J Glob Optim (2008) 42:533–547
DOI 10.1007/s10898-007-9275-5

A geometric characterization of "optimality-equivalent" relaxations

Walid Ben Ameur · **José Neto**

Received: 17 July 2007 / Accepted: 21 December 2007 / Published online: 15 January 2008

Abstract An optimization problem is defined by an objective function to be maximized with respect to a set of constraints. To overcome some theoretical and practical difficulties, the constraint-set is sometimes relaxed and "easier" problems are solved. This led us to study relaxations providing exactly the same set of optimal solutions. We give a complete characterization of these relaxations and present several examples. While the relaxations introduced in this paper are not always easy to solve, they may help to prove that some mathematical programs are equivalent in terms of optimal solutions. An example is given where some of the constraints of a linear program can be relaxed within a certain limit.

Keywords Convex relaxation · Convex geometry · Sensitivity analysis

1 Introduction

An approach that is commonly taken when dealing with optimization problems consists in relaxing some constraints and solving easier problems. Lagrangean relaxations [17], linear relaxations for integer [18,20,22] and convex problems [3,11,14], semidefinite relaxations [24], convex relaxations [6,7,13,19] are often used to get either optimal or approximate solutions of the original problem.

Given an optimization problem, a relaxation will be said to be optimality-equivalent if it has the same set of optimal solutions as the original problem. We will assume that the set of feasible solutions S of the original problem is convex and the objective function is linear. Notice that S is not necessarily given in an explicit way. It can be, for example, the convex hull of the integer solutions of a linear system. Then we are looking for sets T containing the

W. Ben Ameur (✉) · J. Neto
Institut National des Télécommunications, GET/INT CNRS/SAMOVAR, 9 rue Charles Fourier, 91011 Evry, France
e-mail: walid.benameur@int-edu.eu

J. Neto
e-mail: neto@isima.fr

convex set S such that optimizing a given objective function over S is equivalent to optimizing over T. We will give a full characterization of optimality-equivalent relaxations using some geometric arguments. It is important to notice that the concept of optimality-equivalent relaxations is defined for a given objective function: if we change the objective function, the relaxation is no longer optimality-equivalent.

A lot of valuable work was done in different areas to improve the "quality" of relaxations. For example, when we deal with linear relaxations, we generally look for deep valid inequalities. There is also a lot of research related to combinatorial problems where we try to formulate the problem such that the bound given by the relaxation is "good". A systematic study of the polyhedral structure of combinatorial problems allowed the solution of very large size problems. In some cases, the polyhedron is fully described and the corresponding optimization problems are solved in polynomial time. Valid inequalities can also be automatically generated in the spirit of the Chvátal–Gomory method [8,12]. Another line of research consists in lifting the optimization problem in an appropriate space and projecting on the original one to get tight relaxations [2,15,23].

The set of solutions S can also be a non-polyhedral convex set. Polyhedral relaxations are then obtained using the subgradients of the convex functions defining S. Several kinds of cutting plane algorithms are derived depending on how the relaxation is updated [3,4,11,13,14,19,25].

While some relaxations do not depend on the objective function such as the semidefinite relaxation of the maximum cut problem [10], most of them directly or indirectly depend on the objective function. This occurs, for example, when a cutting plane algorithm is used. It is even known that some hard combinatorial problems become easy for some objective functions: the longest path problem with negative weights, the lot-sizing problem with the Wagner–Whitin property [21], the maximum cut problem with negative weights, etc.

The characterization of optimality-equivalent relaxations presented in this paper will also depend on the objective function.

While the relaxation technique described in the paper does not seem to provide a general algorithmic tool to solve hard problems, it gives more insight on optimality-equivalent relaxations and it helps to prove that some mathematical programs are equivalent. An example is given where some of the constraints of a linear program can be relaxed within a certain limit.

The rest of the paper is organized as follows. Some notation and simple examples are provided in Sect. 2. Some preliminary results are presented in Sect. 3. Section 4 is devoted to the main results of the paper. Then we review more examples and applications in Sect. 5. Finally, a conclusion with some research directions are presented in Sect. 6.

2 Notation and simple examples

We consider the Euclidean affine space $\mathbb{R}^n$ where $n \geq 2$ is the dimension of the space. For $(x, y) \in \mathbb{R}^n \times \mathbb{R}^n$, $d(x, y)$ denotes the Euclidean distance between x and y. Given a vector $(\alpha, \lambda) \in \mathbb{R}^n \times \mathbb{R}$, $H_{(\alpha,\lambda)}$ stands for the hyperplane $H_{(\alpha,\lambda)} = \{x \in \mathbb{R}^n \mid \alpha^t x = \lambda\}$. We also use $d(A, B)$ to denote the distance between two subsets A and B defined by $d(A, B) = \inf_{x \in A, y \in B} d(x, y)$. If A and B are closed and one of them is bounded, the infimum is attained.

Consider a convex program of the following kind to be solved:

$$P \begin{cases} \max c^t x \\ x \in S \end{cases}$$

where S is a nonempty compact convex set of $\mathbb{R}^n$. Recall that a subset of $\mathbb{R}^n$ is compact if and only if it is bounded and closed. Notice that this assumption implies that P has a finite optimal solution. We assume without loss of generality that the norm of $c \in \mathbb{R}^n$ is equal to one, i.e., $\|c\| = 1$: when $c = 0$, there is nothing to optimize.

Given any function $f : [0, 2] \to \mathbb{R}_+$, let S_f denote the set of points $x \in \mathbb{R}^n$ satisfying all the inequalities $\pi^t x - \pi_0 \le f(1 - \pi^t c)$ where $\pi^t x - \pi_0 \le 0$ is any valid inequality of S such that $\|\pi\| = 1$. This implies that S_f contains S. Notice that without loss of information one can focus here only on the inequalities $\pi^t x - \pi_0 \le 0$ such that $\pi_0 = \max_{x \in S} \pi^t x$. The optimization problem corresponding to S_f is denoted by P_f.

$$P_f \begin{cases} \max c^t x \\ x \in S_f \end{cases}$$

Given any function $f : [0, 2] \to \mathbb{R}_+$, we will write $\liminf_{\epsilon \to 0+} \frac{f(\epsilon)}{\sqrt{\epsilon}} = 0$ if for any $\epsilon' > 0$ and any $k > 0$, there exists ϵ such that $0 < \epsilon \le \epsilon'$ and $\frac{f(\epsilon)}{\sqrt{\epsilon}} \le k$. Said another way, we say that $\liminf_{\epsilon \to 0+} \frac{f(\epsilon)}{\sqrt{\epsilon}} = 0$ if one can build a sequence $(\epsilon_i)_{i \in \mathbb{N}} > 0$ such that $\lim_{i \to \infty} \epsilon_i = 0$ and $\lim_{i \to \infty} \frac{f(\epsilon_i)}{\sqrt{\epsilon_i}} = 0$.

