Journal of Global Optimization **16:** 245–255, 2000.

A Note on a Penalty Function Approach for Solving Bilevel Linear Programs

MANOEL CAMPÊLO[1,*], SIMONE DANTAS[2] and
SUSANA SCHEIMBERG[2,†]
[1]*Universidade Federal do Ceará, Ceaná, Brazil;* [2]*Universidade Federal do Rio de Janeiro – COPPE/Sistemas, Rio de Janeiro, Brazil*
(e-mail: susana@cos.ufrj.br)

(Received 24 June 1998; accepted in revised form 13 December 1999

Abstract. We have identified some trouble in the article 'A Penalty Function Approach for Solving Bi-Level Linear Programms' (*J. Global Optimization* 3: 397–419). The primal and dual compactness assumption considered is not valid. The set of cuts used in the algorithm to discard local optima is not well-defined. The test to identify possible remaining better solutions is not accurate. We redefine the cut set and correct the test. We obtain good properties for the penalized problem without assuming compactness. However, we note that the global algorithm even needs a dual compactness assumption to be well-defined. Examples are given to illustrate the remarks in the article.

Key words: Bilevel linear programming; Global optimization; Penalty function

1. Introduction

The Linear Bilevel Programming Problem (LBPP) is a strongly NP-Hard problem [8] that has attracted much interest lately. Several solution methods have been proposed in the literature (see p.e. [3, 13]).

In this note we focus our attention on the article of White and Anandalingam [14]. The authors consider the following linear bilevel problem:

$$(P)\ \max_{(x,y)} F(x,y) = ax + by \tag{1}$$

$$\text{s.t.}: x \geqslant 0,\ y \text{ solves:} \tag{2}$$

$$\max_{y}\ f(x,y) = cx + dy \tag{3}$$

$$\text{s.t.}: Ax + By \leqslant p,\ \ y \geqslant 0 \tag{4}$$

where $a, c, x \in R^{n_1}$, $b, d, y \in R^{n_2}$, $p \in R^m$, $A \in R^{m \times n_1}$ and $B \in R^{m \times n_2}$. The general formulation of the LBPP may add to (2) linear constraints involving x and y.

---

* Partially supported by PICDT-CAPES at the Universidade Federal do Rio de Janeiro – COPPE/Sistemas.

† Instituto de Matemética – UFRJ. Partially supported by CNPq Grant 302393185.

The higher level problem, which controls the decision variable x, is called the leader's problem. For each value of x, the lower level problem, which controls the decision variable y, is the follower's problem. Once x is given the follower's objective function (3) can be reduced to dy. Thus the follower's dual problem, defined in the dual variable $w \in R^m$, is:

$$(D) \min_w w(p - Ax) \tag{5}$$

$$\text{s.t.} : wB \geqslant d, \ \ w \geqslant 0 \tag{6}$$

Problem (P) is nonconvex, since its feasible region is nonconvex [4]. Furthermore, it may have local optima.

In [14] White and Anandalingam develop a penalty function approach for solving (P) globally. In order to obtain theoretical results and the well-definition of the algorithm, they introduce two hypotheses:

[A1] If x^* is an optimal solution for the leader, then $\arg\min\{dy : By \leqslant p - Ax^*, y \geqslant 0\}$ is a singleton.

[A2] The following sets are non-empty bounded polyhedra:

$$Z = \{(x, y) \geqslant 0 : Ax + By \leqslant p\}, \quad W = \{w \geqslant 0 : wB \geqslant d\}. \tag{7}$$

We have identified some problems with the approach developed in [14]. As it is shown in the next section, polyhedra Z and W cannot be simultaneously bounded. Therefore assumption [A2] is not valid, which causes both theoretical and numerical troubles in [14]. Another problem arises from the set of cuts used to discard local optima. It is not well-defined even under assumptions [A1] and [A2]. Moreover, the test to identify possible remaining better solutions is not accurate.

In the following section we describe the White and Anandalingam's approach. In Section 3 we replace the hypotheses given in [14] by a weaker one and obtain the same theoretical properties. In Section 4 we punctuate the trouble with the algorithm. Problems with the cuts are overcome by modifying the set definition and correcting the test to recognize the attainment of a solution to the penalized problem. However the complete well-definition of the algorithm is only assured under a compactness assumption. The examples of Section 5 illustrate all these observations. We finish with a conclusion section.

