

An Augmented Penalization Algorithm for the Equality Constrained Minimization Problem¹

M.C. MACIEL²; Department of Mathematics, Southern National University, Bahía Blanca, Argentina

G.N. SOTTOSANTO³; Department of Mathematics, Economy and Administration Faculty, Comahue National University, Neuquén, Argentina.

Abstract. In this contribution an iterative method for solving the nonlinear minimization problem with equality constraints is presented. The method is based on the sequential minimization of the differentiable penalization function known as augmented Lagrangian. Each unconstrained minimization subproblem is solved by using a conjugate-gradient technique combined with a trust-region strategy of globalization, which is especially efficient for large-scale problems. The update of the multipliers and the penalty parameter is done by using standard schemes. The theoretical properties and the behavior of the algorithm are discussed. Details of the implementation are presented, the algorithm is tested with a set of classic problems and with a minimax formulation to the problems which belong to the well known family of *Hard-Sphere Problems*.

1. Introduction

Let us consider the nonlinear optimization problem with equality constraints

$$\text{Minimize } f(x) \text{ subject to } c(x) = 0, \quad (1.1)$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $c : \mathbb{R}^n \rightarrow \mathbb{R}^m$ are functions at least twice differentiable and $m \leq n$.

We propose an algorithm for solving (1.1) based on a sequential penalization method which involves the solution to unconstrained minimization subproblems of the *augmented Lagrangian function* [1],

$$L_\mu(x, \lambda) = f(x) + \lambda^T c(x) + \frac{1}{2\mu} \|c(x)\|_2^2,$$

where $\lambda \in \mathbb{R}^m$ is an estimate of the Lagrange multiplier vector and $\mu > 0$ is the penalty parameter.

The penalization methods based on the augmented Lagrangian function (multipliers methods) were introduced by M.R. Hestenes [6] and M.J.D. Powell [9] to avoid numerical instabilities of the algorithms based on exterior and interior points methods for solving mathematical programming problems.

¹This work has been partially supported by UNCo E24 and UNS 24/L041 projects

²immaciel@criba.edu.ar

³gsottos@uncoma.edu.ar

Several successful algorithms for solving problems of nonlinear programming based on the augmented Lagrangian function have been reported in the last years. Among them, the works of N. Gould [5], A. Conn, N. Gould and Ph. Toint [2], J.P. Dussault [4] and N. Krejić et al. [8] should be mentioned.

In our approach, each unconstrained subproblem is solved by using a conjugate-gradient technique especially efficient for large-scale problems and the trust-region strategy is incorporated as strategy of globalization [12].

In order to show the efficiency of this combination of techniques the algorithm is tested by using two family of problems: a set of test problems from the collection proposed by W. Hock and K. Schittkowski [7] and K. Schittkowski [10] and a set of problems belonging to the *Hard-Spheres Problems* family.

The paper is organized as follows: in §2 the algorithm is described and a complete formulation of it is stated in §3. In §4 the properties of the algorithm are discussed and §5 is devoted to numerical experiences. The final remarks and conclusions are given in §6.

2. Description of the Algorithm

The algorithm has two main iterations, the outer iteration updates the Lagrange multipliers and the penalty parameters, decides stopping criteria and the rules for declaring convergence of the associate procedure and the inner iteration minimizes the augmented Lagrangian function with a trust region strategy of globalization. Now we describe each one of them. Assume that the problem (1.1) has at least a local solution x_* .

At each outer iteration, given x_k and λ_k , a new estimate x_{k+1} of x_* is found by solving the unconstrained subproblem

$$\min_x L_{\mu_k}(x, \lambda_k). \quad (2.1)$$

This subproblem is solved by constructing a quadratic model of L_{μ_k} around x_k , where a trust-region constraint has been added. Then a step is obtained by solving the following trust-region subproblem

$$\min_{\|\delta\| \leq \delta_k} Q_k(s, \lambda_k, \mu_k),$$

where Q_k is a quadratic model of L_{μ_k} around x_k and $\lambda_k \in \mathbb{R}^m$ is the multiplier vector estimate associated to x_k in the previous iteration, μ_k is the penalty parameter and δ_k is the trust-region radius.

