Branch-and-bound methods
- Cutting plane methods

Convex hull

Smallest convex polyhedral which contains all integer points



feasible solutions	$\{x \in \mathbb{N}^n : Ax \leq b\}$
linear relaxation	$\{x \in \mathbb{R}^n : Ax \leq b\}$
convex hull	$\text{conv}\{x \in \mathbb{R}^n : Ax \leq b\}$

If constraints of an IP-model define the convex hull, then we can solve the problem efficiently.

Totally Unimodularity

Definition 1 An $m \times n$ integral matrix A is called *totally unimodular* (TU) if the determinant of each square submatrix of A is equal to 0, 1 or -1.

Obviously a_{ij} must be 0, 1, -1

Recognising whether A is TU demands an exponential number of steps

Proposition 1 If A is TU then A_B is also TU

Proof: If A is TU then the determinant of each square submatrix of A is equal to 0, 1 or -1. This also holds when restricted to columns in A_B.

Proposition 2 If A is TU then A^{-1} is integral

Proof: From Cramer's rule $A^{-1}_{ij} = C_{ji} / \det(A)$ where C_{ji} is the adjoint matrix

$$C_{ji} = (-1)^{i+j} \det(A_{\text{row } i, \text{ column } j \text{ removed}})$$

Proposition 3 If A is TU and b is integral then any basis solution x_B is integral

Proof:

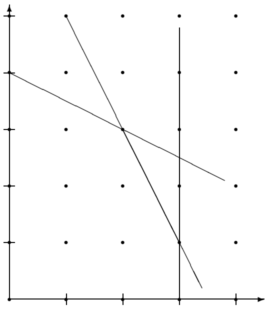
$$x_B = A_B^{-1} b$$

Totally Unimodularity

Proposition 4 If A is TU and b is an integral vector, then the polyhedral defined by

$$\{x \in \mathbb{R}^n : Ax \leq b\}$$

is integeral (i.e. all corner points are integral), or empty.



Proof: The corner points are the basis solutions

Totally Unimodularity

Proposition 5 ("*Property P*") Let A be a $(0,1,-1)$ matrix with no more than two nonzero elements in each column. Then A is TU if and only if the rows of A can be divided in two subsets P_1 and P_2 such that if a column contains two nonzero elements, the following statements are true:

1 If both nonzero elements have the same sign, then one is in a row contained in P_1 and the other is in a row contained in P_2.

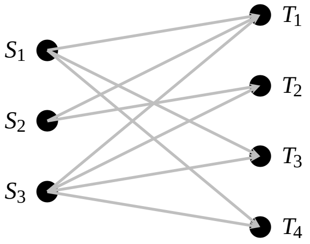
2 If the two nonzero elements have opposite sign, then both are in rows contained in the same subset.

Example

$$\begin{pmatrix}
 & & -1 & & & & & \\
 & 1 & & 1 & & & 1 & \\
-1 & & & -1 & 1 & & & \\
 & & & & & 1 & & \\
 & 1 & -1 & & & & & -1 \\
 & & & & & -1 & &
\end{pmatrix}$$

Example of TU

Three suppliers (S_1, S_2, S_3) should provide four customers (T_1, T_2, T_3, T_4) with a particular commodity.



Commodities cannot be split, hence IP problem.

Supplier	S_1	S_2	S_3
Capacity	135	56	93

Customer	T_1	T_2	T_3	T_4
Requirements	62	83	39	91

	Customer			
Supplier	T_1	T_2	T_3	T_4
S_1	132	—	97	103
S_2	85	91	—	—
S_3	106	89	100	98

Example of TU

Variables: x_{ij} is flow from supplier i to customer j

IP-model

minimize

$$132x_{11}+Mx_{12}+97x_{13}+103x_{14}+85x_{21}+91x_{22}+Mx_{23}+Mx_{24}+106x_{31}+89x_{32}+100x_{33}+98x_{34}$$

subject to

$$\begin{array}{cccccccccccccc}
x_{11} & +x_{12} & +x_{13} & +x_{14} & & & & & & & & & \leq & 135 \\
 & & & & x_{21} & +x_{22} & +x_{23} & +x_{24} & & & & & \leq & 56 \\
 & & & & & & & & x_{31} & +x_{32} & +x_{33} & +x_{34} & \leq & 93 \\
x_{11} & & & & +x_{21} & & & & +x_{31} & & & & \geq & 62 \\
 & x_{12} & & & & +x_{22} & & & & +x_{32} & & & \geq & 83 \\
 & & x_{13} & & & & +x_{23} & & & & +x_{33} & & \geq & 39 \\
 & & & x_{14} & & & & +x_{24} & & & & +x_{34} & \geq & 91
\end{array}$$