2. A penalty function approach

Let us denote by Z_v, W_v and $(Z \times W)_v$ the sets of extreme points of Z, W and $Z \times W$ respectively. For every $(x, y, w) \in Z \times W$ let us define the nonnegative duality gap $\pi(x, y, w) = w(p - Ax) - dy$.

Given $(x, y, w) \in Z \times W$, if $\pi(x, y, w) = 0$ then (x, y) is a feasible solution to (P). Conversely, if $(x, y) \in Z$ is feasible to (P) then $\pi(x, y, w) = 0$ for some $w \in W$. Thus, a penalty problem for (P) is given by:

$$P(K)\ \max_{(x,y,w)} \hat{F}(x, y, w, K) = ax + by - K\pi(x, y, w) \tag{8}$$

$$\text{s.t.} : (x, y) \in Z\ ,\ \ w \in W \tag{9}$$

where the duality gap is introduced in the leader's objective function by a penalty parameter $K \in R_+$.

Under assumptions [A1] and [A2], White and Anandalingam [14] state Theorems 1–4 below. In fact assumption [A1] is only used in Theorem 3 (the proofs can be found in [1]).

THEOREM 1. *For a given value of* $w \in W$ *and fixed* $K \in R_+$, *define:*

$$\Theta(w, K) = \max_{x,y} \{\hat{F}(x, y, w, K) : (x, y) \in Z\}\ . \tag{10}$$

Then $\Theta(\cdot, K)$ *is convex on* R^m *and the solution to the problem*

$$\max_{w} \{\Theta(w, K) : w \in W\} \tag{11}$$

will occur at some $w^* \in W_v$.

THEOREM 2. *For a fixed* $K \in R_+$, *an optimal solution to* $P(K)$ *is achievable in* $Z_v \times W_v$ *and* $Z_v \times W_v = (Z \times W)_v$.

THEOREM 3. *There exists a finite value* $K^* \in R_+$ *of* K *for which an optimal solution to the penalty problem* $P(K)$ *yields an optimal solution to the problem* (P), *for all* $K \geqslant K^*$.

THEOREM 4. *If* $(x(K), y(K), w(K))$ *solves* $P(K)$ *as a function of* K, *then both the leader's objective* $F(x(K), y(K))$ *and the duality gap* $\pi(x(K), y(K), w(K))$ *of the follower's problem are monotonically nonincreasing in the value of the penalty parameter* K.

Nevertheless, assumption [A2] is not suitable as it can be seen from the next result given in [6]:

THEOREM 5. *If* $Y = \{y : By \leqslant q,\ \ y \geqslant 0\}$ *is nonempty and bounded, then* $W = \{w : wB \geqslant d,\ \ w \geqslant 0\}$ *is nonempty and unbounded; also if* W *is nonempty and bounded then* Y *is nonempty and unbounded.*

3. A weaker assumption

In this section we claim to obtain the same results stated in Theorems 1–4 under a different assumption. Initially note that the first part of Theorem 1 is valid independently of [A1] or [A2], i.e. $\Theta(\cdot, K)$ is convex on R^m for each $K \geqslant 0$ (see [10], Theorem 5.5). Its second part is included in Theorem 2.

Let us consider the leader relaxation:

$$(\overline{P}) \quad \max\{ax + by : (x, y) \in Z\}$$

From now on, instead of [A1] and [A2] we assume that:
[A] $W \neq \emptyset$ and $(\overline{P})$ has an optimal solution.

THEOREM 6. *For a fixed $K \in R_+$, there exists a solution to $P(K)$ which is achievable in $(Z \times W)_v$.*

Proof. For the moment we replace 'max' by 'sup' in (8), (10) and (11). As $P(K)$ is also defined by (11), we have:

$$\begin{aligned}&\sup\{\Theta(w, K) : w \in W\}\\ &\quad \sup\{ax + by - K\pi(x, y, w) : (x, y, w) \in Z \times W\}\\ &\quad \leqslant \sup\{ax + by : (x, y, w) \in Z \times W\} = \max\{ax + by : (x, y) \in Z\}\,.\end{aligned}$$

Moreover $\Theta(\cdot, K)$ is convex and W is a non-empty polyhedron. Then the supremum is attained at some $w^* \in W_v$ (see [10], Corollary 32.3.4). Applying the same argumentation to $\hat{F}(\cdot, \cdot, w^*, K)$ and Z, a solution to (10) occurs at some $(x^*, y^*) \in Z_v$. And since $Z_v \times W_v = (Z \times W)_v$, the result follows.