Let us now look at some simple examples. We take $n = 2$. First, we consider a set S which is a polytope (under the two hyperplanes represented with dotted lines in Figs. 1 and 2). We also consider a set T given by $T = S_f$. We take $f(\epsilon) = \epsilon$ for any $\epsilon \in [0, 2]$ in Fig. 1, while $f(\epsilon) = \sqrt{\epsilon}$ in Fig. 2. Notice that both figures show only what happens around the optimal solution of P. The set of optimal solutions of P and P_f are clearly the same when $f(\epsilon) = \epsilon$. On the other hand, when the set T is defined using the function $f(\epsilon) = \sqrt{\epsilon}$, then T contains more optimal solutions than S.

We do the same with another domain S. Figures 3 and 4 use the unit disk. We take $f(\epsilon) = \epsilon$ in the first case and $f(\epsilon) = \sqrt{\epsilon}$ in the second one. One can easily show that the set T is also a disk when $f(\epsilon) = \epsilon$. The set of optimal solutions does not change. However, when $f(\epsilon) = \sqrt{\epsilon}$, the set of optimal solutions of P_f contains the whole segment $[AB]$ (Fig. 4).

Fig. 1 S is a polytope and $f(\epsilon) = \epsilon$

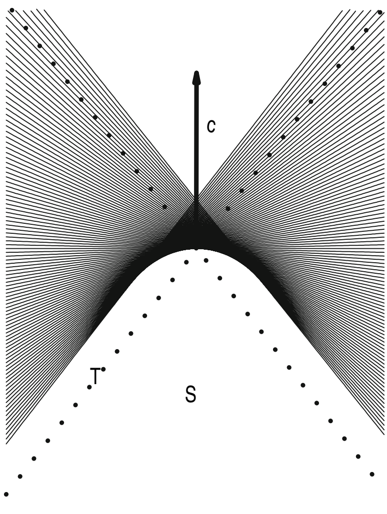


Fig. 2 S is a polytope and $f(\epsilon) = \sqrt{\epsilon}$

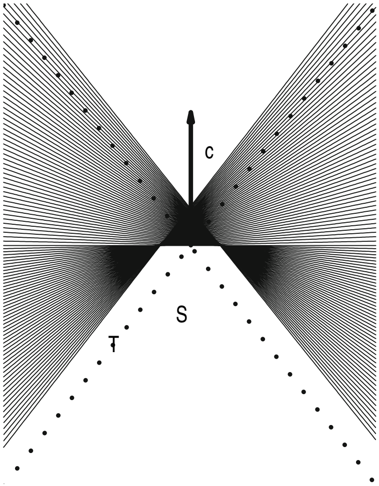
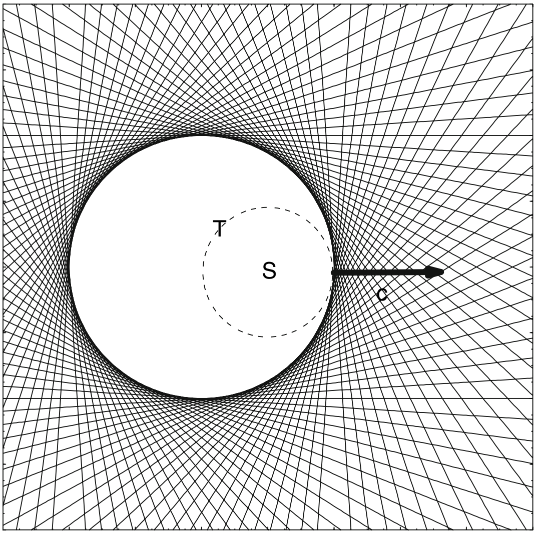


Fig. 3 S is the unit disk and $f(\epsilon) = \epsilon$



While these geometric transformations are interesting on their own, the goal of this paper is to characterize the relaxations that do not change the set of optimal solutions. The two previous examples suggest that the function $f(\epsilon) = \epsilon$ provides such kind of relaxations while $f(\epsilon) = \sqrt{\epsilon}$ is not a suitable function.

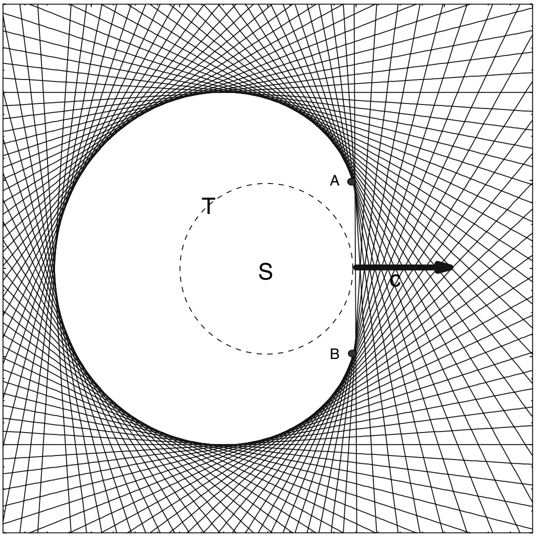
3 Preliminary results

A first straightforward lemma is given below.

Lemma 3.1 *Given any function $f : [0, 2] \to \mathbb{R}_+$, S_f is a compact convex set.*

Proof Any converging sequence of points $(x_t)_{t\in\mathbb{N}}$ belonging to S_f satisfies all valid inequalities $\pi^t x - \pi_0 - f(1 - \pi c) \leq 0$. By continuity of linear inequalities, the limit will also

Fig. 4 S is the unit disk and $f(\epsilon) = \sqrt{\epsilon}$



satisfy the same constraints, so it belongs to S_f. Convexity of S_f is a consequence of the convexity of linear functions. As S is bounded, for any variable x_j ($j \in \{1, \ldots, n\}$) there are two numbers α_j and β_j such that $x_j \leq \alpha_j$ and $-x_j \leq \beta_j$ are two valid inequalities for S. Considering S_f, we get $x_j - \alpha_j - f(1 - c_j) \leq 0$ and $-x_j - \beta_j - f(1 + c_j) \leq 0$. Said another way, S_f is bounded. □

The previous lemma implies that P_f has optimum solutions (they are attained).

We also give the proof of the next lemma for sake of completeness. It only says that the support function is continuous.

Lemma 3.2 *Let T be any compact set and α be any vector of $\mathbb{R}^n$. Then,* $\lim_{\sigma \to \alpha} \max_{x \in T} \sigma^t x = \max_{x \in T} \alpha^t x$.

Proof Consider x_α and x_σ in T such that $\alpha^t x_\alpha = \max_{x \in T} \alpha^t x$ and $\sigma^t x_\sigma = \max_{x \in T} \sigma^t x$. Both x_α and x_σ exist because T is compact. We can write $\sigma^t x_\sigma \geq \sigma^t x_\alpha = (\sigma^t - \alpha^t) x_\alpha + \alpha^t x_\alpha$. We also have $\sigma^t x_\sigma = (\sigma^t - \alpha^t) x_\sigma + \alpha^t x_\sigma \leq (\sigma^t - \alpha^t) x_\sigma + \alpha^t x_\alpha$. Combining the two inequalities and using the fact that T is bounded leads to $\lim_{\sigma \to \alpha} \max_{x \in T} \sigma^t x = \max_{x \in T} \alpha^t x$. □

Lemma 3.3 *Let $f : [0, 2] \to \mathbb{R}_+$ be a function such that* $\liminf_{\epsilon \to 0+} f(\epsilon) = 0$. *Then* $\max_{y \in S} c^t y = \max_{y \in S_f} c^t y$.

