Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 2, No. 3, pp. 657 – 665, September, 2006

A STRONGLY CONVERGENT AUGMENTED LAGRANGIAN METHOD

Rolando Gárciga Otero*

Abstract: This paper deals with the general optimization problem $\min g(x)$ subject to $-G(x) \in K$, with $g : X \to \mathbb{R}$, $G : X \to Y$, where X and Y are real reflexive Banach spaces and K is a nonempty closed convex cone in Y. An augmented Lagrangian method is proposed for this problem, which allows for inexact solutions of the primal subproblems and guarantees strong convergence of the primal-dual sequence of iterates to an optimal pair. Moreover, the relation between the initial primal-dual iterate and the strong limit is established.

Key words: *augmented Lagrangian, cone-constrained optimization, inexact solutions, strong convergence*

Mathematics Subject Classification: *90C25, 90C30, 49J40, 46M37*

---

1 Introduction

Augmented Lagrangian methods are among the main tools for solving optimization problems. These methods started with [6, 10, 15] and were further studied in [1, 14, 16, 18]. Its connection with the Proximal Point method was established in [17]. Since then proximal like methods have inspired augmented Lagrangian methods and have being used to study its convergence properties. Examples are in [5, 11, 12] and also in [13], where a more general structure is considered in the spirit of [19]: an optimization problem with constraints described by a functional taking values in a cone, i.e., the problem

$$(P) \quad \begin{cases} \min & g(x) \\ s.t. & -G(x) \in K, \end{cases} \qquad \text{or equivalently,} \qquad \begin{cases} \min & g(x) \\ s.t. & G(x) \precsim 0, \end{cases} \tag{1}$$

where $g : X \to \mathbb{R}$, $G : X \to Y$, X and Y are real reflexive Banach spaces and "$\precsim$" denotes the *cone ordering* on Y (see e.g. [8]) induced by a nonempty closed and convex cone K in Y, i.e.,

$$z \precsim z' \qquad \text{if and only if} \qquad z' - z \in K.$$

In [13] there is introduced an augmented Lagrangian functional for problem (1). It uses an auxiliary mapping $M : X \times Y^* \times \mathbb{R}_{++} \to Y$ given by

$$M(x, y, \rho) = h'(y) + \rho^{-1} G(x), \tag{2}$$

*The work of this author was partially supported by FAPERJ and CNPq. The current version of this manuscript was prepared while the author was visiting the University of Limoges, France, as a post doctoral fellow of Professor M. Théra.

where $h = \frac{1}{r}||\cdot||_{Y^*}^r$, for some $r \in (1, \infty)$, Y^* is the topological dual of Y and $||\cdot||_{Y^*}$ its norm. With this notation, the augmented Lagrangian $\bar{L} : X \times Y^* \times \mathbb{R}_{++} \to \mathbb{R}$ is defined by

$$\bar{L}(x, y, \rho) = g(x) + \rho \frac{1}{s} d(M(x, y, \rho), -K)^s, \tag{3}$$

with $s = r/(r-1)$ and $d(\cdot, -K)$ representing the metric distance to the convex set $-K$ in Y. The (doubly) augmented Lagrangian method in [13] uses an exogenous bounded sequence $\{\lambda_k\} \subset \mathbb{R}_{++}$ and the Bregman distance (see section 2) associated to a strictly convex function $f : X \to \mathbb{R}$. It generates a sequence $\{(x^k, y^k)\} \subset X \times Y^*$ through the iterative formulae

1. Choose $(x^0, y^0) \in X \times K^*$.
2. Given (x^k, y^k), choose $\lambda_k > 0$ and define x^{k+1} as

$$x^{k+1} = \arg\min_{x \in X} \hat{L}(x, x^k, y^k, \lambda_k) = \arg\min_{x \in X} \left[\bar{L}(x, y^k, \lambda_k) + \lambda_k D_f(x, x^k)\right]. \tag{4}$$

3. Define y^{k+1} as

$$y^{k+1} = J_s\left(M(x^{k+1}, y^k, \lambda_k) - P_{-K}(M(x^{k+1}, y^k, \lambda_k))\right), \quad (1 \le j \le m). \tag{5}$$

Here J_s is the duality map of weight $\varphi(t) = t^{s-1}$ in Y^*, and K^* is the positive polar cone of K.