THEOREM 7. *There exists a finite value $K^* \in R_+$ of K for which an extreme optimal solution to the penalty function problem $P(K)$ yields an extreme optimal solution to the problem (P), for all $K \geqslant K^*$.*

Proof. From Theorem 6 there is a solution to $P(K)$ in $(Z \times W)_v$ for each $K \in R_+$. Split $(Z \times W)_v$ into two finite sets $S^0 = \{(x, y, w) \in (Z \times W)_v : \pi(x, y, w) = 0\}$ and $S^1 = (Z \times W)_v \backslash S^0$. Since $W \neq \emptyset$ and $Z \neq \emptyset$, we have $S^0 \neq \emptyset$. Let $L = \max\{\hat{F}(x, y, w, K) : (x, y, w) \in S^0\}$. Suppose $S^1 \neq \emptyset$ and take

$$K^* > \max\{(ax + by - L)/\pi(x, y, w) : (x, y, w) \in S^1\}\,. \tag{12}$$

Therefore, $ax + by - K^*\pi(x, y, w) < L$, for all $(x, y, w) \in S^1$. Thus, for $K \geqslant K^*$,

$$\begin{aligned}&\max\{\hat{F}(x, y, w, K) : (x, y, w) \in (Z \times W)_v\}\\ &\quad = \max\{L, \max\{\hat{F}(x, y, w, K) : (x, y, w) \in S^1\}\} = L\,.\end{aligned} \tag{13}$$

Moreover, from the strict inequality in (12), any solution to (13), for $K \geqslant K^*$, must belong to S^0. Therefore, it is a solution to (P) (see [2], Corollary of Theorem 9.2.2). The same conclusions are trivially obtained for $K^* \geqslant 0$ when $S^1 = \emptyset$.

Finally, Theorem 4 holds as a consequence of Theorem 6. It is a known result of penalty methods (see [2], Lemma 9.2.1).

4. The algorithm

For given $K \in R_+$ and $w \in W$, let $(x(w, K), y(w, K))$ be a solution to (10). And for each $u \in W$ define:

$$\begin{aligned}\Phi(u, w, K) &= (u - w)(p - Ax(w, K)) \\ &= \pi(x(w, K), y(w, K), u) - \pi(x(w, K), y(w, K), w)\,.\end{aligned} \tag{14}$$

The algorithm in [14] is defined as follows:

Step 0:
Choose K (large) and $w^1 \in W_v$, $\Theta^1 = -\infty$, $\overline{w}^1 = w^1$.

Step 1:
Find $\Theta(w^1, K)$.
Obtain $(x(w^1, K), y(w^1, K))$ and set $\Theta^1 = \max\{\Theta^1, \Theta(w^1, K)\}$ and

$$\overline{w}^1 = \begin{cases} w^1 & \text{if } \Theta(w^1, K) > \Theta^1 \\ \overline{w}^1 & \text{if } \Theta(w^1, K) \leqslant \Theta^1 \end{cases}$$

Step 2:
Let $\{w^{1s}\}$ be the adjacent vertices of w^1, $1 \leqslant s \leqslant N(w^1)$.
If $\Theta(w^{1s}, K) > \Theta^1$ for some s, then
set $w^1 = w^{1s}$, $\Theta^1 = \Theta(w^1, K)$ and repeat Step 2.

Step 3:
If $\Theta(w^{1s}, K) \leqslant \Theta^1$, $\forall$s,
Find $\Gamma(w^1, K) = \min\{\Phi(w, w^1, K) : w \in W\}$.
Obtain $w^*(w^1, K)$.

Step 4:
If $\Gamma(w^1, K) < 0$ then set $w^1 = w^*(w^1, K)$, and
Go to Step 1.

Step 5:
If $\Gamma(w^1, K) \geqslant 0$
extend unit rays $\{t^{1s}\}$ along the edges from w^1, and find
$\alpha_s = \max\{\alpha \geqslant 0 : \Theta(w^1 + \alpha t^{1s}, K) \leqslant \Theta^1\}, 1 \leqslant s \leqslant N(w^1)$.