Proof $\liminf_{\epsilon \to 0+} f(\epsilon) = 0$ implies the existence of a sequence $(\epsilon_i)_{i \in \mathbb{N}} > 0$ such that $\lim_{i \to \infty} \epsilon_i = 0$ and $\lim_{i \to \infty} f(\epsilon_i) = 0$. Since S is compact, one can build a sequence of constraints $\sigma_i^t x - \max_{y \in S} \sigma_i^t y \leq 0$ such that $1 - \sigma_i^t c = \epsilon_i$ and $||\sigma_i|| = 1$. We also have $\max_{y \in S} \sigma_i^t y \leq \max_{y \in S_f} \sigma_i^t y \leq \max_{y \in S} \sigma_i^t y + f(\epsilon_i)$. Taking the limit when $i \to \infty$ and using Lemma 3.2 clearly leads to $\max_{y \in S} c^t y = \max_{y \in S_f} c^t y$. □

The next lemma will be used to prove Proposition 4.3.

Lemma 3.4 *Let T be any compact set, then*

$$\lim_{\sigma\to c} d\left(H_{(c,\max\limits_{x\in T} c^t x)} \cap H_{(\sigma,\max\limits_{x\in T} \sigma^t x)}, H_{(c,\max\limits_{x\in T} c^t x)} \cap T\right) = 0.$$

Proof Given any sequence $(\sigma_i)_{i\in\mathbb{N}}$ converging to c, we will build a subsequence for which the distance mentioned in the Lemma converges to 0. This is clearly enough to prove the wanted result.

Consider again the points $x_{\sigma_i} \in T$ such that $\sigma_i^t x_{\sigma_i} = \max\limits_{x\in T} \sigma_i^t x$. Since T is a compact set and $x_{\sigma_i} \in T$, one can build a converging subsequence of $(x_{\sigma_i})_{i\in\mathbb{N}}$. Then, we can assume that $(x_{\sigma_i})_{i\in\mathbb{N}}$ is a converging sequence (if not, we replace it by a converging subsequence). Let z denote the limit of the sequence.

We already know from Lemma (3.2) that $\lim\limits_{i\to\infty} \sigma_i^t x_{\sigma_i} = c^t x_c$. Moreover, since $\lim\limits_{i\to\infty} \sigma_i = c$ and $\lim\limits_{i\to\infty} x_{\sigma_i} = z$, we can write $\lim\limits_{i\to\infty} \sigma_i^t x_{\sigma_i} = c^t z$. As a consequence, z is a point of $H_{(c,\max\limits_{x\in T} c^t x)} \cap T$.

Let y_{σ_i} be the orthogonal projection of x_{σ_i} on $H_{(c,\max\limits_{x\in T} c^t x)}$. By definition, $y_{\sigma_i} - x_{\sigma_i}$ is a vector proportional to c. Since $c^t x_{\sigma_i} \le \max\limits_{x\in T} c^t x = c^t y_{\sigma_i}$, we can write $y_{\sigma_i} - x_{\sigma_i} = ||y_{\sigma_i} - x_{\sigma_i}||c$. Moreover, for i sufficiently large, σ_i is close to c, implying that $\sigma_i^t y_{\sigma_i} - \sigma_i^t x_{\sigma_i} = ||y_{\sigma_i} - x_{\sigma_i}||\sigma_i^t c \ge 0$.

On the other hand, $z \in T$ implies that $\sigma_i^t z - \sigma_i^t x_{\sigma_i} \le 0$. Said another way, one can find a point w_{σ_i} lying on the segment $[y_{\sigma_i}, z]$ such that $\sigma_i^t w_{\sigma_i} - \sigma_i^t x_{\sigma_i} = 0$ and $c^t w_{\sigma_i} - c^t x_c = 0$. This means that $w_{\sigma_i} \in H_{(c,\max\limits_{x\in T} c^t x)} \cap H_{(\sigma_i,\max\limits_{x\in T} \sigma_i^t x)}$. We also have

$$\begin{aligned} d\left(H_{(c,\max\limits_{x\in T} c^t x)} \cap H_{(\sigma_i,\max\limits_{x\in T} \sigma_i^t x)}, H_{(c,\max\limits_{x\in T} c^t x)} \cap T\right) &\le d(w_{\sigma_i}, z) \\ &\le d(y_{\sigma_i}, z) \\ &\le d(x_{\sigma_i}, z) \end{aligned}$$

where the last inequality is given by Pythagoras' theorem.

Since the limit of $d(x_{\sigma_i}, z)$ is 0, one can deduce that

$$\lim_{i\to\infty} d\left(H_{(c,\max\limits_{x\in T} c^t x)} \cap H_{(\sigma_i,\max\limits_{x\in T} \sigma_i^t x)}, H_{(c,\max\limits_{x\in T} c^t x)} \cap T\right) = 0.$$

□

4 Main results

We are now ready to present our first important result.

Proposition 4.1 *Let $f : [0, 2] \to \mathbb{R}_+$ be a function such that $\liminf\limits_{\epsilon\to 0+} \frac{f(\epsilon)}{\sqrt{\epsilon}} = 0$, then any optimal solution of P_f belongs to S.*

Proof Let x' (resp. x^*) stand for an optimal solution of P_f (resp. P). Let $z^* = c^t x^*$ be the objective value of an optimal solution of P. Since $\liminf\limits_{\epsilon\to 0+} \frac{f(\epsilon)}{\sqrt{\epsilon}} = 0$ implies that $\liminf\limits_{\epsilon\to 0+} f(\epsilon) = 0$, we deduce by Lemma 3.3 that both x' and x^* belong to the hyperplane $H_{(c,z^*)}$.

Let us assume that $x' \notin S$. This means that there exists an hyperplane $H_{(\sigma,\sigma_0)}$ separating x' from S ($\|\sigma\| = 1$). Let B stand for the point corresponding to the intersection between $H_{(\sigma,\sigma_0)}$ and the segment $[x', x^*]$. Let A denote the orthogonal projection of x' onto $H_{(\sigma,\sigma_0)}$.

Let us consider the affine subspace of dimension two containing the points x', A and B. Then we have:

$$\sigma^t x' - \sigma_0 = d(x', H_{(\sigma,\sigma_0)}) = d(x', A) = d(x', B) sin(\theta), \tag{1}$$

where θ stands for the angle between the vectors $\overrightarrow{BA}$ and $\overrightarrow{Bx'}$. Notice that the vector $\overrightarrow{x'A}$ is proportional to σ. We can decompose the vector c into two vectors c_1 and c_2 such that c_2 is orthogonal to the affine subspace considered above and c_1 is the projection of c on the corresponding vector subspace. We will also use β to denote the angle between c and σ. We clearly have,

$$c^t \sigma = cos(\beta) = c_1^t \sigma = ||c_1|| cos(\theta). \tag{2}$$

The second equality comes from the fact that the angle between c_1 and σ is the same as the angle between $\overrightarrow{Bx'}$ and $\overrightarrow{BA}$. To see this, we only have to remember that both x' and $x^\star$ belong to $H_{(c,z^*)}$ implying that c is orthogonal to $\overrightarrow{Bx'}$. Consequently, c_1 is also orthogonal to $\overrightarrow{Bx'}$. Combining this with the fact that σ is orthogonal to $\overrightarrow{BA}$ leads to (2).

