Augmented Lagrangian methods for equilibrium problems

Alfredo N. Iusem * and Mostafa Nasri †

June 25, 2008

Abstract

We introduce Augmented Lagrangian methods for solving finite dimensional equilibrium problems whose feasible sets are defined by convex inequalities, generalizing the proximal Augmented Lagrangian method for constrained optimization. At each iteration, primal variables are updated by solving an unconstrained equilibrium problem, and then dual variables are updated through a closed formula.

Keywords: Augmented Lagrangian method, Equilibrium problem, Inexact solutions, Proximal point method.

1 Introduction

Let K be a non-empty, closed and convex subset of $\mathbb{R}^n$. Given $f : K \times K \to \mathbb{R}$ such that

P1: $f(x, x) = 0$ for all $x \in K$,

P2: $f(x, \cdot) : K \to \mathbb{R}$ is convex and lower semicontinuous for all $x \in K$,

P3: $f(\cdot, y) : K \to \mathbb{R}$ is upper semicontinuous for all $y \in K$,

the equilibrium problem $\mathrm{EP}(f, K)$ consists of finding $x^* \in K$ such that $f(x^*, y) \geq 0$ for all $y \in K$. The set of solutions of $\mathrm{EP}(f, K)$ will be denoted by $S(f, K)$.

*Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina 110, Rio de Janeiro, RJ, CEP 22460-320, Brazil, e-mail: iusp@impa.br

†Instituto de Matemática Pura e Aplicada, Estrada Dona Castorina 110, Rio de Janeiro, RJ, CEP 22460-320, Brazil, e-mail: mostafa@impa.br

The equilibrium problem encompasses, among its particular cases, convex minimization problems, fixed point problems, complementarity problems, Nash equilibrium problems, variational inequality problems, and vector minimization problems (see, e.g., [4], [13]).

The equilibrium problem has been rather widely studied, but most of the work on the issue deals with conditions for the existence of solutions (see, e.g., [2], [3], [5], [7], [11], and [12]).

In terms of computational methods for equilibrium problems, only a few references can be found in the literature. Among those of interest, we mention the algorithms introduced in [13], [14], [15], and [19].

In the current paper we introduce exact and inexact versions of Augmented Lagrangian methods for solving $\mathrm{EP}(f, K)$ in $\mathbb{R}^n$, for the case in which the feasible set K is of the form

$$K = \{x \in \mathbb{R}^n : h_i(x) \leq 0 \ \ (1 \leq i \leq m)\},$$

where all the h_i's are convex. These methods generate a sequence $\{(x^j, \lambda^j)\} \subseteq \mathbb{R}^n \times \mathbb{R}^m_+$ such that at iteration j, x^j is the unique solution of an unconstrained equilibrium problem and then λ^j is obtained through a closed formula. We comment next on Augmented Lagrangian methods.

The augmented Lagrangian method for equality constrained optimization problems (non-convex, in general) was introduced in [8] and [20]. Its extension to inequality constrained problems started with [6] and was continued in [1], [16], [21], [22], and [23].

We describe next the Augmented Lagrangian method for convex optimization, which is the departure point for the methods in this paper. Consider the problem

$$\min h_0(x) \tag{1}$$

$$\text{s.t. } h_i(x) \le 0 \ (1 \le i \le m), \tag{2}$$

where $h_i : \mathbb{R}^n \to \mathbb{R}$ is convex $(0 \le i \le m)$.