Step 6:
Let $\nu^{1s} = w^1 + \alpha_s t^{1s}$, $1 \leqslant s \leqslant N(w^1)$,
$\Lambda(w^1) = \{\lambda = (\mu, \sigma) \in R^{m+1} : \sigma \in \{1, -1\},\ \mu w^1 - \sigma \leqslant 0,$
$\mu \nu^{1s} - \sigma \geqslant 0,\ 1 \leqslant s \leqslant N(w^1)\}$
and, for $w \in W$:
$G(w, w^1) = \min\{\mu w - \sigma : \lambda \in \Lambda(w^1)\}$.

Step 7:
Let $w^{1*} \in \arg\max\{G(w, w^1) : w \in W\}$.

Step 8:
If $G(w^1*, w^1) \leqslant 0$ then
$\overline{w}^1 \in \arg\max\{\Theta(w, K) : w \in W\}$, and the optimal value of $\Theta(\cdot, K)$ is reached for the particular K, with the solution $(x(\overline{w}^1, K),\ y(\overline{w}^1, K))$.
Then Go to Step 10.

Step 9:
If $G(w^1*, w^1) > 0$, set $w^1 = w^1*$, and
Go to Step 1.

Step 10:
If $\pi(x(\overline{w}^1, y(\overline{w}^1), \overline{w}^1) > 0$, set $K = K + \Delta$, and Go to Step 1.
Otherwise $\pi(x(\overline{w}^1), y(\overline{w}^1), \overline{w}^1) = 0$ and $(x(\overline{w}^1), y(\overline{w}^1))$ solves $P(K)$.

The first four steps find a local optimum to $\Theta(\cdot, K)$ as described in [14]. Note that they are still well-defined if we consider [A] instead of [A2]. The minimum in Step 3 is always attained by (14) and since $\pi(x, y, w) \geqslant 0$ for all $(x, y, w) \in Z \times W$. In particular $\Gamma(w^1, K) = 0$ with $w*(w^1, K) = w^1$ when $\pi(x(w^1, K),\ y(w^1, K), w^1) = 0$.

Steps 5–9 are modifications of the original algorithm proposed by Tuy [11]. Steps 5 and 6 define cuts to discard the local optimum w^1. The points ν^{1s} are Θ^1-extensions of the vertices w^{1s} with respect to w^1 (Def. V.1 in [9]). Each steplength α_s along the ray t^{1s} $(1 \leqslant s \leqslant N(w^1))$ is equal to the optimal value of the following linear programming problem (see Prop. IX.3 in [9]):

$$\alpha_s = \min_{(r,u,v)} (\Theta^1 + Kw^1p)r - (a + Kw^1A)u - (b + Kd)v \tag{15}$$

$$\text{s.t.:}\ K(t^{1s}p)r - K(t^{1s}A)u = -1 \tag{16}$$

$$-pr + Au + Bv \leqslant 0 \tag{17}$$

$$r \geqslant 0, u \geqslant 0, v \geqslant 0 \tag{18}$$

where $u \in R^{n_1}$, $v \in R^{n_2}$, $r \in R$.

It is worth mentioning that $\alpha_s > 0$ for all $s \in [1, N(w^1)]$ (see [5]). Furthermore α_s may be infinite for some s, where t^{1s} is a recession direction of W or not (see next section). In this case it is not possible to define some points ν^{1s}. Actually, such situation does not depend on the boundness of W or Z. To overcome this difficulty it is necessary to redefine the set $\Lambda(w^1)$, for example as follows:

Step 5′:
If $\Gamma(w^1, K) \geqslant 0$
extend unit rays $\{t^{1s}\}$ along the edges from w^1, and find
$\alpha_s = \sup\{\alpha \geqslant 0 : \Theta(w^1 + \alpha t^{1s}, K) \leqslant \Theta^1\}$, $1 \leqslant s \leqslant N(w^1)$.

Step 6′:
Let $I = \{s : 1 \leqslant s \leqslant N(w^1),\ \alpha_s < +\infty\}$,
$\nu^{1s} = w^1 + \alpha_s t^{1s},\ \forall s \in I$,
$\Lambda(w^1) = \{\lambda = (\mu, \sigma) \in R^{m+1} : \sigma \in \{1, -1\},\ \mu w^1 - \sigma \leqslant 0,$
$\mu\nu^{1s} - \sigma \geqslant 0,\ s \in I,\ \mu t^{1s} \geqslant 0,\ s \not\in I\}$
and, for $w \in W$:
$G(w, w^1) = \min\{\mu w - \sigma : \lambda \in \Lambda(w^1)\}$.

