can be found in Lovász and Plummer (1986). For an analysis and a computational investigation of the directed versus the undirected formulation for Steiner tree problems we refer to the papers by Chopra and Rao (1994a,b). Cut covering problems have been considered in Goemans and Williamson (1995a).

1.4. The cutting stock problem was introduced in Gilmore and Gomory (1961). For a survey on combinatorial auctions see de Vries and Vohra (2003).

Chapter 2

Methods to enhance formulations

Contents

2.1. Methods to generate valid inequalities
2.2. Methods to generate facet defining inequalities
2.3. Valid inequalities in independence systems
2.4. On the strength of valid inequalities
2.5. Nonlinear relaxations
2.6. Summary
2.7. Exercises
2.8. Notes and sources

In the previous chapter, we argued that proximity of the linear relaxation to the convex hull of integer solutions determines the quality of an integer optimization formulation. In this chapter, we introduce methods to enhance formulations of discrete optimization problems, and thus obtain stronger relaxations. We first outline methods to enhance formulations by generating linear inequalities that all integer solutions satisfy, called valid inequalities, and then describe techniques to generate the best possible such valid inequalities. As an illustration, we apply these methods to independence systems, a very general class of discrete optimization problems. We then introduce methods to measure the strength of valid inequalities. We finally introduce nonlinear methods to enhance formulations including semidefinite optimization. In this chapter, we assume that the reader is familiar with the notions of dimension of polyhedra, valid inequalities and facets from polyhedral theory, which we review in Appendix A. The central message of the chapter is that continuous (polyhedral and nonlinear) optimization methods are powerful tools to enhance formulations of integer optimization problems.

2.1 Methods to generate valid inequalities

We have argued that strong formulations are important for solving integer optimization problems. If $\mathcal{F}$ is the set of feasible integer solutions and P is the feasible region of the linear relaxation for a formulation of the problem, then unless $\text{conv}(\mathcal{F}) = P$, there will be valid inequalities for $\text{conv}(\mathcal{F})$ that are not valid for P. Adding such valid inequalities will strengthen the formulation and, as we will see in later chapters, will lead to more efficient solution methods. In this section, we outline techniques to generate valid inequalities.

Rounding

Let $\mathbf{A}_1, \ldots, \mathbf{A}_n$ be the columns of the matrix $\mathbf{A} \in \mathbb{Q}^{m \times n}$. For $\mathbf{b} \in \mathbb{Q}^m$ let

$$\mathcal{F} = \left\{ \mathbf{x} \in \mathbb{Z}_+^n \;\middle|\; \sum_{j=1}^n \mathbf{A}_j x_j \leq \mathbf{b} \right\}$$

be the set of feasible integer solutions. The method is based on the following two elementary observations:

(a) If variable x is restricted to be an integer and the constraint $x \leq a$ is part of the formulation, then we can strengthen the inequality to $x \leq \lfloor a \rfloor$, where $\lfloor a \rfloor$ is the largest integer less than or equal to a.

(b) If the inequalities $\mathbf{Ax} \leq \mathbf{b}$ are valid for a set of integer points $\mathcal{F}$ and $\mathbf{u} \geq \mathbf{0}$, then clearly $\mathbf{u}'\mathbf{Ax} \leq \mathbf{u}'\mathbf{b}$ is also valid for $\text{conv}(\mathcal{F})$.

We can combine the previous observations to propose the following systematic way of generating valid inequalities:

Generating valid inequalities through rounding

1. We choose an m-vector $\mathbf{u} = (u_1, \ldots, u_m)' \geq \mathbf{0}$. Multiplying the ith constraint with u_i and summing, we obtain the inequality

$$\sum_{j=1}^n (\mathbf{u}'\mathbf{A}_j) x_j \leq \mathbf{u}'\mathbf{b}.$$

2. Since $\lfloor \mathbf{u}'\mathbf{A}_j \rfloor \leq \mathbf{u}'\mathbf{A}_j$ and $x_j \geq 0$, the following inequality is valid for $\text{conv}(\mathcal{F})$

$$\sum_{j=1}^n (\lfloor \mathbf{u}'\mathbf{A}_j \rfloor) x_j \leq \mathbf{u}'\mathbf{b}.$$

3. As $\mathbf{x}$ is restricted to take integer values, we can strengthen the previous inequality to

$$\sum_{j=1}^n (\lfloor \mathbf{u}'\mathbf{A}_j \rfloor) x_j \leq \lfloor \mathbf{u}'\mathbf{b} \rfloor. \tag{2.1}$$

The valid inequality (2.1) can be added to the original inequalities $\mathbf{Ax} \leq \mathbf{b}$ to strengthen the formulation. By varying the vector $\mathbf{u}$ we obtain **the first Chvátal closure** of the polyhedron $P = \{\mathbf{x} \in \mathbb{R}_+^n \mid \mathbf{Ax} \leq \mathbf{b}\}$:

$$P_1 = \left\{ \mathbf{x} \in \mathbb{R}_+^n \;\middle|\; \mathbf{Ax} \leq \mathbf{b},\; \sum_{j=1}^n (\lfloor \mathbf{u}'\mathbf{A}_j \rfloor) x_j \leq \lfloor \mathbf{u}'\mathbf{b} \rfloor,\; \forall\, \mathbf{u} \in \mathbb{R}_+^m \right\}. \tag{2.2}$$

It turns out (Theorem 9.4) that the set P_1 is indeed a polyhedron, that is a finite collection of the inequalities (2.1) suffice to define it.

Example 2.1 The set $\mathcal{F} = \{(x_1, x_2)' \in \mathbb{Z}_+^2 \mid x_1 \leq 3,\ -x_1 \leq -1,\ -x_1 + 2x_2 \leq 4,\ 2x_1 + x_2 \leq 8,\ -x_1 - 2x_2 \leq -3\}$ includes the following integer points: $\mathcal{F} = \{(3,0)', (1,1)', (2,1)', (3,1)', (1,2)', (2,2)', (3,2)', (2,3)'\}$. See Figure 2.1.

By multiplying inequalities $-x_1 \leq -1,\ -x_1 + 2x_2 \leq 4$ by $1/2$, adding and then rounding, we obtain the inequality $-x_1 + x_2 \leq 1$, which is a valid inequality. Note that it is also a facet, and thus needed to describe $\text{conv}(\mathcal{F})$.

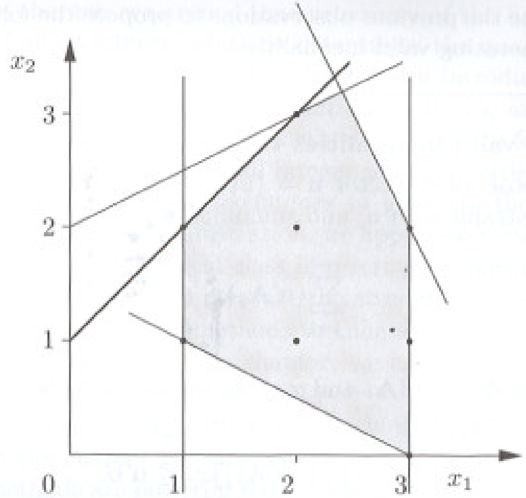


Figure 2.1: The process of rounding in Example 2.1 produces the facet defining inequality $-x_1 + x_2 \leq 1$.

Example 2.2 (The matching problem) In Chapter 1, we introduced the perfect matching problem. A related problem is the matching problem on an undirected graph $G = (V, E)$. A matching is a pairing of nodes in V, so that each node is matched with at most one other node. (In the perfect matching problem we require that each node is matched with exactly one other node). The set of feasible solutions of the matching problem is:

$$\mathcal{F} = \left\{ \mathbf{x} \in \{0,1\}^{|E|} \;\middle|\; \sum_{e \in \delta(\{i\})} x_e \leq 1,\ i \in V \right\}.$$

Let S be a subset of V of odd cardinality. For each $i \in S$, we multiply the inequality $\sum_{e \in \delta(\{i\})} x_e \leq 1$ by 1/2, and we add the resulting inequalities. We obtain that

$$\sum_{e \in E(S)} x_e + \frac{1}{2} \sum_{e \in \delta(S)} x_e \leq \frac{1}{2}|S|.$$

Since $x_e \geq 0$, we obtain

$$\sum_{e \in E(S)} x_e \leq \frac{1}{2}|S|.$$

Performing the rounding step, we obtain

$$\sum_{e \in E(S)} x_e \leq \left\lfloor \frac{1}{2}|S| \right\rfloor = \frac{|S|-1}{2},$$

since $|S|$ is odd. Notice that this inequality was derived in Chapter 1 through combinatorial arguments.

Example 2.3 (The comb inequalities for the traveling salesman problem) The traveling salesman problem on an undirected complete graph on n nodes was formulated in Eq. (1.5):

$$\begin{aligned} \sum_{e \in \delta(\{i\})} x_e &= 2, && i \in V, \\ \sum_{e \in E(S)} x_e &\leq |S| - 1, && S \subset V,\ S \neq \emptyset, V, \\ x_e &\in \{0,1\}. \end{aligned}$$

Let $\mathcal{F}$ be the set of vectors corresponding to feasible tours. In this example, we derive valid inequalities of conv($\mathcal{F}$) by rounding. Consider a subgraph G generated by a node set $H, T_1, \ldots, T_t$ with the following properties:

(a)	$\lvert H \cap T_i \rvert \geq 1$,	$i = 1, \ldots, t$,
(b)	$\lvert T_i \setminus H \rvert \geq 1$,	$i = 1, \ldots, t$,
(c)	$2 \leq \lvert T_i \rvert \leq n - 2$,	$i = 1, \ldots, t$,
(d)	$T_i \cap T_j = \emptyset$,	$i \neq j$,
(e)	t is odd and at least 3.	

Any such graph G is called a **comb**, see Figure 2.2. Given a comb G, we use rounding to show that the **comb inequality** is valid for conv($\mathcal{F}$):

$$\sum_{e \in E(H)} x_e + \sum_{i=1}^{t} \sum_{e \in E(T_i)} x_e \leq |H| + \sum_{i=1}^{t} (|T_i| - 1) - \frac{t+1}{2}. \tag{2.3}$$

We first multiply the degree constraints for all $i \in H$ by 1/2 and sum them to obtain

$$\sum_{e \in E(H)} x_e + \frac{1}{2} \sum_{e \in \delta(H)} x_e = |H|. \tag{2.4}$$

We add $-\frac{1}{2}x_e \leq 0$, for all $e \in \delta(H) \setminus \cup_{i=1}^{t} E(T_i)$ to the previous inequality, and obtain

$$\sum_{e \in E(H)} x_e + \frac{1}{2} \sum_{i=1}^{t} \sum_{e \in \delta(H) \cap E(T_i)} x_e \leq |H|. \tag{2.5}$$

We consider next the subtour elimination constraints for T_i, $H \cap T_i$, and $T_i \setminus H$, respectively:

$$\begin{aligned} \sum_{e \in E(T_i)} x_e &\leq |T_i| - 1, && i = 1, \ldots, t, \\ \sum_{e \in E(H \cap T_i)} x_e &\leq |H \cap T_i| - 1, && i = 1, \ldots, t, \\ \sum_{e \in E(T_i \setminus H)} x_e &\leq |T_i \setminus H| - 1, && i = 1, \ldots, t. \end{aligned}$$

We multiply each of the above inequalities by 1/2, and add them to Eq. (2.5) (see also Figure 2.2). Then, since

$$E(T_i) = E(T_i \cap H) \cup E(T_i \setminus H) \cup (E(T_i) \cap \delta(H)),$$

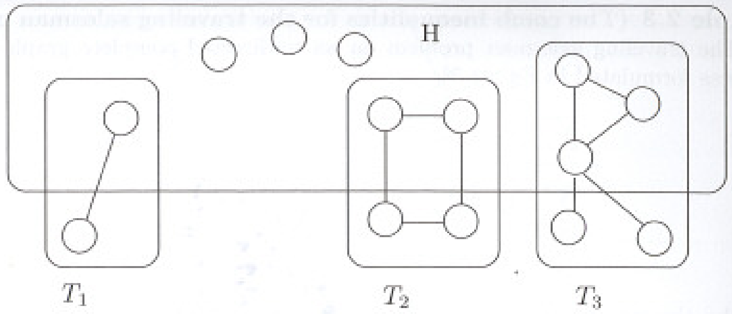


Figure 2.2: A comb subgraph. Note that the set of edges shown is $\cup_{i=1}^{t} E(T_i)$, which is also given by $\cup_{i=1}^{t} ((\delta(H) \cap E(T_i)) \cup E(H \cap T_i) \cup E(T_i \setminus H))$.

we obtain

$$\begin{aligned}\sum_{e \in E(H)} x_e + \sum_{i=1}^{t} \sum_{e \in E(T_i)} x_e &\leq |H| + \frac{1}{2} \sum_{i=1}^{t} \Big[(|T_i| - 1) + (|H \cap T_i| - 1) \\ &\qquad + (|T_i \setminus H| - 1) \Big] \\ &= |H| + \sum_{i=1}^{t} (|T_i| - 1) - \frac{t}{2}.\end{aligned}$$

Using rounding we obtain, since t is odd, that Eq. (2.3) is valid. Note that an alternative representation of Eq. (2.3) is

$$\sum_{e \in \delta(H)} x_e + \sum_{i=1}^{t} \sum_{e \in \delta(T_i)} x_e \geq 3t + 1, \tag{2.6}$$

which can be derived from Eq. (2.4) and

$$\sum_{e \in E(T_i)} x_e + \frac{1}{2} \sum_{e \in \delta(T_i)} x_e = |T_i|, \qquad i = 1, \ldots, t.$$

Superadditivity

For a set $\mathcal{F} = \{\mathbf{x} \in \mathbb{Z}_+^n \mid \mathbf{Ax} \leq \mathbf{b}\}$ the rounding method produces valid inequalities of the type

$$\sum_{j=1}^{n} \lfloor \mathbf{u}'\mathbf{A}_j \rfloor x_j \leq \lfloor \mathbf{u}'\mathbf{b} \rfloor,$$

which can be written as

$$\sum_{j=1}^{n} F(\mathbf{A}_j) x_j \leq F(\mathbf{b}),$$

with $F(\mathbf{a}) = \lfloor \mathbf{u}'\mathbf{a} \rfloor$. This function is a special case of the more general class of superadditive functions, which we define next.

