JOURNAL OF OPTIMIZATION THEORY AND APPLICATIONS: Vol. 73, No. 3, JUNE 1992

Asymptotic Properties of the Fenchel Dual Functional and Applications to Decomposition Problems

A. AUSLENDER[1]

Communicated by O. L. Mangasarian

Abstract. We study dual functionals which have two fundamental properties. Firstly, they have a good asymptotical behavior. Secondly, to each dual sequence of subgradients converging to zero, one can associate a primal sequence which converges to an optimal solution of the primal problem. Furthermore, minimal conditions for the convergence of the Gauss-Seidel methods are given and applied to such kinds of functionals.

Key Words. Convex analysis, duality, Gauss-Seidel methods, decomposition methods.

1. Introduction

Recently, three particular extremum problems have been investigated by Tseng and Bertsekas (Ref. 1), Bertsekas, Hosein, and Tseng (Ref. 2), Han (Ref. 3), Han and Lou (Ref. 4), and Censor and Lent (Ref. 5). In these papers, to solve the primal problem, the authors consider the dual problem for which they propose a parallel decomposition algorithm that leads to an attractive algorithm. In Refs. 1, 2, and 5, the proposed method is the Gauss-Seidel method with either exact (Ref. 5) or inexact (Refs. 1 and 2) minimization along the coordinates. Gauss-Seidel methods for unconstrained optimization have been studied extensively (Refs. 6-10). Typical conditions for convergence are the strict convexity of the objective function or the boundedness of its level sets. Unfortunately, in Refs. 1-5, the dual level sets are unbounded and it is one of the merits of the authors in Refs. 1-2 and 5 to overcome this fact and to prove convergence by using the particular structure of their problems.

[1]Professeur, Département de Mathématiques Appliquées, Université Blaise Pascal (Clermont II), Aubière, France.

In fact, in all these papers, the dual functional has two fundamental properties, which do not appear clearly in the previous papers. Firstly, it is asymptotically well behaved; secondly, each dual sequence of subgradients converging to zero induces a primal sequence converging to an optimal solution. In this paper, we clarify these notions and show that proving primal convergence of various methods reduces, for these problems, to proving convergence to zero of dual sequences of subgradients generated by the method.

The notion of convex function which is asymptotically well behaved was recently introduced by Auslender and Crouzeix in Ref. 11. We recall this notion in Section 2 and give a characterization of such functions required in other sections. In the same section, we introduce the general problem and its dual and recall the main properties resulting from Fenchel's duality theory.

In Section 3, we prove that the dual functional of the former problems possesses the two fundamental properties. As a consequence, we prove that Han and Lou's algorithm converges under very general assumptions, thus improving on their convergence results.

Finally, in Section 4, we propose a general variant of the Gauss–Seidel method whose convergence is obtained without the usual assumptions of strict convexity and boundedness of the level sets, the latter assumption being replaced by the asymptotical behavior condition. Specialized to the particular problems studied in Refs. 1–5, we obtain an implementable parallel algorithm close to those proposed therein. In the sequel, we assume that the reader is familiar with the theory of convex analysis. All notations and classical definitions follow from Rockafellar (Ref. 12).

2. Preliminaries

2.1. Asymptotically Well-Behaved Convex Functions. Let us consider a closed proper convex function defined on $\mathbb{R}^N$; denote by $\partial f(x)$ the subdifferential of f at x and by $d(y|S)$ the distance from the point y to the set S. We say that a sequence $\{x_k\}$ is stationary for f if it satisfies

$$\lim_{k\to\infty} d(0|\partial f(x_k))=0, \tag{1}$$

that is, if we can find $x_k^* \in \partial f(x_k)$ such that $x_k^* \to 0$.

Many algorithms for the minimization of f generate such a sequence, and it is natural to ask oneself whether such a sequence is minimizing or not. The answer is certainly yes if the function is inf-compact; i.e., for each λ, the level set $\{x: f(x) \leq \lambda\}$ is bounded, a condition not always satisfied (see Ref. 13).

Definition 2.1. A closed proper convex function f on $\mathbb{R}^N$ is asymptotically well-behaved if every stationary sequence $\{x_n\}$ for f is minimizing, that is,

$$\lim_{n\to\infty} f(x_n) = m = \inf(f(x) \mid x \in \mathbb{R}^N). \tag{2}$$

Denote by $\mathscr{F}$ this class of functions. Several characterizations of $\mathscr{F}$ and examples of such functionals are given in Ref. 13. Let us now state a new characterization of $\mathscr{F}$ which will be useful in the following.

We say that a sequence $\{x_k\}$ is strongly stationary for f if it is a stationary sequence and if the sequence $\{f(x_n)\}$ is bounded above, that is, there exist M such that $f(x_n) \leq M$ for each n. We denote by $\mathscr{F}_1$ the class of closed proper convex functions f on $\mathbb{R}^N$ for which each strongly stationary sequence for f is minimizing. Then, we have the following proposition.

Proposition 2.1. Let ri dom f denote the relative interior of the domain of f (dom f). If ri dom f = dom f, then $\mathscr{F} = \mathscr{F}_1$.

Proof. Obviously, $\mathscr{F} \subset \mathscr{F}_1$. Assume that $\mathscr{F}_1 \not\subset \mathscr{F}$. Then, by Theorem 2.2 in Ref. 13, there exist $f \in \mathscr{F}_1$, $\lambda > m$ such that $r(\lambda) = 0$, with

$$r(\lambda) = \inf(\|c\| \mid c \in \partial f(x), f(x) = \lambda).$$

This implies the existence of sequences $\{x_k\}$ and $\{x_k^*\}$ such that

$$x_k^* \in \partial f(x_k), \qquad f(x_k) = \lambda, \qquad \lim_{k\to\infty} x_k^* = 0,$$

a contradiction with $f \in \mathscr{F}_1$. □

2.2. General Problem and Its Dual. Let q be a closed proper convex function from $\mathbb{R}^n$ to $]-\infty, +\infty]$, and let C_i, $i = 1, 2, \ldots, m$, be closed convex sets in $\mathbb{R}^n$. Denote by $\delta(\cdot \mid C)$ the indicator function of a set C,

$$\delta(x \mid C) = \begin{cases} 0, & \text{if } x \in C, \\ +\infty, & \text{otherwise.} \end{cases}$$

Its conjugate $\delta^*(\cdot \mid C)$ is the support function of the set and is given by

$$\delta^*(y \mid C) = \sup\{(x, y) \mid x \in C\}.$$

Consider the fundamental problem

$$\text{(P)} \quad \alpha = \inf(q(x) \mid x \in C), \quad \text{with } C = \bigcap_{i=1}^{m} C_i$$

under the following assumptions:

(H1) $\bigcap_{i=1}^{m}$ ri $C_1 \cap$ ri dom q is nonempty;

(H2) there exists $y_i \in$ ri dom $\delta^*(\cdot \mid C_i)$ such that $-\sum_{i=1}^{m} y_i \in$ dom q^*, $\forall i$.