The following proposition shows that the set $\Lambda(w^1)$ is well-defined.

PROPOSITION 1. *The set* $\Lambda(w^1)$ *is non-empty. In particular, there exists* $(\mu, \sigma) \in \Lambda(w^1)$ *with* $\mu w^1 - \sigma < 0$.

Proof. Let $U = \{w = \Sigma_{s\in I}\ \rho_s\alpha_s t^{1s} + \Sigma_{s\not\in I}\ \rho_s t^{1s} : \Sigma_s\ \rho_s \geqslant 1,\ \ \rho_s \geqslant 0\ \ \forall s\}$. We have $0 \not\in U$, since $\alpha_s > 0$ for all $s \in I$, $\rho_s > 0$ for some s, and the convex cone generated by $\{t^{1s}\}$ contains no lines. Consider the closed convex set $V = w^1 + U$. Therefore, $w^1 \not\in V$. Thus, there exists a hiperplane defined by (μ, σ), with $\sigma \neq 0$, such that $\mu w^1 < \sigma$ and $\mu w \geqslant \sigma$ for all $w \in V$. For each $s \in I$, $\nu^{1s} = w^1 + \alpha_s t^{1s} \in V$, and so $\mu\nu^{1s} \geqslant \sigma$. Now let $s \not\in I$. Since $w^1 + t^{1s} \in V$, it is $\mu(w^1 + t^{1s}) \geqslant \sigma$. Then it must be $\mu t^{1s} > 0$. Hence, $(\mu, \sigma) \in \Lambda(w^1)$.

The set $\Lambda(w^1)$ establishes cuts as shown by the next proposition, which justifies the definition of $G(w, w^1)$ in Step 6.

PROPOSITION 2. *Let* $w \in W$ *and* $(\mu, \sigma) \in \Lambda(w^1)$. *If* $\mu w - \sigma < 0$ *then* $\Theta(w, K) \leqslant \Theta^1$.

Proof. Let $w \in W$. Then there are $\rho_s \geqslant 0$ such that:

$$w = w^1 + \sum_s \rho_s t^{1s} = \left(1 - \sum_{s\in I} \frac{\rho_s}{\alpha_s}\right) w^1 + \sum_{s\in I} \frac{\rho_s}{\alpha_s}\, \nu^{1s} + \sum_{s\not\in I} \rho_s t^{1s}\ .$$

And the stated condition to $(\mu, \sigma) \in \Lambda(w^1)$ gives:

$$\begin{aligned} 0 > \mu w - \sigma &= \left(1 - \sum_{s\in I} \frac{\rho_s}{\alpha_s}\right)(\mu w^1 - \sigma) + \sum_{s\in I} \frac{\rho_s}{\alpha_s}\,(\mu\nu^{1s} - \sigma) + \sum_{s\not\in I} \rho_s \mu t^{1s} \\ &\geqslant \left(1 - \sum_{s\in I} \frac{\rho_s}{\alpha_s}\right)(\mu w^1 - \sigma) \end{aligned}$$

Therefore, $0 \leqslant \Sigma_{s\in I}(\rho_s/\alpha_s) < 1$. Define the convex set $D = \{w \in W : \Theta(w, K) \leqslant \Theta^1\}$. Since $w^1 \in D$, $\nu^{1s} \in D$ for $s \in I$, and t^{1s} are recession directions of D for $s \not\in I$, it must be $w \in D$.

Steps 7–9 evaluate whether a solution to $\Theta(\cdot, K)$ is already reached for the fixed K. However some trouble may happen to Step 7 when W is unbounded or in Step 8 if the equality is verified.

In fact the maximization problem in Step 7 can be formulated in terms of the vertex set $\{(\mu^j, \sigma^j) : 1 \leqslant j \leqslant T(w^1)\}$ of $\Lambda(w^1)$. Let M be the matrix with columns

$\{\mu^j\}$ and q be the row-vector with components $\{\sigma^j\}$. Then the maximization problem is equivalent to (see [14]):

$$\tau = \min_{(\beta,\gamma)} -q\beta - d\gamma \tag{19}$$

$$\text{s.t.}: e\beta = 1 \tag{20}$$

$$M\beta + B\gamma \leqslant 0 \tag{21}$$

$$\beta \geqslant 0,\ \ \gamma \geqslant 0 \tag{22}$$

where $\beta \in R^{T(w^1)}$, $\gamma \in R^{n_2}$ and $e = (1, 1, \ldots, 1) \in R^{T(w^1)}$. This formulation can be solved by the column generation method.