Definition 2.1 *A function $F : D \subset \mathbb{R}^n \mapsto \mathbb{R}$ is* **superadditive** *if*

$$F(\mathbf{a}_1) + F(\mathbf{a}_2) \leq F(\mathbf{a}_1 + \mathbf{a}_2), \text{ for all } \mathbf{a}_1, \mathbf{a}_2 \in D, \text{ such that } \mathbf{a}_1 + \mathbf{a}_2 \in D.$$

It is **nondecreasing** *if*

$$F(\mathbf{a}_1) \leq F(\mathbf{a}_2), \text{ if } \mathbf{a}_1 \leq \mathbf{a}_2 \text{ for all } \mathbf{a}_1, \mathbf{a}_2 \in D.$$

The next theorem shows that superadditive and nondecreasing functions yield valid inequalities.

Theorem 2.1 *If $F : \mathbb{R}^m \mapsto \mathbb{R}$ is superadditive and nondecreasing with $F(\mathbf{0}) = 0$, the inequality*

$$\sum_{j=1}^{n} F(\mathbf{A}_j) x_j \leq F(\mathbf{b}) \tag{2.7}$$

is valid for $\mathrm{conv}(\mathcal{F})$ with $\mathcal{F} = \{\mathbf{x} \in \mathbb{Z}_+^n \mid \mathbf{Ax} \leq \mathbf{b}\}$.

Proof.
Let $\mathbf{x} \in \mathcal{F}$. We first show by induction on x_j that $F(\mathbf{A}_j)x_j \leq F(\mathbf{A}_j x_j)$. For $x_j = 0$ it is clearly true. Assuming it is true for $x_j = k - 1$, we obtain

$$\begin{aligned} F(\mathbf{A}_j)k &= F(\mathbf{A}_j) + F(\mathbf{A}_j)(k-1) \\ &\leq F(\mathbf{A}_j) + F(\mathbf{A}_j(k-1)) \\ &\leq F(\mathbf{A}_j + \mathbf{A}_j(k-1)), \end{aligned}$$

by superadditivity, and the induction is complete. Therefore,

$$\sum_{j=1}^{n} F(\mathbf{A}_j) x_j \leq \sum_{j=1}^{n} F(\mathbf{A}_j x_j).$$

By superadditivity,

$$\sum_{j=1}^{n} F(\mathbf{A}_j x_j) \leq F\left(\sum_{j=1}^{n} \mathbf{A}_j x_j\right) = F(\mathbf{Ax}).$$

Since $\mathbf{Ax} \leq \mathbf{b}$ and F is nondecreasing

$$F(\mathbf{Ax}) \leq F(\mathbf{b}).$$

Combining these inequalities we obtain that the inequality (2.7) is valid for all $\mathbf{x} \in \mathcal{F}$. $\square$

The previous theorem provides a general technique to generate valid inequalities starting from superadditive and nondecreasing functions. In the exercises we show ways to generate such functions. In Chapter 4, we use superadditive functions to formulate a dual problem associated with a given integer optimization problem.

Modular arithmetic

We consider the set

$$\mathcal{F} = \left\{ \mathbf{x} \in \mathbb{Z}_+^n \;\middle|\; \sum_{j=1}^n a_j x_j = a_0 \right\},$$

where a_j, $j = 0, 1, \ldots, n$ are given integers. Let $d \in \mathbb{Z}_+$. We write $a_j = b_j + u_j d$, where b_j $(0 \le b_j < d,\ b_j \in \mathbb{Z})$ is the remainder when a_j is divided with d. Then, all points in $\mathcal{F}$ satisfy

$$\sum_{j=1}^n b_j x_j = b_0 + rd, \quad \text{for some integer } r.$$

Since $\sum_{j=1}^n b_j x_j \ge 0$ and $b_0 < d$, we obtain $r \ge 0$. Then, the inequality

$$\sum_{j=1}^n b_j x_j \ge b_0$$

is valid for $\text{conv}(\mathcal{F})$.

Example 2.4 Let $\mathcal{F} = \{\mathbf{x} \in \mathbb{Z}_+^4 \mid 27x_1 + 17x_2 - 64x_3 + x_4 = 203\}$. Applying modular arithmetic for $d = 13$, we obtain that the inequality $x_1 + 4x_2 + x_3 + x_4 \ge 8$ is valid for $\text{conv}(\mathcal{F})$.

An important set of valid inequalities arises when $d = 1$, and a_j are not integers. In this case, since $\mathbf{x} \ge \mathbf{0}$, we obtain $\sum_{j=1}^n \lfloor a_j \rfloor x_j \le a_0$. Since $\mathbf{x} \in \mathbb{Z}$, $\sum_{j=1}^n \lfloor a_j \rfloor x_j \le \lfloor a_0 \rfloor$, and thus the following inequality is valid for $\text{conv}(\mathcal{F})$

$$\sum_{j=1}^n (a_j - \lfloor a_j \rfloor) x_j \ge a_0 - \lfloor a_0 \rfloor.$$

These inequalities are called **Gomory cuts** and form the basis of the Gomory cutting plane algorithm we discuss in Section 9.2.

Disjunctions

The idea of the method is to partition the set of feasible integer solutions $\mathcal{F}$ into two or more parts and derive inequalities for each element of the partition. The technique provides a way to turn the inequalities for the parts into valid inequalities for the enire set $\mathcal{F}$.

Proposition 2.1 *If the inequality $\sum_{j=1}^n a_j x_j \le b$ is valid for $\mathcal{F}_1 \subset \mathbb{R}_+^n$, and the inequality $\sum_{j=1}^n c_j x_j \le d$ is valid for $\mathcal{F}_2 \subset \mathbb{R}_+^n$, then the inequality*

$$\sum_{j=1}^n \min(a_j, c_j) x_j \le \max(b, d)$$

is valid for $\mathcal{F}_1 \cup \mathcal{F}_2$.

Proof.
Let $\mathbf{x} \in \mathcal{F}_1 \cup \mathcal{F}_2$. Without loss of generality we can assume that $\mathbf{x} \in \mathcal{F}_1$, and therefore, $\sum_{j=1}^n a_j x_j \le b$, which implies, as $\mathbf{x} \ge \mathbf{0}$ that

$$\sum_{j=1}^n \min(a_j, c_j) x_j \le \sum_{j=1}^n a_j x_j \le b \le \max(b, d). \qquad \square$$

We next apply Proposition 2.1 to derive valid inequalities for the set $\mathcal{F} = \{\mathbf{x} \in \mathbb{Z}_+^n \mid \mathbf{A}\mathbf{x} \le \mathbf{b}\}$. Let α be a nonnegative integer.

Proposition 2.2 *If the inequality $\sum_{j=1}^n a_j x_j - d(x_k - \alpha) \le b$ is valid for $\mathcal{F}$ for some $d \ge 0$, and the inequality $\sum_{j=1}^n a_j x_j + c(x_k - \alpha - 1) \le b$ is valid for $\mathcal{F}$ for some $c \ge 0$, then the inequality $\sum_{j=1}^n a_j x_j \le b$ is valid for $\mathcal{F}$.*

Proof.
Notice that for $x_k \le \alpha$,

$$\sum_{j=1}^n a_j x_j \le \sum_{j=1}^n a_j x_j - d(x_k - \alpha) \le b.$$

Thus, the inequality $\sum_{j=1}^n a_j x_j \le b$ is valid for $\mathcal{F}_1 = \mathcal{F} \cap \{\mathbf{x} \in \mathbb{Z}_+^n \mid x_k \le \alpha\}$. Similarly, the inequality $\sum_{j=1}^n a_j x_j \le b$ is valid for $\mathcal{F}_2 = \mathcal{F} \cap \{\mathbf{x} \in \mathbb{Z}_+^n \mid x_k \ge \alpha + 1\}$. From Proposition 2.1 the inequality $\sum_{j=1}^n a_j x_j \le b$ is valid for $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$. $\square$

Example 2.5 Revisiting Example 2.1, we write the two inequalities $-x_1 + 2x_2 \leq 4$ and $-x_1 \leq -1$ as follows:

$$(-x_1 + x_2) + (x_2 - 3) \leq 1, \quad (-x_1 + x_2) - (x_2 - 2) \leq 1.$$

Applying the previous proposition with $\alpha = 2$, we obtain that the inequality $-x_1 + x_2 \leq 1$ is valid.

Mixed integer rounding and aggregation

We next introduce an extension of integer rounding by using an idea from mixed integer optimization to derive valid inequalities for integer optimization problems. The application of mixed integer rounding to mixed integer optimization is illustrated in Chapter 13.

For $v \in \mathbb{R}$ we denote by $f(v) = v - \lfloor v \rfloor$, the fractional part of v, and by $v^+ = \max\{0, v\}$. Our point of departure is a derivation of a valid inequality for the special mixed integer set

$$\mathcal{F} = \{(x, y)' \in \mathbb{Z} \times \mathbb{R}_+ \mid x - y \leq b\},$$

where $b \in \mathbb{Q}$. The feasible region $\mathcal{F}$ for the special case of $b = 3/2$ is illustrated in Figure 2.3.

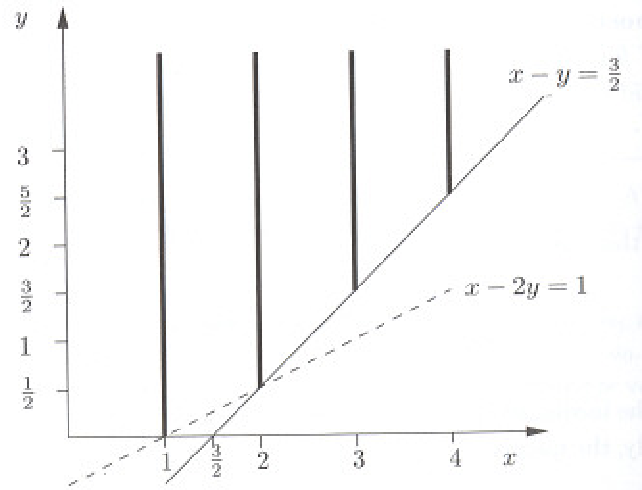


Figure 2.3: The feasible region $\mathcal{F}$ for the special case of $b = 3/2$.

Proposition 2.3 *The inequality*

$$x - \frac{1}{1 - f(b)}\, y \leq \lfloor b \rfloor \tag{2.8}$$

is valid for $\operatorname{conv}(\mathcal{F})$.

Proof.
We consider the disjunction

$$\mathcal{F}_1 = \mathcal{F} \cap \{(x, y)' \mid x \leq \lfloor b \rfloor\} \quad \text{and} \quad \mathcal{F}_2 = \mathcal{F} \cap \{(x, y)' \mid x \geq \lfloor b \rfloor + 1\}.$$

By adding $1 - f(b)$ times the inequality $x - \lfloor b \rfloor \leq 0$ and the inequality $0 \leq y$, we immediately see that the inequality

$$(x - \lfloor b \rfloor)\,(1 - f(b)) \leq y$$

is valid for $\mathcal{F}_1$. For $\mathcal{F}_2$ we combine $-(x - \lfloor b \rfloor) \leq -1$ and $x - y \leq b$ with multipliers $f(b)$ and 1, respectively to obtain

$$(x - \lfloor b \rfloor)(1 - f(b)) \leq y.$$

Applying Proposition 2.1 we obtain that $(x - \lfloor b \rfloor)(1 - f(b)) \leq y$ is valid for $\operatorname{conv}(\mathcal{F}_1 \cup \mathcal{F}_2) = \operatorname{conv}(\mathcal{F})$. □

We next show that Proposition 2.3 generates stronger valid inequalities for integer optimization problems than those generated using the integer rounding principle.

Proposition 2.4 *Let* $\mathbf{A} \in \mathbb{Q}^{m \times n}$ *and* $\mathbf{b} \in \mathbb{Q}^m$. *We set*

$$\mathcal{F} = \left\{ \mathbf{x} \in \mathbb{Z}_+^n \;\middle|\; \sum_{j=1}^{n} \mathbf{A}_j x_j \leq \mathbf{b} \right\}.$$

For every $\mathbf{u} \in \mathbb{Q}_+^m$ *the inequality*

$$\sum_{j=1}^{n} \left(\lfloor \mathbf{u}'\mathbf{A}_j \rfloor + \frac{[f(\mathbf{u}'\mathbf{A}_j) - f(\mathbf{u}'\mathbf{b})]^+}{1 - f(\mathbf{u}'\mathbf{b})} \right) x_j \leq \lfloor \mathbf{u}'\mathbf{b} \rfloor \tag{2.9}$$

is valid for $\operatorname{conv}(\mathcal{F})$.

Proof.
Let $N = \{1, \ldots, n\}$ and for $\mathbf{u} \in \mathbb{Q}_+^m$ we denote by $N_1 = \{j \in N \mid f(\mathbf{u}'\mathbf{A}_j) \leq$

$f(\mathbf{u}'\mathbf{b})\}$ and $N_2 = N \setminus N_1$. Note that if $\mathbf{u}'\mathbf{A}_j \in \mathbb{Z}$, then $j \in N_1$. For all $\mathbf{x} \in \mathcal{F}$ we have

$$\sum_{j\in N_1} \lfloor \mathbf{u}'\mathbf{A}_j \rfloor x_j + \sum_{j\in N_2} (\mathbf{u}'\mathbf{A}_j) x_j \le \mathbf{u}'\mathbf{b}. \tag{2.10}$$

Since $\mathbf{u}'\mathbf{A}_j = \lfloor \mathbf{u}'\mathbf{A}_j \rfloor + f(\mathbf{u}'\mathbf{A}_j)$, and $\lceil \mathbf{u}'\mathbf{A}_j \rceil - \lfloor \mathbf{u}'\mathbf{A}_j \rfloor = 1$ for all $j \in N_2$, we obtain $\mathbf{u}'\mathbf{A}_j = \lceil \mathbf{u}'\mathbf{A}_j \rceil + f(\mathbf{u}'\mathbf{A}_j) - 1$, for all $j \in N_2$. Substituting to Eq. (2.10) we obtain that $w - z \le \mathbf{u}'\mathbf{b}$, for all $\mathbf{x} \in \mathcal{F}$ with

$$w = \sum_{j\in N_1} \lfloor \mathbf{u}'\mathbf{A}_j \rfloor x_j + \sum_{j\in N_2} \lceil \mathbf{u}'\mathbf{A}_j \rceil x_j \in \mathbb{Z}, \qquad z = \sum_{j\in N_2} (1 - f(\mathbf{u}'\mathbf{A}_j)) x_j \ge 0.$$

Applying Proposition 2.3 to $w - z \le \mathbf{u}'\mathbf{b}$ we obtain the valid inequality

$$w - \frac{z}{1 - f(\mathbf{u}'\mathbf{b})} \le \lfloor \mathbf{u}'\mathbf{b} \rfloor. \tag{2.11}$$

Substituting w and z in Eq. (2.11) gives Eq. (2.9). □

Inequality (2.9) is indeed stronger than the pure integer rounding cut

$$\sum_{j=1}^{n} (\lfloor \mathbf{u}'\mathbf{A}_j \rfloor) x_j \le \lfloor \mathbf{u}'\mathbf{b} \rfloor.$$

Note, however, that the coefficients of the strengthened inequality are not integral anymore.

