# Inverse Linear Programming

S. Dempe[1] and S. Lohse[2]

[1] Technical University Bergakademie Freiberg, Department of Mathematics and Computer Sciences, Akademiestr. 6, 09596 Freiberg, Germany, corresponding author, dempe@math.tu-freiberg.de
[2] Technical University Bergakademie Freiberg, Department of Mathematics and Computer Sciences, Akademiestr. 6, 09596 Freiberg, Germany

**Summary.** Let $\Psi(b,c)$ be the solution set mapping of a linear parametric optimization problem with parameters $b$ in the right hand side and $c$ in the objective function. Then, given a point $x^0$ we search for parameter values $\bar{b}$ and $\bar{c}$ as well as for an optimal solution $\bar{x} \in \Psi(\bar{b},\bar{c})$ such that $\|\bar{x} - x^0\|$ is minimal. This problem is formulated as a bilevel programming problem. Focus in the paper is on optimality conditions for this problem. We show that, under mild assumptions, these conditions can be checked in polynomial time.

## 1 Introduction

Let $\Psi(b,c) = \operatorname{argmax}\{c^\top x : \ Ax = b, \ x \geq 0\}$ denote the set of optimal solutions of a linear parametric optimization problem

$$\max\left\{c^\top x : \ Ax = b, \ x \geq 0\right\} , \tag{1}$$

where the parameters of the right hand side and in the objective function are elements of given sets

$$\mathcal{B} = \{b : \ Bb = \tilde{b}\} , \quad \mathcal{C} = \{c : \ Cc = \tilde{c}\} ,$$

respectively. Throughout this note, $A \in \mathbb{R}^{m\times n}$ is a matrix of full row rank $m$, $B \in \mathbb{R}^{p\times m}, C \in \mathbb{R}^{q\times n}, \tilde{b} \in \mathbb{R}^p$ and $\tilde{c} \in \mathbb{R}^q$. This data is fixed once and for all.

Let $x^0 \in \mathbb{R}^n$ also be fixed. Our task is to find values $\bar{b}$ and $\bar{c}$ for the parameters, such that $x^0 \in \Psi(\bar{b},\bar{c})$ or, if this is not possible, $x^0$ is at least close to $\Psi(\bar{b},\bar{c})$. Thus we consider the following bilevel programming problem

$$\min\left\{\|x - x^0\| : \ x \in \Psi(b,c), \ b \in \mathcal{B}, \ c \in \mathcal{C}\right\} , \tag{2}$$

which has a convex objective function $x \in \mathbb{R}^n \mapsto f(x) := \|x - x^0\|$, but not necessarily a convex feasible region. We consider in this note an arbitrary (semi)norm $\|\cdot\|$, not necessarily the Euclidean norm. In fact, we are specially thinking in a polyhedral norm like, for instance, the $l_1$-norm.

Bilevel programming problems have been intensively investigated, see the monographs [2, 3] and the annotated bibliography [4]. Inverse linear programming problems have been investigated in the paper [1], where it is shown that the inverse problem to e.g. a shortest path problem can again be formulated as a shortest path problem and there is no need to solve a bilevel programming problem. However, the main assumption in [1] that there exist parameter values $\overline{b} \in \mathcal{B}$ and $\overline{c} \in \mathcal{C}$ such that $x^0 \in \Psi(\overline{b}, \overline{c})$ seems to be rather restrictive. Hence, we will not use this assumption.

Throughout the paper the following system is supposed to be infeasible:

$$\begin{aligned} A^\top y &= c \\ Cc &= \tilde{c}\,. \end{aligned} \tag{3}$$

Otherwise every solution of

$$\begin{aligned} Ax &= b \\ x &\geq 0 \\ Bb &= \tilde{b}\,, \end{aligned}$$

would be feasible for (2), which means that (2) reduces to

$$\min\left\{\|x - x^0\| :\ Ax = b,\ x \geq 0,\ Bb = \tilde{b}\right\}\ ,$$

which is a convex optimization problem.