When W is bounded $\tau = G(w^{1*}, w^1)$ for some $w^{1*} \in W$. If $\tau < 0$ a solution to $P(K)$ is already reached for the fixed K by Proposition 2. But if $\tau = 0$ it may exist a better solution yet (see Example 1). In this case Step 8 will return a wrong result. Thus, the test $G(w^{1*}, w^1) \leqslant 0$ must be replaced by the strict inequality. In the case that $\tau \geqslant 0$, we always have $w^{1*} \neq w^1$, since $G(w^1, w^1) < 0$ by Proposition 1. Therefore, the algorithm can continue.

If W is unbounded it may be $\tau = +\infty$ (see Example 2). In this case Step 7 fails. A new vertex w^{1*} is not available and the algorithm cannot progress. Therefore, in order to have the algorithm well-defined, the following assumption should be added:

[B] W is compact.
We note that Assumptions [A] and [B] can hold at the same time.

5. Examples

We introduce two examples which illustrate the trouble identified in the algorithm. In the first one W is bounded. It presents the case where the original Step 6 is not well-defined, because there is $\alpha_s = +\infty$ for some s, and the original Step 8 will return a wrong solution to $P(K)$, due to $\tau = 0$. In the second example W is unbounded. We will have again $\alpha_s = +\infty$ for some s, and Step 7 will fail since $\tau = +\infty$.

We now apply the algorithm to these examples. For the calculus details we refer to [5].

EXAMPLE 1

$$(E1)\ \ \max_{(x,y)} -0.4x - 6y_1 - 5y_2$$

$$\text{s.t.}: x \geqslant 0,\ \ y = (y_1, y_2, y_3, y_4) \text{ solves:}$$

$$\max_y 0.5y_2 - y_3 - 2y_4$$

$$\text{s.t.}: -0.1x - y_1 - y_2 \leqslant -1$$

$$0.2x + 1.25y_2 - y_4 \leqslant -1$$

$$-x + 6y_1 + y_2 - 2y_3 \leqslant 1$$
$$y_1, y_2, y_3, y_4 \geqslant 0$$

The follower's dual feasible set and its set of vertices are:

$$W = \{w = (w_1, w_2, w_3) \geqslant 0 : -w_1 + 6w_3 \geqslant 0\,,$$
$$-w_1 + 1.25w_2 + w_3 \geqslant 0.5, -2w_3 \geqslant -1, -w_2 \geqslant -2\}$$
$$W_v = \{2.5, 2, 0.5), (2.4, 2, 0.4), (0, 2, 0.5), (0, 0, 0.5), (0, 2, 0), (0, 0.4, 0)\}.$$

Suppose we start the algorithm with $w^1 = (2.5, 2, 0.5)$ and $K = 10$. At Step 1 we find $\Theta(w^1, K) = -5$ with $(x(w^1, K), \; y(w^1, K)) = (0, 0, 1, 0, 2.25)$. The adjacent vertices of w^1, evaluated at Step 2, are $w^{11} = (2.4, 2, 0.4)$, $w^{12} = (0, 2, 0.5)$ and $w^{13} = (0, 0, 0.5)$. As we have $\Theta(w^{1s}, K) \leqslant -5$, $1 \leqslant s \leqslant 3$, we go to Step 3. We obtain $\Gamma(w^1, K) = 0$, since $\pi(x(w^1), y(w^1), w^1)) = 0$.

At Step 5 we find the rays $t^{11} = (-0.1, 0, -0.1)$, $t^{12} = (-2.5, 0. 0)$ and $t^{13} = (-2.5, -2, 0)$. The steplengths along these rays, calculated by (15)–(18), are respectively $\alpha_1 = 61/55$, $\alpha_2 = 3.16$ and $\alpha_3 = +\infty$. Note that we have an infinite steplength, in spite of W be compact.