The Lagrangian for (1)–(2) is the function $L : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}$ given by

$$L(x, \lambda) = h_0(x) + \sum_{i=1}^{m} \lambda_i h_i(x), \tag{3}$$

and the dual problem associated to (1)–(2) is the convex minimization problem given by

$$\min -\psi(y) \ \text{ s.t. } y \in \mathbb{R}^m_+, \tag{4}$$

where $\psi : \mathbb{R}^m \to \mathbb{R} \cup \{-\infty\}$ is defined as

$$\psi(\lambda) = \inf_{x \in \mathbb{R}^n} L(x, \lambda). \tag{5}$$

The Augmented Lagrangian associated to the problem given by (1)–(2) is the function $\bar{L} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_{++} \to \mathbb{R}$ defined as

$$\bar{L}(x, \lambda, \gamma) = h_0(x) + \gamma \sum_{i=1}^{m} \left[\left(\max \left\{ 0, \lambda_i + \frac{h_i(x)}{2\gamma} \right\} \right)^2 - \lambda_i^2 \right], \tag{6}$$

where $\mathbb{R}_{++}$ is the set of positive real numbers. The Augmented Lagrangian method requires an exogenous sequence of regularization parameters $\{\gamma_j\} \subset \mathbb{R}_{++}$. The method starts with some $\lambda^0 \in \mathbb{R}^m_+$, and, given $x^j \in \mathbb{R}^n$ and $\lambda^j \in \mathbb{R}^m_+$, the algorithm first determines $x^{j+1} \in \mathbb{R}^n$ as any unconstrained minimizer of $\bar{L}(x, \lambda^j, \gamma_j)$ and then it updates λ^j as

$$\lambda_i^{j+1} = \max \left\{ 0, \lambda_i^j + \frac{h_i(x^{j+1})}{2\gamma_j} \right\} \ (1 \le i \le m).$$

Assuming that both the primal problem (1)–(2) and the dual problem (4) have solutions, and that the sequence $\{x^j\}$ is well defined, in the sense that all the unconstrained minimization subproblems are solvable, it has been proved that the sequence $\{\lambda^j\}$ converges to a solution of the dual problem (4) and that the cluster points of the sequence $\{x^j\}$ (if any) solve the primal problem (1)-(2) (see, e.g., [9] or [23]).

Another augmented Lagrangian method for the same problem, with better convergence properties, is the proximal Augmented Lagrangian method (see [23]; this method is called "doubly Augmented Lagrangian" in [9]). In this case, $\bar{L}$ is replaced by $\bar{\bar{L}} : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_{++} \times \mathbb{R}^n \to \mathbb{R}$, defined as

$$\bar{\bar{L}}(x, \lambda, \gamma, z) = \bar{L}(x, \lambda, \gamma) + \gamma \|x - z\|^2 = h_0(x) + \gamma \sum_{i=1}^{m} \left[\left(\max \left\{ 0, \lambda_i + \frac{h_i(x)}{2\gamma} \right\} \right)^2 - \lambda_i^2 \right] + \gamma \|x - z\|^2 .$$

The method uses an exogenous sequence $\{\gamma_j\} \subset \mathbb{R}_{++}$ as before, and it starts with $x^0 \in \mathbb{R}^n$, $\lambda^0 \in \mathbb{R}^m_+$. Given x^j, λ^j, the next primal iterate x^{j+1} is the unique unconstrained minimizer of $\bar{\bar{L}}(x, \lambda^j, \gamma_j, x^j)$ and the next dual iterate is

$$\lambda_i^{j+1} = \max \left\{ 0, \lambda_i^j + \frac{h_i(x^{j+1})}{2\gamma_j} \right\} \ (1 \le i \le m).$$

In this case, the primal unconstrained subproblem always has a unique solution, due to the presence of the quadratic term $\|x - z\|^2$ in $\bar{\bar{L}}$, and assuming that both the primal and the dual problem are solvable, the sequences $\{x^j\}$, $\{\lambda^j\}$ converge to a primal and a dual solution respectively (see, e.g., [9] or [23]).

The main tool used in [23] for establishing the above mentioned convergence results is the proximal point algorithm, whose origins can be traced back to [17] and [18]. It attained its basic formulation in the work of Rockafellar [24], where it is presented as an algorithm for finding zeroes of a maximal monotone point-to-set operator $T : \mathbb{R}^p \to \mathcal{P}(\mathbb{R}^p)$, i.e, for finding $z \in \mathbb{R}^p$ such that $0 \in T(z)$.