Using σ and c, we will build a sequence of valid inequalities $\sigma_i^t x - \sigma_{0i} \le 0$ that are violated by x' and such that $\lim_{i\to\infty} \sigma_i = c$. The assumption $\liminf_{\epsilon\to 0+} \frac{f(\epsilon)}{\sqrt{\epsilon}} = 0$ implies the existence of a sequence $(\epsilon_i)_{i\in\mathbb{N}} > 0$ such that $\lim_{i\to\infty} \epsilon_i = 0$ and $\lim_{i\to\infty} \frac{f(\epsilon_i)}{\sqrt{\epsilon_i}} = 0$. Then we can consider a sequence of constraints $\sigma_i^t x - \sigma_{0i} \le 0$ such that $\epsilon_i = 1 - \sigma_i^t c$ and

$$\sigma_i = \frac{(1-\alpha_i)\sigma + \alpha_i c}{||(1-\alpha_i)\sigma + \alpha_i c||}, \tag{3}$$

where $0 < \alpha_i < 1$. The second member σ_{0i} is also defined in the same way: $\sigma_{0i} = \frac{(1-\alpha_i)\sigma_0 + \alpha_i z^\star}{||(1-\alpha_i)\sigma + \alpha_i c||}$.

All the constraints $\sigma_i^t x - \sigma_{0i} \le 0$ are clearly valid inequalities since they are the combination of two valid inequalities. These constraints are violated by x' (because x' satisfies with equality the constraint $c^t x \le z^\star$ and violates the constraint $\sigma^t x \le \sigma_0$). Notice that this namely implies that $f(\epsilon_i) \neq 0$ which will allow us to consider below the ratio $\frac{1}{f(\epsilon_i)}$.

It is important to notice that the point B previously defined does not change when we consider the sequence of constraints $\sigma_i^t x - \sigma_{0i} \le 0$. In fact, B belongs to all the hyperplanes $H_{(\sigma_i,\sigma_{0i})}$. The point A depends on the constraint and will be denoted A_i.

We will use the index i to denote all the points and angles related to the inequalities $\sigma_i^t x - \sigma_{0i} \le 0$. As we have $\lim_{i\to\infty} \epsilon_i = 0$, we can deduce using (2) that $\lim_{i\to\infty} \sigma_i = c$, $\lim_{i\to\infty} \beta_i = 0$, $\lim_{i\to\infty} \theta_i = 0$, $\lim_{i\to\infty} ||c_{1i}|| = 1$, and $\lim_{i\to\infty} \alpha_i = 1$.

This way we get for ϵ_i close to zero:

$$\frac{\sigma_i^t x' - \sigma_{0i}}{f(\epsilon_i)} = \frac{d(x', B) sin(\theta_i)}{f(1 - cos(\beta_i))} \sim \frac{d(x', B)\theta_i}{f(1 - (1 - \frac{\beta_i^2}{2} + o(\beta_i^3)))} \sim \frac{d(x', B)\theta_i}{f(\frac{\beta_i^2}{2} + o(\beta_i^3))}.$$

From the assumption $\lim_{i\to\infty} \frac{f(\epsilon_i)}{\sqrt{\epsilon_i}} = 0$, we get

$$\frac{\sigma_i^t x' - \sigma_{0i}}{f(1 - \sigma_i^t c)} \sim d(x', B) \frac{\theta_i}{\beta_i} \frac{1}{o(1)}. \tag{4}$$

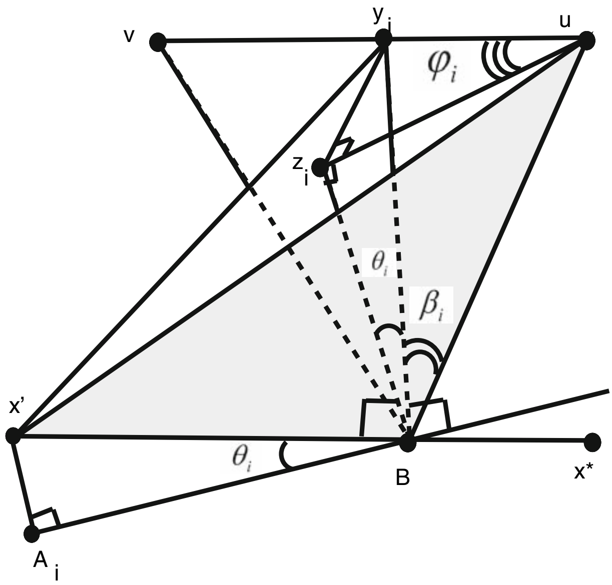


Fig. 5 Illustration relating to the proof of Proposition 4.1

We need to evaluate the ratio $\frac{\theta_i}{\beta_i}$. To do this, we first try to represent the already defined points and angles on Fig. 5.

Assume that the vectors c, σ and $\overrightarrow{x^*x'}$ are linearly independent. We consider the points u, v, y_i such that $\overrightarrow{Bu} = c$, $\overrightarrow{Bv} = \sigma$, $y_i \in [v, u]$, $\overrightarrow{By_i}$ and σ_i are collinear. In other words, $\overrightarrow{By_i}$ is defined by $\overrightarrow{By_i} = (1 - \alpha_i)\sigma + \alpha_i c$ where $0 < \alpha_i < 1$ (see Eq. 3). Let z_i be the projection of u on the plane containing B, x' and y_i. Thus, we have $\overrightarrow{Bz_i} = c_{1i}$. The angle between $\overrightarrow{uz_i}$ and $\overrightarrow{uy_i}$ is denoted by φ_i.

Notice that all the points considered here $(B, x', x^\star, A_i, y_i, z_i, u, v)$ belong to the unique three dimensional affine space containing B, x', u and v.

Using the notation above, when ϵ_i is close to 0 (β_i close to 0) we get $d(u, y_i) \sim d(u, B)\beta_i$ and $d(z_i, y_i) \sim d(u, B)\theta_i$. Combining the two approximation results leads to $\frac{d(z_i, y_i)}{d(u, y_i)} \sim \frac{\theta_i}{\beta_i}$.