If C_i is polyhedral, ri C_i and ri dom $\delta^*(\cdot \mid C_i)$ can be replaced by C_i and dom $\delta^*(\cdot \mid C_i)$, respectively.

To comply with Refs. 1-5, we rewrite problem (P) as

$$\text{(P)} \quad \alpha = \inf(q(x) + k(Ax)),$$

where

$$k(u_1, u_2, \ldots, u_m) = \sum_{i=1}^{m} \delta(u_i \mid C_i)$$

and A is the linear transformation from $\mathbb{R}^n$ to $\mathbb{R}^{n\times m}$, defined by

$$Ax = (x_1, x_2, \ldots, x_m), \qquad \text{with } x_i = x,\ i = 1, 2, \ldots, m.$$

We have

$$A^t y = A^t \cdot (y_1, y_2, \ldots, y_m) = \sum_{i=1}^{m} y_i,\ k^*(x_1^*, \ldots, x_m^*)$$

$$= \sum_{i=1}^{m} \delta^*(x_i^* \mid C_i), \tag{3}$$

and from Theorem 31.2 in Ref. 12, the Fenchel dual (D) of (P) is given by

$$\text{(D)} \quad \beta = \inf(g(y) \mid y = (y_1, y_2, \ldots, y_m) \in \mathbb{R}^{n\times m}),$$

$$g(y) = q^*\left(\sum_{i=1}^{m} y_i\right) + \sum_{i=1}^{m} \delta^*(-y_i \mid C_i).$$

Remark that

$$\text{ri } C_i = \text{ri dom } \delta(\cdot \mid C_i).$$

Then, under Assumptions (H1) and (H2), we can use Theorems 23.9, 6.7, and 23.8 in Ref. 12 to obtain

$$\partial\left(\sum_{i=1}^{m} \delta(\cdot \mid C_i)\right)(x) = \sum_{i=1}^{m} \partial\delta(x \mid C_i),$$

$$\partial_y q^*\left(\sum_{y=1}^{m} y_i\right) = \left(\partial q^*\left(\sum_{i=1}^{m} y_i\right)\right)_{i=1,\ldots,m}, \tag{4}$$

$$\partial\left(q + \sum_{i=1}^{m} \delta(\cdot \mid C_i)\right)(x) = \partial q(x) + \sum_{i=1}^{m} \partial\delta(x \mid C_i), \tag{5a}$$

$$\partial g(y) = \left(\partial q^*\left(\sum_{i=1}^{m} y_i\right) - \partial\delta^*(-y_i \mid C_i)\right)_{i=1,\ldots,m}. \tag{5b}$$

Now, denote by S_P the optimal set of solutions of (P) and by S_D the optimal set of solutions of (D). Then, S_P is the set of vectors x such that

$$0 \in \partial q(x) + \sum_{i=1}^{m} \partial\delta(x \mid C_i), \tag{6}$$

and S_D is the set of vectors $y = (y_1, \ldots, y_m)$ such that

$$0 \in \partial q^*\left(\sum_{i=1}^{m} y_i\right) - \partial\delta^*(-y_i \mid C_i), \qquad \forall i, \tag{7}$$

and we have

$$S_p \neq \varnothing, \qquad S_D \neq \varnothing, \qquad \alpha = -\beta, \tag{8a}$$

$$g(y) + q(x) \geq 0, \qquad \forall x \in C, \forall y \in \mathbb{R}^{n\times m}. \tag{8b}$$

Let us remark that Assumption (H1) does not imply that g is inf-compact. In order to obtain this property, it is easy to see that we must replace (H1) by the following stronger assumption:

(H1′) $\bigcap_{i=1}^{m} \operatorname{int} C_1 \cap \operatorname{int} \operatorname{dom} q$ is nonempty,

which is necessary and sufficient for g to be inf-compact.

3. Asymptotical Behavior of the Dual Functional

In this section, we prove that the dual functional of the extremal problems considered in Refs. 1–5 possesses the two fundamental properties claimed in the introduction.

3.1. The Han and Lou Problem. In Ref. 4, Han and Lou study Problem (P) with the following additional assumption:

(A1) q is strongly convex and differentiable on $\mathbb{R}^n$; that is, there exists $\rho > 0$ such that

$$q(\lambda x + (1-\lambda)y) \leq \lambda q(x) + (1-\lambda)q(y) - \lambda(1-\lambda)\rho\|x-y\|^2,$$

$$\forall x, y, \forall \lambda \;]0, 1[.$$

Under this assumption, it follows from Theorem 26.6 in Ref. 12 that q^* is differentiable on $\mathbb{R}^n$ strictly convex and cofinite. Then, (H2) is automatically satisfied and (H1) becomes:

(A2) $\bigcap_{i=1}^{m} \operatorname{ri} C_i$ is nonempty.

In the remainder of this subsection, we shall only assume (A1) and (A2), and we shall denote by P1, D1, S_{P1}, S_{D1}, α_1, β_1 the symbols P, D, S_p, S_D, α, β. Furthermore, from Assumption (A1), S_{P1} reduces to the singleton x^*.

Now, we begin with a preliminary result.

Proposition 3.1. Let

$$z \in \bigcap_{i=1}^{m} \text{ri } C_i.$$

Set

$$h_i = \delta^*(\cdot \mid C_i - z),$$

and let $h_i 0^+$ be the recession function of h_i. Denote for each i by L_i the constancy space of h_i,

$$L_i = \{y: -(h_i 0^+)(-y) = (h_i 0^+)(y) = 0\},$$

and let $L_i^\perp$ be the orthogonal complement of L_i. Set

$$k_i = h_i + \delta(\cdot \mid L_i^\perp).$$

Then:

(i) k_i is inf-compact; (9a)

(ii) $(y_i, x) = (y_i, z), \quad \forall x \in C_i, \forall y_i \in L_i.$ (9b)

Proof.

(i) Since

$$\text{ri}(C_i) = \text{ri dom } \delta(\cdot \mid C_i),$$

and since

$$z \in \text{ri } C_i,$$

we have

$$0 \in \text{ri dom } \delta(\cdot \mid C_i - z).$$

From Corollary 13.3.4 in Ref. 12, it follows that

$$(h_i 0)^+(y) > 0, \qquad \forall y \neq 0, y \in L_i^\perp. \tag{10}$$

We have now to prove that, for each sequence $\{y_n\}$ with $y_n \in L_i^\perp$ and $\lim_{n\to\infty} \|y_n\| = +\infty$, we have $h_i(y_n) \to +\infty$. If this were not true, there would exist $M > 0$ and a sequence $\{y_n\}$ such that

$$y_n \in L_i^\perp, \qquad \|y_n\| \to +\infty, \qquad \lim_{n\to\infty} y_n/\|y_n\| = y, \qquad h_i(y_n) \leq M.$$

Since

$$(y_n, M) \in \text{epi } h_i \quad \text{and} \quad (y, 0) = \lim_{n\to\infty} (y_n, M)/\|y_n\|,$$

it follows that $(y, 0) \in 0^+$ epi h_i; note that 0^+ epi h_i is the recession cone of epi h_i. Since 0^+ epi $h_i = \text{epi}(h_i 0^+)$, we would have

$$(h_i 0)^+(y) \leq 0.$$

Now, $y \in L_i^\perp$ and $y \neq 0$ contradicts (10).