This subproblem is solved using a conjugate-gradient method [12]. This may be regarded as a generalized dogleg technique to approximate the solution to this subproblem. The advantage of this method is that it can handle quite well semidefinite and indefinite Hessian matrices.

The step is evaluated and the trust-region radius is updated in the standard ways for a method which is based on trust-region approach.

Clearly, the method used to solve the unconstrained trust-region subproblem (2.1) define an inner iteration inside each outer iteration of the penalization algorithm.

Each inner iteration terminates at $(x_*)_{k+1}$ and declares it as an approximate solution to (2.1) if

$$\|\nabla L_{\mu_k}((x_*)_{k+1}, \lambda_k)\| \leq tol,$$

for some tolerance $tol > 0$.

At this point we must remark that the choice of the penalty parameter is a crucial key in the algorithm. A decreasing sequence $\{\mu_k\}$ must be generated to guarantee global convergence.

The iterate $(x_\star)_{k+1}$ and its associate multiplier is declared close to a *Karush-Kuhn-Tucker* (KKT) pair (x_\star, λ_\star) of problem (1.1) if

$$\|\nabla l((x_\star)_{k+1}, \lambda_k)\| \leq \epsilon_1, \quad \|c((x_\star)_{k+1})\| \leq \epsilon_2,$$

where $l(x, \lambda) = f(x) + \lambda^T c(x)$ is the Lagrangian function associated to the problem (1.1) and ϵ_1, ϵ_2 are positive tolerances.

If some of these conditions fail, the Lagrange multipliers vector and the penalty parameter are updated.

The multiplier vector is updated by using the Hestenes and Powell's formula [6],

$$\lambda_k = \lambda_{k-1} + \frac{1}{\mu_k} c(x_k). \quad (2.2)$$

Since our proofs of convergence depend strongly on the boundness of λ_k , formula (2.2) must be modified to avoid the possibility that it does not happen.

3. Description of the Main Model Algorithm

In this section the iterations that define our algorithm are established. First we describe the algorithmic parameters for this specific implementation. In the scheme for the outer iteration, a lower bound $\mu_{min} > 0$ for the penalty parameter is used, a positive constant M will be defined as upper bound for the Lagrange multiplier estimates. In the inner iteration a lower bound $\delta_{min} > 0$ is used for the trust-region radius at the beginning of each iteration. Then, the j -th iteration begins with $\delta_j \geq \delta_{min}$. The constants $0 < \eta_1 < \eta_2 < 1$ are stated for the acceptance step. Positive constants ϵ_1, ϵ_2 and γ are tolerances.

Given $x_k \in \mathbb{R}^n$, the current estimate of the solution to problem (1.1), $\lambda_k \in \mathbb{R}^m$, the estimate of the Lagrange multiplier vector, the following algorithm defines how to obtain the next iterate x_{k+1} and the next estimate of the Lagrange multiplier vector λ_{k+1} .

Algorithm 1. (Framework for the outer iteration) Assume $x_k \in \mathbb{R}^n$ the current iterate, $\lambda_k \in \mathbb{R}^m$, $\|\lambda_k\| \leq M$ and μ_k the penalty parameter such that $0 < \mu_{min} \leq \mu_k$. If x_k is not a stationary point of problem (1.1), the steps for obtaining the next iterate and its associate multiplier are the following.

STEP I. (Inner iteration)

Find $(x_\star)_{k+1}$ an approximate solution to the unconstrained minimization subproblem

$$\min_x L_{\mu_k}(x, \lambda_k)$$

by using algorithm 2.