Moreover, we have $sin(\varphi_i) = \frac{d(z_i, y_i)}{d(u, y_i)}$. When ϵ_i is close to 0, then φ_i becomes close to the angle ψ between $\overrightarrow{uv}$ and the normal vector to the plane containing u, B and x'. Recall that since we are considering the three dimensional (closed) space containing B, x', u and v, the normal vector to the plane containing u, B and x' is well defined. In other words, we have

$$\frac{\theta_i}{\beta_i} \sim \sin(\psi). \tag{5}$$

Notice that $sin(\psi) \neq 0$ because $\overrightarrow{uv}$ cannot be orthogonal to $\overrightarrow{Bu}$ ($||\overrightarrow{Bu}|| = ||\overrightarrow{Bv}|| = 1$). Combining the results (4) and (5) leads to

$$\lim_{i \to \infty} \frac{\sigma_i^t x' - \sigma_{0i}}{f(1 - \sigma_i^t c)} = \infty. \tag{6}$$

The same result holds if we assume that the vectors c, σ and $\overrightarrow{x^*x'}$ are not linearly independent. Indeed, in this case we have $\theta_i = \beta_i$.

Let us summarize what we have proved. When we assume that x' is not in S, we can build a sequence of constraints violated by x' such that $\lim_{i\to\infty} \frac{\sigma_i^t x' - \sigma_{0i}}{f(1-\sigma_i^t c)} = \infty$. But this is not possible since $x' \in S_f$. Then, our assumption is wrong and one can deduce that $x' \in S$. □

Corollary 4.2 *Let $f : [0, 2] \to \mathbb{R}_+$ be a function such that $\liminf_{\epsilon\to 0+} \frac{f(\epsilon)}{\sqrt{\epsilon}} = 0$, and T be a set such that $S \subseteq T \subseteq S_f$. Then any point of T maximizing $c^t x$ belongs to S.*

We can now prove the second important result.

Proposition 4.3 *Let T be a compact set containing S such that any point x of T maximizing $c^t x$ lies in S. Then there exists a function $f : [0, 2] \to \mathbb{R}_+$ such that $\liminf_{\epsilon\to 0+} \frac{f(\epsilon)}{\sqrt{\epsilon}} = 0$ and $T \subseteq S_f$.*

Proof Let us define the function f as follows:

$$f(\epsilon) = \max_{1-\sigma^t c=\epsilon, ||\sigma||=1} \left(\max_{x\in T} \sigma^t x - \max_{x\in S} \sigma^t x \right).$$

Notice that f is well defined because the set $\{\sigma \in \mathbb{R}^n$, such that $1 - \sigma^t c = \epsilon, ||\sigma|| = 1\}$ is a compact set and the function $\sigma \to \max_{x\in T} \sigma^t x - \max_{x\in S} \sigma^t x$ is continuous (by Lemma 3.2). For each $\epsilon \in [0, 2]$ there exists at least one σ_ϵ such that $1 - \sigma_\epsilon^t c = \epsilon$, $||\sigma_\epsilon|| = 1$ and $f(\epsilon) = \max_{x\in T} \sigma_\epsilon^t x - \max_{x\in S} \sigma_\epsilon^t x$.

Suppose that the function f defined above does not satisfy the requirements. This means that there exists $\epsilon' > 0$ and $k > 0$ such that $f(\epsilon) \geq k\sqrt{\epsilon}$ for any $\epsilon \leq \epsilon'$. For each $\epsilon \leq \epsilon'$, we consider the set $H_{(c,\max_{x\in T} c^t x)} \cap H_{(\sigma_\epsilon,\max_{x\in T} \sigma_\epsilon^t x)}$. Using the fact that $\lim_{\epsilon\to 0} \sigma_\epsilon = c$, one can deduce from Lemma 3.4 that $\lim_{\epsilon\to 0} d\left(H_{(c,\max_{x\in T} c^t x)} \cap H_{(\sigma_\epsilon,\max_{x\in T} \sigma_\epsilon^t x)}, H_{(c,\max_{x\in T} c^t x)} \cap T \right) = 0$. This implies that one can find ϵ'' such that $d\left(H_{(c,\max_{x\in T} c^t x)} \cap H_{(\sigma_\epsilon,\max_{x\in T} \sigma_\epsilon^t x)}, H_{(c,\max_{x\in T} c^t x)} \cap T \right) < \frac{k}{2\sqrt{2}}$ for any $0 < \epsilon \leq \epsilon''$. We can of course take $\epsilon'' = \min\{\epsilon', \epsilon''\}$ to simultaneously have $f(\epsilon) \geq k\sqrt{\epsilon}$ and $d\left(H_{(c,\max_{x\in T} c^t x)} \cap H_{(\sigma_\epsilon,\max_{x\in T} \sigma_\epsilon^t x)}, H_{(c,\max_{x\in T} c^t x)} \cap T \right) < \frac{k}{2\sqrt{2}}$ for any $0 < \epsilon \leq \epsilon''$. Let $y_\epsilon \in H_{(c,\max_{x\in T} c^t x)} \cap H_{(\sigma_\epsilon,\max_{x\in T} \sigma_\epsilon^t x)}$ and $y_0 \in H_{(c,\max_{x\in T} c^t x)} \cap T$ be such that $d(y_\epsilon, y_0) < \frac{k}{2\sqrt{2}}$ for $\epsilon \leq \epsilon''$. Since $y_\epsilon \in H_{(\sigma_\epsilon,\max_{x\in T} \sigma_\epsilon^t x)}$ we can write $\sigma_\epsilon^t y_\epsilon - \max_{x\in T} \sigma_\epsilon^t x = 0$. By definition of σ_ϵ we have $\max_{x\in T} \sigma_\epsilon^t x = f(\epsilon) + \max_{x\in S} \sigma_\epsilon^t x$. This implies that $\sigma_\epsilon^t(y_\epsilon - y_0) + \sigma_\epsilon^t y_0 - \max_{x\in S} \sigma_\epsilon^t x = f(\epsilon)$. By our assumption that any point x of T maximizing $c^t x$ lies in S, we know that $H_{(c,\max_{x\in T} c^t x)} \cap T = H_{(c,\max_{x\in S} c^t x)} \cap S$. Consequently, we have $y_0 \in S$ and $\sigma_\epsilon^t y_0 - \max_{x\in S} \sigma_\epsilon^t x \leq 0$. In other words, the following inequalities hold for $0 < \epsilon \leq \epsilon''$:

$$k\sqrt{\epsilon} \leq f(\epsilon) \leq \sigma_\epsilon^t(y_\epsilon - y_0). \tag{7}$$

Let θ_ϵ be the angle between σ_ϵ and c. Then we have $\sigma_\epsilon = cos(\theta_\epsilon)c + \overrightarrow{z_\epsilon}$ where $\overrightarrow{z_\epsilon}$ is a vector orthogonal to c. As the norm of σ_ϵ is 1, we should have $||\overrightarrow{z_\epsilon}|| = |sin(\theta_\epsilon)|$. Moreover, both y_0 and y_ϵ belong to $H_{(c,\max_{x\in T} c^t x)}$. Consequently, we have

$$\begin{aligned}\sigma_\epsilon^t(y_\epsilon - y_0) &= \left(\cos(\theta_\epsilon)c + \overrightarrow{z_\epsilon}\right)^t (y_\epsilon - y_0)\\ &= \overrightarrow{z_\epsilon}^t (y_\epsilon - y_0)\\ &\leq ||\overrightarrow{z_\epsilon}||d(y_0, y_\epsilon)\\ &\leq |\sin(\theta_\epsilon)|\frac{k}{2\sqrt{2}}\end{aligned} \tag{8}$$

Combining (7) and (8) and using the fact that $1 - \sigma_\epsilon^t c = \epsilon$, we get $\sqrt{1-\cos(\theta_\epsilon)} \leq |\sin(\theta_\epsilon)|\frac{1}{2\sqrt{2}}$ which is clearly impossible for $\theta_\epsilon \neq 0$ close to 0. □

Combination of Corollary 4.2 and Proposition 4.3 obviously leads to the following theorem (the main contribution of the paper). Recall that S is a nonempty convex compact set.