## 2 Reformulation as an MPEC

First we transform (2) via the Karush-Kuhn-Tucker conditions into a mathematical program with equilibrium constraints (MPEC) [5] and we get

$$\begin{aligned} \|x - x^0\| &\longrightarrow \min_{x,b,c,y} \\ Ax &= b \\ x &\geq 0 \\ A^\top y &\geq c \\ x^\top(A^\top y - c) &= 0 \\ Bb &= \tilde{b} \\ Cc &= \tilde{c}\,. \end{aligned} \tag{4}$$

The next thing which should be clarified is the notion of a local optimal solution.

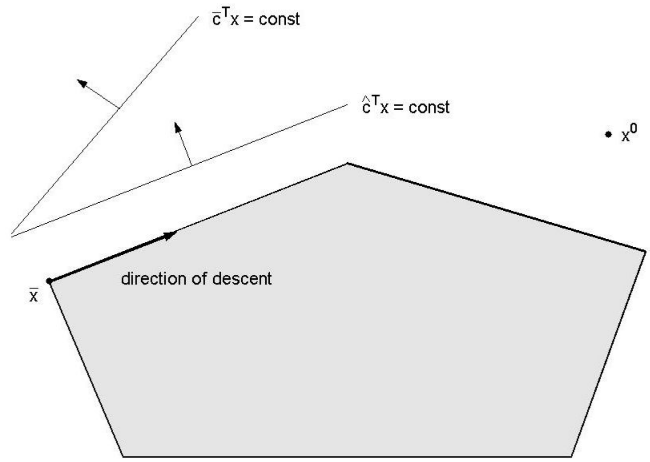


**Fig. 1.** Definition of a local optimal solution

**Definition 1.** *A point* $\overline{x}$ *is a local optimal solution of problem* (2) *if there exists a neighborhood* $U$ *of* $\overline{x}$ *such that* $\|x - x^0\| \geq \|\overline{x} - x^0\|$ *for all* $x, b, c$ *with* $b \in \mathcal{B}$, $c \in \mathcal{C}$ *and* $x \in U \cap \Psi(b, c)$.

Using the usual definition of a local optimal solution of problem (4) it can be easily seen that for each local optimal solution $\overline{x}$ of problem (2) there are $\overline{b}, \overline{c}, \overline{y}$ such that $(\overline{x}, \overline{b}, \overline{c}, \overline{y})$ is a local optimal solution of problem (4), cf. [3]. The opposite implication is in general not true.

**Theorem 1.** *Let* $\mathcal{B} = \{\overline{b}\}$, $\{\overline{x}\} = \Psi(\overline{b}, c)$ *for all* $c \in U \cap \mathcal{C}$, *where* $U$ *is some neighborhood of* $\overline{c}$. *Then,* $(\overline{x}, \overline{b}, \overline{c}, \overline{y})$ *is a local optimal solution of* (4) *for some dual variables* $\overline{y}$.

The proof of Theorem 1 is fairly easy and therefore it is omitted. Figure 1 can be used to illustrate the fact of the last theorem. The points $\overline{x}$ satisfying the assumptions of Theorem 1 are the vertices of the feasible set of the lower level problem given by the dashed area in this figure.

## 3 Optimality via Tangent Cones

Now we consider a feasible point $\overline{x}$ of problem (2) and we want to decide whether $\overline{x}$ is local optimal or not. To formulate suitable optimality conditions certain subsets of the index set of active inequalities in the lower level problem need to be determined. Let

$$I(\overline{x}) = \{i : \overline{x}_i = 0\}$$

be the index set of active indices. Then every feasible solution $x$ of (2) close enough to $\overline{x}$ satisfies $x_i > 0$ for all $i \notin I(\overline{x})$. Complementarity slackness motivates us to define the following index sets, too:

- $I(c,y) = \{i : (A^\top y - c)_i > 0\}$
- $\mathcal{I}(\overline{x}) = \left\{I(c,y) : A^\top y \geq c,\ (A^\top y - c)_i = 0\ \ \forall i \notin I(\overline{x}),\ Cc = \tilde{c}\right\}$
- $I^0(\overline{x}) = \bigcap\limits_{I \in \mathcal{I}(\overline{x})} I.$

*Remark 1.* If an index set $I$ belongs to the family $\mathcal{I}(\overline{x})$ then $I^0(\overline{x}) \subseteq I \subseteq I(\overline{x})$.