(ii) From Theorem 8.8 in Ref. 12, we have, for each $y_i \in L_i$ and for each $x_i^* \in \text{dom } \delta^*(\cdot \mid C_i - z)$,

$$\delta^*(x_i^* + \lambda y_i \mid C_i - z) = \delta^*(x_i^* \mid C_i - z) = \alpha_i, \qquad \forall \lambda \in \mathbb{R},$$

so that, for each $x \in C_i$, we have

$$\lambda (y_i, x - z) \leq \alpha_i - (x_i^*, x - z), \qquad \forall \lambda,$$

which implies that

$$(y_i, x - z) = 0. \qquad \square$$

Theorem 3.1. Under Assumptions (A1) and (A2), g belongs to $\mathscr{F}_1$; i.e., for each sequence $\{y_k, c_k\}$ such that

$$c^k \in \partial g(y^k), \qquad g(y^k) \leq M, \qquad \lim_{k\to\infty} c^k = 0, \tag{11}$$

the sequence $\{y^k\}$ is a minimizing sequence. Furthermore, the sequence $\{x^k = \nabla q^*(\sum y_i^k)\}$ converges to the unique solution x^* of (P1).

Proof.

(i) Fix

$$z \in \bigcap_{i=1}^m \text{ri } C_i$$

and denote, for each $y_i \in \mathbb{R}^n$,

$$y_i = y_{i1} + y_{i2}, \qquad \text{with } y_{i1} \in L_i,\ y_{i2} \in L_i^\perp,$$

$$\bar{y} = \sum_{i=1}^m y_i, \qquad \bar{y}^1 = \sum_{i=1}^m y_{i1}, \qquad \bar{y}^2 = \sum_{i=1}^m y_{i2}.$$

Obviously,

$$\bar{y} = \bar{y}^1 + \bar{y}^2.$$

Let

$$c^k = (c_1^k, c_2^k, \ldots, c_m^k)$$

by definition, and from (11) we have

$$c_i^k = \nabla q^*(\bar{y}^k) - d_i^k, \qquad \lim_{k \to \infty} c_i^k = 0, \tag{12}$$

with

$$d_i^k \in \partial \delta^*(-y_i^k \mid C_i) \Leftrightarrow d_i^k \in C_i, \qquad -y_i^k \in \partial \delta(d_i^k \mid C_i). \tag{13}$$

(ii) We first show that the sequences $\{\bar{y}^k\}$ and $\{x^k\}$ are bounded. From (11), there follows that

$$q^*(\bar{y}^k) + \delta^*(-\bar{y}^k \mid C) \leq q^*(\bar{y}^k) + \sum_{i=1}^{m} \delta^*(-y_i^k \mid C_i) \leq M.$$

Now, let

$$x \in \bigcap_{i=1}^{m} C_i.$$

We have then

$$q^*(\bar{y}^k) - (x, \bar{y}^k) \leq M. \tag{14}$$

Since q^* is cofinite, it follows from Lemma 26.7 in Ref. 12 that $q^*(\cdot) - (x, \cdot)$ is also cofinite and thus is inf-compact. From (14), it follows that $\{\bar{y}^k\}$ is bounded. Since ∇q^* is continuous, this implies that the sequence $\{x^k = \nabla q^*(\bar{y}^k)\}$ is also bounded.

(iii) Now, let us prove that, for each i, the sequence $\{y_{i2}^k\}$ is bounded. From (11), we have

$$q^*(\bar{y}^k) + \sum_{i=1}^{m} \delta^*(-y_i^k \mid C_i - z) - (\bar{y}^k, z) \leq M.$$

Since the sequence $\{\bar{y}^k\}$ is bounded, it follows that there exists $L > 0$ such that

$$\sum_{i=1}^{m} \delta^*(-y_i^k \mid C_i - z) \leq L, \qquad \forall k.$$

Since $z \in C_i$, we have

$$\delta^*(-y_i^k \mid C_i - z) \geq 0, \qquad \text{for each } i,$$

and we obtain then

$$\delta^*(-y_i^k \mid C_i - z) \leq L, \qquad \forall k.$$

But, from Theorem 8.8 in Ref. 12, we have

$$\delta^*(-y_i^k \mid C_i - z) = \delta^*(-y_{i2}^k \mid C_i - z),$$

so that

$$\delta^*(-y_{i2}^k \mid C_i - z) \leq L.$$

Using Proposition 3.1, we conclude that the sequence $\{y_{i2}^k\}$ is bounded.

(iv) From the definition of the conjugate, we have

$$\alpha_k = q(x^k) + g(y^k) = q(x^k) + q^*(\bar{y}^k) + \sum_{i=1}^m \delta^*(-y_i^k \mid C_i)$$

$$= q(x^k) + [-q(x^k) + (\bar{y}^k, x^k)] + \sum_{i=1}^m (-y_i^k, d_i^k)$$

$$= (\bar{y}^k, x^k) + \sum_{i=1}^m (-y_i^k, d_i^k).$$

Let us show that $\{\alpha_k\}$ goes to zero. Without loss of generality, since the sequences $\{x^k\}$, $\{\bar{y}^k\}$, $\{y_{i2}^k\}$ are bounded, we can suppose that they converge to $\tilde{x}$, $\tilde{y}$, $\tilde{y}_{i2}$, respectively. It follows that the sequence $\{\sum_{i=1}^m y_{i1}^k\}$ converges to $\tilde{y} - \sum_{i=1}^m \tilde{y}_{i2}$. From (9) and since $d_i^k \in C_i$, we have

$$(-y_i^k, d_i^k) = -(y_{i2}^k, d_i^k) - (y_{i1}^k, z), \tag{15}$$

and summing over i we obtain

$$\beta_k = \sum_{i=1}^m (-y_i^k, d_i^k) = -\sum_{i=1}^m (y_{i2}^k, d_i^k) - \left(\sum_{i=1}^m y_{i1}^k, z\right).$$

Since $\{d_i^k\}$ converges to $\tilde{x}$, passing to the limit we obtain

$$\lim_{k\to\infty} \beta_k = -(\sum \tilde{y}_{i2}, \tilde{x}) - (\tilde{y} - \sum \tilde{y}_{i2}, z). \tag{16}$$

On the other hand, the sequence $\{\bar{y}^k, x^k\}$ converges to $(\tilde{y}, \tilde{x})$. This implies that

$$\lim_{k\to\infty} \alpha_k = (\tilde{y}, \tilde{x}) - (\sum \tilde{y}_{i2}, \tilde{x}) - (\tilde{y} - \sum \tilde{y}_{i2}, z) = (\tilde{y} - \sum \tilde{y}_{i2}, \tilde{x} - z).$$

Now, since $d_i^k \in C_i$, $\forall i$, it follows that $\tilde{x} \in C$, and from (9) we have

$$(y_{i1}^k, \tilde{x} - z) = 0, \qquad \forall i,$$

so that

$$\left(\sum_{i=1}^m y_{i1}^k, \tilde{x} - z\right) = 0,$$

which implies that

$$\lim_{k\to\infty} \alpha_k = 0.$$

We then have

$$\lim_{k\to\infty} g(y^k) = -q(\tilde{x}), \qquad \tilde{x} \in C.$$

Together with (8) this implies that $\tilde{x} = x^*$ and that $\{y^k\}$ is a minimizing sequence for the dual (D1). □

As a consequence of Theorem 3.1, we prove now that the Han and Lou algorithm converges under Assumptions (A1) and (A2). Since we have taken the same notations as those of Han and Lou, we can state the forthcoming theorem.