STEP II. (Convergence test)

- If $\|\nabla l((x_\star)_{k+1}, \lambda_k)\| \leq \epsilon_1$ and $\|c((x_\star)_{k+1})\| \leq \epsilon_2$ then **TERMINATE**, $(x_\star)_{k+1} = x_\star$ is declared as solution to the subproblem,

- **otherwise**, update the penalty parameter so that $0 < \mu_{k+1} < \mu_k$.
 - **If** $\mu_{k+1} \geq \mu_{min}$ **update** the multiplier estimate and go to STEP I,
 - **otherwise TERMINATE** with the approximation $(x_*)_{k+1} \cong x_*$.

The following scheme describes the inner iteration algorithm of the **STEP I** of Algorithm 1.

Algorithm 2. (Framework for the inner iteration)

Given $\delta_k \geq \delta_{min} > 0$, $x_k \in \mathbb{R}^n$, $\lambda_k \in \mathbb{R}^m$, $\mu_k \geq \mu_{min} > 0$, a quadratic model Q_k of the augmented Lagrangian function around x_k is defined.

STEP 1. (Compute a trial step)

Find s_k an approximate solution to the trust-region subproblem

$$\min_{\|s\| \leq \delta_k} Q_k(s, \lambda_k, \mu_k)$$

by using conjugate gradient and trust-region method (T. Steihaug, 1983).

STEP 2. (Evaluate the trial step)

Compute the quantities

$$Ared_k(s_k, \lambda_k, \mu_k) = L_{\mu_k}(x_k, \lambda_k) - L_{\mu_k}(x_k + s_k, \lambda_k),$$

$$Pred_k(s_k, \lambda_k, \mu_k) = -Q_k(s_k, \lambda_k, \mu_k)$$

and

$$\rho_k(s_k, \lambda_k, \mu_k) = \frac{Ared_k(s_k, \lambda_k, \mu_k)}{Pred_k(s_k, \lambda_k, \mu_k)}.$$

STEP 3. (Update the trust-region radius and state the new iterate)

- **If** $\rho_k > \eta_2$ then $x_{k+1} = x_k + s_k$, increase the trust-region radius and go to STEP 4.
- **If** $\eta_1 \leq \rho_k \leq \eta_2$ then $x_{k+1} = x_k + s_k$, keep the trust-region radius and go to STEP 4.
- **If** $\rho_k < \eta_1$, reduce the trust-region radius and go to STEP 1.

STEP 4. (Convergence test)

- **If** $\|\nabla L_{\mu_k}(x_{k+1}, \lambda_k)\| \leq \gamma \mu_k$ then **TERMINATE**, set $x_{k+1} = (x_*)_{k+1}$ and go to STEP II of algorithm 1.
- **otherwise**, define a new quadratic model and go to STEP 1.

4. Properties of the Algorithm

In this section two theoretical results concerning to the convergence of the algorithm are stated.

Theorem 1. *Consider the problem (1.1). Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ and $c : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be at least twice continuously differentiable and let the sequences of matrices $\{\|\nabla^2 L_{\mu_k}(x_k, \lambda_k)\|\}$ be uniformly bounded. Then Algorithm 1 is well defined.*

Proof. Clearly the well definition of the algorithm results from the way that Algorithm 1 and Algorithm 2 for the inner iteration have been defined. On the other hand, in order to select the trial step, algorithm 2 starts with the direction of the negative gradient of the augmented Lagrangian at the initial point. More details can be found in [12] and [11]. \square

Theorem 2. *Under the assumptions of Theorem 1, and*

- i) the sequence $\{(x_\star)_k\}$ generated by the Algorithm 1 is contained in a compact set $X \subset \mathbf{R}^n$,*
- ii) x_\star is an accumulation point of the sequence such that the matrix $\nabla c(x_\star)$ has full rank,*
- iii) the sequence of penalty parameters $\{\mu_k\}$ goes to zero when k goes to infinity,*
- iv) the sequence $\{(\lambda_\star)_k\}$, is uniformly bounded,*

then x_\star is a stationary point of (1.1) and the sequence $\{(\lambda_\star)_k\}$ converges to the Lagrange multiplier λ_\star associated to x_\star .