$x_{ij} \in \mathbb{N}$

The model is TU, hence LP returns IP solutions

Notice

1 Problem can be formulated by a TU matrix

2 Problem can be solved efficiently

1 ⇒ 2 is obvious, but frequently also 2 ⇒ 1.
(e.g. flow conservation)

Good and bad formulations (Williams)

i) The straightforward formulation results in an IP model where the feasible region is already the convex hull of integer points.

ii) The problem can fairly easily be reformulated to give a feasible region corresponding to the convex hull of integer points.

iii) By reformulation it is possible to reduce the feasible region of the LP problem to nearer that of the convex hull of integer points.

Cutting Planes – introduction

- branch-and-bound: divide and conquer.
- cutting plane: add inequalities which separate fractional solution from solution space.

Development

- 50’s cutting plane (Gomory: simplex, no $\mathcal{NP}$-hardness)
- 70’s tighten formulation in preprocessing
- 80-90’s branch-and-cut (Padberg, Rinaldi)

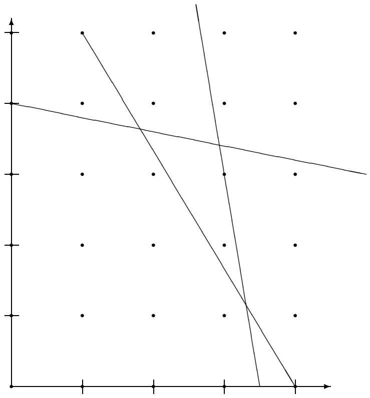
Preprocessing → part of solution process

Definitions

- valid inequality
- cuts: valid inequalities which separates LP-solution
- facets: inequalities defining convex hull

Cuts and facets are redundant for IP formulation
Tighten formulation for LP relaxation

Cuts and facets



Example

Preprocessing, integer variables

$$
\begin{array}{lrcrcrcrcr}
\text{maximize} & \dots & & & & & & & & \\
\text{subject to} & 7x_1 & + & 3x_2 & - & 4x_3 & - & 2x_4 & \le & 1 \\
 & -2x_1 & + & 7x_2 & + & 3x_3 & + & 4x_4 & \le & 6 \\
 & & - & 2x_2 & - & 3x_3 & - & 6x_4 & \le & -5 \\
 & 3x_1 & & & - & 2x_3 & & & \ge & -1 \\
 & x & \in & \mathbb{B}^4 & & & & & &
\end{array}
$$

Generating logical inequalities

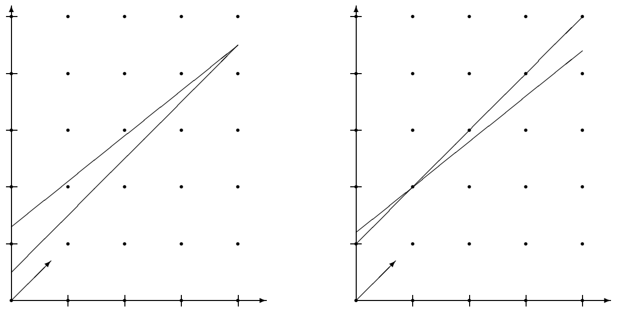
From constraint 1 we see that

- if $x_1 = 1$ and $x_2 = 1$ then infeasible, thus

$$x_1 + x_2 \le 1$$

Example

$$\begin{array}{ll} \text{maximize} & x_1+x_2 \\ \text{subject to} & -2x_1+\ \ 2x_2 \geq 1 \\ & -8x_1+10x_2 \leq 13 \\ & x_1,x_2 \geq 0, \text{integer} \end{array}$$



Tightening formulation

$$\begin{array}{rcl} -2x_1+2x_2 & \geq & 1 \\ -x_1+x_2 & \geq & 1/2 \\ -x_1+x_2 & \geq & 1 \end{array} \qquad \begin{array}{rcl} -8x_1+10x_2 & \leq & 13 \\ -4x_1+5x_2 & \leq & 13/2 \\ -4x_1+5x_2 & \leq & 6 \end{array}$$