Convergence results in [13] can be resumed as follows: if Y is a uniformly convex and uniformly smooth Banach spaces and $f : X \to \mathbb{R}$ satisfies some technical assumptions (See Section 2 for H1-H4) and there exist *KKT-pairs* for problem (1) and its dual (D) (see Section 2 for the appropriate definitions), then the sequence $\{z^k\} = \{(x^k, y^k)\}$ is bounded and all its weak accumulation points are optimal pairs. Moreover, unicity of the weak accumulation point can be provided when $Y = \mathcal{L}^p(\Omega)$ and Ω is countable or Y is a Hilbert space and $f' : X \to X^*$ is sequentially weak-to-weak* continuous (e.g. the squared norm in a Hilbert space or a $p-th$ norm in a l^p space).

Up to now the question concerning the relation between the initial iterate and the weak accumulation points is unanswered. Moreover, only weak convergence is guaranteed (in infinite dimension). Thus, the main objective of this paper is to present an inexact version of the doubly augmented Lagrangian method that guarantees strong convergence of the primal-dual sequence of iterates. Moreover, it is established that the strong limit will always be the closest point to the initial iterate in the sense of a Bregman distance associated to a separable convex function, which is the sum of the regularizing functions for the primal and dual variables.

Section 2 resumes the preliminaries and describes the main algorithm. Convergence properties are then stated on Section 3, through the relation with a proximal like method.

2 Preliminaries and Algorithm

The main problem (1), also called *primal problem* and denoted by (P), is assumed to be smooth and convex, that is, described with functions $g : X \to \mathbb{R}$ and $G : X \to Y$ satisfying

(A1) g is convex, and G is K-convex (i.e., $\alpha M(x) + (1-\alpha)M(x') - M(\alpha x + (1-\alpha)x') \in K$, for all $x,\ x' \in X$ and all $\alpha \in [0, 1]$).

(A2) g and G are Fréchet differentiable functions with Gâteaux derivatives denoted by g' and G', respectively.

The dual of (P), will be defined as

$$(D)\ \begin{cases}\max\ \Phi(y)\\ \text{s.t. } y\in K^*,\end{cases}\qquad \text{or equivalently,}\qquad \begin{cases}\max\ \Phi(y)\\ \text{s.t. } y\succsim_* 0,\end{cases}$$

where "$\succsim_*$" denotes the cone ordering induced in Y^* by the closed convex cone K^* (i.e $y\succsim_* y'$ if and only if $y-y'\in K^*$). The dual objective $\Phi: Y^*\to\mathbb{R}\cup\{-\infty\}$ is defined as $\Phi(y)=\inf_{x\in X}L(x,y)$ with the *Lagrangian* $L: X\times Y^*\to\mathbb{R}$, given by

$$L(x,y)=g(x)+\langle y,G(x)\rangle, \tag{6}$$

A pair $(x,y)\in X\times Y^*$ is *feasible* if x is *primal feasible*, i.e., $G(x)\precsim 0$ and y is *dual feasible*, i.e., $y\succsim_* 0$. A pair $(x,y)\in X\times Y^*$ is a *Karush-Kuhn-Tucker-pair* (KKT for short) if it is feasible and additionally

$$0=L'_x(x,y)=g'(x)+y\circ G'(x)\qquad \text{(Lagrangian condition)}, \tag{7}$$

$$\langle y,G(x)\rangle=0\qquad \text{(complementarity)}. \tag{8}$$

A pair $(x,y)\in X\times Y^*$ is *optimal* if x is an optimal solution of problem (P) and y an optimal solution of problem (D).