Given an exogenous sequence of regularization parameters $\{\gamma_j\} \subset \mathbb{R}_{++}$ and an initial $z^0 \in \mathbb{R}^p$, the proximal point method generates a sequence $\{z^j\} \subset \mathbb{R}^p$ in the following way: given the j-th iterate z^j, the next iterate z^{j+1} is the unique zero of the operator $T_j : \mathbb{R}^p \to \mathcal{P}(\mathbb{R}^p)$ defined as $T_j(z) = T(z) - \gamma_j(z - z^j)$. It has been proved in [23] that if T has zeroes then $\{z^j\}$ converges to a zero of T.

Inexact versions of the method are also available; instead of requiring $\gamma_j(z^j - z^{j+1}) \in T(z^{j+1})$, they compute an auxiliary vector $\tilde{z}^j$ satisfying $e^j + \gamma_j(z^j - \tilde{z}^j) \in T(\tilde{z}^j)$, where $e^j \in \mathbb{R}^p$ is an error vector, whose norm is small enough. The auxiliary vector $\tilde{z}^j$ defines a hyperplane H_j which separates z^j from the set of zeroes of T. The next iterate z^{j+1} is then obtained by projecting orthogonally z^j onto H_j, or by taking a step from x^j in the direction of H_j (see, e.g., [10], [25], and [26]).

The connection between the Augmented Lagrangian method for convex optimization and the proximal point method can be described as follows. Let $\{x^j\}$, $\{\lambda^j\}$ be the sequences generated by the Augmented Lagrangian method. Consider the maximal monotone operator $T : \mathbb{R}^m \to \mathcal{P}(\mathbb{R}^m)$ defined as $T = \partial(-\psi)$, with ψ as in (5). The sequence $\{z^j\}$ generated by the proximal point for finding zeroes of T coincides with $\{\lambda^j\}$, assuming that $\lambda^0 = z^0$, and that the same sequence $\{\gamma_j\}$ is used for both methods (see, e.g.,[9] or [23]). Hence, the convergence of $\{\lambda^j\}$ to some solution of the dual problem (4) follows from the convergence of the sequence $\{z^j\}$, generated by the proximal point method, to a zero of T.

The convergence analysis of the proximal Augmented Lagrangian method proceeds in a similar way. In this case, the proximal point method is used for finding zeroes of $\widehat{T} : \mathbb{R}^n \times \mathbb{R}^m \to \mathcal{P}(\mathbb{R}^n \times \mathbb{R}^m)$ defined as

$$\widehat{T}(z) = (\partial_x L(z), -\partial_\lambda L(z)) + N_{\mathbb{R}^m_+}(z),$$

with $z = (x, \lambda) \in \mathbb{R}^n \times \mathbb{R}^m$, where L is as in (3) and $N_{\mathbb{R}^m_+}$ is the normalizing operator of the non-negative orthant of $\mathbb{R}^m$. In this case, the sequence $\{z^j\}$ generated by the proximal point method coincides with the sequence $\{(x^j, \lambda^j)\}$ generated by the proximal Augmented Lagrangian method, assuming again that $z^0 = (x^0, \lambda^0)$, and that the same regularization sequence $\{\gamma_j\}$ is used in both algorithms (see, e.g., [9] or [23]).

The convergence analysis of the Augmented Lagrangian methods for equilibrium problems to be introduced here invokes the proximal point method for equilibrium problems, presented in [14]. At iteration j of this method, given $x^j \in \mathbb{R}^n$, one solves $\mathrm{EP}(\bar{f}_j, K)$, where the regularized function $\bar{f}_j$ is defined as

$$\bar{f}_j(x, y) = f(x, y) + \gamma_j \langle x - x^j, y - x\rangle. \qquad (7)$$

Two inexact versions of this method in Banach spaces have been recently proposed in [15]. In finite dimensional spaces, the first one can be described as follows: at iteration j, problem $\mathrm{EP}(f^e_j, K)$ is solved, where f^e_j is defined as:

$$f^e_j(x, y) = f(x, y) + \gamma_j \langle x - x^j, y - x\rangle - \langle e^j, y - x\rangle. \qquad (8)$$

Here, $e^j \in \mathbb{R}^n$ is an error vector, whose norm is small, in a sense to be defined below. The solution $\tilde{x}^j$ of $\mathrm{EP}(f^e_j, K)$ makes it possible to construct a hyperplane separating x^j from $S(f, K)$. A step is then taken from x^j in the direction of the separating hyperplane, generating the next iterate x^{j+1}. In the second version, x^{j+1} is the orthogonal projection of x^j onto the separating hyperplane.