Proof. Let us consider those indices $k \in K$ such that there exists a subsequence $\{(x_\star)_k\}$ that converges to x_\star . Let the gradient be

$$\begin{aligned} \nabla L_{\mu_{k-1}}((x_\star)_k, \lambda_{k-1}) &= \nabla f((x_\star)_k) + \nabla c((x_\star)_k)(\lambda_\star)_{k-1} \\ &\quad + \frac{1}{\mu_{k-1}} \nabla c((x_\star)_k)c((x_\star)_k), \end{aligned} \quad (4.1)$$

where $\nabla c(x) \in \mathbf{R}^{n \times m}$ is the Jacobian matrix of $c(x)$.

The convergence test in the inner iteration demands

$$\begin{aligned} \|\nabla L_{\mu_{k-1}}((x_\star)_k, \lambda_{k-1})\|_2 &= \|\nabla f((x_\star)_k) + \nabla c((x_\star)_k)(\lambda_\star)_{k-1} \\ &\quad + \frac{1}{\mu_{k-1}} \nabla c((x_\star)_k)c((x_\star)_k)\|_2 \leq \gamma \mu_{k-1}. \end{aligned} \quad (4.2)$$

Now, by using the Hestenes and Powell update for the multiplier, (4.2) becomes

$$\|\nabla L_{\mu_{k-1}}((x_\star)_k, \lambda_{k-1}, \mu_{k-1})\|_2 = \|\nabla f((x_\star)_k) + \nabla c((x_\star)_k)(\lambda_\star)_k\|_2 \leq \gamma \mu_{k-1}. \quad (4.3)$$

From the continuity of $\nabla c(x)$ and $\text{rank}(\nabla c(x_\star)) = m$, for k large enough the pseudo-inverse matrix of $\nabla c(x_\star)$, denoted by $\nabla c((x_\star)_k)^+$, exists, is bounded and converges to $\nabla c(x_\star)^+ = (\nabla c(x_\star)^T \nabla c(x_\star))^{-1} \nabla c(x_\star)^T$.

Multiplying (4.1) by $\nabla c((x_\star)_k)^+$, from the ℓ_2 -norm property of consistency and inequalities (4.2) and (4.3) we have

$$\begin{aligned} &\|\nabla c((x_\star)_k)^+ \nabla f((x_\star)_k) + \nabla c((x_\star)_k)^+ \nabla c((x_\star)_k)(\lambda_\star)_k\|_2 \\ &= \|\nabla c((x_\star)_k)^+ \nabla f((x_\star)_k) + (\lambda_\star)_k\|_2 \\ &\leq \gamma \mu_{k-1} \|\nabla c((x_\star)_k)^+\|_2. \end{aligned} \quad (4.4)$$

From the first order optimality conditions for the problem (1.1) it is possible to define

$$\lambda_\star = -\nabla c(x_\star)^+ \nabla f(x_\star).$$

Now, let us consider the difference

$$\begin{aligned} (\lambda_\star)_k - \lambda_\star &= (\lambda_\star)_k + \nabla c((x_\star)_k)^+ \nabla f((x_\star)_k) - \nabla c(x_\star)^+ \nabla f(x_\star) \\ &\quad - \nabla c((x_\star)_k)^+ \nabla f((x_\star)_k), \end{aligned}$$

and using (4.4) we obtain the following inequality

$$\begin{aligned} \|(\lambda_\star)_k - \lambda_\star\|_2 &\leq \gamma \mu_{k-1} \|\nabla c((x_\star)_k)^+\|_2 \\ &\quad + \|\nabla c(x_\star)^+ \nabla f(x_\star) - \nabla c((x_\star)_k)^+ \nabla f((x_\star)_k)\|_2. \end{aligned} \quad (4.5)$$

Since $\nabla f(x)$ and $\nabla c(x)$ are continuous functions allowing $k \xrightarrow[k \in K]{} \infty$ in (4.5), results that the sequence $(\lambda_\star)_k$ converges to λ_\star .