Concerning the relation between the Lagrangian and the augmented Lagrangian we recall the facts ([13, Proposition 6]) that taking $\bar L$ as in (3) and any $(y,\rho)\in Y^*\times\mathbb{R}_{++}$, the function $\bar L(\cdot,y,\rho): X\to\mathbb{R}$ is convex. If Y and Y^* are strictly convex reflexive Banach spaces satisfying the Kadec-Klee property, then $\bar L'_x(\cdot,y,\rho)$ is norm-to-norm continuous and $\bar L'_x(x,y,\rho)=L'_x(x,Q(x,y,\rho))$, where $Q(x,y,\rho)\in B^*$ is defined as

$$Q(x,y,\rho)=J_s\big(M(x,y,\rho)-P_{-K}(M(x,y,\rho))\big). \tag{9}$$

Following the approach in [7], [12], [13] and [17], a regularizing term for primal variables is introduced using a strictly convex and Fréchet differentiable function $f: X\to\mathbb{R}$, with Gâteaux derivative denoted by f'. In order to state the required properties of f, recall that the *Bregman distance* related to f, $D_f: X\times X\to\mathbb{R}$, is given by

$$D_f(x,y)=f(x)-f(y)-\langle f'(y),x-y\rangle, \tag{10}$$

and the *modulus of total convexity* $\nu_f: X\times\mathbb{R}_+\to\mathbb{R}$, is defined as

$$\nu_f(x,t)=\inf_{y\in\{y\in X:||y-x||=t\}} D_f(y,x). \tag{11}$$

The function f is said to be *totally convex* if $\nu_f(x,t)>0$ for all $x\in X$ and all $t>0$. Total convexity first appeared (albeit under a different name) on p. 25 of [3]. Each of the convergence results requires some of the following assumptions on f:

H1: The level sets of $D_f(x,\cdot)$ are bounded for all $x\in X$.

H2: $\inf_{x\in C}\nu_f(x,t)>0$, for all bounded set $C\subset X$ and all $t\in\mathbb{R}_{++}$.

H3: f' is uniformly continuous on bounded subsets of X.

H4: f' is onto.

With the help of f, and of D_f as given by (10), the *doubly augmented Lagrangian* $\hat{L} : X \times X \times Y^* \times \mathbb{R}_{++} \to \mathbb{R}$ is defined as

$$\hat{L}(x, z, y, \rho) = \bar{L}(x, y, \rho) + \rho D_f(x, z), \tag{12}$$

which allows to define the doubly augmented Lagrangian method described in (4)-(5) (or *proximal multiplier method* as known in finite dimension for the case of K being the non-negative orthant). It is introduced, in the sequel, an inexact version of the method. But it should be noted first that the augmented Lagrangian (3) uses a regularizing function h, fixed as $1/r\|\cdot\|_{Y^*}^r$, $r > 1$, for the dual variables. Uniform convexity of Y^* is sufficient for H1-H2, reflexivity for H4 and uniform smoothness is necessary and sufficient for H3. The use of a specific h allow us to get explicit formulae for the dual updating. Then we have a separable regularizing function F, for both primal and dual variables, defined in the product space $X \times Y^*$ by $F(x, y) = f(x) + h(y)$. We recall the fact that F satisfies Hi, provided both f and h satisfy Hi, $i = 1, \dots, 4$ ([12, Proposition 3]). Thus, from now on we assume Y to be a uniformly convex and uniformly smooth reflexive real Banach space, so that property Hi of F will be ensured by property Hi of f, $i = 1, \dots, 4$.

Algorithm

1. Choose $z^0 = (x^0, y^0) \in X \times Y^*$.
2. Given $z^k = (x^k, y^k)$, choose $\lambda_k > 0$ and find $\tilde{x}^k \in X$ such that

$$\langle \hat{L}'_x(\tilde{x}^k, x^k, y^k, \lambda_k), \tilde{x}^k - x^k \rangle \leq \sigma \lambda_k D_F(\tilde{z}^k, z^k), \tag{13}$$

where

$$\tilde{z}^k = \left(\tilde{x}^k, Q(\tilde{x}^k, y^k, \lambda_k)\right). \tag{14}$$

3. Set

$$\begin{aligned} v^k &= \left(\bar{L}'_x(\tilde{x}^k, y^k, \lambda_k), -G(\tilde{x}^k) + \lambda_k P_{-K}(M(\tilde{x}^k, y^k, \lambda_k))\right) && (15)\\ H_k &= \{z \in X \times Y^* : \langle v^k, z - \tilde{z}^k \rangle \leq 0\}, && (16)\\ W_k &= \{z \in X \times Y^* : \langle F'(z^0) - F'(z^k), z - z^k \rangle \leq 0\} && (17) \end{aligned}$$

and

$$z^{k+1} = (x^{k+1}, y^{k+1}) = \arg \min_{z \in H_k \cap W_k} D_F(z, z^0). \tag{18}$$

Concerning the projection step, it is worth to mention that the existence of z^{k+1} is ensured by the total convexity of F. For more on Bregman projections see, for example, [4] and also [2].