Theorem 3.2. Let $\{x^k\}$ and $\{y^k\}$ be the sequences defined in Ref. 4. Assume that (A1) and (A2) are satisfied. Then, the sequence $\{x^k\}$ converges to the optimal solution of (P1) and the sequence $\{y^k\}$ is a minimizing sequence for the dual (D1).

Proof. Following Corollary 4.4 and formula (4.2) in Ref. 4, we have

$$-y_i^k \in \partial\delta(x^k + w_i^k \mid C_i), \qquad \lim_{k\to\infty} w_i^k = 0.$$

Then, setting

$$d_i^k = x^k + w_i^k,$$

and since

$$x^k = \nabla q^*\left(\sum_{i=1}^{m} y_i^k\right),$$

we obtain

$$-w^k \in \partial g(y_k), \qquad \lim_{k\to\infty} w^k = 0.$$

Furthermore, from Lemma 4.5 in Ref. 4, there follows that

$$g(y^k) \leq g(y^0).$$

The statement of the theorem is now a direct consequence of Theorem 3.1. □

Remark 3.1. Han and Lou have only proved convergence of the sequence $\{x^k\}$ in the polyhedral case, and when int $C \neq \varnothing$. In this case, they proved also convergence of the sequence $\{y^k\}$. However, in this case, g is inf-compact and this assumption is very restrictive. Indeed, if one of the C_i is a hyperplane, int C_i is empty.

Remark 3.2. Recent contributions by Tseng (Refs. 14 and 15) concerning Han and Lou's result were brought to the author's attention by a referee. Using a radically different way of reasoning, Tseng showed convergence under the weaker assumption

$$\bigcap_{i=1}^{m} \text{ri } C_i \neq \emptyset.$$

Tseng (Corollary 3 in Ref. 14) also sharpened the result of Theorem 3.2, obtaining convergence of the sequence $\{y^k\}$ to an optimal dual solution. Indeed, Theorem 3.2 of this paper, a direct consequence of Theorem 3.1, ignores the particular structure of the algorithm of Han and Lou. Rather, it is a result about the structure of the dual functional.

We close this section with a result that will be used in Section 4.

Proposition 3.2. Under Assumption (A1), the function

$$y = (y_1, \ldots, y_m) \to \nabla q^*\left(\sum_{i=1}^{m} y_i\right)$$

is Lipschitz on $\mathbb{R}^{n \times m}$.

Proof. Set

$$x = \nabla q^*(\textstyle\sum y_i), \qquad x' = \nabla q^*(\textstyle\sum y_i').$$

Then,

$$\sum_{i=1}^{m} y_i = \nabla q(x), \qquad \sum_{i=1}^{m} y_i' = \nabla q(x'),$$

and since ∇q is strongly monotone, we have

$$\left\langle \sum_{i=1}^{m} y_i - \sum_{i=1}^{m} y_i', x - x' \right\rangle \geq 2\rho \|x - x'\|^2,$$

so that

$$\|x - x'\| \leq (1/2\rho) \sum_{i=1}^{m} \|y_i - y_i'\|. \qquad \square$$

3.2. The Bertsekas, Hosein and Tseng Problem (Ref. 2). In Ref. 1 and in Section 3 of Ref. 2, the authors study the following problem:

(P2) $\alpha_2 = \min(q(x) \mid x \in C)$,

where

$$C = \{x \in \mathbb{R}^n : Ax = b\}, \qquad q(x) = \sum_{j=1}^{n} f_j(x_j),$$

A is an (m, n) real matrix, and $b \in \mathbb{R}^m$. They assumed that:

(B1) each function f_j is a closed proper cofinite and essentially strictly convex function;

(B2) $C \cap \operatorname{ri} \operatorname{dom} q$ is nonempty.

Set

$$g_j = f_j^*, \qquad C_i = \{x: a_i^t \cdot x = b_i\},$$

where a_i is the ith row of A. We have

$$\delta^*(x_i^* \mid C_i) = \begin{cases} p_i b_i, & \text{where } p_i \text{ satisfies } x_i^* = p_i a_i, \\ +\infty, & \text{otherwise,} \end{cases}$$

$$q^*\left(\sum_{i=1}^m x_i^*\right) = \sup_x \left(\sum_{i=1}^m \sum_{j=1}^n (x_i^*)_j \cdot x_j - f_j(x_j)\right)$$

$$= \sup_x \sum_{j=1}^m \left(\sum_{i=1}^n (x_i^*)_j x_j - f_j(x_j)\right) = \sum_{j=1}^n g_j\left(\sum_{i=1}^m (x_i^*)_j\right).$$

Performing the scaling $x_i^* = p_i a_i$, we see that the dual becomes

(D2) $\beta_2 = \inf(h(p) \mid p \in \mathbb{R}^m)$,

$$h(p) = \sum_{j=1}^n g_j(A_j^t \cdot p) - \sum_{i=1}^m p_i b_i,$$

where A_j denotes the jth column of A and $A_j^t \cdot p = (A_j, p)$.

It is easy to remark that h is differentiable and that

$$\nabla h(p) = Ax(p) - b,$$

where

$$x_j(p) = \arg\min_{v_j} f_j(v_j) - A_j^t \cdot p v_j.$$

Indeed, since f_j is essentially strictly convex and cofinite, g_j is differentiable everywhere, and we have

$$\partial g_j(t_j) = \{g'(t_j)\}.$$

Furthermore, we have

$$t_j \in \partial f_j(x_j) \Leftrightarrow x_j \in \partial g_j(t_j) \Leftrightarrow f_j(x_j) + g_j(t_j) = t_j x_j,$$

so that

$$x_j = g_j'(t_j), \text{ with } x_j = \arg\min_{v_j}(f_j(v_j) - t_j v_j).$$

Obviously, Assumptions (H1) and (H2) are implied by (B1) and (B2). From the results of Section 2.2, the solution sets S_{P2} and S_{D2} are nonempty, $-\alpha_2 = \beta_2$, and S_{P2} is reduced to a singleton x^*.

Before proving that h possesses the two fundamental properties, we recall some properties of the subgradients of f_j.