2.2 Methods to generate facet defining inequalities

In the previous section, we presented techniques to generate valid inequalities. These general techniques, however, do not necessarily yield high dimensional faces of the convex hull of a set of integral solutions. In this section, we describe proof techniques to check when a valid inequality is indeed facet defining. Unfortunately, the determination of families of such inequalities is more of an art than a formal methodology. For this reason, we will demonstrate the basic concepts through examples.

Applying the definition

Suppose that we have found a valid inequality $\mathbf{a}'\mathbf{x} \le b$ for an integer optimization problem with a feasible set $\mathcal{F}$. We would like to show that this inequality is facet defining of the convex hull of $\mathcal{F}$. Applying Definition A.6 of a facet we need first to calculate the dimension $d = \dim(\text{conv}(\mathcal{F}))$ (see Definition A.3), and then find d affinely independent vectors in $\mathcal{F}$ that satisfy the inequality $\mathbf{a}'\mathbf{x} \le b$ with equality. To show the latter it is sufficient to show that $\dim(\{\mathbf{x} \in \mathcal{F} \mid \mathbf{a}'\mathbf{x} = b\}) = d - 1$. Let us consider an example.

Example 2.6 (The stable set problem) Given an undirected graph $G = (V, E)$ and weights w_i for each $i \in V$, the weighted stable set problem asks for a collection of nodes S of maximum weight, so that no two nodes in S are adjacent. We let $n = |V|$ and introduce the decision variable x_i, which is equal to one, if node i is selected in the stable set, and zero, otherwise. The problem can then be formulated as follows:

$$\begin{array}{lll} \text{maximize} & \sum_{i\in V} w_i x_i & \\ \text{subject to} & x_i + x_j \le 1, & \{i,j\} \in E, \\ & x_i \in \{0,1\}, & i \in V. \end{array}$$

A collection of nodes U, such that for all $i, j \in U$, $\{i,j\} \in E$ is called a **clique**. Clearly, the following inequality is valid for all feasible solutions to the stable set problem:

$$\sum_{i\in U} x_i \le 1, \qquad \text{for any clique } U. \tag{2.12}$$

We next find conditions under which this inequality is facet defining. The set of feasible solutions to the stable set problem is given by

$$\mathcal{F} = \Big\{ \mathbf{x} \in \{0,1\}^n \ \Big|\ x_i + x_j \le 1,\ \forall\, \{i,j\} \in E \Big\}.$$

Since $\mathbf{0} \in \mathcal{F}$ and the n unit vectors $\mathbf{e}_i$ belong to $\mathcal{F}$, there are $n+1$ affinely independent vectors in $\mathcal{F}$, which implies that $\dim(\text{conv}(\mathcal{F})) = n$.

A clique U is maximal if for all $i \in V \setminus U$, $U \cup \{i\}$ is not a clique. We will show next that inequality (2.12) is facet defining if and only if U is a maximal clique. Suppose for simplicity $U = \{1, \dots, k\}$. Then, clearly, the vectors $\mathbf{e}_i$, $i = 1, \dots, k$ satisfy inequality (2.12) with equality. Since U is a maximal clique, for each $i \notin U$, there is a node $j = r(i) \in U$, such that $(i, r(i)) \notin E$. Therefore, the vector $\mathbf{x}_i$ with the ith and $r(i)$th coordinates equal to one and the rest equal to zero is in $\mathcal{F}$, and satisfies inequality (2.12) with equality. The vectors $\mathbf{e}_1, \dots, \mathbf{e}_k, \mathbf{x}_{k+1}, \dots, \mathbf{x}_n$ are linearly independent and, therefore, they are also affinely independent. It follows that if U is a maximal clique, then inequality (2.12) is facet defining.
Conversely, if U is not maximal, inequality (2.12) is not facet defining. Since U is not maximal, there is a node $i \notin U$ such that $U \cup \{i\}$ is a clique, and thus

$$\sum_{j\in U\cup\{i\}} x_j \le 1 \tag{2.13}$$

is valid for conv $(\mathcal{F})$. Since inequality (2.12) is the sum of $-x_i \le 0$ and inequality (2.13), then the inequality (2.12) is not facet defining.

Similarly, consider a cycle $C \subseteq E$. We denote by $V(C)$ the nodes incident to an edge in C. For any such cycle C of odd cardinality, there cannot be more than $(|V(C)| - 1)/2$ nodes in a stable set, otherwise two of these nodes will be adjacent. Therefore, the inequality

$$\sum_{i\in V(C)} x_i \le \frac{|V(C)| - 1}{2}, \qquad \text{for any cycle } C \text{ of odd cardinality} \tag{2.14}$$

is valid for $\mathcal{F}$.

Example 2.7 (Facets of the traveling salesman problem) Let $\mathcal{F}$ be the feasible region of the symmetric traveling salesman problem on an undirected complete graph on n nodes formulated in Eq. (1.5). Since there are n equality constraints, $\dim(\text{conv}(\mathcal{F})) \leq n(n-1)/2 - n$. It turns out that $\dim(\text{conv}(\mathcal{F})) = n(n-1)/2-n$ (Exercise 2.9(a)). In addition, the inequalities $\sum_{e\in E(S)} x_e \leq |S|-1$, define facets of $\text{conv}(\mathcal{F})$, for all S with $2 \leq |S| \leq n/2$, and $n \geq 4$ (see Exercise 2.9(b)). The polyhedron $\text{conv}(\mathcal{F})$ has very complicated structure. While many classes of facet defining inequalities are known, a complete characterization of $\text{conv}(\mathcal{F})$ is unknown.

Lifting a valid inequality

We next outline a technique to derive a higher dimensional face from a lower dimensional one. We will use the stable set problem to illustrate the technique.

Example 2.8 (The stable set problem, continued) Consider the maximal clique inequalities on the graph G of Figure 2.4:

$$\begin{array}{llllllll} x_1 & + x_2 & + x_3 & & & & \leq 1, & \\ x_1 & & + x_3 & + x_4 & & & \leq 1, & \\ x_1 & + & & x_4 & + x_5 & & \leq 1, & (2.15) \\ x_1 & + & & & x_5 & + x_6 & \leq 1, & \\ x_1 & + x_2 & & & & + x_6 & \leq 1. & \end{array}$$

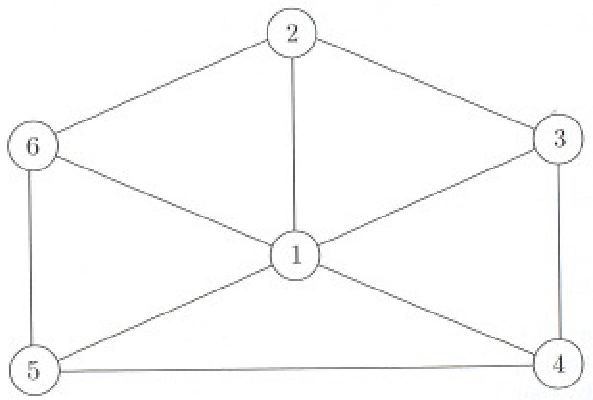


Figure 2.4: In the graph G the maximal clique inequalities do not describe the convex hull of feasible solutions to the stable set problem.

The linear optimization problem of maximizing $\sum_{i=1}^{6} x_i$ subject to the constraints (2.15) and $x_i \geq 0$, $i = 1,\ldots,6$ has the unique optimal solution $\mathbf{x}^0 =$

$(1/2)(0,1,1,1,1,1)'$. This shows that the maximal clique inequalities are not enough to describe the convex hull of solutions. Note that the vector $\mathbf{x}^0$ does not satisfy the valid inequality (2.14) for $V(C) = \{2,3,4,5,6\}$:

$$x_2 + x_3 + x_4 + x_5 + x_6 \leq 2. \tag{2.16}$$

Let $\mathcal{F}$ be the set of feasible solutions to the stable set problem on G. Inequality (2.16) is satisfied with equality by the five linearly independent vectors corresponding to the stable sets $\{2,4\}$, $\{2,5\}$, $\{3,5\}$, $\{3,6\}$, $\{4,6\}$, but is not facet defining, since there are no other stable sets that satisfy (2.16) with equality. Note, however, that inequality (2.16) is facet defining for

$$\text{conv}\left(\mathcal{F} \cap \{\mathbf{x} \in \{0,1\}^6 \mid x_1 = 0\}\right).$$

The idea of lifting is to modify inequality (2.16), which is facet defining for $\text{conv}\left(\mathcal{F} \cap \{\mathbf{x} \in \{0,1\}^6 \mid x_1 = 0\}\right)$, so that it becomes facet defining for $\text{conv}(\mathcal{F})$. We consider the inequality

$$ax_1 + x_2 + x_3 + x_4 + x_5 + x_6 \leq 2, \tag{2.17}$$

for some $a > 0$. We want to select a in order for inequality (2.17) to be still valid, and to define a facet for $\text{conv}(\mathcal{F})$. We next find the values of a for which the inequality is valid. For $x_1 = 0$, inequality (2.17) is valid for all a. If $x_1 = 1$, $a \leq 2 - x_2 - x_3 - x_4 - x_5 - x_6$. Since $x_1 = 1$ implies $x_2 = \cdots = x_6 = 0$, then $a \leq 2$. Therefore, if $0 \leq a \leq 2$, inequality (2.17) is valid. Moreover, for $a = 2$, the previous five vectors satisfy the inequality with equality and, in addition, the solution corresponding to the stable set $\{1\}$ satisfies it with equality. Since these six vectors are linearly independent the inequality

$$2x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \leq 2,$$

is valid and defines a facet of $\text{conv}(\mathcal{F})$.

The general principle of lifting is given in the following theorem.

> **Theorem 2.2** *Suppose $\mathcal{F} \subset \{0,1\}^n$, $\mathcal{F}_i = \mathcal{F} \cap \{\mathbf{x} \in \{0,1\}^n \mid x_1 = i\}$, $i = 0,1$, and that the inequality*
>
> $$\sum_{j=2}^{n} a_j x_j \leq a_0 \tag{2.18}$$
>
> *is valid for $\mathcal{F}_0$.*
>
> (a) *If $\mathcal{F}_1 = \emptyset$, then $x_1 \leq 0$ is valid for $\mathcal{F}$.*

(b) If $\mathcal{F}_1 \neq \emptyset$, then the inequality

$$a_1 x_1 + \sum_{j=2}^{n} a_j x_j \leq a_0 \tag{2.19}$$

is valid for $\mathcal{F}$ for any $a_1 \leq a_0 - Z$, where Z is the optimal value of the problem of maximizing $\sum_{j=2}^{n} a_j x_j$ subject to $\mathbf{x} \in \mathcal{F}_1$.

(c) If $a_1 = a_0 - Z$ and inequality (2.18) defines a face of dimension k of $\text{conv}(\mathcal{F}_0)$, then inequality (2.19) gives a face of dimension $k+1$ of $\text{conv}(\mathcal{F})$.

Proof.

(a) If $\mathcal{F}_1 = \emptyset$, then clearly $x_1 \leq 0$ is valid for $\mathcal{F}$.

(b) If $\mathcal{F}_1 \neq \emptyset$, let $\mathbf{x} \in \mathcal{F}_0$. Then,

$$a_1 x_1 + \sum_{j=2}^{n} a_j x_j \leq a_0.$$

If $\mathbf{x} \in \mathcal{F}_1$, then

$$a_1 x_1 + \sum_{j=2}^{n} a_j x_j \leq a_1 + Z \leq a_0.$$

Therefore, the inequality $a_1 x_1 + \sum_{j=2}^{n} a_j x_j \leq a_0$ is valid for $\mathcal{F} = \mathcal{F}_0 \cup \mathcal{F}_1$.

(c) If inequality (2.18) defines a face of dimension k of $\text{conv}(\mathcal{F}_0)$, then there are $k+1$ affinely independent vectors $\mathbf{x}_1, \ldots, \mathbf{x}_{k+1}$ in $\mathcal{F}_0$ satisfying the inequality with equality. Let $\mathbf{x}^*$ be an optimal solution of the problem of maximizing $\sum_{j=2}^{n} a_j x_j$ subject to $\mathbf{x} \in \mathcal{F}_1$. For $a_1 = a_0 - Z$, we have

$$a_1 x_1^* + \sum_{j=2}^{n} a_j x_j^* = a_1 + Z = a_0.$$

Therefore, the vectors $(\mathbf{x}^*, \mathbf{x}_1, \ldots, \mathbf{x}_{k+1})$ satisfy inequality (2.19) with equality. They are also affinely independent since $\mathbf{x}^* \in \mathcal{F}_1$ and $\mathbf{x}_i \in \mathcal{F}_0$ for $i = 1, \ldots, k+1$. Therefore, inequality (2.19) gives a face of dimension $k+1$ of $\text{conv}(\mathcal{F})$. $\square$

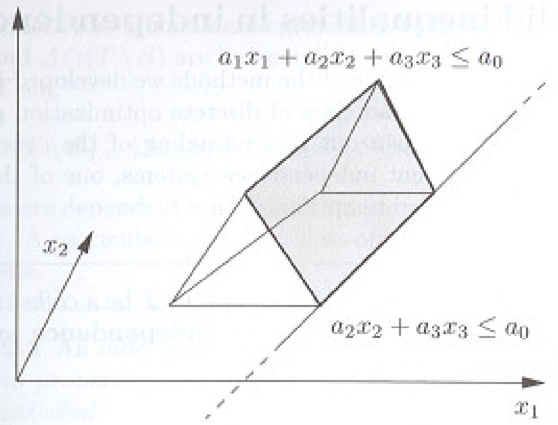


Figure 2.5: A pictorial representation of lifting.

Figure 2.5 illustrates the process of lifting. Theorem 2.2 is meant to be used sequentially. Given a subset $N_1 \subset \{1, \ldots, n\}$ and an inequality $\sum_{j \in N_1} a_j x_j \leq a_0$ that is valid for $\mathcal{F} \cap \{\mathbf{x} \in \{0,1\}^n \mid x_j = 0 \text{ for } j \notin N_1\}$, we lift one variable at a time to obtain a valid inequality of the form

$$\sum_{j \notin N_1} \pi_j x_j + \sum_{j \in N_1} a_j x_j \leq a_0.$$

Notice that the coefficients π_j depend on the order in which the variables are lifted. Different orderings lead to different valid inequalities for $\mathcal{F}$ as we show next.