Motivation

Integer programming problem (IP)

$$\max\{cx : x \in X\}$$

where $X = \{x : Ax \leq b, x \in \mathbb{Z}^n_+\}$. Reformulate to

$$\max\{cx : x \in \text{conv}(X)\}$$

For any c, an optimal solution to LP is also optimal to IP

Valid inequalities (def. 8.1)

Consider the problem:

$$\begin{array}{ll} \text{maximize} & f(x) \\ \text{subject to} & x \in X \end{array}$$

An inequality

$$\pi x \leq \pi_0$$

is a *valid inequality* for $X \subseteq \mathbb{R}^n$ if

$$\pi x \leq \pi_0 \quad \text{for all} \quad x \in X$$

Characterization of valid inequalities (sec. 8.3.2)

Consider the problem:

$$\begin{array}{ll} \text{maximize} & f(x) \\ \text{subject to} & x \in X \end{array}$$

where

$$X = \{y \in \mathbb{Z} : y \leq b\}$$

then the inequality

$$y \leq \lfloor b \rfloor$$

is valid for X

- Simple observation
- Complete characterization

Overview of cuts

- Chvatal cuts
- Gomory cuts (Modular cuts)
- Chvatal-Gomory cuts
- Disjunctive cuts
- Cover inequalities
- Clique inequalities
- Problem specific cuts

Notice

- Cuts and facets are independent of objective function
- A tight formulation can be used for any objective

Example of Facets

The problem

$$\begin{array}{ll} \text{minimize} & 2x_1+7x_2+2x_3 \\ \text{subject to} & \begin{array}{rcl} x_1+4x_2+x_3 & \geq & 10 \\ 4x_1+2x_2+2x_3 & \geq & 13 \\ x_1+x_2-x_3 & \geq & 0 \\ x_1,x_2,x_3 & \geq & 0,\ \text{integer} \end{array} \end{array}$$

has the facets

$$\begin{array}{rcl} x_1+4x_2+x_3 & \geq & 10 \\ 2x_1+x_2+x_3 & \geq & 7 \\ x_1+x_2-x_3 & \geq & 0 \\ x_1+3x_2+x_3 & \geq & 9 \\ 2x_1+4x_2+x_3 & \geq & 13 \\ x_1+x_2+x_3 & \geq & 5 \\ x_1+2x_2 & \geq & 5 \\ 2x_1+x_2 & \geq & 4 \\ x_1 & \geq & 0,\ \text{integer} \\ x_2 & \geq & 0,\ \text{integer} \\ x_3 & \geq & 0,\ \text{integer} \end{array}$$

Using the new formulation we obtain an integer optimal solution by solving the LP-relaxed problem. (For any objective function).

Chvátal Cuts

Valid inequalities for a pure IP-model (minimization)

1 Add constraints, using suitable multipliers

2 Divide through by a common coefficient factor

3 Round up right-hand-side to the next integer

Example

$$\begin{array}{llr} \text{minimize} & 2x_1+7x_2+2x_3 & \\ \text{subject to} & \begin{array}{rcl} x_1+4x_2+x_3 & \geq & 10 \\ 4x_1+2x_2+2x_3 & \geq & 13 \\ x_1+x_2-x_3 & \geq & 0 \\ x_1 & \geq & 0 \\ x_2 & \geq & 0 \\ x_3 & \geq & 0 \end{array} & \begin{array}{r} (1) \\ (2) \\ (3) \\ (4) \\ (5) \\ (6) \end{array} \\ & x_1,x_2,x_3 \text{ integer} & \end{array}$$

1 times (2) is

$$4x_1+2x_2+2x_3 \ \geq\ 13$$

divide by two

$$2x_1+x_2+x_3 \ \geq\ 6\tfrac{1}{2}$$

left hand side is integral, thus round up right-hand

$$2x_1+x_2+x_3 \ \geq\ 7$$

Example (continued)

$$\begin{array}{llr} \text{minimize} & 2x_1+7x_2+2x_3 & \\ \text{subject to} & \begin{array}{rcl} x_1+4x_2+x_3 & \geq & 10 \\ 4x_1+2x_2+2x_3 & \geq & 13 \\ x_1+x_2-x_3 & \geq & 0 \\ x_1 & \geq & 0 \\ x_2 & \geq & 0 \\ x_3 & \geq & 0 \end{array} & \begin{array}{r} (1) \\ (2) \\ (3) \\ (4) \\ (5) \\ (6) \end{array} \\ & x_1,x_2,x_3 \text{ integer} & \end{array}$$