The set of cuts is defined according to the new Step 6 as:

$$\Lambda(w^1) = \{(\mu_1, \mu_2, \mu_3, \sigma) : 2.5\mu_1 + 2\mu_2 + 0.5\mu_3 \leqslant \sigma\,,$$
$$-5.4\mu_1 + 2\mu_2 + 0.5\mu_3 \geqslant \sigma, 13.14\mu_1 + 11\mu_2 + 2.14\mu_3 \geqslant 5.5\sigma\,,$$
$$-2.5\mu_1 - 2\mu_2 \geqslant 0, \sigma \in \{1, -1\}\}\,.$$

Take $(\mu^1, \sigma^1) = (0, -0.5, 0, -1)$ and $(\mu^2, \sigma^2) = (-2, 2.5, 2, 1)$, both belonging to $\Lambda(w^1)$. Let $w^2 = (0, 2, 0)$ and $w^3 = (0, 0.4, 0)$. Then for each vertex of W we get:

$$\mu^1 w^1 = \mu^1 w^{11} = \mu^1 w^{12} = \mu^1 w^2 = \sigma^1$$
$$\mu^2 w^{13} = \mu^2 w^3 = \sigma^2$$

Therefore, $G(w, w^1) \leqslant 0$, for all $w \in W$. Thus, $\tau \leqslant 0$ and the algorithm will conclude at Step 8 that w^1 is a solution to (11). However, this answer is not correct, since $\Theta(w^2, K) = -4$.

EXAMPLE 2. (Adapted from [4])

$$(E2) \quad \max_{(x,y)} x + y_1 - 4y_2$$
$$\text{s.t.} : x \geqslant 0, \;\; y = (y_1, y_2) \text{ solves:}$$
$$\max_y \; y_2$$
$$\text{s.t.} : x + y_1 + y_2 \leqslant 3$$
$$-x - y_1 + y_2 \leqslant -1$$
$$-x + y_1 + y_2 \leqslant 1$$
$$x - y_1 + y_2 \leqslant 1$$

$$y_2 \leqslant 1/2$$
$$y_1 \geqslant 0,\ \ y_2 \geqslant 0$$

The follower's dual feasible set is:

$$W = \{(w_1, w_2, w_3, w_4, w_5) \geqslant 0: \\ w_1 - w_2 + w_3 - w_4 \geqslant 0, w_1 + w_2 + w_3 + w_4 + w_5 \geqslant 1\} .$$

We have the primal and dual set of vertices:

$$Z_v = \{(0, 1, 0), (0.5, 1, 0.5), (1, 0.5, 0.5), (1, 1.5, 0.5), \\ (1.5, 1, 0.5), (2, 1, 0), (1, 2, 0), (1, 0, 0)\}$$
$$W_v = \{(0.5, 0.5, 0, 0, 0), (0.5, 0, 0, 0.5, 0), (0, 0.5, 0.5, 0, 0), \\ (0, 0, 0.5, 0.5, 0), (0, 0, 1, 0, 0), (0, 0, 0, 0, 1), (1, 0, 0, 0, 0)\} .$$

We now start the algorithm with $w^1 = (0, 0.5, 0.5, 0, 0)$ and K large enough. At Step 1 we find $\Theta((w^1, K) = 1$ at $(x(w^1, K), y(w^1, K)) = (0, 1, 0)$. The adjacent vertices of w^1 are $w^{11} = (0.5, 0.5, 0, 0, 0)$, $w^{12} = (0, 0, 0.5, 0.5, 0)$, $w^{13} = (0, 0, 0, 0, 1)$ and $w^{14} = (0, 0, 1, 0, 0)$. We have $\Theta(w^{1s}, K) < 1$, $1 \leqslant s \leqslant 4$. Then we go to Step 3 to get $\Gamma(w^1, K) = 0$, since $\pi(x(w^1), y(w^1), w^1)) = 0$.

We obtain the rays $t^{11} = (1, 0, -1, 0, 0)$, $t^{12} = (0, -1, 0, 1, 0)$, $t^{13} = (0, -1, -1, 0, 2)$, $t^{14} = (0, -1, 1, 0, 0)$, towards the vertices w^{11}, w^{12}, w^{13}, w^{14}, and the recession ray $t^{15} = (0, 1, 1, 0, 0)$. The steplengths, calculated with $K = 10$, are respectively $\alpha_1 = \alpha_2 = 0.9$, $\alpha_3 = 0.525$ and $\alpha_4 = \alpha_5 = +\infty$.