The following facts, obtained in [13, Lemma 1], will be needed in the sequel: if the Banach space Y is strictly convex, smooth and reflexive. Then, for all $(x, y, \rho) \in X \times Y^* \times \mathbb{R}_{++}$ it holds

$$h'\left(Q(x, y, \rho)\right) = h'(y) - \frac{1}{\rho}\left[-G(x) + \rho P_{-K}\big(M(x, y, \rho)\big)\right], \tag{19}$$

$$Q(x, y, \rho) \in K^* \qquad \text{and} \qquad P_{-K}\big(M(x, y, \rho)\big) \in N_{K^*}\left(Q(x, y, \rho)\right). \tag{20}$$

3 Convergence Analysis

The convergence properties of the Algorithm will be consequence of its relation to a hybrid inexact version of the proximal point algorithm, applied to the problem of finding zeros of the saddle point operator. Such method was studied in [9]. The *saddle-point operator*, associated to problems (P)-(D), $T_L : X \times Y^* \to \mathcal{P}(X^* \times Y)$ is defined as

$$T_L(x,y) = \big(L'_x(x,y), -L'_y(x,y) + N_{K^*}(y)\big) = (g'(x) + [G'(x)]^*(y), -G(x) + N_{K^*}(y)), \quad (21)$$

where $N_{K^*} : Y^* \to \mathcal{P}(Y)$ denotes the normalizing operator of the cone K^*, given by

$$N_{K^*}(y) = \begin{cases} \{z \in Y | \ \langle z, y' - y\rangle \le 0, \ \forall y' \in K^*\} & \text{if } y \in K^* \\ \emptyset & \text{otherwise.} \end{cases}$$

Under assumptions (A1) and (A2) the operator T_L, defined in (21), is maximal monotone, $0 \in T_L(x,y)$ if and only if (x,y) is a KKT-pair and if $0 \in T_L(x,y)$, then (x,y) is an optimal pair ([13, Proposition 5]).

Proposition 1 (Proximal behavior). *Take* F, $\{v^k\}$, $\{z^k\}$, $\{\tilde{z}^k\}$, $\{\lambda_k\}$ *as in the Algorithm* (13)-(18). *Then*

i) $v^k \in T_L(\tilde{z}^k)$.

ii) *Let* $e^k = v^k + \lambda_k[F'(\tilde{z}^k) - F'(z^k)] \in X^* \times Y$. *Then*

$$\langle e^k, \tilde{z}^k - z^k\rangle \le \lambda_k D_F(\tilde{z}^k, z^k).$$

iii) $z^{k+1} = \arg\min_{z \in H_k \cap W_k} D_F(z, z^0)$.

Proof. Let $v^k = (u^k, w^k) \in X^* \times Y$. Then, by definition of v^k in (15), it holds

$$u^k = \bar{L}'_x(\tilde{x}^k, y^k, \lambda_k) = L'_x(\tilde{x}^k, Q(\tilde{x}^k, y^k, \lambda_k)) = L'_x(\tilde{z}^k),$$

where the second equality follows from definition of Q, (9); and the last one, from (14), and also

$$w^k = -G(\tilde{x}^k) + \lambda_k P_{-K}(M_r(\tilde{x}^k, y^k, \lambda_k)) \in -G(\tilde{x}^k) + N_{K^*}(Q(\tilde{x}^k, y^k, \lambda_k)),$$

where the inclusion is from the fact in (20). So v^k is an element of $T_L(\tilde{z}^k)$ and (i) holds.