Facets

$$\begin{array}{rclr} x_1+4x_2+x_3 & \geq & 10 & (a) \\ 2x_1+x_2+x_3 & \geq & 7 & (b) \\ x_1+x_2-x_3 & \geq & 0 & (c) \\ x_1+3x_2+x_3 & \geq & 9 & (d) \\ 2x_1+4x_2+x_3 & \geq & 13 & (e) \\ x_1+x_2+x_3 & \geq & 5 & (f) \\ x_1+2x_2 & \geq & 5 & (g) \\ 2x_1+x_2 & \geq & 4 & (h) \\ x_1,x_2,x_3 & \geq & 0,\ \text{integer} & \end{array}$$

Obtained as

(d) : 5 × (1), 1 × (b), 1 × (6), divide 7
(f) : 1 × (1), 3 × (b), 3 × (6), divide 7
(g) : 4 × (1), 1 × (b), 5 × (3), divide 11
(h) : 1 × (b), 1 × (c), 1 × (4), divide 2
(e) : 3 × (d), 1 × (b), 3 × (g), divide 4

Chvátal Cuts (maximization)

$$\begin{array}{rl} \text{maximize} & \displaystyle\sum_{j=1}^{n} c_j x_j \\ \text{subject to} & \displaystyle\sum_{j=1}^{n} a_{1j}x_j \leq b_1 \\ & \vdots \\ & \displaystyle\sum_{j=1}^{n} a_{mj}x_j \leq b_m \\ & x_j \in \mathbb{Z}_+, \quad j=1,\ldots,n \end{array}$$

1 Take a linear combination of the constraints

$$\sum_{j=1}^{n}(\sum_{i=1}^{m} u_i a_{ij})x_j \leq (\sum_{i=1}^{m} u_i b_i)$$

in short

$$\sum_{j=1}^{n} a'_j x_j \leq b'$$

2 Divide through by a common factor $d|a'_j,\ j=1,\ldots,n$

$$\sum_{j=1}^{n} \frac{a'_j}{d} x_j \leq \frac{b'}{d}$$

3 Since all a'_j/d are integers round down b'

$$\sum_{j=1}^{n} \frac{a'_j}{d} x_j \leq \lfloor \frac{b'}{d} \rfloor$$

Chvatal-Gomory (maximization)

$$\begin{aligned} \text{maximize} \quad & \sum_{j=1}^{n} c_j x_j \\ \text{subject to} \quad & \sum_{j=1}^{n} a_{1j} x_j \le b_1 \\ & \vdots \\ & \sum_{j=1}^{n} a_{mj} x_j \le b_m \\ & x_j \in \mathbb{Z}_+, \quad j = 1, \ldots, n \end{aligned}$$

1 Take a linear combination of the constraints

$$\sum_{j=1}^{n} (\sum_{i=1}^{m} u_i a_{ij}) x_j \le (\sum_{i=1}^{m} u_i b_i)$$

in short

$$\sum_{j=1}^{n} a'_j x_j \le b'$$

2 Since $x \ge 0$ implies $\sum_{j=1}^{n} a'_j x_j \ge \sum_{j=1}^{n} \lfloor a'_j \rfloor x_j$ we have

$$\sum_{j=1}^{n} \lfloor a'_j \rfloor x_j \le b'$$

3 Since $x_j \in \mathbb{Z}_+$ implies $\lfloor a'_j \rfloor x_j \in \mathbb{Z}$ we get

$$\sum_{j=1}^{n} \lfloor a'_j \rfloor x_j \le \lfloor b' \rfloor$$

Chvatal-Gomory (Theorem 8.4)

$$X = \{x : Ax \le b, x \in \mathbb{Z}^n_+\}$$

Every valid inequality for X can be obtained by applying the Chvatal-Gomory procedure a finite number of times.