Let us denote by c_j and l_j the right and left endpoints of dom f_j (possibly $+\infty$ or $-\infty$), i.e.,

$$c_j = \sup\{x_j \,|\, f_j(x_j) < +\infty\},\ l_j = \inf\{x_j : f_j(x_j) < +\infty\}.$$

We make the assumption

(B3) $\quad l_j \neq c_j, \qquad \forall j.$

Then, under Assumption (B2), there exists an x such that

$$x \in C, \qquad l_j < x_j < c_j, \quad \forall j. \tag{17}$$

Let us denote by $f_j^-(x_j)$ and $f_j^+(x_j)$ the left and right derivatives of f_j at x_j. For $x_j \in]l_j, c_j[$, the derivatives f_j^+ and f_j^- are finite and $f_j^- \leq f_j^+$. These functions are increasing functions with f_j^+ right-continuous and f_j^- left-continuous. We define

$$f_j^+(l_j) = \lim_{\xi \to l_j^+} f_j^+(\xi), \quad f_j^-(l_j) = -\infty, \qquad \text{if } l_j > -\infty, \tag{18}$$

$$f_j^-(c_j) = \lim_{\xi \to c_j^-} f_j^-(\xi), \quad f_j^+(c_j) = +\infty, \qquad \text{if } c_j < +\infty. \tag{19}$$

For each $x_j \in]l_j, c_j[$, we have

$$u \in \partial f_j(x_j) \Leftrightarrow f_j^-(x_j) \leq u \leq f_j^+(x_j). \tag{20}$$

These inequalities can be extended to the endpoints l_j and c_j by using Theorem 25.6 in Ref. 12, and we have

$$u \in \partial f_j(l_j), \qquad \text{iff } u \leq f_j^+(l_j) \text{ when } f_j(l_j) \in \mathbb{R},$$

$$u \in f_j(c_j), \qquad \text{iff } u \geq f_j^-(c_j) \text{ when } f_j(c_j) \in \mathbb{R}.$$

Theorem 3.3. Assume that (B1), (B2), (B3) are satisfied. Then, $h \in \mathscr{F}$; and if the sequence $\{\nabla h(p^k)\}$ converges to zero, the sequence $\{x^k = x(p^k)\}$ converges to the optimal solution x^*.

Proof.

(i) Let us prove first that the sequence $\{x^k\}$ is bounded. Denote by A^{-1} the pseudoinverse of A, and set

$$d^k = Ax^k - b = \nabla h(p^k), \qquad z^k = A^{-1} d^k, \qquad w^k = x^k - z^k.$$

Since $d^k \in A\mathbb{R}^n$, we have

$$Aw^k = Ax^k - Az^k = Ax^k - AA^{-1}d^k = Ax^k - d^k = b.$$

Since $d^k \to 0$, we have $z^k \to 0$, and we can write the equivalences

$$x_j^k \to +\infty \Leftrightarrow w_j^k \to +\infty, \tag{21a}$$

$$x_j^k \to -\infty \Leftrightarrow w_j^k \to -\infty. \tag{21b}$$

Set $t_j^k = A_j^t \cdot p^k$. Since $x_j^k = g_j'(t_k)$, we have $t_j^k \in \partial f_j(x_j^k)$, and by (20) this implies that

$$f_j^-(x_j^k) \leq t_j^k \leq f_j^+(x_j^k). \tag{22}$$

Let J be the set of indices j for which the sequence $\{x_j^k\}$ is bounded, and let J^c denote its complementary set. If J^c is nonempty, there exists a subsequence $\{x^{k_l}\}$ such that, for each $j \in J^c$, we have

$$\text{either } x_j^{k_l} \to +\infty \quad \text{or} \quad x_j^{k_l} \to -\infty.$$

Then, set

$$J_1 = \{j: x_j^{k_l} \to +\infty\}, \qquad J_2 = \{j: x_j^{k_l} \to -\infty\},$$

and let x satisfy (17). For l sufficiently large, we deduce from (21) and (22) that

$$\sum_{j \in J} t_j^{k_l}(w_j^{k_l} - x_j) + \sum_{j \in J_1} f_j^-(x_j^{k_l})(w_j^{k_l} - x_j) + \sum_{j \in J_2} f_j^+(x_j^{k_l})(w_j^{k_l} - x_j) \leq L, \tag{23}$$

where

$$L = \sum_{n=1}^{n} t_j^{k_l}(w_j^{k_l} - x_j) = (A^t \cdot p^{k_l}, w^{k_l} - x)$$
$$= (p^{k_l}, Aw^{k_l} - Ax) = (p^{k_l}, b - b) = 0.$$

Set

$$f_j^-(x_j^{k_l}) = y_j^l, \qquad j \in J_1,$$
$$f_j^+(x_j^{k_l}) = z_j^l, \qquad j \in J_2.$$

Since f_j^- and f_j^+ are increasing functions,

$$\lim_{l \to +\infty} y_j^l = \alpha_j \quad \text{and} \quad \lim_{l \to +\infty} z_j^l = \beta_j$$

exist. Furthermore, we have

$$x_j^{k_l} = g_j'(y_j^{k_l}), \qquad j \in J_1,$$
$$x_j^{k_l} = g_j'(z_j^l), \qquad j \in J_2.$$

Since f_j is cofinite, g_j is real-valued, and then using Corollary 24.5.1 in Ref. 12 it follows that $\alpha_j = +\infty$ and $\beta_j = -\infty$. For $j \in J$, we can assume without loss of generality that $x_j^{k_l}$ converges to $\tilde{x}_j \in [l_j, c_j]$. If $\tilde{x}_j \in]l_j, c_j[$, the sequence $\{t_j^{k_l}\}$ is bounded. In the other case, $\{t_j^{k_l}(w_j^{k_l} - x_j)\}$ is bounded or converges to $+\infty$. Then, taking the limit in (23) where $L = 0$, it follows that J^c is empty.

(ii) Since the sequence $\{x^k\}$ is bounded, there exists at least one limit point of this sequence. Let $\tilde{x}$ be such a point. Then, there exists a subsequence $\{x^{k_l}\}$ which converges to $\tilde{x}$ and we have

$$A\tilde{x} - b = \lim_{l \to +\infty} Ax^{k_l} - b = 0,$$

so that $\tilde{x} \in C$.