An efficient calculation of the index set $I^0(\overline{x})$ is necessary for the evaluation of the optimality conditions below. By contrast, the knowledge of the family $\mathcal{I}(\overline{x})$ itself is not necessary.

*Remark 2.* We have $j \in I(\overline{x}) \setminus I^0(\overline{x})$ if and only if the system

$$\begin{aligned} (A^\top y - c)_i &= 0 \quad \forall i \notin I(\overline{x}) \\ (A^\top y - c)_j &= 0 \\ (A^\top y - c)_i &\geq 0 \quad \forall i \in I(\overline{x}) \setminus \{j\} \\ Cc &= \tilde{c} \end{aligned}$$

is feasible. Furthermore $I^0(\overline{x})$ is an element of $\mathcal{I}(\overline{x})$ if and only if the system

$$\begin{aligned} (A^\top y - c)_i &= 0 \quad \forall i \notin I^0(\overline{x}) \\ (A^\top y - c)_i &\geq 0 \quad \forall i \in I^0(\overline{x}) \\ Cc &= \tilde{c} \end{aligned}$$

is feasible.

Now we are able to transform (4) into a locally equivalent problem, which does not explicitly depend on $c$ and $y$.

**Lemma 1.** *$\overline{x}$ is a local optimal solution of (2) if and only if $\overline{x}$ is a (global) optimal solution of all problems $(A_I)$*

$$\begin{aligned} \|x - x^0\| &\longrightarrow \min_{x,b} \\ Ax &= b \\ x &\geq 0 \\ x_i &= 0 \quad \forall i \in I \\ Bb &= \tilde{b} \end{aligned} \tag{$A_I$}$$

*with $I \in \mathcal{I}(\overline{x})$.*

*Proof.* Let $\overline{x}$ be a local optimal solution of (2) and assume that there is a set $I \in \mathcal{I}(\overline{x})$ with $\overline{x}$ being not optimal for $(A_I)$. Then there exists a sequence $\{x^k\}_{k\in\mathbb{N}}$ of feasible solutions of $(A_I)$ with $\lim\limits_{k\to\infty} x^k = \overline{x}$ and $\|x^k - x^0\| < \|\overline{x} - x^0\|$ for all $k$. Consequently $\overline{x}$ can not be a local optimal solution to (2) since $I \in \mathcal{I}(\overline{x})$ implies that all $x^k$ are also feasible for (2).

Conversely, let $\overline{x}$ be an optimal solution of all problems $(A_I)$ and assume that there is a sequence $\{x^k\}_{k\in\mathbb{N}}$ of feasible points of (2) with $\lim\limits_{k\to\infty} x^k = \overline{x}$ and $\|x^k - x^0\| < \|\overline{x} - x^0\|$ for all $k$. For $k$ sufficiently large the elements of this sequence satisfy the condition $x_i^k > 0$ for all $i \notin I(\overline{x})$ and due to the feasibility of $x^k$ for (2) there are sets $I \in \mathcal{I}(\overline{x})$ such that $x^k$ is feasible for problem $(A_I)$. Because $\mathcal{I}(\overline{x})$ consists only of a finite number of sets, there is a subsequence $\{x^{k_j}\}_{j\in\mathbb{N}}$ where $x^{k_j}$ are all feasible for a fixed problem $(A_I)$. So we contradict the optimality of $\overline{x}$ for this problem $(A_I)$.

