A note on Fejér-monotone sequences in product spaces and its applications to the dual convergence of augmented Lagrangian methods

M. Marques Alves · B. F. Svaiter

Abstract In a recent *Math. Program.* paper, Eckstein and Silva proposed a new error criterion for the approximate solutions of augmented Lagrangian subproblems. Based on a saddle-point formulation of the primal and dual problems, they proved that dual sequences generated by augmented Lagrangians under this error criterion are bounded and that their limit points are dual solutions. In this note, we prove a new result about the convergence of Fejér-monotone sequences in product spaces (which seems to be interesting by itself) and, as a consequence, we obtain the full convergence of the dual sequence generated by augmented Lagrangians under Eckstein and Silva's criterion.

Let $\mathbb{R}^n$ and $\mathbb{R}^m$ be endowed with the canonical inner products, denoted by $\langle \cdot, \cdot \rangle$, and let $\mathbb{E}$ be the product space $\mathbb{R}^n \times \mathbb{R}^m$ endowed with the usual inner product $\langle (x, x'), (y, y') \rangle = \langle x, y \rangle + \langle x', y' \rangle$. Moreover, the norms in $\mathbb{R}^n$, $\mathbb{R}^m$ and $\mathbb{E}$ are defined by $\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}$. We shall also assume the same notation and results from [1].

Given a proper closed and convex function $F : \mathbb{E} \to (-\infty, \infty]$, we consider the *primal problem*

$$\min_{x \in \mathbb{R}^n} F(x, 0) \tag{1}$$

and its corresponding *dual problem*

$$\max_{p \in \mathbb{R}^m} Q(0, p), \tag{2}$$

where $Q : \mathbb{E} \to (-\infty, \infty]$ is the concave conjugate of F, i.e.,

$$Q(y, p) = \inf_{(x, u) \in \mathbb{E}} F(x, u) - \langle x, y \rangle - \langle u, p \rangle. \tag{3}$$

M. Marques Alves
Departamento de Matemática, Universidade Federal de Santa Catarina, Florianópolis, Brazil, 88040-900.
Tel.: +55-48-3721-9558
Fax: +55-48-3721-9558
E-mail: maicon.alves@ufsc.br

B. F. Svaiter
IMPA, Estrada Dona Castorina 110, 22460-320 Rio de Janeiro, Brazil.
Tel.: +55-21-2529-5000
Fax: +55-21-2112-4115
E-mail: benar@impa.br

Consider also the saddle-point problem

$$0 \in \partial L(x, p), \tag{4}$$

where the Lagrangian $L : \mathbb{E} \to [-\infty, \infty]$ is defined by

$$L(x, p) = \inf_{u \in \mathbb{R}^m} F(x, u) - \langle u, p \rangle$$

and the maximal monotone operator $\partial L : \mathbb{E} \rightrightarrows \mathbb{E}$ is defined as follows

$$(y, u) \in \partial L(x, p) \iff \begin{cases} L(x', p) \geq L(x, p) + \langle y, x' - x \rangle & \forall x' \in \mathbb{R}^n \\ L(x, p') \leq L(x, p) - \langle u, p' - p \rangle & \forall p' \in \mathbb{R}^m. \end{cases} \tag{5}$$

In [1], an algorithm for solving the optimization problem (1) was proposed based on the following scheme: for $\sigma \in [0, 1)$, find sequences $\{c_k\}_{k \geq 1} \subset \mathbb{R}_{++}$, $\{x_k\}_{k \geq 1}, \{y_k\}_{k \geq 1}, \{w_k\}_{k \geq 0} \subset \mathbb{R}^n$ and $\{p_k\}_{k \geq 0} \subset \mathbb{R}^m$ such that

$$\left(y_k, \frac{1}{c_k}(p_{k-1} - p_k)\right) \in \partial L(x_k, p_k); \tag{6}$$

$$2c_k |\langle w_{k-1} - x_k, y_k \rangle| + c_k^2 \|y_k\|^2 \leq \sigma^2 \|p_{k-1} - p_k\|^2; \tag{7}$$

$$w_k = w_{k-1} - c_k y_k. \tag{8}$$

The main convergence result presented in [1] for the iteration defined in (6)–(8) is as follows.

Proposition 1 [1, Proposition 1] *Consider the sequences* $\{c_k\}_{k \geq 1}$, $\{x_k\}_{k \geq 1}$, $\{y_k\}_{k \geq 1}$, $\{w_k\}_{k \geq 0}$ *and* $\{p_k\}_{k \geq 0}$ *defined in* (6)– (8) *and assume that* $\inf_{k \geq 1} c_k > 0$. *Define*

$$u_k = \frac{1}{c_k}(p_{k-1} - p_k) \quad \forall k \geq 1.$$

If there exists any saddle-point of L, *i.e., any solution of* (4), *then the following statements hold:*

(a) the sequences $\{p_k\}$ *and* $\{w_k\}$ *are bounded;*
(b) $\{u_k\}$ *and* $\{y_k\}$ *converge to zero;*
(c) the sequences $\{F(x_k, u_k)\}$ *and* $\{Q(y_k, p_k)\}$ *both converge to the common optimal value of* (1) *and* (2);
(d) all the limit points of $\{x_k\}$ *and* $\{p_k\}$ *are solutions of* (1) *and* (2), *respectively;*

If no saddle-point exists, then at least one of the sequences $\{p_k\}$ *and* $\{x_k\}$ *is unbounded.*

Moreover, the following property was proved inside the proof of [1, Proposition 1].