Theorem 4.4 *Let T be a compact set containing S, then the following properties are equivalent:*

(i) any point x of T maximizing $c^t x$ lies in S
(ii) there exists a function $f : [0, 2] \to \mathbb{R}_+$ such that $\liminf_{\epsilon \to 0+} \frac{f(\epsilon)}{\sqrt{\epsilon}} = 0$ and $T \subseteq S_f$.

Suppose now that we are dealing with a multi-objective problem where we have r objective functions defined by the vectors $c_1, \ldots, c_r$. We look for relaxations that do not change the optimal solutions related to any objective function. We assume that $||c_i|| = 1$ for any $i \in \{1, \ldots, r\}$. Using the previous theorem, we can deduce what follows.

Corollary 4.5 *Let T be a compact set containing S, then the following properties are equivalent:*

(i) any point x of T maximizing $c_i^t x$, $i \in \{1, \ldots, r\}$, lies in S
(ii) there exist r functions $f_i : [0, 2] \to \mathbb{R}_+$, $i \in \{1, \ldots, r\}$ such that $\liminf_{\epsilon \to 0+} \frac{f_i(\epsilon)}{\sqrt{\epsilon}} = 0$ and $T \subseteq S' = \{y \in \mathbb{R}^n, \pi^t y - \pi_0 \leq \min_{i=1,\ldots,r} f_i(1 - \pi^t c_i), ||\pi|| = 1, \pi^t x - \pi_0 \leq 0$ valid for $S\}$.

Proof Theorem 4.4 tells us that condition (i) is equivalent to $T \subseteq S_{f_1} \cap \ldots \cap S_{f_r}$. This directly leads to condition (ii). □

Remarks

- Theorem 4.4 gave necessary and sufficient conditions to get a relaxation without changing the set of optimal solutions. We know that this holds if we take $T = S_f$ where $\liminf_{\epsilon \to 0+} \frac{f(\epsilon)}{\sqrt{\epsilon}} = 0$. However, Theorem 4.4 does not say anything about a set $T = S_f$ where f is a function not satisfying the requirement above. In fact, one can build a function f such that $\liminf_{\epsilon \to 0+} \frac{f(\epsilon)}{\sqrt{\epsilon}} > 0$ and the set $T = S_f$ contains exactly the same set of optimal solutions as S. In other words, it is possible to have two functions f and g such that $\liminf_{\epsilon \to 0+} \frac{f(\epsilon)}{\sqrt{\epsilon}} > 0$, $\liminf_{\epsilon \to 0+} \frac{g(\epsilon)}{\sqrt{\epsilon}} = 0$ and $S_f \subseteq S_g$. An example is given by $S = \{(a, b) \in \mathbb{R}^2, a \geq 0, b \geq 0, a+b \leq 1\}$ with $f(2) = 0$, $f(1 - \frac{1}{\sqrt{2}}) = 0$ and $f(\epsilon) = 1$ otherwise. The objective function is defined by $c = (-1/\sqrt{2}, -1/\sqrt{2})^t$. We clearly have $\liminf_{\epsilon \to 0+} \frac{f(\epsilon)}{\sqrt{\epsilon}} > 0$ and $S_f \subseteq S_0$ ($S_0 = S$ is given by $g = 0$).
- The set S considered in the paper is bounded. One may ask whether this condition is necessary to prove Theorem 4.4. Let us take $S = \{y \in \mathbb{R}^n, c^t y = z^\star\}$. The only non-dominated

valid inequalities for S (up to multiplicative factor) are $c^t x \leq z^\star$ and $-c^t x \leq -z^\star$. Let us define f by $f(0) = f(2) = 1$ and $f(\epsilon) = 0$ elsewhere. Although f satisfies the second requirement of Theorem 4.4, the set S_f defined by $\{y \in \mathbb{R}^n, -1 \leq c^t y - z^\star \leq 1\}$ has a different set of optimal solutions than S.

In fact, Theorem 4.4 is valid even if we assume that S is a closed convex set and $S \cap H_{(c,z^\star)}$ is bounded, but the proofs are slightly more complicated.

5 More examples

More elaborate examples will be presented in this section. Theorem 4.4 is used to deduce the equivalence (in terms of optimal solutions) between different mathematical programs. Notice that we generally need a full description of all valid inequalities. While this is not always possible, there are some cases where we have such descriptions. This is straightforward at least for polyhedrons and ellipsoids. We will focus on polyhedrons in this section. It is clear that the relaxations defined in this paper are generally not useful to solve a linear program. However, these relaxations (at least in the linear case) lead to more "complicated" mathematical programs that are equivalent to the initial problems. Then if such a "complicated" program is given, we know that it can be replaced by a simple linear program. An application is given at the end of this section.

We assume here that P is a linear problem denoted LP

$$LP \begin{cases} \max c^t x \\ Ax - b \leq 0 \\ x \in \mathbb{R}^n \end{cases} \tag{9}$$

where A is a matrix with n columns and m rows. Recall that $||c|| = 1$. We also assume that the set of feasible solutions is a nonempty bounded set (a nonempty polytope).

The relaxation LP_f where $f(x) = \lambda x$ ($\lambda \geq 0$) is given by

$$LP_f \begin{cases} \max c^t x \\ \alpha^t (Ax - b) \leq \lambda(1 - \alpha^t Ac); \forall \alpha \in \mathbb{R}^m, \alpha \geq 0, ||\alpha^t A|| = 1 \\ x \in \mathbb{R}^n \end{cases}$$

The equality constraint $||\alpha^t A|| = 1$ can be replaced by $||\alpha^t A|| \leq 1$ without any change in the relaxation. Let us first consider that $\alpha^t A = 0$. Then $\alpha^t(Ax - b) \leq 0$ is valid for S. This leads to $\alpha^t b \geq 0$ which implies that the inequality $\alpha^t(Ax - b) \leq \lambda(1 - \alpha^t Ac)$ is valid for S_f. Suppose now that $\alpha^t A \neq 0$ and $||\alpha^t A|| < 1$. By the definition of S_f, the constraint $\beta^t(Ax - b) \leq \lambda(1 - \beta^t Ac)$ is valid where $\beta = \frac{\alpha}{||\alpha^t A||}$. Given $x \in S_f$, we have

$$\begin{aligned} \alpha^t(Ax - b) &= ||\alpha^t A|| \beta^t (Ax - b) \\ &\leq ||\alpha^t A|| \lambda(1 - \beta^t Ac) \\ &= ||\alpha^t A|| \lambda - \lambda \alpha^t Ac \\ &\leq \lambda - \lambda \alpha^t Ac \\ &= \lambda(1 - \alpha^t Ac). \end{aligned}$$

Said another way, $\alpha^t(Ax - b) \leq \lambda(1 - \alpha^t Ac)$ is valid even if $\alpha^t A \neq 0$ and $||\alpha^t A|| < 1$. LP_f becomes

$$LP_f \begin{cases} \max c^t x \\ \left(x^t A^t - b^t + \lambda c^t A^t\right)\alpha \leq \lambda; \forall \alpha \in \mathbb{R}^m, \alpha \geq 0, ||\alpha^t A|| \leq 1 \\ x \in \mathbb{R}^n \end{cases}$$

For any x satisfying all the constraints of LP_f, we use y^t to denote $x^t A^t - b^t + \lambda c^t A^t$. This means that we should have $y^t\alpha \leq \lambda$ for any α such that $||A^t\alpha|| \leq 1$ and $\alpha \geq 0$.