Example 2.9 (The knapsack problem) Consider the knapsack problem

$$\mathcal{F} = \{\mathbf{x} \in \{0,1\}^6 \mid 5x_1 + 5x_2 + 5x_3 + 5x_4 + 3x_5 + 8x_6 \leq 17\}.$$

The inequality $x_1 + x_2 + x_3 + x_4 \leq 3$ is valid for $\text{conv}(\mathcal{F} \cap \{x_5 = x_6 = 0\})$. Applying lifting first on variable x_5 and then on variable x_6, yields the inequality $x_1 + x_2 + x_3 + x_4 + x_5 + x_6 \leq 3$. If we apply lifting first on variable x_6 and then on variable x_5, the resulting inequality is $x_1 + x_2 + x_3 + x_4 + 2x_6 \leq 3$.

Note that in order to perform the lifting procedure we need to solve a couple of integer optimization problems that might appear as difficult as the original problem. Sometimes they are not. In Exercise 2.18 we present an example in which the lifting coefficients can be computed in polynomial time, although the original problem is $\mathcal{NP}$-hard. It is however true that for many integer optimization problems the lifting procedure can hardly be implemented in the way we presented it, because computing the coefficients step by step is just too expensive. In such cases, we resort to lower bounds on the coefficients that we obtain from relaxations.

2.3 Valid inequalities in independence systems

In this section, we apply some of the methods we developed in the previous two sections to a very broad class of discrete optimization problems. Our objective is both to deepen our understanding of the concepts we introduced, but also to present independence systems, one of the most fundamental structures in discrete optimization.

> **Definition 2.2** *Let N be a finite set and let $\mathcal{I}$ be a collection of subsets from N. The pair $(N, \mathcal{I})$ is called an* **independence system** *if the following two properties are satisfied:*
>
> (a) $\emptyset \in \mathcal{I}$,
>
> (b) *if $A \subseteq B$ and $B \in \mathcal{I}$, then $A \in \mathcal{I}$.*

Independence systems try to capture combinatorial structures that exhibit hereditary properties, in the sense that if a structure has such a property, then all subsets of the structure inherit the property. We next present several examples of independence systems.

(a) **Node disjoint paths.** Let $G = (V, E)$ be an undirected graph and let $\mathcal{I}_1$ be the collection of node disjoint paths in G. In particular, each element of $\mathcal{I}_1$ is the disjoint union of paths in G. The pair $(E, \mathcal{I}_1)$ is an independence system, since subpaths of a collection of node disjoint paths are again node disjoint paths.

(b) **Acyclic subgraphs.** Let $\mathcal{I}_2$ be the collection of acyclic subgraphs (forests) in an undirected graph $G = (V, E)$. Then, $(E, \mathcal{I}_2)$ is an independence system, since subgraphs of acyclic graphs are also acyclic.

(c) **Linear independence.** Let $\mathbf{A}$ be an $m \times n$ matrix, let N be the index set of columns of $\mathbf{A}$, and let $\mathcal{I}_3$ be the collection of linearly independent columns of $\mathbf{A}$. Clearly, $(N, \mathcal{I}_3)$ is also an independence system.

(d) **Feasible solutions to packing problems.** Let $\mathcal{F} = \{\mathbf{x} \in \{0,1\}^n \mid \mathbf{Ax} \leq \mathbf{b}\}$ with $\mathbf{A} \geq \mathbf{0}$. Let $N = \{1, 2, \ldots, n\}$. Associated with a feasible solution $\mathbf{x} \in \mathcal{F}$, we consider the set $A(\mathbf{x}) = \{i \mid x_i = 1\}$. Let $\mathcal{I}_4 = \cup_{\mathbf{x} \in \mathcal{F}} A(\mathbf{x})$. Then, $(N, \mathcal{I}_4)$ is an independence system. Special cases include the set of feasible solutions of the knapsack, the stable set problem, among many others.

Let $(N, \mathcal{I})$ be an independence system. Each member of $\mathcal{I}$ is called **independent** and every set $T \subseteq N$, $T \notin \mathcal{I}$ is called **dependent**. A set $S \in \mathcal{I}$ is **maximal** with respect to another set $T \subset N$, if $S \cup \{j\} \notin \mathcal{I}$, $\forall$ $j \in T \setminus S$. Such a set S is called **a basis** of T. The maximum cardinality of a basis of T, denoted by $r(T)$, is called **the rank** of T, that is

$$r(T) = \max\{|S| : S \in \mathcal{I}, S \subset T\}.$$

Let $S \subseteq T$, and let A be a maximum cardinality basis of T, i.e., $|A| = r(T)$. Then, $A \cap S$ and $A \cap (T \setminus S)$ are independent sets contained in S and $T \setminus S$ respectively. Then,

$$r(S) + r(T \setminus S) \geq |A \cap S| + |A \cap (T \setminus S)| = |A| = r(T). \tag{2.20}$$

Circuits are dependent sets $F \subseteq N$ such that, for every $i \in F$, $F \setminus \{i\}$ is independent. A particular important class of independence systems is the class of matroids.

> **Definition 2.3** *An independence system $(N, \mathcal{I})$ is called a* **matroid** *if in addition to properties* **(a)** *and* **(b)** *of Definition 2.2, the following condition is satisfied:*
>
> (c) *Every maximal independent set contained in F has the same cardinality $r(F)$, for all $F \subset N$.*

We next examine whether the four independent systems we introduced earlier are matroids.

Example 2.10 (Node disjoint paths) We consider the independence system $(E, \mathcal{I}_1)$ of node disjoint paths in a complete graph on 6 nodes denoted by K_6. Consider the following subset $F = \{(1,2), (2,3), (2,4), (4,5), (4,6)\}$ of E. Two maximal independent sets contained in F are $\{(1,2), (2,4), (4,5)\}$ corresponding to the path $1-2-4-5$ and $\{(1,2), (2,3), (4,5), (4,6)\}$ corresponding to the paths $1-2-3$, $5-4-6$. Since these independent sets have different cardinalities, $(E, \mathcal{I}_1)$ is not a matroid.

Example 2.11 (Acyclic subgraphs) Given a graph $G = (V, E)$, we consider the independence system $(E, \mathcal{I}_2)$ of forests, i.e., subsets of edges that do not contain a cycle. Dependent sets are edge sets that contain a cycle. Given, a set F of edges that forms one connected component and spans all the nodes in V, maximal independent sets contained in F correspond to spanning trees on F, and thus have the same cardinality. If (V, F) has k connected components and spans all the nodes in V, then maximal independent sets contained in F correspond to collections of spanning trees of the components of F and have the same cardinality $r(F) = |V| - k$. Therefore, $(E, \mathcal{I}_2)$ is a matroid.

Example 2.12 (Linear independence) Given an $m \times n$ matrix $\mathbf{A}$ with N denoting the index set of columns of $\mathbf{A}$, we consider the independence system $(N, \mathcal{I}_3)$ of linearly independent columns of $\mathbf{A}$. If $T \subset N$ is a collection of indices corresponding to columns of the matrix $\mathbf{A}$, we denote by $\mathbf{A}_T = [\mathbf{A}_j]_{j \in T}$. It is a central result in linear algebra that all maximal independent sets contained in T have the same cardinality, which is equal to the rank of the matrix $\mathbf{A}_T$, i.e., $r(T) = \text{rank}(\mathbf{A}_T)$. Thus, the independence system $(N, \mathcal{I}_3)$ is a matroid.

Example 2.13 (Feasible solutions to packing problems) Let $\mathcal{F} = \{\mathbf{x} \in \{0,1\}^4 \mid x_1 + 2x_2 + 3x_3 + 4x_4 \leq 4\}$ and $N = \{1,2,3,4\}$. We consider the independence system $(N, \mathcal{I}_4)$ of feasible solutions to $\mathcal{F}$. The set $A = \{1,3,4\}$ is a dependent set. The sets $\{1,3\}$ and $\{4\}$ are maximal independent sets, but they do not have the same cardinality, thus $(N, \mathcal{I}_4)$ is not a matroid.

Our goal in this section is to give valid inequalities for independence systems $(N, \mathcal{I})$. Given a set $F \in \mathcal{I}$, we define a vector $\mathbf{x}^F$ as follows: x_i^F is equal to one, if $i \in F$, and zero, otherwise. Let

$$\mathcal{F} = \{\mathbf{x}^F \in \{0,1\}^{|N|} \mid F \in \mathcal{I}\}.$$

Note that all elements of N can be assumed to be independent sets; otherwise, we can delete them from N. Thus,

$$\dim(\text{conv}(\mathcal{F})) = |N|.$$

Let $C \subseteq N$ define a circuit in the independence system $(N, \mathcal{I})$. Then, the inequality (called circuit inequality)

$$\sum_{i \in C} x_i \leq |C| - 1$$

is valid for conv($\mathcal{F}$). Using the set $\mathcal{C}$ of all circuits of independence system $(N, \mathcal{I})$, we formulate the problem of optimizing a linear objective function $\mathbf{c}'\mathbf{x}$ over the independence system $(N, \mathcal{I})$ as the integer optimization problem:

$$\begin{aligned} \text{maximize} \quad & \mathbf{c}'\mathbf{x} \\ \text{subject to} \quad & \sum_{i \in C} x_i \leq |C| - 1, \quad \text{for all } C \in \mathcal{C}, \\ & \mathbf{x} \in \{0,1\}^n. \end{aligned}$$

Example 2.14 (Circuit constraints for the knapsack problem) Let $\mathcal{F} = \{\mathbf{x} \in \{0,1\}^4 \mid 9x_1 + 17x_2 + 25x_3 + 33x_4 \leq 50\}$. The circuits correspond to the sets $\{1,2,3\}$, $\{3,4\}$, $\{1,2,4\}$. The corresponding valid inequalities are:

$$\begin{aligned} x_1 + x_2 + x_3 &\leq 2, \\ x_3 + x_4 &\leq 1, \\ x_1 + x_2 + x_4 &\leq 2. \end{aligned}$$

A natural generalization of the family of circuit inequalities is the family of rank inequalities. Let $T \subseteq N$. Then, the rank inequality

$$\sum_{i \in T} x_i \leq r(T) \tag{2.21}$$

is valid for conv($\mathcal{F}$), since the right hand side reflects the maximum cardinality of an independent set contained in T. Note that for T being a circuit, $r(T) = |T| - 1$ and we obtain the circuit inequalities as a special case.

A rank inequality does not necessarily define a facet of conv($\mathcal{F}$). To capture the properties that are needed for this to happen, we give the following definition.

Definition 2.4 (Closed and nonseparable sets)

(a) *A set T is called* **closed** *if $r(T \cup \{i\}) = r(T) + 1$, for all $i \in N \setminus T$.*

(b) *A set T is called* **nonseparable** *if for every $S \subseteq T$, $\emptyset \neq S \neq T$ we have that $r(T) < r(S) + r(T \setminus S)$.*

Example 2.15 Let $N = \{1, \ldots, 5\}$ and $\mathcal{I}$ the set of all subsets of N that do not contain $T = \{2,3,4\}$, which is a circuit. In addition, T is closed and nonseparable. It is closed since $r(T \cup \{i\}) = 3 = 2 + 1 = r(T) + 1$, for $i = 1, 5$. It is nonseparable, since $r(T) = 2$, $r(S) = |S|$ and $r(T \setminus S) = |T \setminus S|$, for all $S \subseteq T$, $S \neq \emptyset, T$, and thus $2 = r(T) < 3 = r(S) + r(T \setminus S)$.

Proposition 2.5 *Suppose the rank inequality $\sum_{i \in T} x_i \leq r(T)$ is facet defining for conv($\mathcal{F}$). Then, T is nonseparable and closed.*

Proof.
In order to derive a contradiction, suppose first that T is separable, that is, there exists an $S \subseteq T$, such that $r(T) = r(S) + r(T \setminus S)$ (see also Eq. (2.20)). Then, every integral point $\mathbf{x} \in \text{conv}(\mathcal{F})$ satisfying $\sum_{i \in T} x_i = r(T)$ also satisfies the two valid linearly independent inequalities $\sum_{i \in S} x_i \leq r(S)$ and $\sum_{i \in T \setminus S} x_i \leq r(T \setminus S)$ with equality. From Proposition A.2(b), the maximum number of affinely independent points in conv($\mathcal{F}$) satisfying $\sum_{i \in T} x_i = r(T)$ is less than or equal to

$$\dim(\text{conv}(\mathcal{F})) + 1 - 2 = \dim(\text{conv}(\mathcal{F})) - 1.$$

This contradicts the definition of a facet.
If T is not closed, then there exists $j \in N \setminus T$ such that $r(T \cup \{j\}) \leq r(T)$. Therefore,

$$\sum_{i \in T} x_i + x_j \leq r(T \cup \{j\}) \leq r(T)$$

is a valid inequality for conv($\mathcal{F}$). Moreover, $-x_j \leq 0$ is also a valid inequality for conv($\mathcal{F}$). Since every integral point $\mathbf{x} \in \text{conv}(\mathcal{F})$ satisfying $\sum_{i \in T} x_i = r(T)$ also satisfies $\sum_{i \in T} x_i + x_j = r(T)$ and $-x_j = 0$, we derive a contradiction using the previous reasoning. □

In Section 3.2, we show that if the independence system $(N, \mathcal{I})$ is a matroid, then the rank inequalities completely characterize conv($\mathcal{F}$), that is,

$$\text{conv}(\mathcal{F}) = \left\{ \mathbf{x} \in \mathbb{R}^n_+ \,\middle|\, \sum_{i \in T} x_i \leq r(T),\ T \subseteq N \right\}.$$

Moreover, in this case we can identify exactly those rank inequalities that are facet defining. The proof of the proposition is outlined in Exercise 2.11.

> **Proposition 2.6** *Let $(N, \mathcal{I})$ be a matroid. The rank inequality*
>
> $$\sum_{i \in T} x_i \leq r(T)$$
>
> *defines a facet of conv($\mathcal{F}$) if and only if T is nonseparable and closed.*

2.4 On the strength of valid inequalities

We have argued so far that the proximity of a relaxation to the convex hull of integer solutions determines its quality. We used the dimension of the face that a valid inequality represents as an indication of its strength. Since it is difficult to compute the dimension of a face, we introduce, in this section, an alternative measure of strength for a valid inequality of a particular type of polyhedra, which is relatively easy to compute and leads to new insights about the quality of formulations.

Let $\mathcal{F}$ be the set of integer feasible vectors of an integer optimization problem. Let P_1 and P_2 be two relaxations. We assume that the polyhedra P_i, $i = 1, 2$ are nonempty and are given as follows:

$$P_i = \{\mathbf{x} \in \mathbb{R}^n_+ \mid \mathbf{A}_i \mathbf{x} \geq \mathbf{b}_i\}, \qquad i = 1, 2,$$

where all entries in $\mathbf{A}_i$, $\mathbf{b}_i$ are nonnegative and $\mathbf{0} \notin P_i$. We will say that the polyhedra P_i, $i = 1, 2$ are of **covering type**. Let $\alpha > 0$, and let P be a polyhedron of covering type. We denote by $\alpha P = \{\alpha \mathbf{x} \mid \mathbf{x} \in P\}$, the dilation of P by a factor of α. We give the following definition.