Now, set

$$I_1 = \{j\colon \tilde{x}_j = c_j\}, \qquad I_2 = \{j\colon \tilde{x}_j = l_j\}, \qquad I_3 = \{j\colon l_j < \tilde{x}_j < c_j\}.$$

If the sequence $\{t_j^{k_l}\}$ is bounded, we have

$$\lim_{l \to \infty} t_j^{k_l}(\tilde{x}_j - x_j^{k_l}) = 0. \tag{24}$$

From Corollary 24.5.1 in Ref. 12, this is the case for $j \in I_3$. Furthermore, if the sequence $\{t_j^{k_l}\}$ is unbounded, we have

$$t_j^{k_l} \to +\infty, \qquad j \in I_1,$$
$$t_j^{k_l} \to -\infty, \qquad j \in I_2.$$

It follows that, for l sufficiently large,

$$t_j^{k_l} \cdot (\tilde{x}_j - x_j^{k_l}) \geq 0, \qquad j \in I_1 \cup I_2. \tag{25}$$

Now, from the definition of x_j^k, we have

$$\alpha_k = q(x^k) + h(p^k) = \sum_{j=1}^{n} t_j^k \cdot x_j^k - (p^k, b).$$

Since

$$\sum_{j=1}^{n} t_j^k \cdot \tilde{x}_j = \sum_{j=1}^{n} A_j^t p^k \cdot \tilde{x}_j = (p^k, A\tilde{x}) = (p^k, b),$$

then

$$\alpha_k = \sum_{j=1}^{n} t_j^k (x_j^k - \tilde{x}_j),$$

and from (8), (25), and (24), it follows that

$$\lim_{l \to \infty} h(p^{k_l}) = -q(\tilde{x}).$$

Together with (8), this implies that $\tilde{x}$ is an optimal solution and that the subsequence $\{p^{k_l}\}$ is a minimizing sequence for the dual (D2). Now, since S_{P2} is reduced to a singleton x^* and since $\tilde{x}$ was an arbitrary limit point of the sequence $\{x^k\}$, it follows that the whole sequence $\{x^k\}$ converges to x^*. Furthermore, taking again the same arguments as before, it follows that the whole sequence $\{p^k\}$ is a minimizing sequence for (D2). □

Remark 3.3. In Ref. 1 and in Ref. 2, Section 2, it is only assumed that $C \cap \text{dom}\, q \neq \emptyset$ instead of (B2). In addition, we must note that the proof of Theorem 3.3 uses essentially the technical tools introduced in the proof of Lemma 5 in Ref. 1 and the latter part of Proposition 1 in Ref. 1.

In the next section, we replace Assumption (B1) by the stronger condition

(B1′) f_j is strongly convex for each j.

As a consequence of (B1′), we see that (B1) is satisfied, and there exists $\delta > 0$ such that

$$(c_j - d_j) \times (x_j - y_j) \geq \delta |x_j - y_j|^2,$$
$$\forall c_j \in \partial f_j(x_j), \quad \forall d_j \in \partial f_j(y_j), \quad \forall x_j, y_j \in \text{dom}\, \partial f_j. \tag{26}$$

The following theorem is needed for Section 5.

Proposition 3.3. If Assumptions (B1′) and (B2) are satisfied, then ∇h is Lipschitz on $\mathbb{R}^m$.

Proof. By characterization of $x_j(p)$, there exists $c_j(p) \in \partial f_j(x_j(p))$ such that

$$0 = A_j^t \cdot p - c_j(p).$$

Using (26), this implies that

$$(A_j^t \cdot (p - p')) \times (x_j(p) - x_j(p')) \geq \delta |x_j(p) - x_j(p')|^2,$$

which yields

$$|x_j(p) - x_j(p')| \leq \|A_j / \delta\| \, \|p - p'\|;$$

and since

$$\nabla h(p) = Ax(p) - b,$$

it follows that ∇h is Lipschitz on $\mathbb{R}^m$. □

3.3. The Censor and Lent Problem. In Ref. 5, Censor and Lent consider the following extremal problem:

$$\text{(P3)} \quad \alpha_3 = \min\left(-\sum_{j=1}^{n} \log x_j \,\middle|\, x \in C \right), \qquad C = \{x \colon Ax = b\},$$

where A is an (m, n) real matrix and $b \in \mathbb{R}^m$.

As in Section 3.2, setting $x_i^* = p_i a_i$, we can write the dual problem as

$$\text{(D3)} \quad \beta_3 = \inf(h(p) \,|\, p \in \mathbb{R}^m),$$

$$h(p) = \sum_{j=1}^{n} g_j(A_j^t \cdot p) - \sum_{i=1}^{m} p_i b_i,$$

$$g_j(t_j) = \begin{cases} \sup(t_j \cdot x_j + \log x_j \,|\, x_j > 0), & \text{if } t_j < 0, \\ = +\infty, & \text{otherwise.} \end{cases}$$

Denote

$$x_j(p) = \arg\max_{v_j}(A_j^t \cdot p \cdot v_j + \log v_j), \qquad \text{if } A_j^t \cdot p < 0. \tag{27}$$

Then, we have

$$x_j(p) = -1/t_j, \qquad \text{with } t_j = A_j^t \cdot p, \text{ if } t_j < 0, \tag{28a}$$

$$g_j(A_j^t \cdot p) = \begin{cases} -\log(-A_j^t \cdot p) - 1, & \text{if } A_j^t \cdot p < 0, \\ +\infty, & \text{otherwise,} \end{cases} \tag{28b}$$

so that

$$\text{dom } h = \{p = A^t \cdot p < 0\}, \tag{29a}$$

$$h(p) = -\sum_{j=1}^{n} \log(-A_j^t \cdot p) - \sum_{i=1}^{m} p_i b_i - n, \qquad \forall p \in \text{dom } h. \tag{29b}$$

In Ref. 5, the following assumptions were made:

(C1) $N(A) \cap \mathbb{R}_+^n = \{0\}$, where $N(A) = \{x \colon Ax = 0\}$;

(C2) $C \cap \text{int } \mathbb{R}_+^n \neq \varnothing$.

As pointed out in Ref. 5, it follows from Assumption (C1) and the Gordan transposition theorem that dom h is nonempty. Furthermore, Assumption (C1) implies also that (H2) is satisfied and Assumption (H1) is a direct consequence of (C2). Then, $\alpha_3 = -\beta_3$, and the optimal sets S_{P3} and S_{D3} are nonempty. Since $-\log$ is a strictly convex function, S_{P3} is reduced to a singleton x^*. Also, from (27), (28), and (29), it follows that

$$\text{dom } h = \text{int dom } h, \tag{30}$$

h is continuously differentiable on its domain, and

$$\nabla h(p) = Ax(p) - b. \tag{31}$$

Theorem 3.4. Assume that (C1) and (C2) hold. Then:

(i) let $M \geq \beta$ and set $L_M = \{p: h(p) \leq M\}$; then, there exist two real $r \geq s > 0$ such that

$$s \leq x_j(p) \leq r, \qquad \forall p \in L_M \cap \operatorname{dom} h, \tag{32}$$

and ∇h is Lipschitz on L_M;

(ii) $h \in \mathscr{F}$ and, for each sequence $\{p^k\}$ in L_M such that $\nabla h(p^k) \to 0$, the sequence $\{x^k = x(p^k)\}$ converges to x^*.

Proof.

(i) Since $x_j(p) = -1/A_j^t \cdot p$, we have

$$-\log(-A_j^t \cdot p) = \log x_j(p),$$

and for $z \in C \cap \operatorname{int} \mathbb{R}_+^n$, it follows that

$$\sum_{j=1}^n z_j/x_j(p) = -\sum_{j=1}^n A_j^t \cdot p \cdot z_j = -(A^t p, z) = -(Az, p) = -(b, p),$$

which implies that

$$h(p) = \alpha_z(p) - n, \qquad \alpha_z(p) = \sum_{j=1}^n \log x_j(p) + z_j/x_j(p).$$

Then, we have

$$\alpha_z(p) \leq M + n, \qquad \forall p \in L_M \cap \operatorname{dom} h. \tag{33}$$

Now, since $z_j > 0$, we have

$$\lim_{\xi \to +\infty} \log \xi + z_j/\xi = +\infty, \qquad \lim_{\eta \to 0^+} \log \eta + z_j/\eta = +\infty.$$

From (33), it follows that there exist r and s such that (32) holds. Furthermore, it follows from formula (28) that ∇h is Lipschitz on L_M.