It has been proved in [15] that the sequences $\{x^j\}$ generated by these methods converge to a solution of $\mathrm{EP}(f, K)$ under appropriate assumptions on f, when $\mathrm{EP}(f, K)$ has solutions.

2 Augmented Lagrangian methods for equilibrium problems

We will assume that the function f can be extended to $\mathbb{R}^n \times \mathbb{R}^n$, while preserving P1–P3. In addition, we assume that the closed convex set K in $\mathrm{EP}(f, K)$ is defined as

$$K = \{x \in \mathbb{R}^n : h_i(x) \le 0 \ \ (1 \le i \le m)\}, \qquad (9)$$

where $h_i : \mathbb{R}^n \to \mathbb{R}$ is convex $(1 \le i \le m)$. We will also assume that this set of constraints satisfies any standard constraint qualification, for instance the following Slater's condition.

CQ: If I is the (possibly empty) set of indices i such that the function h_i is affine, then there exists $w \in \mathbb{R}^n$ such that $h_i(w) \le 0$ for $i \in I$, and $h_i(w) < 0$ for $i \notin I$.

We define next our Lagrangian bifunction for $\mathrm{EP}(f, K)$, $\mathcal{L} : (\mathbb{R}^n \times \mathbb{R}^m) \times (\mathbb{R}^n \times \mathbb{R}^m) \to \mathbb{R}$ as

$$\mathcal{L}((x, \lambda), (y, \mu)) = f(x, y) + \sum_{i=1}^m \lambda_i h_i(y) - \sum_{i=1}^m \mu_i h_i(x). \qquad (10)$$

It is worthwhile to mention that when we consider the optimization problem (1)–(2) as a particular case of $\mathrm{EP}(f,K)$ by taking $f(x,y) = h_0(y) - h_0(x)$, (10) reduces to

$$\mathcal{L}((x,\lambda),(y,\mu)) = h_0(y) - h_0(x) + \sum_{i=1}^{m} \lambda_i h_i(y) - \sum_{i=1}^{m} \mu_i h_i(x) = L(y,\lambda) - L(x,\mu),$$

where L is the usual Lagrangian for optimization problems, defined in (3). We introduce now the proximal Augmented Lagrangian for $\mathrm{EP}(f,K)$. Define $s_i : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}_{++} (1 \le i \le m)$, $\widetilde{\mathcal{L}} : \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^n \times \mathbb{R}_{++} \to \mathbb{R}$ as

$$s_i(x,y,\lambda,\gamma) = \frac{\gamma}{2}\left[\left(\max\left\{0, \lambda_i + \frac{h_i(y)}{\gamma}\right\}\right)^2 - \left(\max\left\{0, \lambda_i + \frac{h_i(x)}{\gamma}\right\}\right)^2\right], \tag{11}$$

$$\widetilde{\mathcal{L}}(x,y,\lambda,z,\gamma) = f(x,y) + \gamma\langle x - z, y - x\rangle + \gamma \sum_{i=1}^{m} s_i(x,y,\lambda,\gamma). \tag{12}$$

Now we present Algorithm EALM (*Exact Augmented Lagrangian Method*) for $\mathrm{EP}(f,K)$. Take a bounded sequence $\{\gamma_j\} \subset \mathbb{R}_{++}$. The algorithm is initialized with a pair $(x^0,\lambda^0) \in \mathbb{R}^m \times \mathbb{R}^m_+$.