**Corollary 1.** *We can also consider*

$$\begin{aligned} \|x - x^0\| &\longrightarrow \min_{x,b,I} \\ Ax &= b \\ x &\geq 0 \\ x_i &= 0 \quad \forall i \in I \\ Bb &= \tilde{b} \\ I &\in \mathcal{I}(\overline{x}) \end{aligned} \tag{5}$$

*to check if $\overline{x}$ is a local optimal solution of (2). Here the index set $I$ is a minimization variable. Problem (5) combines all the problems $(A_I)$ into one problem and means that we have to find a best one between all the optimal solutions of the problems $(A_I)$ for $I \in \mathcal{I}(\overline{x})$.*

In what follow we use the notation

$$T_I(\overline{x}) = \{d |\ \exists r :\ Ad = r,\ Br = 0,\ d_i \geq 0\ \ \forall i \in I(\overline{x}) \setminus I,\ d_i = 0\ \ \forall i \in I\}\ .$$

This set corresponds to the tangent cone (relative to $x$ only) to the feasible set of problem $(A_I)$ at the point $\overline{x}$. The last lemma obviously implies the following necessary and sufficient optimality condition.

**Lemma 2.** *$\overline{x}$ is a local optimal solution of (5) if and only if $f'(\overline{x}, d) \geq 0$ for all*

$$d \in T(\overline{x}) := \bigcup_{I \in \mathcal{I}(\overline{x})} T_I(\overline{x})\ .$$

*Remark 3.* $T(\overline{x})$ is the (not necessarily convex) tangent cone (relative $x$) of problem (5) at the point $\overline{x}$.

**Corollary 2.** *The condition $I^0(\overline{x}) \in \mathcal{I}(\overline{x})$ implies $T_{I^0(\overline{x})}(\overline{x}) = T(\overline{x})$.*

*Remark 4.* If $f$ is differentiable at $\overline{x}$, then saying that $f'(\overline{x},\cdot)$ is nonnegative over $T(\overline{x})$ is obviously equivalent to saying that

$$f'(\overline{x},d) \geq 0 \quad \forall d \in \operatorname{conv} T(\overline{x}) \ , \tag{6}$$

where the "conv" indicates the convex hull operator.

As shown in the next example, without differentiablility assumption, (6) is sufficient for optimality but not necessary.

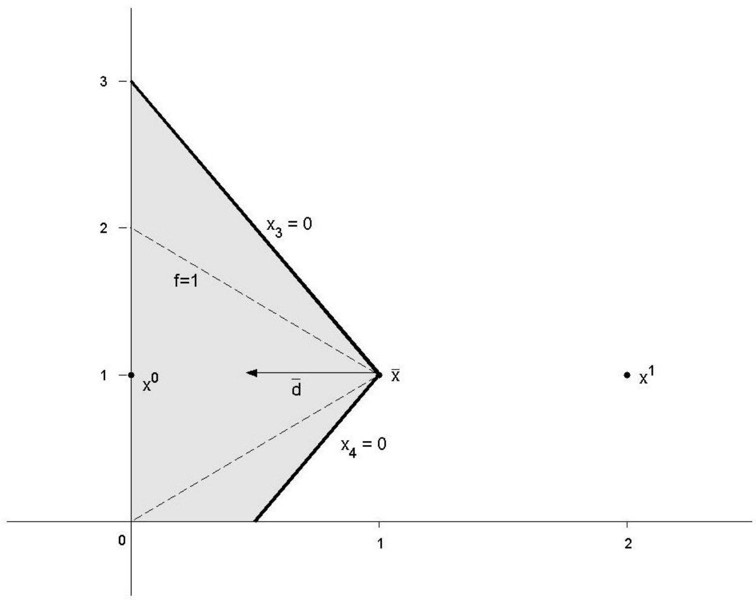


**Fig. 2.** Illustration of Example 1

*Example 1.* Let us consider a problem with the $l_1$-norm restricted to the first two components of $x$ as objective function and

$$A = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 2 & -1 & 0 & 1 \end{pmatrix}, \ \mathcal{B} = \left\{ \begin{pmatrix} 3 \\ 1 \end{pmatrix} \right\}, \ \mathcal{C} = \left\{ 2e_1^{(4)} + te_2^{(4)} : \ t \in \mathbb{R} \right\},$$

$$x^0 = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 2 \end{pmatrix}, \ x^1 = \begin{pmatrix} 2 \\ 1 \\ -2 \\ -2 \end{pmatrix} \quad \text{and} \quad \overline{x} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} .$$