In order to prove (ii), let $e^k = (\epsilon^k, \eta^k) \in X^* \times Y$. Then

$$e^k = v^k + \lambda_k[F'(\tilde{z}^k) - F'(z^k)]$$

if and only if

$$\begin{aligned} \epsilon^k &= u^k + \lambda_k[f'(\tilde{x}^k) - f'(x^k)], \\ \eta^k &= w^k + \lambda_k[h'(Q(\tilde{x}^k, y^k, \lambda_k)) - h'(y^k)], \end{aligned}$$

if and only if

$$\begin{aligned} \epsilon^k &= \bar{L}'_x(\tilde{x}^k, y^k, \lambda_k) + \lambda_k[f'(\tilde{x}^k) - f'(x^k)], \\ \eta^k &= -G(\tilde{x}^k) + \lambda_k P_{-K}(M(\tilde{x}^k, y^k, \lambda_k)) + \lambda_k[h'(Q(\tilde{x}^k, y^k, \lambda_k)) - h'(y^k)], \end{aligned}$$

if and only if, using (12) and fact in (19),

$$\begin{aligned} \epsilon^k &= \hat{L}'_x(\tilde{x}^k, x^k, y^k, \lambda_k), \\ \eta^k &= 0, \end{aligned}$$

if and only if

$$e^k = \left(\hat{L}'_x(\tilde{x}^k, x^k, y^k, \lambda_k), 0\right).$$

Thus, $\langle e^k, \tilde{z}^k - z^k\rangle = \langle \hat{L}'_x(\tilde{x}^k, x^k, y^k, \lambda_k), \tilde{x}^k - x^k\rangle$ and the result follows from (13). Item (*iii*) is just equation (18). □

Proposition 1 shows that the proposed algorithm inherits the properties of the Proximal Algorithm studied in [9]. Let S represents the set of zeros of T_L, i.e., the set of KKT-pairs. The next result not only establishes that the algorithm is well defined and that the set S is in the desired side of the hyperplanes when projecting, but also that the error criterion is robust, in the sense that any point sufficiently close to the exact solution of the primal subproblem satisfies the error criterion. As a consequence, if the subproblems are solved with any algorithm guaranteed to converge to the (unique) solution of the subproblem, then a finite number of iterations of such an inner loop will be enough to generate a point satisfying the error criterion.

Proposition 2 (Good definition). *Let f be a totally convex function satisfying H4. Then the algorithm* (13)-(18)*, is well defined, i.e., for each k there exists a unique exact primal solution, the projection step is well defined and $S \subset H_k \cap W_k$. Moreover, if z^k is not a KKT-pair for (P)–(D), then there exists an open subset $U_k \subset X$ such that any $x \in U_k$ solves* (13)-(14).

Proof. Proposition 1 and [9, Proposition 3.1] ensure good definition of the algorithm. In fact, for existence of primal solutions it is enough to assume that g is bounded from below [13, Proposition 7]. For the last part, let $\bar{x}^k$ denote the exact solution of (4) whose existence is ensured by the first part of the proposition, and $\bar{z}^k = (\bar{x}^k, Q(\bar{x}^k, y^k, \lambda_k))$, where Q is as in (9). Then $\bar{z}^k \neq z^k$, because otherwise, by Proposition 1, $0 \in T_L(z^k)$, in contradiction with the assumption that z^k is not a KKT-pair. Hence, $D_F(\bar{z}^k, z^k) > 0$, with $F(x, y) = f(x) + h(y)$. Corollary 1 of [13] establishes continuity of $Q(\cdot, y^k, \lambda_k)$ and then the assumptions on the data functions of problem (P) and Fréchet differentiability of f ensure continuity of the function $\psi_k : X \to \mathbb{R}$ defined as

$$\psi_k(x) = \langle \hat{L}'_x(x, x^k, y^k, \lambda_k), x - x^k\rangle - \lambda_k[D_f(x, x^k) + D_h(Q(x, y^k, \lambda_k), y^k)].$$

Also, $\psi_k(\bar{x}^k) = 0 - \lambda_k D_F(\bar{z}^k, z^k) < 0$, and consequently there exists $\delta_k > 0$ such that $\psi_k(x) \leq 0$ for all $x \in U_k := \{x \in X \ : ||x - \bar{x}^k|| < \delta_k\}$. □

Theorem below states the main convergence results for the Algorithm. Essentially says that when the sequence of errors converges strongly to zero, the sequence of iterates converges strongly to the solution pair which is closest to the initial primal-dual iterate, over the set of KKT-pairs, in the sence of the Bregman distance associated to the regularizing function F. It will be used the notation $\hat{L}_k$ for $\hat{L}(\cdot, x^k, y^k, \lambda_k)$.