> **Definition 2.5** *Let P_1, P_2 be nonempty polyhedra of covering type. The* **strength** *of P_1 with respect to P_2 denoted by $t(P_1, P_2)$ is the minimum value of $\alpha > 0$ such that $\alpha P_1 \subset P_2$.*

Note that given a pair P_1, P_2 of polyhedra of covering type, the strength of P_1 with respect to P_2 need not exist. Consider, for example, $P_1 = \{x \in \mathbb{R} \mid x \geq 0\}$, and $P_2 = \{x \in \mathbb{R} \mid x \geq 1\}$. We can interpret the strength $t(P_1, P_2)$ from an optimization point of view.

> **Proposition 2.7** *Let P_1, P_2 be nonempty polyhedra of covering type. Then, $\alpha P_1 \subset P_2$ if and only if for all $\mathbf{c} \geq \mathbf{0}$, we have that*
>
> $$Z_2 \leq \alpha Z_1,$$
>
> *where Z_i, $i = 1, 2$ is the optimal cost of the problem of minimizing $\mathbf{c}'\mathbf{x}$ subject to $\mathbf{x} \in P_i$.*

Proof.
If $\alpha P_1 \subset P_2$, then clearly $Z_2 \leq \alpha Z_1$, for all $\mathbf{c} \geq \mathbf{0}$. For the converse result, we assume that $Z_2 \leq \alpha Z_1$, for all $\mathbf{c} \geq \mathbf{0}$, and, for the purpose of deriving a contradiction, we assume that there exists an $\mathbf{x}_0 \in \alpha P_1$, but $x_0 \notin P_2$. By the separating hyperplane theorem, there exists a vector $\mathbf{c}$ such that $\mathbf{c}'\mathbf{x}_0 < \mathbf{c}'\mathbf{x}$, for all $\mathbf{x} \in P_2$, i.e., $\mathbf{c}'\mathbf{x}_0 < Z_2$. Since $\mathbf{x}_0 \in \alpha P_1$, $\mathbf{x}_0 = \alpha \mathbf{y}_0$, $\mathbf{y}_0 \in P_1$. Clearly, with respect to this cost vector $\mathbf{c}$, $Z_1 \leq \mathbf{c}'\mathbf{y}_0$, i.e., $\alpha Z_1 \leq \mathbf{c}'\mathbf{x}_0$, and thus $\alpha Z_1 < Z_2$. We note that $\mathbf{c} \geq \mathbf{0}$, since otherwise $Z_1 = Z_2 = -\infty$. This is a contradiction. □

Proposition 2.7 leads to the following corollary.

> **Corollary 2.1**
>
> $$t(P_1, P_2) = \sup_{\mathbf{c} \geq \mathbf{0}} \frac{Z_2}{Z_1}, \tag{2.22}$$
>
> *where Z_i, $i = 1, 2$ is the optimal cost of the problem of minimizing $\mathbf{c}'\mathbf{x}$ subject to $\mathbf{x} \in P_i$.*

We next illustrate how to compute $t(P_1, P_2)$ if the polyhedron P_2 is known explicitly.

> **Theorem 2.3** *Let P_1, P_2 be nonempty polyhedra of covering type, with $P_2 = \{\mathbf{x} \in \mathbb{R}^n_+ \mid \mathbf{a}_i'\mathbf{x} \geq b_i,\ i = 1, \ldots, m\}$, and $\mathbf{a}_i \geq \mathbf{0}$, $b_i \geq 0$, for all $i = 1, \ldots, m$. Then,*
>
> $$t(P_1, P_2) = \max_{i=1,\ldots,m} \frac{b_i}{d_i},$$
>
> *where d_i is the optimal cost of minimizing $\mathbf{a}_i'\mathbf{x}$ subject to $\mathbf{x} \in P_1$. If $d_i = 0$, then $t(P_1, P_2)$ is defined to be $+\infty$.*

Proof.
From Eq. (2.22), it follows

$$t(P_1,P_2) \geq \max_{i=1,\ldots,m} \frac{\min\{\mathbf{a}_i'\mathbf{x} \;:\; \mathbf{x}\in P_2\}}{\min\{\mathbf{a}_i'\mathbf{x} \;:\; \mathbf{x}\in P_1\}} \geq \max_{i=1,\ldots,m} \frac{b_i}{d_i}.$$

Let $\mathbf{c}\geq \mathbf{0}$. We consider the problem

$$\begin{array}{ll} \text{minimize} & \mathbf{c}'\mathbf{x} \\ \text{subject to} & \mathbf{A}\mathbf{x}\geq \mathbf{b}, \\ & \mathbf{x}\geq \mathbf{0}, \end{array}$$

and its dual

$$\begin{array}{ll} \text{maximize} & \mathbf{b}'\mathbf{p} \\ \text{subject to} & \mathbf{A}'\mathbf{p}\leq \mathbf{c}, \\ & \mathbf{p}\geq \mathbf{0}. \end{array} \tag{2.23}$$

Since $\mathbf{c}\geq\mathbf{0}$, the primal problem has bounded cost, and thus its cost Z_2 is equal to the optimal cost of the dual problem. Let $\mathbf{p}^*$ be a dual optimal solution to problem (2.23). Let $\mathbf{x}^*$ be an optimal solution of the problem

$$\begin{array}{ll} \text{minimize} & \mathbf{c}'\mathbf{x} \\ \text{subject to} & \mathbf{x}\in P_1 \end{array}$$

with optimal cost equal to Z_1. Then,

$$Z_1 = \mathbf{c}'\mathbf{x}^* \geq (\mathbf{p}^*)'\mathbf{A}\mathbf{x}^* = \sum_{i=1}^m (\mathbf{a}_i'\mathbf{x}^*)p_i^* \geq \sum_{i=1}^m d_i p_i^*,$$

since d_i is the optimal cost of minimizing $\mathbf{a}_i'\mathbf{x}$ subject to $\mathbf{x}\in P_1$ and $\mathbf{x}^*\in P_1$. Therefore,

$$\frac{\min\{\mathbf{c}'\mathbf{x} \;:\; \mathbf{x}\in P_2\}}{\min\{\mathbf{c}'\mathbf{x} \;:\; \mathbf{x}\in P_1\}} \leq \frac{\sum_{i=1}^m b_i p_i^*}{\sum_{i=1}^m d_i p_i^*} = \sum_{i=1}^m \left(\frac{d_i p_i^*}{\sum_{i=1}^m d_i p_i^*}\right)\frac{b_i}{d_i}.$$

Since $p_i^*\geq 0$, and $d_i\geq 0$ (because $\mathbf{a}_i\geq\mathbf{0}$ and $P_1\subset\mathbb{R}_+^n$), the latter expression is a convex combination of b_i/d_i, and therefore

$$\frac{\min\{\mathbf{c}'\mathbf{x} \;:\; \mathbf{x}\in P_2\}}{\min\{\mathbf{c}'\mathbf{x} \;:\; \mathbf{x}\in P_1\}} \leq \max_{i=1,\ldots,m}\frac{b_i}{d_i}.$$

Taking the supremum over all $\mathbf{c}\geq\mathbf{0}$, we obtain

$$t(P_1,P_2)\leq \max_{i=1,\ldots,m}\frac{b_i}{d_i},$$

and the theorem follows. □

Let P be a relaxation of the covering type to an integer optimization problem. We want to quantify the effect of adding a valid inequality $\mathbf{f}'\mathbf{x}\geq g$ to P. This motivates the following definition.

Definition 2.6 *The strength of an inequality $\mathbf{f}'\mathbf{x}\geq g$, $\mathbf{f}\geq\mathbf{0}$, $g>0$ with respect to a polyhedron $P=\{\mathbf{x}\in\mathbb{R}_+^n \mid \mathbf{A}\mathbf{x}\geq\mathbf{b}\}$ of covering type is defined as g/d, where $d=\min_{\mathbf{x}\in P}\mathbf{f}'\mathbf{x}$.*

It is generally difficult to compute explicitly the optimal cost d. Instead, by strong duality, d is also equal to the optimal cost of the dual problem:

$$\begin{array}{ll} \text{maximize} & \mathbf{b}'\mathbf{p} \\ \text{subject to} & \mathbf{A}'\mathbf{p}\leq \mathbf{f}, \\ & \mathbf{p}\geq \mathbf{0}. \end{array}$$

Let $\overline{\mathbf{p}}$ be a feasible dual solution. By weak duality, $\mathbf{b}'\overline{\mathbf{p}}\leq d$. Then, the strength of inequality $\mathbf{f}'\mathbf{x}\geq g$ with respect to P is at most $g/(\mathbf{b}'\overline{\mathbf{p}})$. In other words, any dual feasible solution provides an upper bound on the strength. We use this method to find the strength of several valid inequalities for relaxations of the traveling salesman problem that are of covering type.

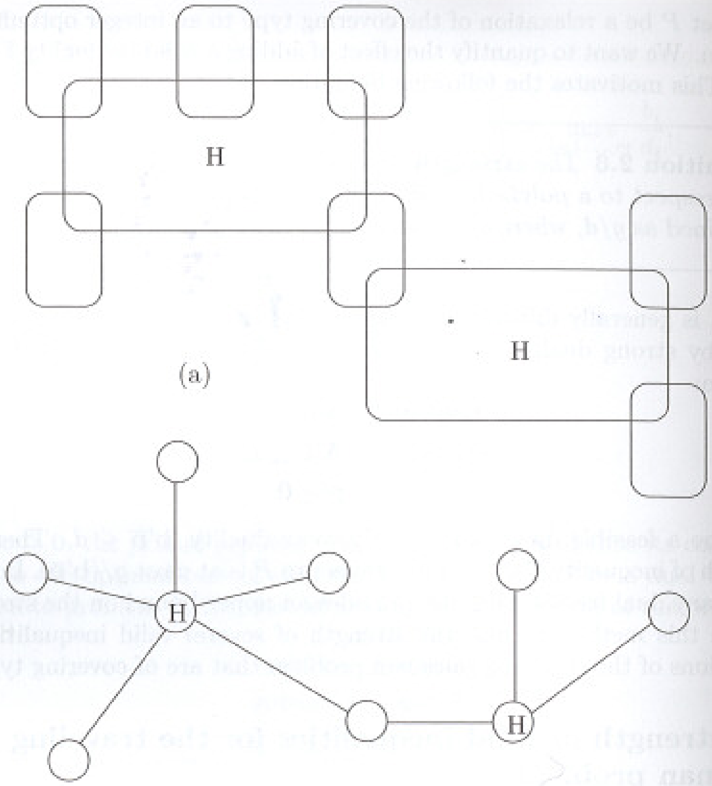
The strength of valid inequalities for the traveling salesman problem

A **clique tree** C (see Figure 2.6(a)) consists of a collection of subsets of nodes partitioned into handles $H_1,\ldots,H_h$ and teeth $T_1,\ldots,T_t$. We define the intersection graph G of the handles and teeth as follows: each handle H_i and each tooth T_j is represented by a node and there exists an edge $\{i,j\}$, if $H_i\cap T_j\neq\emptyset$ (see Figure 2.6(b)). We assume that the sets $H_1,\ldots,H_h$ and teeth $T_1,\ldots,T_t$ satisfy:

(a)	$T_i\cap T_j=\emptyset$,	$i\neq j$,
(b)	$H_i\cap H_j=\emptyset$,	$i\neq j$,
(c)	$\lvert T_i\setminus\cup_{j=1}^h H_j\rvert\geq 1$,	$i=1,\ldots,t$,
(d)	$\lvert i:\ T_i\cap H_j\neq\emptyset\rvert\geq 3$ and odd,	$j=1,\ldots,h$,
(e)	G is a tree.	

When the number of handles h is equal to one, the clique tree is called a comb (recall Example 2.3). The following inequality generalizes the comb inequality (2.6):

$$\sum_{i=1}^h\sum_{e\in\delta(H_i)}x_e+\sum_{i=1}^t\sum_{e\in\delta(T_i)}x_e\geq 3t+2h-1. \tag{2.24}$$



(b)

Figure 2.6: (a) A clique tree with two handles and seven teeth. (b) The intersection graph corresponding to the clique tree.

Theorem 2.4 *The strength of the clique tree inequality with h handles and t teeth with respect to the relaxation of the cutset formulation, in which we do not enforce the degree constraints*

$$\left\{ \mathbf{x} \in \mathbb{R}_+^{|E|} \;\middle|\; \sum_{e\in\delta(S)} x_e \geq 2,\ \forall S \subset V,\ S \neq \emptyset, V \right\} \tag{2.25}$$

is at most

$$\frac{3t+2h-1}{3t+h-1}.$$

Proof.
Using strong duality, the strength is equal to $(3t+2h-1)/d$, where d is the optimal cost of the dual problem

$$\begin{aligned} \text{maximize} \quad & 2 \sum_{S\subset V,\ S\neq\emptyset,V} p(S) & \\ \text{subject to} \quad & \sum_{S:\ e\in\delta(S)} p(S) \leq a_e, & \forall\, e \in E, \\ & p(S) \geq 0, & \forall\, S \subset V, \end{aligned}$$

where a_e denotes the number of occurences of edge $e \in E$ in the set $\cup_{i=1}^h \delta(H_i) \cup_{i=1}^t \delta(T_i)$. One can verify that the solution

$$\begin{aligned} p(H_i \cap T_j) &= \frac{1}{2}, \quad \forall\, i,j \in V,\ H_i \cap T_j \neq \emptyset, \\ p(T_j) &= \frac{1}{2}, \quad \forall j \in V, \\ p(T_j \setminus \cup_{i=1}^h H_i) &= \frac{1}{2}, \quad \forall j \in V, \\ p(S) &= 0, \quad \text{otherwise}, \end{aligned} \tag{2.26}$$

is dual feasible [see Figure 2.6(a)]. Note that, by definition of a clique tree, the sets in Eq. (2.26) are nonempty. We next calculate the cost of this solution. Since the intersection graph G is a tree, the number of (i,j) with $H_i \cap T_j \neq \emptyset$, is equal to $t+h-1$. The cost is then

$$2\sum_{S\subset V} p(S) = (t+h-1) + t + t = 3t + h - 1,$$

where the term $(t+h-1)$ is equal to the number of (i,j) with $H_i \cap T_j \neq \emptyset$, the second term t is equal to the number of T_j, and the third term is equal to the number of $T_j \setminus \cup_{i=1}^h H_i$. Hence, $d \leq 3t+h-1$, and the theorem follows. □

Since a comb inequality corresponds to the case of a clique tree inequality with $h=1$, we obtain that the strength of a comb inequality with t teeth with respect to the relaxation of the cutset polyhedron (2.25) is at most $(3t+1)/3t$, and since $t \geq 3$ it is at most $10/9$.