We consider the point $\overline{x}$. The bold marked lines in Fig. 2 are the feasible set of our problem and the dashed lines are iso-distance-lines with the value 1. So we get the convexified tangent cone as

$$\operatorname{conv} T(\overline{x}) = \{d : \ 2d_1 + d_2 + d_3 = 0;\ 2d_1 - d_2 + d_4 = 0;\ d_3, d_4 \geq 0\}\ .$$

Finally $\overline{d} = (-1\ 0\ 2\ 2)^\top \in \operatorname{conv} T(\overline{x})$ is a direction of descent with $f'(\overline{x}, \overline{d}) = -1$ although $\overline{x}$ is obviously the global optimal solution. If we choose $x^1$ (instead of $x^0$ ) and the objective function $|x_1 - x_1^1| + |x_2 - x_2^1|$, condition (6) implies the optimality of $\overline{x}$.

*Remark 5.* Because it is a matter of illustration, we considered the problem with inequality constraints in the lower level. For that reason we used the $l_1$-norm restricted to the first two components of $x$ as objective function and not the $l_1$-norm over the whole space $\mathbb{R}^4$. By the way, in this case $\overline{x}$ would not be a local optimal solution.

## 4 A Formula for the Tangent Cone

For the verification of the optimality condition (6) an explicit formula for the tangent cone $\operatorname{conv} T(\overline{x})$ is essential. For notational simplicity we suppose $I(\overline{x}) = \{1, \ldots, k\}$ and $I^0(\overline{x}) = \{l+1, \ldots, k\}$ with $l \leq k \leq n$. Consequently all feasible points of (2) sufficiently close to $\overline{x}$ satisfy $x_i = 0$ for all $i \in I^0(\overline{x})$. We pay attention to this fact and consider the following relaxed problem:

$$\begin{aligned} \|x - x^0\| &\longrightarrow \min_{x,b} \\ Ax &= b \\ x_i &\geq 0 \quad i = 1, \ldots, l \\ x_i &= 0 \quad i = l+1, \ldots, k \\ Bb &= \tilde{b}\ . \end{aligned} \tag{7}$$

In what follow we use the notation

$$T_R(\overline{x}) = \{d |\ \exists r : Ad = r,\ Br = 0,\ d_i \geq 0\ i = 1, \ldots, l,\ d_i = 0\ \ i = l+1, \ldots, k\}.$$

This set corresponds to the tangent cone (relative $x$) of (7) at the point $\overline{x}$. Since $I^0 \subseteq I$ for all $I \in \mathcal{I}(\overline{x})$, it follows immediately that

$$\operatorname{conv} T(\overline{x}) = \operatorname{cone} T(\overline{x}) \subseteq T_R(\overline{x})\ . \tag{8}$$

The point $\overline{x}$ is said to satisfy the *full rank condition*, if

$$\operatorname{span}(\{A_i :\ i \notin I(\overline{x})\}) = \mathbb{R}^m, \tag{FRC}$$

where $A_i$ denotes the $i$th column of the matrix $A$.

*Example 2.* All non-degenerate vertices of $Ax = b$, $x \geq 0$ satisfy (FRC).

This condition allows us now to establish equality between the cones above.

**Theorem 2.** *Let (FRC) be satisfied at the point $\overline{x}$. Then equality holds in (8).*

*Proof.* Let $\overline{d}$ be an arbitrary element of $T_R(\overline{x})$, that means there is a $\overline{r}$ with $A\overline{d} = \overline{r},\ B\overline{r} = 0,\ \overline{d}_i \geq 0\ \ i = 1, \dots, l,\ \overline{d}_i = 0\ \ i = l+1, \dots, k$. We consider the following linear systems

$$\begin{aligned} Ad &= \delta_{1,j}\overline{r} \\ d_j &= \overline{d}_j \\ d_i &= 0 \quad i = 1, \dots, k,\ i \neq j \end{aligned} \tag{$S_j$}$$

for $j = 1, \dots, l$, where $\delta_{1,j} = 1$ if $j = 1$ and $\delta_{1,j} = 0$ if $j \neq 1$. These systems are all feasible because of (FRC).