Lemma 1 *Consider the sequences* $\{w_k\}_{k \geq 0}$ *and* $\{p_k\}_{k \geq 0}$ *given in* (6)–(8) *and define* $\{z_k\}_{k \geq 0} \subset \mathbb{E}$ *by* $z_k = (w_k, p_k)$ *for all* $k \geq 0$. *Then, for any solution* $z^* = (x^*, p^*)$ *of* (4) *we have:*

$$\|z_k - z^*\|^2 + (1 - \sigma^2)\|p_k - p_{k-1}\|^2 \leq \|z_{k-1} - z^*\|^2 \quad \forall k \geq 1. \tag{9}$$

It follows from Lemma 1 that the sequence $\{z_k\}_{k \geq 0}$ is Fejér-monotone with respect to the solution set of (4). Motivated by this, we prove the following result concerning the convergence of Fejér-monotone sequences in product spaces, which we believe is interesting by itself.

Lemma 2 *Let* $W \times V \subset \mathbb{E}$ *be a nonempty set and let* $\{z_k = (w_k, v_k)\}_{k \geq 0}$ *be a sequence in* $\mathbb{E}$ *such that:*

(a) $\{\|z_k - z\|\}_{k \geq 0}$ *is nonincreasing for all* $z \in W \times V$;
(b) every limit point of $\{v_k\}_{k \geq 0}$ *belongs to* V.

Then, $\{v_k\}_{k \geq 0}$ *converges to some element in* V.

Proof Since $W \times V$ is nonempty, assumption (a) implies that the sequence $\{(w_k, v_k)\}_{k \geq 0}$ is bounded. Define

$$\alpha = \liminf_{k \to \infty} d(w_k, W),$$

where $d(x, W) = \inf_{w \in W} \|x - w\|$. Observe that $0 \leq \alpha < \infty$. There exists a subsequence $\{w_{k_j}\}_{j \geq 0}$ such that

$$\lim_{j \to \infty} d(w_{k_j}, W) = \alpha.$$

Without loss of generality (refining the subsequence, if necessary) we may assume that $\{v_{k_j}\}_{j \geq 0}$ converges to some v which, in view of assumption **(b)**, belongs to V. Since $v \in V$, it follows from assumption **(a)** that if $i \geq k_j$ then, for any $w \in W$

$$\|w_i - w\|^2 + \|v_i - v\|^2 \leq \|w_{k_j} - w\|^2 + \|v_{k_j} - v\|^2.$$

As $0 \leq \alpha \leq \liminf_{i \to \infty} \|w_i - w\|$ (for $w \in W$), taking the lim sup when $i \to \infty$ on the left-hand side of the above inequality we conclude that for any $w \in W$

$$\alpha^2 + \limsup_{i \to \infty} \|v_i - v\|^2 \leq \limsup_{i \to \infty} \left(\|w_i - w\|^2 + \|v_i - v\|^2 \right) \leq \|w_{k_j} - w\|^2 + \|v_{k_j} - v\|^2 \quad \forall j \geq 0.$$

Taking the inf in $w \in W$ on the right-hand side of the last inequality in the above equation we conclude that

$$\limsup_{i \to \infty} \|v_i - v\|^2 \leq d(w_{k_j}, W)^2 - \alpha^2 + \|v_{k_j} - v\|^2$$

To complete the proof, observe that the right-hand side of the above inequality converges to 0 as $j \to \infty$.

We now present the main result of this note. We prove that if the solution set of (4) is nonempty, then the (dual) sequence $\{p_k\}_{k \geq 0}$ generated in (6)–(8) converges to a solution of (2). This is stronger than the convergence result presented in Proposition 1(d).

Proposition 2 *If the solution set of* (4) *is nonempty, then the sequence* $\{p_k\}_{k \geq 0}$ *defined in* (6)–(8) *converges to a solution of* (2).

Proof As the solution set of (4) is nonempty, there is no duality gap in the primal-dual pair of problems (1)-(2) and the solution set of (4) can be written as $W \times V \subset \mathbb{E}$, where W is the solution set of (1) and V the solution set of (2). Then, it follows from Lemma 1 that the sequence $\{z_k = (w_k, p_k)\}_{k \geq 0}$ and the set $W \times V$ satisfy condition (a) of Lemma 2. Moreover, using Proposition 1(d) we conclude that condition (b) of Lemma 2 is satisfied by $\{p_k\}_{k \geq 0}$ and V. Therefore, the result follows as a direct application of Lemma 2 for $\{z_k = (w_k, p_k)\}_{k \geq 0}$ and V.

References

1. J. Eckstein and P. J. S. Silva. A practical relative error criterion for augmented Lagrangians. *Math. Program.*, 141(1-2, Ser. A):319–348, 2013.