The strength of several known classes of valid inequalities for the symmetric traveling salesman problem with respect to polyhedron (2.25) is at most $4/3$ (Exercises 2.16 and 2.17).

2.5 Nonlinear formulations

In this section, we use nonlinear formulations to produce stronger relaxations to an integer optimization problem. Our analysis is based on two observations:

(a) The requirement that a variable x takes values in $\{0,1\}$, can be expressed as $x^2 = x$.

(b) It is often simple to express either the objective function or the constraints of a discrete optimization problem as a nonlinear optimization problem. For example, in the stable set problem, we can express the requirement that no adjacent nodes i, j should be selected in the stable set as $x_i x_j = 0$, for $\{i,j\} \in E$.

Given a nonlinear formulation we introduce two methods to obtain stronger relaxations, one based on semidefinite optimization, and the other based on linearizations of the nonlinear formulation.

Semidefinite relaxations

So far, we have used linear relaxations to approximate the convex hull of solutions of an integer optimization problem. We demonstrate how to obtain stronger relaxations by using semidefinite relaxations. We consider the integer optimization problem:

$$\begin{aligned} \text{minimize} \quad & \sum_{j=1}^{n} c_j x_j \\ \text{subject to} \quad & \sum_{j=1}^{n} \mathbf{A}_j x_j = \mathbf{b}, \\ & x_j \in \{0,1\}. \end{aligned} \tag{2.27}$$

We multiply each of the constraints in Eq. (2.27) by x_i and obtain

$$\sum_{j=1}^{n} \mathbf{A}_j x_j x_i = \mathbf{b} x_i, \qquad i = 1, \ldots, n.$$

We introduce the variables $z_{ij} = x_i x_j$. Then, the requirement $x_j \in \{0,1\}$ implies that the following relations hold:

$$\begin{aligned} & z_{ii} = x_i^2 = x_i, && \forall\, i = 1, \ldots, n, \\ & x_i x_j \geq 0 \iff z_{ij} \geq 0, && \forall\, i,j = 1, \ldots, n,\ i \neq j, \\ & x_i(1-x_j) \geq 0 \iff z_{ij} \leq z_{ii}, && \forall\, i,j = 1, \ldots, n,\ i \neq j, \\ & (1-x_i)(1-x_j) \geq 0 \iff z_{ii} + z_{jj} - z_{ij} \leq 1, && \forall\, i,j = 1, \ldots, n,\ i \neq j. \end{aligned}$$

Moreover, the matrix $\mathbf{Z} = \mathbf{x}\mathbf{x}'$ is positive semidefinite, i.e., for all vectors $\mathbf{u} \in \mathbb{R}^n$,

$$\mathbf{u}'\mathbf{Z}\mathbf{u} = \|\mathbf{u}'\mathbf{x}\|^2 \geq 0.$$

The semidefiniteness of $\mathbf{Z}$ is denoted by $\mathbf{Z} \succeq \mathbf{0}$. We consider the optimization problem in the variables $\mathbf{Z} = [z_{ij}]_{i,j=1,\ldots,n}$:

$$\begin{aligned} \text{minimize} \quad & \sum_{j=1}^{n} c_j z_{jj} \\ \text{subject to} \quad & \sum_{j=1}^{n} \mathbf{A}_j z_{ij} - \mathbf{b} z_{ii} = \mathbf{0}, && i = 1, \ldots, n, \\ & \sum_{j=1}^{n} \mathbf{A}_j z_{jj} = \mathbf{b}, \\ & 0 \leq z_{ij} \leq z_{ii}, && i,j = 1, \ldots, n, i \neq j, \\ & 0 \leq z_{ij} \leq z_{jj}, && i,j = 1, \ldots, n, i \neq j, \\ & 0 \leq z_{ii} \leq 1, && i = 1, \ldots, n, \\ & z_{ii} + z_{jj} - z_{ij} \leq 1, && i,j = 1, \ldots, n, i \neq j, \\ & \mathbf{Z} \succeq \mathbf{0}. \end{aligned} \tag{2.28}$$

Let Z_{SD} be the optimal cost of problem (2.28), Z_{IP} be the optimal cost of problem (2.27), and Z_{LP} be the optimal cost of the linear relaxation. Clearly, an optimal solution $\mathbf{x}^*$ to problem (2.27) creates a feasible solution to problem (2.28) through the relations $\mathbf{Z} = \mathbf{x}^*(\mathbf{x}^*)'$, $x_i^* = z_{ii}$. This leads to $Z_{\text{SD}} \leq Z_{\text{IP}}$, i.e., problem (2.28) is a valid relaxation to the initial integer optimization problem. Note that if we require that in addition to $\mathbf{Z} \succeq \mathbf{0}$, $\text{rank}(\mathbf{Z}) = 1$, we obtain an integral solution. Formulation (2.28) is a relaxation of problem (2.27), since we relax the constraint $\text{rank}(\mathbf{Z}) = 1$. Moreover, since the z_{ii} variables of an optimal solution to problem (2.28) define a feasible solution to the linear relaxation, $Z_{\text{LP}} \leq Z_{\text{SD}}$, which implies that the semidefinite relaxation (2.28) is stronger than the linear relaxation. In Section 3.5, we explore the related method of lift and project.

Relaxation (2.28) is an instance of a semidefinite optimization problem in the variables $\mathbf{Z} = [z_{ij}]_{i,j=1,\ldots,n}$. The feasible set of problem (2.28) is convex, since a convex combination of two semidefinite matrices is semidefinite. Thus, problem (2.28) is a convex optimization problem that is efficiently solvable. Indeed, interior point algorithms, similar to the ones proposed for linear optimization, have been proposed for semidefinite optimization problems. These algorithms combine polynomial complexity with excellent empirical performance.

Example 2.16 (The stable set problem, revisited) Given an undirected graph $G = (V, E)$ and weights w_i for each $i \in V$, recall that the weighted stable set problem asks for a collection of nodes S of maximum weight so that no two nodes in S are adjacent. Let $n = |V|$. We introduce the decision variable x_i, $i = 1, \ldots, n$ which is equal to one, if node i is selected in the stable set, and zero, otherwise. The problem can then be formulated as follows:

$$\begin{aligned}
\text{maximize} \quad & \sum_{i=1}^{n} w_i x_i \\
\text{subject to} \quad & x_i + x_j \leq 1, && \forall \{i,j\} \in E, \\
& x_i \in \{0,1\}, && \forall i \in V.
\end{aligned}$$

The semidefinite relaxation (2.28) in this case is

$$\begin{aligned}
\text{maximize} \quad & \sum_{i=1}^{n} w_i z_{ii} \\
\text{subject to} \quad & z_{ij} = 0, && \forall \{i,j\} \in E, \\
& z_{ii} + z_{jj} \leq 1, && \forall \{i,j\} \in E, \\
& z_{ik} + z_{kj} \leq z_{kk}, && \forall \{i,j\} \in E,\ k \neq i,j, \\
& z_{ii} + z_{jj} + z_{kk} \leq 1 + z_{ik} + z_{jk}, && \forall \{i,j\} \in E,\ k \neq i,j, \\
& 0 \leq z_{ij} \leq z_{ii}, && i,j = 1,\ldots,n,\ i \neq j, \\
& 0 \leq z_{ij} \leq z_{jj}, && i,j = 1,\ldots,n,\ i \neq j, \\
& 0 \leq z_{ii} \leq 1, && i = 1,\ldots,n, \\
& z_{ii} + z_{jj} - z_{ij} \leq 1, && i,j = 1,\ldots,n,\ i \neq j, \\
& \mathbf{Z} \succeq \mathbf{0}.
\end{aligned}$$

From our previous discussion $Z_{\mathrm{LP}} \leq Z_{\mathrm{SD}} \leq Z_{\mathrm{IP}}$.

Example 2.17 (The maximum $s-t$ cut problem) Given an undirected graph $G = (V,E)$ and weights w_{ij}, for all $\{i,j\} \in E$, the maximum $s-t$ cut problem asks for a set $S \subset V$ containing node s, but not node t, such that the total weight $\sum_{\{i,j\}\in\delta(S)} w_{ij}$ is maximized. We introduce decision variables x_i equal to one, if node i is selected in the set S, and zero, otherwise. Thus, $x_s = 1$ and $x_t = 0$.

The edge $\{i,j\}$ contributes to the cut if node $i \notin S$ and $j \in S$, or $j \notin S$ and $i \in S$. Therefore, the contribution of edge $\{i,j\}$ can be written as

$$w_{ij}\Big(x_i(1-x_j) + x_j(1-x_i)\Big).$$

Hence, the problem can be formulated as follows:

$$\begin{aligned}
\text{maximize} \quad & \sum_{\{i,j\}\in E} w_{ij}(x_i + x_j - 2x_i x_j) \\
\text{subject to} \quad & x_s = 1, \quad x_t = 0, \\
& x_i \in \{0,1\},\ \forall\, i \in V.
\end{aligned}$$

This leads to the semidefinite relaxation

$$\begin{aligned}
\text{maximize} \quad & \sum_{\{i,j\}\in E} w_{ij}(z_{ii} + z_{jj} - 2z_{ij}) \\
\text{subject to} \quad & z_{ss} = 1, \quad z_{tt} = 0, \quad z_{st} = 0, \\
& 0 \leq z_{ij} \leq z_{ii}, && i,j \in V, i \neq j, \\
& 0 \leq z_{ij} \leq z_{jj}, && i,j \in V, i \neq j, \\
& 0 \leq z_{ii} \leq 1, && i \in V, \\
& z_{ii} + z_{jj} - z_{ij} \leq 1, && i,j \in V, i \neq j, \\
& \mathbf{Z} \succeq \mathbf{0}.
\end{aligned}$$

More general nonlinear models

In the previous subsection, we have introduced semidefinite relaxations of an integer optimization problem. This section is devoted to more general nonlinear models. We present an example of a quadratic formulation of an integer optimization problem. The example emphasizes the idea of using a relaxation that is efficiently solvable. In a second example we introduce a nonlinear formulation that we linearize in order to derive a stronger linear relaxation.

Example 2.18 (A scheduling problem) We are given a set $J = \{1,\ldots,n\}$ of n jobs and m machines. Each job must be processed in any of the m machines; however, if job j is processed on machine i, its processing time is p_{ij}. Every machine processes one job at a time. The completion time of job j in a schedule is denoted by C_j. Given a nonnegative weight w_j for every job j, the objective is to assign jobs to machines, and schedule each machine in order to minimize the weighted completion time $\sum_{j\in J} w_j C_j$.

In Proposition 3.6 we show that if jobs j and k are assigned to machine i, then job j is scheduled before job k on machine i, denoted by $j \prec_i k$ if and only if

$$\frac{w_k}{p_{ik}} > \frac{w_j}{p_{ij}}.$$

In order to model the problem we define decision variables x_{ij} to be one, if job j is assigned to machine i, and zero, otherwise. The completion time of job j is given by

$$C_j = \sum_{i=1}^{m} x_{ij}\left(p_{ij} + \sum_{k \prec_i j} x_{ik} p_{ik}\right),$$

that is, if job j is assigned to machine i, its completion time is the sum of its own processing time p_{ij} and the processing times of all other jobs k that are assigned to machine i, and precede job j on machine i ($k \prec_i j$). In addition, every job j needs to be assigned to a machine, i.e.,

$$\sum_{i=1}^{m} x_{ij} = 1, \qquad \forall j \in J.$$

Thus, the problem can be formulated as follows:

$$\begin{aligned} \text{minimize} \quad & \sum_{j\in J} w_j \sum_{i=1}^{m} x_{ij}\left(p_{ij} + \sum_{k\prec_i j} x_{ik}p_{ik}\right) \\ \text{subject to} \quad & \sum_{i=1}^{m} x_{ij} = 1, \qquad \forall j \in J, \\ & x_{ij} \in \{0,1\}. \end{aligned}$$

Defining the vector $\mathbf{c} \in \mathbb{R}^{nm}$ with $c_{ij} = w_j p_{ij}$, the vector $\mathbf{x} = (x_{ij})$ and the $nm \times nm$ symmetric matrix $\mathbf{D}$ with entries:

$$d_{(ij),(hk)} = \begin{cases} 0, & \text{if } i \neq h \text{ or } j = k, \\ w_j p_{ik}, & \text{if } i = h \text{ and } k \prec_i j, \\ w_k p_{ij}, & \text{if } i = h \text{ and } j \prec_i k, \end{cases}$$

we can write the problem as

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}'\mathbf{x} + \frac{1}{2}\mathbf{x}'\mathbf{D}\mathbf{x} \\ \text{subject to} \quad & \sum_{i=1}^{m} x_{ij} = 1, \qquad \forall j \in J, \\ & x_{ij} \in \{0,1\}. \end{aligned}$$

Unfortunately, the matrix $\mathbf{D}$ is not positive semidefinite, and thus, the quadratic relaxation in which we replace $x_{ij} \in \{0,1\}$ with $0 \leq x_{ij} \leq 1$ is not convex. However, since $x_{ij}^2 = x_{ij}$, we can write

$$\mathbf{c}'\mathbf{x} + \frac{1}{2}\mathbf{x}'\mathbf{D}\mathbf{x} = \frac{1}{2}\mathbf{c}'\mathbf{x} + \frac{1}{2}\mathbf{x}'\left(\mathbf{D} + \operatorname{diag}(\mathbf{c})\right)\mathbf{x},$$

where $\operatorname{diag}(\mathbf{c})$ denotes the diagonal matrix, whose diagonal entries coincide with the entries of $\mathbf{c}$. It can be verified that the matrix $\mathbf{D} + \operatorname{diag}(\mathbf{c})$ is positive definite, and thus, the relaxation

$$\begin{aligned} Z_{\mathrm{CP}} = \min \quad & \frac{1}{2}\mathbf{c}'\mathbf{x} + \frac{1}{2}\mathbf{x}'\left(\mathbf{D} + \operatorname{diag}(\mathbf{c})\right)\mathbf{x} \\ \text{s.t.} \quad & \sum_{i=1}^{m} x_{ij} = 1, \qquad \forall j \in J, \\ & 0 \leq x_{ij} \leq 1. \end{aligned} \tag{2.29}$$

involves a convex quadratic optimization problem, which is efficiently solvable.