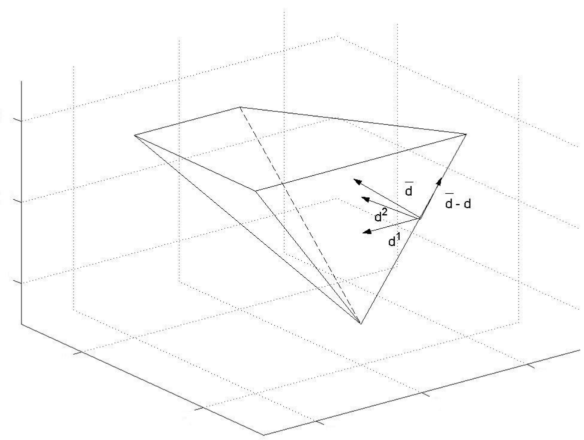


**Fig. 3.** Illustration of the proof of Theorem 2

Furthermore let $d^1, \dots, d^l$ be (arbitrary) solutions of the systems $(S_1), \dots, (S_l)$ respectively. We define now the direction $d = \sum\limits_{j=1}^{l} d^j$ and get $d_i = \overline{d}_i$ for $i = 1, \dots, k$ as well as $Ad = A\overline{d} = \overline{r}$. Because we chose arbitrary vectors $d^1, \dots, d^l$ it is possible that $d \neq \overline{d}$. But we can achieve equality with a translation of the solution $d^1$ by a specific vector of $\mathcal{N}(A) = \{z :\ Az = 0\}$. Therefore we define $\hat{d}^1 := d^1 + \overline{d} - d$, and because $d^1$ is feasible for $(S_1)$ and $d_i = \overline{d}_i$ for $i = 1, \dots, k$ as well as $Ad = A\overline{d} = \overline{r}$ we get $\hat{d}^1_i = 0$ for all $i = 2, \dots, k$ and $A\hat{d}^1 = A(d^1 + \overline{d} - d) = \overline{r} + \overline{r} - \overline{r} = \overline{r}$. Hence $\hat{d}^1$ is also a solution of $(S_1)$. Thus we have $\hat{d}^1 + \sum\limits_{j=2}^{l} d^j = \overline{d} - d + \sum\limits_{j=1}^{l} d^j = \overline{d}$. As a result of the definition of the set $I^0(\overline{x})$ there are index sets $I_j \in \mathcal{I}(\overline{x})$ with $j \notin I_j$ for all $j \in \{1, \ldots, l\} = I(\overline{x}) \setminus I^0(\overline{x})$. So $\hat{d}^1$ is an element of the tangent cone of problem $(A_{I_1})$ and $d^j$ are elements of the tangent cones of the problems $(A_{I_j})$ for $j = 2, \ldots, l$, see the definition of these cones. Finally $\overline{d}$ is the sum of a finite number of elements of $T(\overline{x})$ and therefore $T_R(\overline{x}) \subseteq \operatorname{cone} T(\overline{x})$.

By combining Lemma 2 and Remarks 2 and 4, one obtains:

**Corollary 3.** *Let $\overline{x}$ be a point of differentiability of $f$. Then, at most $n$ systems of linear equalities\inequalities are needed to be investigated in order to compute the index set $I^0(\overline{x})$. Furthermore, verification of local optimality of a feasible point of problem* (2) *is possible in polynomial time.*

*Example 3.* This example will show that (FRC) is not necessary for equality in (8).