Again we have an infinite steplength. Consequently, we apply the modified Step 6 to define:

$$\Lambda(w^1) = \{(\mu_1, \mu_2, \mu_3, \mu_4, \mu_5, \sigma) : \sigma \in \{1, -1\}, \mu_2 + \mu_3 \leqslant 2\sigma , \\ 9\mu_1 + 5\mu_2 - 4\mu_3 \geqslant 10\sigma, -4\mu_2 + 5\mu_3 + 9\mu_4 \geqslant 10\sigma , \\ -\mu_2 - \mu_3 + 42\mu_5 \geqslant 40\sigma, -\mu_2 + \mu_3 \geqslant 0, \mu_2 + \mu_3 \geqslant 0\} .$$

The maximization problem at Step 7 can be solved by (19)–(22). If we apply the first phase of the simplex method, we get an infeasibility. Therefore, $\tau = +\infty$, and the algorithm cannot progress, since a new vertex w^{1*} is not available.

6. Conclusions

We presented modifications to the paper given by White and Anandalingam [14] in such a way that it preserves its general philosophy but overcomes its difficulties. The desired theoretical properties of the penalty function approach were obtained under Assumption [A]. It substitutes the nonvalid hypothesis [A2] and avoids hypothesis [A1]. The new assumption is used in the literature [4, 7, 8, 12]. The procedure to find an optimal solution to $P(K)$ was redefined according to Tuy [9]. Assumption [B] was considered to guarantee its convergence. We would note that assumptions [A] and [B] could be replaced by the single hypothesis that Z is nonempty and compact. In fact, instead of function $\Theta(\ )$, the algorithm could use function $\Phi(x, y, K) = \max\{\hat{F}(x, y, w, K) : w \in W\}$, defined for all $(x, y) \in Z$ and $K \in R_+$.

References

1. Anandalingam, G. and White, D.J. (1990), A Solution Method for the Linear Static Stackelberg Problem using Penalty Functions, *IEEE Transactions on Automatic Control* 35(10): 1170–1173.
2. Bazaraa, M.S., Sherali, H.D. and Shetty, C.M. (1979), *Nonlinear Programming: Theory and Algorithms*, Wiley, New York.
3. Ben-Ayed, O. (1993), Bilevel Linear Programming, *Computers and Operations Research* 20: 485–501.
4. Bialas, W.F. and Karwan, M.H. (1984), Two-Level Linear Programming, *Management Science* 30(8): 1004–1020.
5. Campêlo, M., Dantas, S. and Scheimberg, S. (1998), A Note on a Penalty Function Approach for Solving Bi-Level Linear Programs, *Tech. Report ES-467/98*, Engenharia de Sistemas e Computação, Universidade Federal do Rio de Janeiro.
6. Charnes, A., Cooper, W.W. and Thompson, G.L. (1962), Some Properties of Redundant Constraints and Extraneous Variables in Direct and Dual Linear Programming Problems, *Operations Reserach* 10(5): 711–723.
7. Dempe, S. (1987), a Simple Algorithm for the Linear Bilevel Programming Problem, *Optimization* 18: 373–385.
8. Hansen, P., Jaumard, B. and Savard, G. (1992), New Branch-and-Bound Rules for Linear Bilevel Programming, *SIAM J. on Scientific and Statistical Computing* 13: 1194–1217.
9. Horst, R. and Tuy, H. (1990), *Global Optimization: Deterministic Approaches*, Springer-Verlag, Berlin.
10. Rockafeller, R.T. (1970), *Convex Analysis*, Princeton University Press, Princeton, NJ.
11. Tuy, H. (1964), Concave Programming under Linear Constraints, *Doklady Academii Nauk SSSR* 158: 32–35 [English version: *Soviet Mathematics* 5: 1437–1440].
12. Tuy, H., Migdalas, A. and Värbrand, P. (1993), A Global Optimization Approach for the Linear Two-Level Program, *Journal of Global Optimization* 3: 1–23.
13. Vicente, L.N. and Calamai, P.H. (1994), Bilevel and Multilevel Programming: a Bibliography Review, *Journal of Global Optimization* 5(3): 291–306.
14. White, D.J. and Anandalingam, G. (1993), A Penalty Function Approach for Solving Bi-Level Linear Programs, *Journal of Global Optimization* 3: 397–419.