At iteration j, x^{j+1} is computed as the unique solution of the unconstrained regularized equilibrium problem $\mathrm{EP}(\tilde{\mathcal{L}}_j, \mathbb{R}^n)$ with $\tilde{\mathcal{L}}_j$ given by

$$\widetilde{\mathcal{L}}_j(x,y) = \widetilde{\mathcal{L}}(x,y,\lambda^j,x^j,\gamma_j) = f(x,y) + \gamma_j\langle x - x^j, y - x\rangle + \sum_{i=1}^{m} s_i(x,y,\lambda^j,\gamma_j). \tag{13}$$

Then, the dual variables are updated as

$$\lambda_i^{j+1} = \max\left\{0, \lambda_i^j + \frac{h_i(x^{j+1})}{\gamma_j}\right\} \quad (1 \le i \le m). \tag{14}$$

We introduce now our inexact Augmented Lagrangian method for solving $\mathrm{EP}(f,K)$.

Algorithm IALEM: Inexact Augmented Lagrangian-Extragradient Method for $\mathrm{EP}(f,K)$

1. Take an exogenous bounded sequence $\{\gamma_j\} \subset \mathbb{R}_{++}$ and a relative error tolerance $\sigma \in (0,1)$. Initialize the algorithm with $(x^0,\lambda^0) \in \mathbb{R}^n \times \mathbb{R}^m_+$.
2. Given (x^j,λ^j), find a pair $(\tilde{x}^j, e^j) \in \mathbb{R}^n$ such that $\tilde{x}^j$ solves $\mathrm{EP}(\widetilde{\mathcal{L}}^e_j, \mathbb{R}^n)$, where $\widetilde{\mathcal{L}}^e_j$ is defined as

$$\widetilde{\mathcal{L}}^e_j(x,y) := f(x,y) + \gamma_j\langle x - x^j, y - x\rangle + \sum_{i=1}^{m} s_i(x,y,\lambda^j,\gamma_j) - \langle e^j, y - x\rangle, \tag{15}$$

with s_i as given by (11), and e^j satisfies

$$\left\|e^j\right\| \le \sigma\gamma_j \left\|\tilde{x}^j - x^j\right\|. \tag{16}$$

3. Define λ^{j+1} as

$$\lambda_i^{j+1} = \max\left\{0, \lambda_i^j + \frac{h_i(\tilde{x}^j)}{\gamma_j}\right\} \quad (1 \le i \le m). \tag{17}$$

4. If $(x^j,\lambda^j) = (\tilde{x}^j, \lambda^{j+1})$, then stop. Otherwise,

$$x^{j+1} = \tilde{x}^j - \frac{1}{\gamma_j} e^j. \tag{18}$$

We mention that EALM can be realized as a particular instance of IALEM by taking $e^j = 0$ for all $j \in \mathbb{N}$.

The convergence properties of Algorithm IALEM, whose proof is omitted for the sake of conciseness, are established in the following theorem.

Theorem 2.1. *Consider* $\mathrm{EP}(f,K)$. *Assume that*

i) f satisfies P1–P4 and P4" on $\mathbb{R}^n \times \mathbb{R}^n$,

ii) K is given by (9),

iii) the constraint qualification CQ *stated in Section 2 holds for K,*

iv) $\{\gamma_j\} \subset (\theta, \bar{\gamma}]$ for some $\bar{\gamma} > \theta$, where θ is the undermonotonicity constant of f in P4.

Let $\{(x^j,\lambda^j)\}$ be the sequence generated by Algorithm IALEM *for solving* $\mathrm{EP}(f,K)$. *If* $\mathrm{EP}(f,K)$ *has solutions then the sequence $\{(x^j,\lambda^j)\}$ converges to some optimal pair (x^*,λ^*) for* $\mathrm{EP}(f,K)$, *and consequently $x^* \in S(f,K)$.*

References

[1] Bertsekas, D.P. On penalty and multiplier methods for constrained optimization problems. *SIAM Journal on Control and Optimization* **14** (1976) 216-235.