$$\begin{array}{rcrcrcrcrcrcrcrcl}
 & & x_2 & & & - & x_4 & & & & & & & & & = & 1 \\
2x_1 & + & 2x_2 & - & x_3 & & & + & x_5 & & & & & & & = & 3 \\
 & & 2x_2 & - & x_3 & & & & & + & x_6 & & & & & = & 1 \\
2x_1 & & & + & x_3 & & & & & & & - & x_7 & & & = & 3 \\
 & & & & x_3 & & & & & & & & & + & x_8 & = & 3 \\
 & & & & & & & & & & & & & & x_i & \geq & 0
\end{array}$$

$\mathcal{B} = \{(1\ 3\ 1\ 3\ 3)^\top\}$ and $\mathcal{C} = \{c = -e_2^{(8)} + t(2e_1^{(8)} + 3e_2^{(8)} - e_3^{(8)}) + s(3e_2^{(8)} - e_3^{(8)}) :\ t, s \in \mathbb{R}\}$.

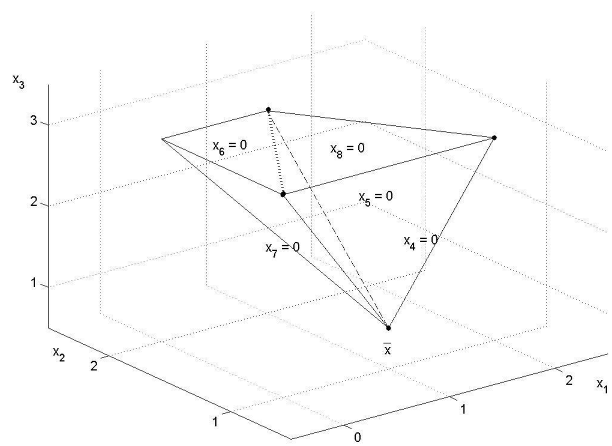


**Fig. 4.** Illustration of Example 3

Consider the point $\overline{x} = (1, 1, 1, 0, 0, 0, 0, 2)^\top$. Hence we get $I(\overline{x}) = \{4, 5, 6, 7\}$, $I^0 = \emptyset$ and $T_R(\overline{x}) = \{d :\ Ad = 0,\ d_i \geq 0 \quad \forall i \in I(\overline{x})\}$. The feasible region of (5) consists of the four faces $x_4 = 0$, $x_5 = 0$, $x_6 = 0$ and $x_7 = 0$ ($t = s = 0$; $t = 1, s = 0$; $t = 0, s = 1$ respectively $t = -\frac{1}{3}, s = \frac{2}{3}$). Obviously we have $T_R(\overline{x}) = \operatorname{cone} T(\overline{x})$.

Now delete the second vector in $\mathcal{C}$, that means $\mathcal{C} = \{c = -e_2^{(8)} + t(2e_1^{(8)} + 3e_2^{(8)} - e_3^{(8)}) : \ t \in \mathbb{R}\}$. Then we also get $I^0 = \emptyset$. That is why the tangent cone of the relaxed problem is the same as above. But the convexified tangent cone $\operatorname{conv} T(\overline{x})$ of (5) is a proper subset of this cone. Because the feasible set consists only of the two faces $x_4 = 0$ and $x_5 = 0$, the cone $\operatorname{conv} T(\overline{x})$ is spanned by the four bold marked vertices where the apex of the cone is $\overline{x}$, see Fig. 4.

**Acknowledgement.** The authors sincerely thank the anonymous referee, whose comments led to an improvement of the note.

## References

1. Ahuja RK, Orlin JB (2001) Inverse Optimization. Operations Research 49:771–783
2. Bard JF (1998) Practical Bilevel Optimization: Algorithms and Applications. Kluwer Academic Publishers, Dordrecht
3. Dempe S (2002) Foundations of Bilevel Programming. Kluwer Academic Publishers, Dordrecht
4. Dempe S (2003) Annotated Bibliography on Bilevel Programming and Mathematical Programs with Equilibrium Constraints. Optimization 52:333–359
5. Outrata J, Kočvara M, Zowe J (1998) Nonsmooth Approach to Optimization Problems with Equilibrium Constraints. Kluwer Academic Publishers, Dordrecht