Theorem 1 (Strong convergence). *Take $f : X \to \mathbb{R}$ satisfying H1-H3 and $\lambda_k \leq \bar{\lambda}$. Let $\{z^k\}$ be the sequence generated by the Algorithm* (13)-(18)*. If there exist KKT-pairs for problems (P) and (D), then*

$$\lambda_k^{-1}\hat{L}'_k(\tilde{x}^k) \overset{s}{\longrightarrow} 0$$

implies that the primal-dual sequence of iterates $\{z^k\} = \{(x^k, y^k)\}$ *converges strongly to an optimal pair* $\bar{z} = (\bar{x}, \bar{y})$. *Moreover,*

$$\bar{z} = \arg\min_{z \in S} D_F(z, z^0).$$

Proof. By Proposition 1 the sequence $\{z^k\}$ is a particular instance of the sequences generated by the Proximal Algorithm, described in [9], for finding zeros of the operator T_L, with regularizing function $F : X \times Y^* \to \mathbb{R}$ given by $F(x, y) = f(x) + h(y)$. By assumptions on f and h (through Y), F satisfies H1–H3. By the assumption on existence of KKT-pairs, T_L is a maximal monotone operator with zeros. The result then follows from [9, Theorem 4.3]. □

The next theorem consider the case of no KKT-pairs, i.e. $S = \emptyset$; which includes the case of no solution pairs (we are not assuming constraint qualifications).

Theorem 2 (Case of no solutions). *Take* $f : X \to \mathbb{R}$ *satisfying H2-H4 and* $\lambda_k \leq \bar{\lambda}$. *Let* $\{z^k\}$ *be the sequence generated by the Algorithm* (13)-(18). *If* $S = \emptyset$ *and*

$$\lambda_k^{-1} \hat{L}'_k(\tilde{x}^k) \xrightarrow{s} 0$$

then $\{z^k\}_k$ *is unbounded and* $D_F(z^k, z^0) \to +\infty$.

Proof. Follow the proof of Theorem 1, but use [9, Proposition 4.2] to conclude. □

Observe that assumption H3 on f together with uniform convexity of Y imply

$$\lim_{z \to w} \frac{D_F(z, w)}{||z - w||} = 0.$$

Then, fixed w, the function defined by

$$\Psi_w(z) = \begin{cases} \frac{D_F(z,w)}{||z-w||}, & \text{when } z \neq w \\ 0, & \text{when } z = w, \end{cases} \tag{22}$$

is continuous at w. Thus, $\Psi_w(z)$ can be explored as an upper bound for a measure of the error with $w = z^k$ at iteration k. Denote it by Ψ_k. The next result is then devoted to an alternative (more practical) error criterion.

Corollary 1 (Alternative error criterion). *Let* $f : X \to \mathbb{R}$ *be a regularizing function satisfying assumptions H1, H2 and H3 and suppose that* $\lambda_k \leq \overline{\lambda}$ *for all* k *and some* $\overline{\lambda}$. *Assume that there exist KKT-pairs for problems (P) and (D) and that for all* k *it is chosen the error criterion*

$$\|\hat{L}'_k(\tilde{x}^k)\|_* \leq \lambda_k \Psi_k(\tilde{z}^k) \tag{23}$$

instead of (13). *Then the algorithm remains well defined. Moreover, if* $\{\Psi_k(\tilde{z}^k)\}_k$ *is bounded then* $\lambda_k^{-1} \hat{L}'_k(\tilde{x}^k) \xrightarrow{s} 0$ *and* $\{z^k\}$ *converges strongly to* $\hat{z} = \Pi_S^F(z^0) = \arg\min_{z \in S} D_F(z, z^0)$ *and* $\{D_F(z^k, z^0)\}$ *converges to* $D_F(\hat{z}, z^0)$.