(ii) Consider now a sequence $\{p^k\}$ in L_M such that $\nabla h(p^k) \to 0$, and let us prove that $\{p^k\}$ is a minimizing sequence for the dual problem and that $\{x^k\}$ converges to x^*. From (32), it follows that $\{x^k\}$ is bounded and then there exists at least one limit point of the sequence $\{x^k\}$. Let $\tilde{x}$ be an arbitrary limit point of this sequence, and let $\{x^{k_l}\}$ be a subsequence converging to $\tilde{x}$. From (32), it follows that $s \leq \tilde{x} \leq r$ and we have

$$A\tilde{x} - b = \lim_{l \to +\infty} \nabla h(p^{k_l}) = 0,$$

so that

$$\tilde{x} \in C \cap \operatorname{int} \mathbb{R}_+^n.$$

Furthermore, since

$$A_j^t p^k x_j^k = -1,$$

it follows that

$$h(p^{k_l}) = \sum_{j=1}^{n} \log x_j^{k_l} - \sum_{i=1}^{n} p_i^{k_l} b_i + \sum_{j=1}^{n} A_j^t p^{k_l} x_j^{k_l}$$

$$= \sum_{j=1}^{n} \log x_j^{k_l} + (A^t p^{k_l}, x^{k_l} - \tilde{x}). \tag{34}$$

Now, from (32) and (28), it follows that the sequence $\{A^t p^k\}$ is bounded, and taking the limit in (34), we obtain

$$\lim_{l \to \infty} h(p^{k_l}) = \sum_{j=1}^{m} \log \tilde{x}_j.$$

Then, from (8), it follows that $\tilde{x} = x^*$, and we deduce obviously that the whole sequence $\{x^k\}$ converges to x^* and that $\{p^k\}$ is a minimizing sequence for the dual problem. Then, $h \in \mathscr{F}_1$ and from Proposition 2.1, h belongs to $\mathscr{F}$. □

4. Variant of the Gauss–Seidel Method

Let n_i, $i = 1, 2, \ldots, n$, be positive integers; set $N = \sum_{i=1}^{n} n_i$; and for each i, let h_i be a closed proper convex function defined on $\mathbb{R}^{n_i}$. Let h_0 be a closed proper convex function defined on $\mathbb{R}^N$; set

$$h(y) = h_0(y_1, \ldots, y_n) + \sum_{i=1}^{n} h_i(y_i),$$

and consider the extremum problem

(R) $\quad \beta = \inf(h(y) \mid y \in \mathbb{R}^N)$.

All the dual problems introduced in Refs. 1-5 are special instances of (R). Indeed, for (D1),

$$h_i(y_i) = \delta^*(-y_i \mid C_i), \qquad h_0(y) = q^*(\textstyle\sum y_i);$$

for (D2) and (D3),

$$h = h_0.$$

We shall assume that:

(E0) $\quad \beta$ is finite;

(E1) $\quad h \in \mathscr{F}_1$;

(E2) for each sequence $\{y^k\}$ converging to a boundary point of $\operatorname{dom} h_0$, we have $\lim_{k\to+\infty} h(y^k)=+\infty$; then, (E2) trivially holds when $\operatorname{dom} h_0=\mathbb{R}^N$;

(E3) $\operatorname{dom} h_0=\operatorname{int}\operatorname{dom} h_0$, h_0 is continuously differentiable on $\operatorname{dom} h_0$, and for each $M>\beta$, ∇h_0 is Lipschitz on the level set $\{y: h(y)\leq M\}$;

(E4) $(\prod_{i=1}^n \operatorname{dom} h_i)\subset \operatorname{dom} h_0$.

It is easy to see that all these assumptions are satisfied for the former dual problems (D1), (D2), (D3) if the basic assumptions of these problems hold. Consider now the following variant of the Gauss-Seidel method.

Algorithm 4.1. Start with $y^0\in \operatorname{dom} h$. Suppose that y^k has been computed. Then, for $i=1,2,\ldots,n$, find successively y_i^{k+1} which solves

$$(\mathrm{S}_i^k)\quad \min h^{i,k}(y_i)+(1/2)\|y_i-y_i^k\|^2,$$

where $h^{i,k}$ is defined by

$$h^{i,k}(y_i)=h(y_1^{k+1},\ldots,y_{i-1}^{k+1},y_i,y_{i+1}^k,\ldots,y_n^k). \tag{35}$$

Proposition 4.1. Set

$$y^{i,k}=(y_1^{k+1},\ldots,y_{i-1}^{k+1},y_i^{k+1},y_{i+1}^k,\ldots,y_n^k).$$

Suppose that Assumptions (E0), (E1), (E2), (E3), (E4) are satisfied. Then, for each k and for each $i=1,2,\ldots,n$, there exists a point y_i^{k+1} which solves (S_i^k) such that $y^{i,k}$ belongs to $\operatorname{dom} h$.

Proof. Suppose that y^k and $y^{j,k}$, $j=1,\ldots,i-1$, belong to $\operatorname{dom} h$. Let $\{y_i^l\}$ be a minimizing sequence for problem (S_i^k),

$$\lim_{l\to\infty} h^{i,k}(y_i^l)=d_{i,k}=\inf(h^{i,k}(y_i)+(1/2)\|y_i-y_i^k\|^2\,|\,y_i\in\mathbb{R}^{n_i}).$$

From Assumption (E2), and since $h^{i,k}(\cdot)+(1/2)\|\cdot-y_i^k\|^2$ is strongly convex, it follows that the sequence

$$\{y^{i,k,l}=(y_i^{k+1},\ldots,y_{i-1}^{k+1},y_i^l,y_{i+1}^k,\ldots,y_n^k)\}$$

belongs to a compact set included in $\operatorname{dom} h_0$. Let then

$$\bar{y}^{i,k}=(y_1^{k+1},\ldots,y_{i-1}^{k+1},\bar{y}_i,y_{i+1}^k,\ldots,y_n^k)$$

be a limit point of this sequence. Since h is closed, it follows that

$$(1/2)\|\bar{y}_i-y_i^k\|^2+h(\bar{y}^{i,k})\leq d_{i,k}\quad\text{and}\quad y_i^{k+1}=\bar{y}_i. \qquad \square$$

From Proposition 4.1, we have that the algorithm is well defined. There remains to prove that it converges.

Theorem 4.1. Under Assumptions (E0), (E1), (E2), (E3), (E4), $\{y^k\}$ is a minimizing sequence.