Said another way, the following inequality holds

$$-\lambda \leq \min_{\alpha \in \mathbb{R}^m_+, ||A^t\alpha|| \leq 1} -y^t\alpha. \tag{10}$$

Moreover, the minimization problem used in (10) and written below is a second-order cone program (see, [1,7] for results about second-order cone programming).

$$\begin{cases} \min -y^t\alpha \\ ||A^t\alpha|| \leq 1 \\ \alpha \in \mathbb{R}^m_+ \end{cases}$$

The dual of this second-order cone program is given by:

$$\begin{cases} \max -v \\ Au + y \leq 0 \\ ||u|| \leq v \\ u \in \mathbb{R}^n, v \in \mathbb{R} \end{cases}$$

Slater's conditions are clearly satisfied, so strong duality holds and we can replace Eq. 10 by

$$-\lambda \leq \max_{u \in \mathbb{R}^n, v \in \mathbb{R}, Au + y \leq 0, ||u|| \leq v} -v. \tag{11}$$

(11) is equivalent to the existence of $u \in \mathbb{R}^n$ and $v \in \mathbb{R}$ such that $Au + y \leq 0$, $||u|| \leq v$ and $-\lambda \leq -v$. It is easy to see that v can be eliminated and the last two inequalities can be replaced by $||u|| \leq \lambda$.

Dividing u by λ and replacing y^t by $x^t A^t - b^t + \lambda c^t A^t$, we get a new formulation for LP_f

$$LP_f \begin{cases} \max c^t x \\ Ax - b + \lambda(Ac - Au) \leq 0 \\ ||u|| \leq 1 \\ u \in \mathbb{R}^n, x \in \mathbb{R}^n \end{cases} \tag{12}$$

Theorem 4.4 tells us that the problem LP defined in (9) and the problem LP_f reformulated in (12) have exactly the same set of optimal solutions. Notice that problem (12) is a second-order cone program.

The equivalence between (9) and (12) can also be shown using linear programming duality. In fact, assuming that the linear program (9) has a finite optimal solution, it is easy to prove by duality for any fixed $u \neq c \in \mathbb{R}^n$, $||u|| \leq 1$, and $\lambda > 0$ the following inequalities:

$$\max_{Ax-b+\lambda(Ac-Au)\leq 0} c^t x < \max_{Ax \leq b} c^t x < \max_{Ax-b+\lambda(Au-Ac)\leq 0} c^t x. \tag{13}$$

In fact, given a vector x maximizing $c^t x$ with respect to constraints $Ax - b + \lambda(Ac - Au) \leq 0$, we can build a solution $x' = x + \lambda(c - u)$ satisfying constraints $Ax' - b \leq 0$. We clearly have $c^t x' > c^t x$ which proves the first inequality of (13).

It is also easy to see that for any vector $\delta \in \mathbb{R}^n$ for which $\max_{Ax \le b + A\delta} c^t x < \max_{Ax \le b} c^t x$ there exists $\lambda > 0$ and $u \neq c \in \mathbb{R}^n$, $||u|| \le 1$ such that $\delta = \lambda(u - c)$. A similar result holds if $\max_{Ax \le b + A\delta} c^t x > \max_{Ax \le b} c^t x$: we should have $\delta = \lambda(c - u)$. We can also show in the same way that $\max_{Ax \le b + A\delta} c^t x = \max_{Ax \le b} c^t x$ if and only if $\delta^t c = 0$. These straightforward results may be useful in the context of sensitivity analysis. They are summarized in the following proposition.

Proposition 5.1 *Assuming that* $\max_{Ax \le b} c^t x$ *is finite and* $||c|| = 1$*, then we have:*

(i) $\max_{Ax \le b + A\delta} c^t x < \max_{Ax \le b} c^t x$ *holds if and only if there exists* $\lambda > 0$ *and* $u \neq c \in \mathbb{R}^n$, $||u|| \le 1$ *such that* $\delta = \lambda(u - c)$*; i.e.,* $\delta^t c < 0$.

(ii) $\max_{Ax \le b + A\delta} c^t x > \max_{Ax \le b} c^t x$ *holds if and only if there exists* $\lambda > 0$ *and* $u \neq c \in \mathbb{R}^n$, $||u|| \le 1$ *such that* $\delta = \lambda(c - u)$*; i.e.,* $\delta^t c > 0$.

(iii) $\max_{Ax \le b + A\delta} c^t x = \max_{Ax \le b} c^t x$ *holds if and only if* $\delta^t c = 0$.

We give below a possible application of the equivalence between problems (9) and (12). Consider a linear problem with two sets of constraints: a first set that should be respected by x and a second set where a limited violation can be accepted. This can be expressed by writing that there is a vector u whose norm is small ($||u|| \le \lambda$) such that $x + u$ satisfies the second set of constraints.

$$\begin{cases} \max c^t x \\ Dx \le e \\ A(x + u) \le b \\ ||u|| \le \lambda \\ u \in \mathbb{R}^n, x \in \mathbb{R}^n \end{cases} \tag{14}$$

Relaxing the first set of constraints and using ϕ to denote the Lagrange multipliers, the dual problem can be written as below:

$$\min_{\phi \ge 0} \left(\phi^t e + \max_{A(x+u) \le b, ||u|| \le \lambda} (c - D^t \phi)^t x \right). \tag{15}$$

The inner maximization problem has the same form as problem (12). Then we can deduce that if the set $Ax \le b$ is bounded, the inner problem is equivalent to

$$\max_{||c - D^t \phi|| Ax \le ||c - D^t \phi|| b + \lambda A(c - D^t \phi)} (c - D^t \phi)^t x. \tag{16}$$

In other words, the convex problem (14) can be handled (under constraint qualification) by solving a series of linear programs. Several approaches can be considered here. One can use a classical subgradient optimization algorithm where ϕ is moved in the direction of the current subgradient given by $Dx - e$. We can also use a cutting plane algorithm where linear cuts are given by the subgradients obtained when the inner maximization problems (16) are solved. A linear approximation of the dual function is then improved at each iteration. The next ϕ can be computed in different ways (a Kelly's algorithm [14], a bundle type method [16], central cutting plane algorithms [11,9,25], a multiple-points separation algorithm [3], an improvement of Kelly's algorithm [4], etc.).