Notice

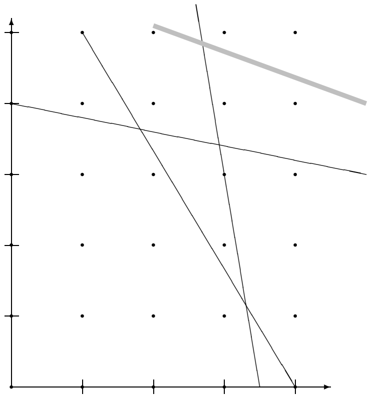
- No stronger inequalities than Chvatal-Gomory exists.
- Even the facet constraints can be generated as Chvatal-Gomory cuts.
- No constructive (polynomial) algorithm for how the linear combination of constraints should be chosen.
- In practice, the derivation of Chvatal-Gomory cuts must rely on specific features of a given application.

Gomory cuts is a systematical way of deriving cutting planes.

Only 0-1 case

All bounded integer variables can be expressed as sum of binary variables.

The set $X = P \cap \mathbb{Z}^n$



$P = \{x \in \mathbb{R}^n : Ax \le b, 0 \le x \le 1\} \ne \emptyset$
$X = P \cap \mathbb{Z}^n$

Nemhauser and Wolsey, Proposition 1.1 page 208:

If inequality $\pi x \le \pi_0$ is valid for P then it can be obtained as a C-G cut (*)

- LP-redundant constraints are C-G inequalities
- Theorem 8.4 deals with IP-constraints

Proof (0-1 case)

Assume that

$$\pi x \le \pi_0 \text{ where } \pi, \pi_0 \text{ integers}$$

is a valid inequality for X. We will show that this inequality can be obtained by using the C-G procedure a finite number of times.

- Step 1: Find a large number $t \in \mathbb{Z}_+$ such that
  $$\pi x \le \pi_0 + t$$
  is a valid C-G inequality
- Step 2: Prove that if
  $$\pi x \le \pi_0 + \tau + 1$$
  for $\tau \in \mathbb{Z}_+$ is a C-G inequality for X then also
  $$\pi x \le \pi_0 + \tau$$
  is a C-G inequality for X.
- Step 3: Use step 2 for $\tau = t - 1, \ldots, 0$ each time getting a new C-G inequality

(Proof by induction)

Step 1

The inequality
$$\pi x \leq \pi_0 + t$$
is valid for P for some $t \in \mathbb{Z}_+$.

Proof

We have the inequality
$$x \leq 1$$
derive C-G inequality using multipliers $u = \pi$
$$\pi x \leq \pi 1$$
choosing $t = \pi 1 - \pi_0$ (π, π_0 is integer) we get the form
$$\pi x \leq \pi 1 = \pi_0 + t$$
for some $t \in \mathbb{Z}_+$

Note that $t < \infty$ as $P \subseteq [0,1]^n$ is bounded so $\max\{\pi x \mid x \in P\} < \infty$.

Step 2

Difficult part

a) Prove that if $\pi x \leq \pi_0 + \tau + 1$ with $\tau \in \mathbb{Z}_+$ is a C-G cut then
$$\pi x \leq \pi_0 + \tau + \sum_{j \in N^0} x_j + \sum_{j \in N^1} (1 - x_j)$$
is a C-G inequality for X for every partition (N^0, N^1) of $N = \{1, \ldots, n\}$.

b) Use partitionings $(T^0 \cup \{n\}, T^1)$ and $(T^0, T^1 \cup \{n\})$ to obtain a new inequality for (T^0, T^1).

c) Derive all valid inequalities for partitionings of $N' = \{1, \ldots, n-1\}$

d) Repeating this procedure n times implies that we eliminate the sums on the right side and thus
$$\pi x \leq \pi_0 + \tau$$
is a C-G cut

Time complexity

- part (c) takes $O(2^n)$,
  part (d) is performed n times,
  in total $O(n2^n)$
- we run Step 2 $O(t)$ times, thus in total $O(tn2^n)$.

Step 2, a)

An inequality is valid for P if it is valid for all extreme points $\{x^1, \ldots, x^m\}$ of P

Assume that $\pi x \leq \pi_0 + \tau + 1$ with $\tau \in \mathbb{Z}_+$ is a valid cut. Let (N^0, N^1) be any partitioning of $N = \{1, \ldots, n\}$. Consider an extreme point x^k of P

- x^k integer: then $\pi x^k \leq \pi_0$ (since $\pi x \leq \pi_0$ valid for X)
- x^k fractional: exists $\varepsilon > 0$ such that
$$\varepsilon^k \leq \sum_{j \in N^0} x_j^k + \sum_{j \in N^1} (1 - x_j^k)$$

Choose $\alpha = \min_{x^k \text{ vertex in } P} \varepsilon^k$

Using $M \geq (\tau + 1)/\alpha$, we have
$$\tau + 1 \leq M\alpha \leq M \left(\sum_{j \in N^0} x_j + \sum_{j \in N^1} (1 - x_j) \right)$$
adding π_0 at both sides we get valid inequality for P
$$\pi x \leq \pi_0 + \tau + 1 \leq \pi_0 + M \left(\sum_{j \in N^0} x_j + \sum_{j \in N^1} (1 - x_j) \right)$$
Due to (*) the inequality is a C-G cut.