Proof. Since y_i^{k+1} is an optimal solution of (S_i^k), we have

$$(1/2)\|y_i^{k+1}-y_i^k\|^2 \leq h^{i,k}(y_i^k)-h^{i,k}(y_i^{k+1}), \tag{36}$$

from which it follows that

$$h(y^{k+1}) \leq \cdots \leq h^{i,k}(y_i^{k+1}) \leq h^{i,k}(y_i^k) \leq \cdots \leq h(y^k). \tag{37}$$

This implies that the sequence $\{h(y^k)\}$ is decreasing. Then, if we set

$$S_0=\{y: h(y)\leq h(y_0)\},$$

we have

$$y^{i,k}\in S_0, \qquad \gamma=\lim_{k\to\infty} h(y^k) \text{ exists.} \tag{38}$$

From Assumption (E0), γ is finite, and it follows from (36), (37), and (38) that

$$\lim_{k\to\infty}\|y_i^k-y_i^{k+1}\|=0, \qquad \forall i. \tag{39}$$

Set

$$\tilde{h}^{i,k}(v_i)=h^{i,k}(v_i)+(1/2)\|v_i-y_i^k\|^2.$$

From the necessary optimality conditions,

$$0\in\partial\tilde{h}^{i,k}(y_i^{k+1}),$$

we obtain the existence of $c_i^k\in\partial h_i(y_i^{k+1})$ such that

$$\nabla_i h_0(y^{i,k})+c_i^k+y_i^{k+1}-y_i^k=0,$$

where $\nabla_i h_0$ denotes the partial gradient of h_0 relative to $y_i\in\mathbb{R}^{n_i}$. Then, from Assumption (E3), (38), and (39), we get

$$\nabla_i h_0(y^{k+1})+c_i^k+y_i^{k+1}-y_i^k+\epsilon_i^k=0, \qquad \text{with } \lim_{k\to\infty}\epsilon_i^k=0. \tag{40}$$

Finally, from (39), we have

$$\lim_{k\to\infty}\nabla_i h_0(y^{k+1})+c_i^k=0.$$

Since $h\in\mathscr{F}_1$, we conclude that $\{y^k\}$ is a minimizing sequence. □

Remark 4.1. Assume that $h = h_0$ and $n_i = 1$, for each i. Then, the algorithm can be made implementable by requiring y_i^{k+1} to satisfy Ineq. (36) and

$$\|\nabla_i h^{i,k}(y_i^{k+1}) + y_i^{k+1} - y_i^k\| \leq \epsilon_i^k, \qquad \text{with } \epsilon_i^k \to 0,$$

rather than being the exact minimum.

The proof of Proposition 4.2 holds with minor modifications, and such a point can be computed in a finite number of steps.

Remark 4.2. In fact, this algorithm may be viewed as the Gauss-Seidel method applied to minimizing in an augmented linear space the function w given by

$$w(y_1, \ldots, y_n, z_1, \ldots, z_n) = h(y_1, \ldots, y_n) + \sum_{i=1}^{n} \|y_i - z_i\|^2.$$

By letting

$$w_0(y_1, y_2, \ldots, y_n, z_1, \ldots, z_n) = h_0(y_1, \ldots, y_n) + \sum_{i=1}^{n} \|y_i - z_i\|^2,$$

we see that w has the same form as h. The above reasoning is based on that in Ref. 6, page 232.

Remark 4.3. In view of the results given in Section 4 concerning Problems (P1), (P2), (P3), we observe, as mentioned previously, that all assumptions of Theorem 4.2 are satisfied by the duals (D1), (D2), (D3). Applying the Gauss-Seidel variant to these duals produces a dual minimizing sequence $\{y^k\}$ and a corresponding primal sequence $\{x^k\}$ converging to the optimal solution of the primal problem.

Remark 4.4. For the nonpolyhedral case, in order to obtain convergence results, Han and Lou assumed in Ref. 4 that $\bigcap_{i=1}^{n}$ int dom C_i was nonempty. This is equivalent to asserting that the dual level sets are bounded. Such an assumption is unnecessary, and we may only assume that $\bigcap_{i=1}^{n}$ ri dom C_i is nonempty.

Remark 4.5. In Ref. 5, the result of Censor and Lent relies on an exact minimization at each iteration. We showed (Remark 4.1) that an inexact minimization works as well and yields also dual convergence results not obtained in Ref. 5.

References

1. TSENG, P., and BERTSEKAS, D. P., *Relaxation Methods for Problems with Strictly Convex Separable Costs and Linear Constraints*, Mathematical Programming, Vol. 38, pp. 303–321, 1987.
2. BERTSEKAS, D. P., HOSEIN, P. A., and TSENG, P., *Relaxation Methods for Network Flow Problems with Convex Arc Costs*, SIAM Journal on Control and Optimization, Vol. 25, pp. 1219–1243, 1987.
3. HAN, S. P., *A Successive Projection Method*, Mathematical Programming, Vol. 40, pp. 1–14, 1988.
4. HAN, S. P., and LOU, G., *A Parallel Algorithm for a Class of Convex Programs*, SIAM Journal on Control and Optimization, Vol. 26, pp. 344–355, 1988.
5. CENSOR, Y., and LENT, A., *Optimization of log x Entropy over Linear Equality Constraints*, SIAM Journal on Control and Optimization, Vol. 25, pp. 921–933, 1987.
6. BERTSEKAS, D. P., and TSITSIKLIS, J. N., *Parallel and Distributed Computation Numerical Methods*, Prentice-Hall, Englewood Cliffs, New Jersey, 1989.
7. ZANGWILL, W. I., *Nonlinear Programming: A Unified Approach*, Prentice-Hall, Englewoods Cliff, New Jersey, 1969.
8. AUSLENDER, A., *Méthodes Numériques pour la Décomposition de Fonctions Differentiables*, Numerische Mathematik, Vol. 18, pp. 213–223, 1971.
9. AUSLENDER, A., and MARTINET, B., *Méthodes de Décomposition pour la Minimization d'une Fonction sur un Espace Produit*, SIAM Journal on Control, Vol. 12, pp. 635–654, 1974.
10. ORTEGA, J. M., and RHEINBOLDT, W. C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, New York, 1970.
11. AUSLENDER, A., and CROUZEIX, J. P., *Well-Behaved Asymptotical Convex Functions*, Analyse Nonlinéaire, Edited by Attouch, Gauthiers-Villars, Paris, France, pp. 101–122, 1989.
12. ROCKAFELLAR, R. T., *Convex Analysis*, Princeton University Press, Princeton, New Jersey, 1970.
13. ROCKAFELLAR, R. T., *Monotone Operators and the Proximal Point Algorithm*, SIAM Journal on Control and Optimization, Vol. 14, pp. 877–898, 1976.
14. TSENG, P., *Applications of a Splitting Algorithm to Decomposition in Convex Programming and Variational Inequalities*, Technical Report LIDS 1836, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, 1988.
15. TSENG, P., *Coordinate Ascent for Maximizing Nondifferentiable Concave Functions*, Technical Report LIDS 1840, Laboratory for Information and Decision Systems, Massachusetts Institute of Technology, 1988.