Example 2.19 (The multicut problem)

Given an undirected graph $G = (V, E)$, edge weights c_{ij}, for all edges $\{i,j\} \in E$, and a subset H of V, the multicut problem asks for a minimum weight set of edges $S \subset E$, whose removal disconnects each node $i \in H$ from any other node $j \in H$. In particular, after the set of edges in S is removed, the graph is partitioned into components such that every component contains at most one node in H. If $H = \{s,t\}$, the multicut problem becomes the usual minimum $s-t$ cut problem.

Let $H = \{1,\ldots,k\}$. For all $j \in H$ and $u \in V$, we introduce the decision variable y_{ju}, which is equal to one, if node $u \in V$ belongs to the same component as $j \in H$ in a multicut, and zero, otherwise. Obviously, $y_{jj} = 1$, for all $j \in H$, and $y_{ij} = 0$, for all $i,j \in H$. Since node u should be in exactly one component, we have

$$\sum_{j=1}^{k} y_{ju} = 1, \qquad u \in V.$$

The edge $\{u,v\} \in E$ contributes to the cost of the multicut if there is a node $j \in H$ such that j is in the same component with exactly one of the nodes u or v. In this case, the contribution of edge $\{u,v\}$ is:

$$c_{uv}\left(1 - \sum_{j=1}^{k} y_{ju}y_{jv}\right).$$

Therefore, the problem can be formulated as follows:

$$\begin{aligned} \text{minimize} \quad & \sum_{\{u,v\}\in E} c_{uv}\left(1 - \sum_{j=1}^{k} y_{ju}y_{jv}\right) \\ \text{subject to} \quad & \sum_{j=1}^{k} y_{ju} = 1, && \forall\, u \in V, \\ & y_{jj} = 1, && \forall\, j \in H, \\ & y_{ij} = 0, && \forall\, i,j \in H,\ i \neq j, \\ & y_{ju} \in \{0,1\}, && \forall\, j \in H, u \in V. \end{aligned} \tag{2.30}$$

We first prove that even if we relax the integrality constraint $y_{ju} \in \{0,1\}$ in the nonlinear formulation (2.30), the formulation is exact. The proof illustrates the use of probabilistic methods. We will revisit probabilistic arguments in Chapter 3.

Theorem 2.5 *An exact formulation of the multicut problem is obtained if we replace the constraints $y_{ju} \in \{0,1\}$ in Eq. (2.30) by $0 \leq y_{ju} \leq 1$.*

Proof.
Let $\overline{y}_{ju}$ be an optimal solution to the relaxation of problem (2.30). With probability $\overline{y}_{ju}$, we assign node $u \in V$ to the same component as node

$j \in H$. This creates a feasible solution to the multicut problem. Let Z_A be the cost of the resulting solution A. Note that Z_A is a random variable. We next calculate the expected value of Z_A. Let E_{uv} be the event that edge $\{u, v\}$ contributes to the multicut and let $\overline{E}_{uv}$ be its complement. Then, $\overline{E}_{uv} = \cup_{j=1}^{k} B_{uv}^{j}$, where B_{uv}^{j} is the event that nodes u and v are in the same component with node j.

$$\begin{aligned} E[Z_A] &= \sum_{\{u,v\}\in E} c_{uv} P(E_{uv}) \\ &= \sum_{\{u,v\}\in E} c_{uv}(1 - P(\overline{E}_{uv})) \\ &= \sum_{\{u,v\}\in E} c_{uv}(1 - P(\cup_{j=1}^{k} B_{uv}^{j})) \\ &= \sum_{\{u,v\}\in E} c_{uv}\Big(1 - \sum_{j=1}^{k} \overline{y}_{ju}\overline{y}_{jv}\Big), \end{aligned}$$

since the event that node u is assigned to the same component as j is independent from the event that node v is assigned to the same component as j. Since the solution A is always feasible, and its expected value is equal to the value of the relaxation, it follows that the relaxation is exact. □

The previous theorem suggests that the multicut problem can be modeled exactly as a quadratic optimization problem over linear constraints. Unfortunately, the objective function is not convex, and, thus, the relaxation is difficult to solve. For this reason, we obtain valid inequalities by linearizing the nonlinear objective function. For any set $S \subset H$, we have

$$\sum_{j\in S} y_{ju}(1 - y_{jv}) + \sum_{j\notin S} y_{jv}(1 - y_{ju}) \geq 0,$$

and thus, $\sum_{j=1}^{k} y_{ju}y_{jv} \leq \sum_{j\in S} y_{ju} + \sum_{j\notin S} y_{jv}$. We introduce the auxiliary variables x_{uv}, and obtain the linear relaxation:

$$\begin{aligned} \text{minimize} \quad & \sum_{(u,v)\in E} c_{uv}x_{uv} \\ \text{subject to} \quad & \sum_{j=1}^{k} y_{ju} = 1, && \forall\, u \in V, \\ & x_{uv} + \sum_{j\in S} y_{ju} + \sum_{j\notin S} y_{jv} \geq 1, && \forall\, S \subset H,\ \forall\, (u,v) \in E, \\ & y_{jj} = 1, && \forall\, j \in H, \\ & y_{ij} = 0, && \forall\, i,j \in H,\ i \neq j, \\ & 0 \leq y_{ju} \leq 1, && \forall\, j \in H,\ u \in V, \\ & 0 \leq x_{uv} \leq 1, && \forall\, u,v \in V. \end{aligned}$$

Another linearization can be obtained by introducing, in addition to the variables x_{uv}, the auxiliary variables $z_{juv} = y_{ju}y_{jv}$, and use the linearization $z_{juv} \leq y_{ju}$, $z_{juv} \leq y_{jv}$ and $z_{juv} \leq y_{ju} + y_{jv} - 1$. This linearization results in the linear relaxation:

$$\begin{aligned} \text{minimize} \quad & \sum_{(u,v)\in E} c_{uv}x_{uv} \\ \text{subject to} \quad & \sum_{j=1}^{k} y_{ju} = 1, && \forall\, u \in V, \\ & x_{uv} + \sum_{j=1}^{k} z_{juv} = 1, && \forall\, (u,v) \in E, \\ & z_{juv} \leq y_{ju}, && \forall\, j \in H,\ u,v \in V, \\ & z_{juv} \leq y_{jv}, && \forall\, j \in H,\ u,v \in V, \\ & z_{juv} \leq y_{ju} + y_{jv} - 1, && \forall\, j \in H,\ u,v \in V, \\ & y_{jj} = 1, && \forall\, j \in H, \\ & y_{ij} = 0, && \forall\, i,j \in H,\ i \neq j, \\ & 0 \leq y_{ju} \leq 1, && \forall\, j \in H,\ u \in V, \\ & 0 \leq z_{juv} \leq 1, && \forall\, j \in H,\ u,v \in V, \\ & 0 \leq x_{uv} \leq 1, && \forall\, u,v \in V. \end{aligned}$$

2.6 Summary

In this chapter, we demonstrated methods to improve formulations. The unifying idea in this chapter is the use of continuous (linear and nonlinear) relaxations of the convex hull of integer feasible solutions. While only facet defining inequalities are needed to describe the convex hull, they are hard to identify, and it is often difficult to prove that they are facet defining. For this reason, in designing practical algorithms, we often use valid inequalities that are not facet defining. Moreover, improved formulations sometimes involve semidefinite constraints, and an exponential number of linear inequalities. Therefore, the need arises to be able to solve efficiently such problems. This is addressed in Chapter 5.

2.7 Exercises

Exercise 2.1 Apply the methods of this chapter to find valid inequalities for the integer optimization problem

$$\begin{aligned} \text{maximize} \quad & 9x_1 + 2x_2 + 5x_3 + 7x_4 + 8x_5 + 3x_6 \\ \text{subject to} \quad & 7x_1 + 12x_2 + 5x_3 + 8x_4 + 7x_5 + 11x_6 \leq 26, \\ & x_1, \ldots, x_6 \in \{0, 1\}. \end{aligned}$$

By adding valid inequalities, solve the problem using a linear optimization solver.

Exercise 2.2 Let C be the convex hull of all 0/1 solutions of the knapsack problem $\sum_{i=1}^n a_i x_i \le b$ where $a \in \mathbb{Z}_+^n$ and $b \in \mathbb{Z}_+$.

(a) Prove that for every $\lambda \in \mathbb{Z}_+$ the inequality

$$\sum_{i=1}^n \left\lfloor \frac{a_i}{\lambda} \right\rfloor x_i \le \left\lfloor \frac{b}{\lambda} \right\rfloor$$

is valid for C.

(b) Determine at least two different conditions on $a_1, \dots, a_n$, b and λ such that the inequality induces a facet of C.

Exercise 2.3 Let $G = (V, E)$ be an undirected graph where each node has degree at least three. Let

$$\mathcal{F} = \left\{ \mathbf{x} \in \{0,1\}^n \;\middle|\; x_{e'} - \sum_{e \in \delta(\{v\}) \setminus \{e'\}} x_e \le 0, \ \forall\, e' \in \delta(\{v\}),\ v \in V \right\}.$$

(a) Show that the inequalities $x_e \le 1$ are facet defining for $\operatorname{conv}(\mathcal{F})$.

(b) Show that the inequalities $x_{e'} - \sum_{e \in \delta(\{v\}) \setminus \{e'\}} x_e \le 0$ are facet defining for $\operatorname{conv}(\mathcal{F})$.

Exercise 2.4* (The Steiner tree problem) Given an undirected graph $G = (V, E)$, a set of nodes $T \subset V$ and costs c_{ij} for $\{i, j\} \in E$, the Steiner tree problem asks for a tree of minimum cost, such that all nodes in T are connected. Let $\mathcal{F}$ the set of feasible 0/1 vectors corresponding to Steiner trees.

(a) Formulate the problem as an integer optimization problem using the decision variables $x_e, e \in E$, which is equal to one, if edge e is contained in the solution, and zero, otherwise.

(b) Let $C_1, \dots, C_k$ be a partition of V with $T \cap C_i \neq \emptyset$ for $i = 1, \dots, k$. Show that the inequality

$$\sum_{e \in \delta(C_1, \dots, C_k)} x_e \ge k - 1 \qquad (2.31)$$

is valid for $\operatorname{conv}(\mathcal{F})$.

(c) Show that inequality (2.31) is facet defining for $\operatorname{conv}(\mathcal{F})$ if

(i) The graph defined by shrinking each node set C_i into a single node is two connected, i.e., there exist at least two edge disjoint paths between every pair of nodes.

(ii) For $i = 1, \dots, k$, the subgraph induced by each C_i is connected.

Exercise 2.5 Let $a_0 \in \mathbb{R} \setminus \mathbb{Z}$, $a_1, \dots, a_n > 0$ and

$$\mathcal{F} = \left\{ \mathbf{x} \in \mathbb{R}_+^n \;\middle|\; x_0 = a_0 - \sum_{j=1}^n a_j x_j \in \mathbb{Z} \right\}.$$

Prove that the inequality

$$\sum_{j=1}^n \frac{a_j}{a_0 - \lfloor a_0 \rfloor} x_j \ge 1$$

is valid for $\operatorname{conv}(\mathcal{F})$.

Exercise 2.6

(a) If $h : \mathbb{R}^k \mapsto \mathbb{R}$ and $f_i : \mathbb{R}^m \mapsto \mathbb{R}$ for $i = 1, \dots, n$, are superadditive and nondecreasing functions, then show that the function $h(f_1, \dots, f_k)$ is superadditive and nondecreasing.

(b) Let $f, g : \mathbb{R}^k \mapsto \mathbb{R}$ be superadditive and nondecreasing functions. Show that the following functions are superadditive and nondecreasing:

(i) λf with $\lambda \ge 0$,

(ii) $\lfloor f \rfloor$,

(iii) $f + g$,

(iv) $\min(f, g)$.

Exercise 2.7 Show that the following functions are superadditive and nondecreasing:

(a) Given $\alpha \ge 0$, $F_\alpha : \mathbb{R} \mapsto \mathbb{R}$ such that

$$F_\alpha(d) = \begin{cases} \lfloor d \rfloor, & \text{if } d - \lfloor d \rfloor \le \alpha, \\ \lfloor d \rfloor + \dfrac{d - \lfloor d \rfloor - \alpha}{1 - \alpha}, & \text{otherwise.} \end{cases}$$

(b) $H_\alpha : \mathbb{R}^2 \mapsto \mathbb{R}$

$$H_\alpha(d_1, d_2) = \frac{d_1}{1 - \alpha} + F_\alpha(d_2 - d_1).$$

(c) If F_1, $F_2 : \mathbb{R} \mapsto \mathbb{R}$ are superadditive and nondecreasing, then prove that the function

$$G(d) = \frac{1}{1 - \alpha} F_1(d) + F_\alpha(F_2(d) - F_1(d))$$

is superadditive and nondecreasing.

Exercise 2.8 Given $0 < \alpha \le 1$, show that the function

$$f_\alpha(d) = \begin{cases} d - \lfloor d \rfloor, & \text{if } 0 \le d - \lfloor d \rfloor \le \alpha, \\ \dfrac{(d - \lfloor d \rfloor - 1)\alpha}{1 - \alpha}, & \alpha < d - \lfloor d \rfloor < 1, \end{cases}$$

is superadditive, but not nondecreasing.

Exercise 2.9* (The convex hull of symmetric traveling salesman tours) Let $\mathcal{F}$ be the set of vectors corresponding to feasible tours on an undirected complete graph on n nodes in formulation (1.5).

(a) Show that $\dim(\operatorname{conv}(\mathcal{F})) = n(n-1)/2 - n$.

(b) Show that the inequalities $\sum_{e \in E(S)} x_e \le |S| - 1$, are facet defining in $\operatorname{conv}(\mathcal{F})$ for all S with $2 \le |S| \le n/2$, and $n \ge 4$.

Exercise 2.10 Let $G = (V, E)$ be an undirected graph. Let $C = \{1, 2, \ldots, 2k + 1\} \subseteq V$ be a cycle of odd cardinality, such that there are no edges $\{i, j\}$ for $j \neq i + 1$. Let h be a node such that $\{i, h\} \in E$, for all $i \in C$. Use the method of disjunctions to show that the inequality

$$\sum_{i \in C} x_i + \frac{|C| - 1}{2} x_h \leq \frac{|C| - 1}{2}$$

is valid for the convex hull of stable sets in G.

Exercise 2.11 Consider a matroid $(N, \mathcal{I})$, and let $\text{conv}(\mathcal{F})$ be the convex hull of solutions. In this exercise, we outline a proof of Proposition 2.6. Let $\mathbf{c}'\mathbf{x} \leq \gamma$ be a facet defining inequality of $\text{conv}(\mathcal{F})$ such that

$$\left\{ \mathbf{x} \in \text{conv}(\mathcal{F}) \,\middle|\, \sum_{i \in T} x_i = r(T) \right\} \subseteq \{\mathbf{x} \in \text{conv}(\mathcal{F}) \mid \mathbf{c}'\mathbf{x} = \gamma\}.$$

(a) Show that if $j \in N \setminus T$, then $c_j = 0$.