[2] Bianchi, M., Pini, R. Coercivity conditions for equilibrium problems. *Journal of Optimization Theory and Applications* **124** (2005) 79-92.

[3] Bianchi, M., Schaible, S. Generalized monotone bifunctions and equilibrium problems. *Journal of Optimization Theory and Applications* **90** (1996) 31-43.

[4] Blum, E. Oettli, W. From optimization and variational inequalities to equilibrium problems. *The Mathematics Student* **63** (1994) 123-145.

[5] Brezis, H., Nirenberg, L., Stampacchia, S. A remark on Ky Fan minimax principle. *Bolletino della Unione Matematica Italiana* **6** (1972) 293-300.

[6] Buys, J.D. *Dual algorithms for constrained optimization problems.* Ph.D. Thesis. University of Leiden, Leiden, The Netherlands, (1972).

[7] Flores-Bazán, F. Existence theorems for generalized noncoercive equilibrium problems: quasiconvex case. *SIAM Journal on Optimization* **11** (2000) 675-790.

[8] Hestenes, M.R. Multiplier and gradient methods. *Journal of Optimization Theory and Applications* **4** (1969) 303-320.

[9] Iusem, A.N. Augmented Lagrangian methods and proximal point methods for convex optimization. *Investigación Operativa* **8** (1999) 11-49.

[10] Iusem, A.N., Gárciga Otero, R. Inexact versions of proximal point and augmented Lagrangian algorithms in Banach spaces. *Numerical Functional Analysis and Optimization* **22** (2001) 609-640.

[11] Iusem, A.N., Kassay, G., Sosa, W. On certain conditions for the existence of solutions of equilibrium problems (to be published in *Mathematical Programming*).

[12] Iusem, A.N., Sosa, W. New existence results for equilibrium problems. *Nonlinear Analysis* **52** (2003) 621-635.

[13] Iusem, A.N., Sosa, W. Iterative algorithms for equilibrium problems. *Optimization* **52** (2003) 301-316.

[14] Iusem, A.N., Sosa, W. A proximal point method for equilibrium problems in Hilbert spaces (to be published).

[15] Iusem, A.N., Nasri, M. Inexact proximal point methods for equilibrium problems in Banach spaces (to be published in *Numerical Functional Analysis and Optimization*).

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

[17] Krasnoselskii, M.A. Two observations about the method of succesive approximations. *Uspekhi Matematicheskikh Nauk* **10** (1955) 123-127.

[18] Moreau, J. Proximité et dualité dans un espace hilbertien. *Bulletin de la Societé Mathématique de France* **93** (1965).

[19] Muu, L.D., Oettli, W. Convergence of an adaptive penalty scheme for finding constraint equilibria. *Nonlinear Analysis* **18** (1992) 1159-1166.

[20] Powell, M.J.D. Method for nonlinear constraints in minimization problems. In *Optimization* (R. Fletcher, editor). Academic Press, London (1969).

[21] Rockafellar, R.T. A dual approach to solving nonlinear programming problems by unconstrained optimization. *Mathematical Programming* **5** (1973) 354-373.

[22] Rockafellar, R.T. The multiplier method of Hestenes and Powell applied to convex programming. *Journal of Optimization Theory and Applications* **12** (1973) 555-562.

[23] Rockafellar, R.T. Augmented Lagrangians and applications of the proximal point algorithm in convex programming. *Mathematics of Operations Research* **1** (1976) 97-116.

[24] Rockafellar, R.T. Monotone operators and the proximal point algorithm. *SIAM Journal on Control and Optimization* **14** (1976) 877-898.

[25] Solodov, M.V., Svaiter, B.F. A hybrid projection-proximal point algorithm. *Journal of Convex Analysis* **6** (1999) 59-70.

[26] Solodov, M.V., Svaiter, B.F. An inexact hybrid extragardient-proximal point algorithm using the enlargement of a maximal monotone operator. *Set-Valued Analysis* **7** (1999) 323-345.