Proof. See [9, Corollary 4.4] to get $\lambda_k^{-1} \hat{L}'_k(\tilde{x}^k) \xrightarrow{s} 0$ and apply Theorem 1. □

References

[1] D.P. Bertsekas, On penalty and multiplier methods for constrained minimization, *SIAM J. Control Optim.* 14 (1976) 216–235.

[2] L.M. Bregman, A relaxation method of finding a common point of convex sets and its application to the solution of problems in convex programming, *Ž. Vyčisl. Mat. i Mat. Fiz.* 7 (1967) 620–631.

[3] D. Butnariu, Y. Censor and S. Reich, Iterative averaging of entropic projections for solving stochastic convex feasibility problems, *Comput. Optim. Appl.* 8 (1997) 21–39.

[4] D. Butnariu, A. Iusem and E. Resmerita, Total convexity for powers of the norm in uniformly convex Banach spaces, *J. Convex Anal.* 7 (2000) 319–334.

[5] D. Butnariu and A.N Iusem, *Totally Convex Functions for Fixed Points Computation and Infinite Dimensional Optimization*, Kluwer Academic Publishers, Dordrecht, 2000.

[6] J.D. Buys, *Dual Algorithms for Constrained Optimization Problems*, Doctoral dissertation, Faculty of Science, University of Leiden, Rijksuniversiteit te Leiden, Leiden, 1972.

[7] J. Eckstein, Nonlinear proximal point algorithms using Bregman functions, with applications to convex programming, *Math. Oper. Res.* 18 (1993) 202–226.

[8] J.B.G. Frenk and G. Kassay, On classes of generalized convex functions, Gordan-Farkas type theorems, and Lagrangian duality, *J. Optim. Theory Appl.* 102 (1999) 315–343.

[9] R. Gárciga Otero and B.F. Svaiter, A strongly convergent Hybrid Proximal method in Banach spaces., *J. Math. Anal. Appl.* 289 (2004) 700–711.

[10] M.R. Hestenes, Multiplier and gradient methods, *J. Optim. Theory Appl.* 4 (1969) 303–320.

[11] C. Humes Jr., P.J.S. Silva and B.F. Svaiter, Some inexact hybrid proximal augmented Lagrangian algorithms, *Numer. Algorithms* 35 (2004) 175–184.

[12] A.N. Iusem, and R. Gárciga Otero, Inexact versions of proximal point and augmented Lagrangian algorithms in Banach spaces, *Numer. Funct. Anal. Optim.* 22 (2001) 609–640.

[13] A.N. Iusem and R. Gárciga Otero, Augmented Lagrangian methods for cone-constrained convex optimization in Banach spaces, *J. Nonlinear Convex Anal.* 3 (2002) 155–172

[14] B.W. Kort and D.P. Bertsekas, Combined primal-dual and penalty methods for convex programming, *SIAM J. Control Optim.* 14 (1976) 268–294.

[15] M.J.D. Powell, A Method for Nonlinear Constraints in Minimization Problems, in *Optimization (Sympos., Univ. Keele, Keele, 1968)*, Academic Press, London, 1969 pp. 283–298.

[16] R.T. Rockafellar, The multiplier method of Hestenes and Powell applied to convex programming, *J. Optim. Theory Appl.* 12 (1973) 555–562.

[17] R.T. Rockafellar, Augmented Lagrangians and applications of the proximal point algorithm in convex programming, *Math. Oper. Res.* 1 (1976) 97–116.

[18] R.T. Rockafellar, Augmented Lagrange multiplier functions and duality in nonconvex programming, *SIAM J. Control* 12 (1974) 268–285.

[19] A.P. Wierzbicki and S. Kurcyusz, Projection on a cone, penalty functionals and duality theory for problems with inequality constraints in Hilbert space, *SIAM J. Control Optim.* 15 (1977) 25–56.

Manuscript received 11 January 2006
revised 22 June 2006
accepted for publication 22 June 2006

ROLANDO GÁRCIGA OTERO
Instituto de Economia da Universidade Federal de Rio de Janeiro, Avenida Pasteur 250,
Rio de Janeiro, RJ, Brazil
E-mail address: `rgarciga@ie.ufrj.br`