(b) Let $i \in T$, and $N(i) = \{j \in T \mid$ there exists a basis B of T, $i \notin B$ and $B \setminus \{j\} \cup \{i\}$ is a basis of $T\}$. Let $k \in N(i)$. Show that $c_i = c_k$.

(c) Consider the graph $G = (T, E)$ where $E = \{(i, j) \mid j \in N(i),\ i \in T\}$. Show that G is connected.

(d) Show that $c_i = c_k$, for all $i, k \in T$.

(e) Use the previous parts to complete the proof of Proposition 2.6.

Exercise 2.12 Let $n, t, q \in V$, $n \geq t \geq q \geq 2$ and $N = \{0, \ldots, n - 1\}$. For $i \in N$ we denote by $N_i = \{i, i + 1, \ldots, i + t - 1\}$ the set of t consecutive elements in N, starting with i. All indices are taken modulo n. Let

$$\mathcal{I} = \{I \subseteq N \mid |I \cap N_j| \leq q - 1, \text{ for all } j = 0, \ldots, n - 1\}.$$

(a) Show that $(N, \mathcal{I})$ is an independence system.

(b) Characterize the set $\mathcal{C}$ of circuits of $(N, \mathcal{I})$.

(c) Show that the inequality

$$\sum_{i \in N} x_i \leq \left\lfloor \frac{n(q - 1)}{t} \right\rfloor \tag{2.32}$$

is valid for $\text{conv}(\mathcal{F})$.

(d) There exists a solution in $\text{conv}(\mathcal{F})$ that satisfies (2.32) with equality, i.e., there exists an independent set in $\mathcal{I}$ of cardinality $\alpha = \lfloor (n(q - 1))/t \rfloor$.

Exercise 2.13 Let $G = (V, E)$ where E is an odd cycle and $|V| = n$.

(a) Express the family of stable sets in G as an independence system of the type given in Exercise 2.12.

(b) By applying Exercise 2.12(c), show that the inequality

$$\sum_{i \in V} x_i \leq \left\lfloor \frac{n}{2} \right\rfloor$$

is valid for the convex hull of stable sets.

Exercise 2.14 Express the set of circuits in Example 2.14 as an independence system of the type given in Exercise 2.12, and use this characterization to show that the inequality

$$x_1 + x_2 + x_3 + x_4 \leq 2$$

is valid.

Exercise 2.15* Prove that the class of clique tree inequalities defined in Eq. (2.24):

$$\sum_{i=1}^{h} \sum_{e \in \delta(H_i)} x_e + \sum_{i=1}^{t} \sum_{e \in \delta(T_i)} x_e \geq 3t + 2h - 1,$$

is valid for the convex hull of solutions of the symmetric traveling salesman problem.

Exercise 2.16 (The path inequalities for the symmetric traveling salesman problem)** A **k-path configuration** is defined by an odd integer $k \geq 3$, integers $n_i \geq 2$ for $i = 1, \ldots, k$ and a partition of the node set V into $\{A, B, B_j^i$ for $i = 1, \ldots, k$ and $j = 1, \ldots, n_i\}$, where A and/or B could be empty. For notational convenience we let $B_0^i = A$ and $B_{n_i+1}^i = B$ for $i = 1, \ldots, k$. Let $\delta(S, T)$ be $\delta(S) \cap \delta(T)$. The k-path inequality is $\mathbf{a}'\mathbf{x} \geq b$, where

$$b = 1 + \sum_{i=1}^{k} \frac{n_i + 1}{n_i - 1},$$

$$a_e = \begin{cases} 1, & e \in \delta(A, B), \\ \dfrac{|j - i|}{n_i - 1}, & e \in \delta(B_j^i, B_l^i),\ \forall\, i,\ |j - l| \leq n_i, \\ \dfrac{1}{n_h - 1} + \dfrac{1}{n_i - 1} + \left| \dfrac{j - 1}{n_h - 1} - \dfrac{l - 1}{n_i - 1} \right|, & e \in \delta(B_j^h, B_l^i),\ j = 1, \ldots, n_h,\ h \neq i,\ l = 1, \ldots, n_i, \\ 0, & \text{otherwise}. \end{cases}$$

(a) When $n_i = p$, for all i, the k-path inequality is called p-regular. Show that the class of 2-regular path inequalities is equivalent with the class of comb inequalities.

(b) Show that the class of k-path inequalities is valid for the convex hull of solutions of the symmetric traveling salesman problem.

(c) The strength of a path inequality with respect to the subtour elimination polyhedron is

$$\frac{1 + \sum_{i=1}^{k} r_i}{\sum_{i=1}^{k} r_i} \leq \frac{k + 1}{k} \leq \frac{4}{3},$$

where $r_i = (n_i + 1)/(n_i - 1)$.

Exercise 2.17 ** (The crown inequalities for the symmetric traveling salesman problem) Consider a partition of V into $4k$ sets $U_1, \ldots, U_{4k}$. Given $i, j = 1, \ldots, 4k$, let $d(i,j)$ be the distance on the cycle $(1, 2, \ldots, 4k, 1)$, i.e., $d(i,j) = \min(|i-j|, 4k - |i-j|)$. The **crown inequality** is $\mathbf{a}'\mathbf{x} \geq b$, where

$$b = 12k(k-1) - 2,$$

$$a_e = \begin{cases} 4k - 6 + d(i,j), & e \in \delta(U_i, U_j),\ 1 \leq d(i,j) \leq 2k-1, \\ 2k - 2, & e \in \delta(U_i, U_{i+2k}),\ 1 \leq i \leq 2k, \\ 0, & \text{otherwise.} \end{cases}$$

(a) Show that the class of crown inequalities is valid for the convex hull of solutions of the symmetric traveling salesman problem.

(b) Show that the strength of a crown inequality with respect to the subtour elimination polyhedron is

$$\frac{6k(k-1) - 1}{6k(k-1) - k} \leq \frac{11}{10}.$$

Exercise 2.18 (Extended weight inequalities for the knapsack problem)** Let $N = \{1, 2, \ldots, n\}$ and consider the set of feasible solutions to the knapsack problem

$$\mathcal{F} = \left\{ x \in \{0,1\}^n \,\middle|\, \sum_{i=1}^n a_i x_i \leq b \right\}.$$

Let $T \subseteq N$. For every $v \in \mathbb{Z}_+$, the covering number of v associated with T is defined

$$\phi(v) = \min \left\{ |A| : \ A \subseteq T,\ \sum_{i \in A} a_i \geq v \right\}.$$

Suppose that $\sum_{i \in T} a_i \leq b$ and let $r = b - \sum_{i \in T} a_i$. We denote by P the subset of $N \backslash T$ such that $a_i \geq r$, for all $i \in P$. We consider the permutation $(\pi_1, \ldots \pi_{|P|})$ of the set P. We define $c_{\pi_1} = \phi(a_{\pi_1} - r)$, and for $i = 2, \ldots, |P|$,

$$\begin{aligned} c_{\pi_i} = |T| - \ \max \quad & \sum_{j \in T \cup \{\pi_1, \ldots, \pi_{i-1}\}} c_j z_j \\ \text{s.t.} \quad & \sum_{j \in T \cup \{\pi_1, \ldots, \pi_{i-1}\}} a_j z_j \leq b - a_{\pi_i}, \\ & z_j \in \{0,1\}, \quad \forall j \in T \cup \{\pi_1, \ldots, \pi_{i-1}\}. \end{aligned}$$

(a) Show that the inequality

$$\sum_{i \in T} x_i + c_{\pi_1} x_{\pi_1} \leq |T|$$

is valid for $\text{conv}(\mathcal{F})$.

(b) By applying lifting to the inequality from part (a) show that the inequality

$$\sum_{i \in T} x_i + \sum_{i \in P} c_i x_i \leq |T|$$

is valid for $\text{conv}(\mathcal{F})$.

(c) Show that

$$\phi(a_{\pi_i} - r) - 1 \leq c_{\pi_i} \leq \phi(a_{\pi_i} - r).$$

(d) Show that the exact value of the coefficients c_i can be computed in time that is polynomial in the encoding length of the input of the original knapsack problem.

Exercise 2.19 (The node cover problem) Given an undirected graph $G = (V, E)$ and weights w_i, for all $i \in E$, the node cover problem asks for a set $S \subset V$ with the property that for every edge $\{i, j\} \in E$, either $i \in S$ or $j \in S$, so that the total weight $\sum_{i \in S} w_i$ is minimized.

(a) Propose a formulation of the problem.

(b) Propose a semidefinite relaxation of the problem.

Exercise 2.20 (Computational Exercise: The multi-dimensional knapsack problem) We consider the N-dimensional knapsack problem

$$\begin{aligned} \text{maximize} \quad & \sum_{i=1}^n c_i x_i \\ \text{subject to} \quad & \sum_{i=1}^n w_{ij} x_i \leq C_j, \quad j = 1, \ldots, N, \\ & x_i \in \{0,1\}. \end{aligned}$$

Our objective is to study computationally the importance of applying valid inequalities in solving this problem. All coefficients w_{ij} are uniformly and independently distributed between 100 to 101 and $C_j = 1000$, $j = 1, \ldots, N$.

(a) **(The one-dimensional knapsack problem)** Study the provided matlab code *knapsack.m*. Execute *knapsack.m* to generate the data file *knapsack.dat*. Formulate *knapsack.mod* and solve the one-dimensional knapsack instance without utilizing in the solver cover cuts. Note that you should include the instruction *setting covers = -1;* in the model file. Ensure that the log file is printed as well. You may want to take a look at the provided file *knapsackN.mod*.

   (i) How long does it take to solve the problem? How many nodes were evaluated?

   (ii) Is the problem solved to optimality or did the solver terminate because the gap tolerance limit was met? What is the (default) gap tolerance limit used? Explain what the gap tolerance limit means, and what is its impact on the solution found. In particular, if the value is reduced, does it take longer or shorter to solve a problem?

(b) **(The two-dimensional knapsack problem)** Study the formulation file *knapsackN.mod* that is used to solve a N-dimensional knapsack instance. Note that N is defined as NbAttribute, $N > 1$ and it is to be read from the appropriate datafile. Modify *knapsack.m* so that it can generate a N-dimensional knapsack problem that can be read from *knapsackN.mod*.

(i) Generate 5 instances of the two-dimensional knapsack problem. Still with the cuts generation disabled, solve these instances. What can you observe regarding the run-time taken to solve these 5 instances?

(ii) In all the instances, is the problem solved to optimality or did we terminate because the terminating tolerance was met?

(c) (The four-dimensional knapsack problem without cuts)

(i) Generate and try solving a four-dimensional knapsack instance, still with cuts generation disabled. What is the gap achieved after a two minute solution time? If you can solve this instance to optimality generate another instance.

(ii) At the end of two minutes, is the number of nodes left in the enumeration tree increasing or decreasing? Can you estimate how much longer it will take to termination?

(iii) During the solver runs, how many improving (improving on previous solutions) integral solutions were found? What prevents us from termination?

(d) (The four-dimensional knapsack problem with cuts)

(i) Enable cuts generation. How long does it take to solve the same instance as in part (c) now? How many nodes do we have to evaluate? How many cover cuts were applied?

(ii) Can you explain why there is a difference in the performance with and without cuts?

(iii) Compare the objective function obtained with and without cuts. Is the optimal solution obtained when cover cuts were disabled?

2.8 Notes and sources

2.1. The idea of using rounding arguments for deriving valid inequalities goes back to Gomory (1960b). Hu (1969) provides a geometric illustration of Gomory's method. He recognized that a Gomory cutting plane is always derived from a supporting hyperplane of the polyhedron associated with the linear optimization problem by rounding the right hand side. Independently, Chvátal (1973a) showed that the iterative process of rounding the right hand sides of supporting hyperplanes indeed yields a description of the integer polyhedron, in case it is bounded. Gomory cuts are a special case of generating cuts from superadditive functions. This work and the group approach for integer optimization started with the pioneering paper by Gomory (1969) about corner polyhedra, generalized in Gomory and Johnson (1972a,b), and revisited in Gomory and Johnson (2003), Gomory et al. (2003) and Aráoz et al. (2003). The principle of using disjunctions for enhancing a formulation is a central topic in the theory of disjunctive optimization, the problem of optimizing over the union

of a finite number of sets, proposed by Balas (1975). Basic references on disjunctive optimization are Blair (1976), Jeroslow (1977), Balas (1979, 1985). The idea of mixed integer rounding has been introduced by Nemhauser and Wolsey (1988).

2.2. One fundamental reference to stable set polyhedra is the early work of Padberg (1973). He introduced the family of clique inequalities and the odd cycle constraints. The family of odd cycle inequalities for the matching polyhedron were proposed by Edmonds (1965). A thorough investigation of the polyhedral structure of the traveling salesman problem is given in Grötschel and Padberg (1985). The method of lifting was first introduced by Padberg (1975) and generalized by Wolsey (1976). For an application of lifting to the traveling salesman problem see Grötschel and Padberg (1979a).

2.3. For an introduction to independence systems, in particular matroids, we recommend the book of Welsh (1976) and the survey Faigle (1991). Polyhedral aspects on matroids are summarized in Grötschel et al. (1988). The paper Edmonds (1971) contains a proof that the rank inequalities together with the lower bound constraints describe the convex hull of all incidence vectors of independent sets of a matroid. This paper also shows that the nonseparability and closedness condition for a rank inequality is necessary and sufficient to define a facet of the associated matroid polyhedron. A generalization of the result to the intersection of two matroids can be found in Edmonds (1979).

2.4. Goemans (1995) introduced the geometric concept of comparing the strength of different classes of valid inequalities. Clique tree inequalities were introduced by Grötschel and Pulleyblank (1986). Comb inequalities were introduced by Chvátal (1973b) for the case of at most one common vertex between a tooth and a handle and by Grötschel and Padberg (1979b) for the general case. Path inequalities were introduced in Cornuéjols et al. (1985), who also proved that they are facet defining. Crown inequalities were introduced in Naddef and Rinaldi (1992), who also proved that they are facet defining.

2.5. The idea to use nonlinear methods for discrete optimization problems has been introduced in Sherali and Adams (1990) and Lovász and Schrijver (1991). Semidefinite optimization methods in discrete optimization problems start with the work of Lovász (1979) on the Shannon capacity of a graph. The semidefinite relaxations for the maximum cut problem are from Delorme and Poljak (1993); Goemans and Williamson (1994a); Poljak and Rendl (1995). The nonlinear formulation for the multicut problem and its linearization is from Bertsimas et al. (1996).