Remarks

- A similar optimality-equivalent formulation can be derived if we use the function $f(\epsilon) = \lambda\epsilon^2$, where $\lambda \geq 0$. We prove in [5] that program (17) is optimality-equivalent to program (9).

$$LP_f \begin{cases} \max c^t x \\ Ax - b \leq (2\lambda Acc^t A^t + 2vAA^t)w - 2\lambda Ac \\ w^t(\lambda Acc^t A^t + vAA^t)w - \lambda + v \leq 0 \\ v \in \mathbb{R}, v \geq 0 \\ w \in \mathbb{R}^m, x \in \mathbb{R}^n \end{cases} \tag{17}$$

- For any function f such that $\liminf_{\epsilon \to 0+} \frac{f(\epsilon)}{\sqrt{\epsilon}} = 0$, we know that any point x maximizing $c^t x$ and satisfying $\frac{\sigma^t x - \max_{y \in S} \sigma^t y}{f(1-\sigma^t c)} \leq 1$ (when $f(1-\sigma^t c) \neq 0$) for all vectors σ is necessarily an optimal solution. This may provide a new kind of rules to generate cuts: instead of generating the most violated cut, one can look for a cut maximizing the violation divided by $f(1-\sigma^t c)$. The same can be done when we are dealing with column generation. New pricing rules can be tried for example to choose the next vertex to be visited when the simplex algorithm is applied. We only have to normalize the right-hand size vector of the linear problem and consider the angle between this vector and the column to be generated. Computational experiments are needed to study the performances of this approach.
- We can also try to generate either cuts or columns maximizing $\sigma^t x - \max_{y \in S} \sigma^t y - f(1 - \sigma^t c)$. If we take $f(\epsilon) = \lambda\epsilon$, then the separation problem becomes $\max_{\sigma} \left(\sigma^t (x + \lambda c) - \max_{y \in S} \sigma^t y \right)$. Said another way, instead of separating x, we separate $x + \lambda c$ (more precisely, we compute the most violated inequality by $x + \lambda c$). Notice that this seems to be quite natural: when $\lambda \gg 1$, a vector σ defining an inequality violated by $x + \lambda c$ will be close to c. Computational experiments are again needed to study the performances of this approach.

6 Conclusion and further research

Given a linear objective function and a convex set of feasible solutions, we have a full characterization of the relaxations that do not modify the set of optimal solutions. It turns out that the angle between the objective vector c and the vector σ defining a valid inequality has some importance. Roughly speaking, when this angle becomes closer to 0, there exists a sequence of valid inequalities that are relaxed in a limited way (in the sense of Theorem 4.4). We can even skip all the valid inequalities for which the angle between σ and c is larger than any given small number.

The general form of relaxations given in this paper may be useful to simplify calculation. Given a problem similar to either (12) or (17), it can be replaced by a linear program without changing the set of optimal solutions. This was used in Sect. 5 to solve a linear problem where some constraints can be violated within a certain limit. Notice that the model (14) is interesting in its own right since it introduces a flexibility in a simple way.

A promising research direction consists in looking for some other transformations of "difficult" problems into simpler ones.

Some connections with sensitivity analysis were pointed out in Proposition 5.1. We also mentioned some potential applications in the context of both cut and column generation. A deeper study and computational experiments are needed to evaluate the performances of this approach.

References

1. Alizadeh, F., Goldfarb, D.: Second-order cone programming. Math. Program. Ser. B **95**, 3–51 (2003)
2. Balas, E., Ceria, S., Cornuéjols, G.: A lift-and-project cutting plane algorithm for mixed 0-1 programs. Math. Program. **58**, 295–324 (1993)
3. Ben-Ameur, W., Neto, J.: A constraint generation algorithm for large scale linear programs using multiple-points separation. Math. Program. Ser. A **107**, 517–537 (2006)
4. Ben-Ameur, W., Neto, J.: Acceleration of cutting plane and column generation algorithms: applications to network design. Networks **49**, 3–17 (2007)
5. Ben-Ameur, W., Neto, J.: A geometric characterization of good relaxations. Technical report, Institut National des Telecommunications (2007)
6. Ben-Ameur, W., Ouorou, A.: Mathematical models of the delay constrained routing problem. Algorithmic Oper. Res. **1**(2), 94–103 (2006)
7. Boyd, S., Vandenberghe, L.: Convex Optimization. Cambridge University Press (2004)
8. Chvátal, V.: Edmonds polytopes and a hierarchy of combinatorial problems. Discrete Math. **4**, 305–337 (1973)
9. Elzinga, J., Moore, T.G.: A central cutting plane algorithm for the convex programming problem. Math. Program. **8**, 134–145 (1973)
10. Goemans, M., Williamson, D.: Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. J. ACM **42**, 1115–1145 (1995)
11. Goffin, J.L., Vial, J.P.: Convex nondifferentiable optimization: a survey focussed on the analytic center cutting plane method. Optim. Method Softw. **17**, 805–867 (2002)
12. Gomory, R.E.: Outline of an algorithm for integer solutions to linear programs. Bull. Am. Soc. **64**, 275–278 (1958)
13. Hiriart-Urruty, J., Lemaréchal, C.: Convex Analysis and Minimization Algorithms. Springer (1993)
14. Kelley, J.E.: The cutting-plane method for solving convex programs. J. SIAM **8**, 703–712 (1960)
15. Lovász, L., Schrijver, A.: Cones of matrices and set-functions and 0-1 optimization. SIAM J. Optim. **1**, 166–190 (1991)
16. Lemaréchal, C.: An extension of Davidon methods to nondifferentiable problems. Math. Program. **3**, 95–109 (1975)
17. Lemaréchal, C.: The omnipresence of Lagrange. 4OR **1**(1), 7–25 (2003)
18. Nemhauser, G.L., Wolsey, L.A.: Integer and Combinatorial Optimization. Wiley, New-York (1988)
19. Ouorou, A.: Epsilon-proximal decomposition method. Math. Program. **99**(1), 89–108 (2004)
20. Padberg, M.: Linear Optimization and Extensions. Springer (1995)
21. Pochet, Y., Wolsey, L.: Algorithms and reformulations for lot-sizing problems in combinatorial optimization. In: Cook, W., Lovasz, L., Seymour, P. (eds.) Papers from the DIMACS Special Year, DIMACS Series in Discrete Mathematics and Theoretical Computer Science, vol. 20. American Mathematical Society (1995)
22. Schrijver, A.: Combinatorial Optimization: Polyhedra and Efficiency. Springer (2003)
23. Sherali, H., Adams, W.: A hierarchy of relaxations between the continuous and convex hull representations for zero-one programming problems. SIAM J. Discrete Math. **3**, 411–430 (1990)
24. Todd, M.: Semidefinite optimization. Acta Numer. **10**, 515–560 (2001)
25. Vaidya, P.M.: A new algorithm for minimizing convex functions over convex sets. Math. Program. **73**, 291–341 (1996)