Step 2, a)

We have just shown that the following is a C-G inequality
$$\pi x \leq \pi_0 + M \left(\sum_{j \in N^0} x_j + \sum_{j \in N^1} (1 - x_j) \right)$$
By assumption we had the C-G inequality
$$\pi x \leq \pi_0 + \tau + 1$$
use weights $1/M$ and $(M-1)/M$ for the two inequalities getting C-G inequality
$$\pi x \leq \pi_0 + \tau + \sum_{j \in N^0} x_j + \sum_{j \in N^1} (1 - x_j)$$

Step 2, b)

Use partitions $(T^0 \cup \{n\}, T^1)$ and $(T^0, T^1 \cup \{n\})$

$$\pi x \le \pi_0 + \tau + \sum_{j \in T^0 \cup \{n\}} x_j + \sum_{j \in T^1} (1 - x_j)$$

and

$$\pi x \le \pi_0 + \tau + \sum_{j \in T^0} x_j + \sum_{j \in T^1 \cup \{n\}} (1 - x_j)$$

using multipliers $1/2$ and $1/2$ we get C-G inequality

$$\pi x \le \pi_0 + \tau + \sum_{j \in T^0} x_j + \sum_{j \in T^1} (1 - x_j)$$

Chvatal-Gomory rank-k inequalities

Problem

$$\max\{cx \mid Ax \le b, x \in \{0,1\}^n\}$$

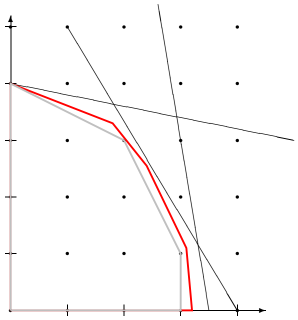
$$I^0 = \{\text{original inequalities } Ax \le b\}$$

Chvatal-Gomory rank-1 inequalities

$$I^1 = \left\{\text{CG inequalities obtained using } I^0\right\}$$

Chvatal-Gomory rank-k inequalities

$$I^k = \left\{\text{CG inequalities obtained using } I^{k-1}\right\}$$



$$P^k = \left\{x \in \mathbb{R}^n \mid ax \le b, \quad (ax \le b \in I^k)\right\}$$

Chvatal-Gomory rank-k inequalities

Convex hull

$$P_{IP} = \text{conv}\left\{Ax \le b, x \in \{0,1\}^n\right\}$$

The smallest k such that

$$P^k = P_{IP}$$

is called the *Chvatal-Gomory rank* of the problem.

Chvatal-Gomory rank is a measure of complexity of problem

Eisenbrand and Schulz (1999) showed, maximum rank k

$$(1 + \varepsilon)n \le k \le O(n^2 \log n)$$

Rank of inequality

A valid inequality has CG rank k if it can be generated as a CG inequality based on inequalities of rank $0, 1, \ldots, k-1$, but does not have rank less than k.

- Often considered when new valid inequalities are proposed for a problem.
- Determining an upper bound on the rank is often done by trial-and-error.
- Lower bound proofs are generally even harder.
- Not aware of any existing computational method for testing rank 2.

Separating rank 1 CG inequalities is NP-hard (Eisenbrand 1999).

Branch-and-cut algorithms

Combines best properties from Branch-and-bound and cutting plane.

- Basically a branch-and-bound algorithm
- at each node solve LP-relaxation to find bound
- generate valid inequalities which separate the LP-solution, and which are *valid for the whole problem*
- maintain pool of valid inequalities
- branch when cuts become weak
- convergence ensured by branch-and-bound

Improvements

- Heuristic for generating cut
- Problem specific cuts
- Heuristic for removing cuts

If separation problem is “easy” the cut is not tight