Therefore when $k \xrightarrow[k \in K]{} \infty$ in (4.3) we get

$$\nabla f(x_\star) + \nabla c(x_\star) \lambda_\star = 0.$$

Since

$$(\lambda_\star)_k = (\lambda_\star)_{k-1} + \frac{1}{\mu_{k-1}} c((x_\star)_k) \quad (4.6)$$

results

$$\|\mu_{k-1}((\lambda_\star)_k - (\lambda_\star)_{k-1})\|_2 = \|c((x_\star)_k)\|_2,$$

and using the assumption that the sequence of multipliers is bounded, when $k \xrightarrow[k \in K]{} \infty$ we have

$$c(x_\star) = 0.$$

Therefore, x_\star satisfies the first order necessary conditions for the problem (1.1) and the subsequence $\{(\lambda_\star)_k\}$ converges to the Lagrange multiplier λ_\star associated to the solution. \square

5. Numerical Experiments

In this section a specific implementation of the algorithm is described. A MATLAB code was written corresponding to this implementation. It is worth pointing out that the preliminary numerical implementation produced reliable results in a large number of test problems. Parameters and specific procedures used are as follows:

- (i) the constants $\gamma = 10^{-3}$, $\eta_1 = 10^{-4}$, $\eta_2 = 10^{-1}$,
- (ii) the trust-region radius at each iteration is at least $\delta_{min} = 10^{-4}$,
- (iii) the sequences of penalty parameters has been generated following the scheme $\mu_{k+1} = 0.1\mu_k$ and the lower bound $\mu_{min} = 10^{-10}$,
- (iv) the tolerances for the convergence tests were fixed as $\varepsilon_1 = \varepsilon_2 = 10^{-6}$.

The behavior of the algorithm has been analyzed by solving two different sets of problems. The first set was taken from the books of Hock and Schittkowski [7] and Schittkowski [10]. A total of 38 equality constrained optimization problems were selected. The results obtained with this first set of test problems are shown

in table I. The starting guess x_0 is the one indicated in [7], [10] for each problem.

The following abbreviations are used:

HS-number stands for the number of problem from Hock and Schittkowski [7]

S-number stands for the number of problem from Schittkowski [10]

m : number of constraints

n : number of variables

NOI: number of outer iterations

NII: number of inner iterations

μ_* : indicates the value of the penalty parameter at which convergence has been declared.

We report the number of outer and inner iterations required for two different choices of Lagrange multiplier formula. Besides the theoretical results obtained by using the Hestenes and Powell (λ_{HP}) update formula, we also show the numerical results by using the least-squares approximation formula (λ_{LS}), which is expressed as

$$\lambda_k = -[\nabla c(x_k)^T \nabla c(x_k)]^{-1} \nabla c(x_k)^T \nabla f(x_k).$$

Table I: Numerical results

Problems			$\lambda = \lambda_{HP}$			$\lambda = \lambda_{LS}$		
Prob #	n	m	NOI	NII	μ_*	NOI	NII	μ_*
HS-40	4	3	5	46	5×10^{-5}	5	46	5×10^{-5}
HS-46	5	2	6	113	5×10^{-6}	6	117	5×10^{-6}
HS-47	5	3	4	27	5×10^{-4}	5	28	5×10^{-5}
HS-51	5	3	1	2	5×10^{-1}	1	2	5×10^{-1}
HS-52	5	3	6	42	5×10^{-6}	6	42	5×10^{-6}
HS-56	7	4	6	137	5×10^{-6}	5	104	5×10^{-5}
HS-77	5	2	5	25	5×10^{-5}	4	25	5×10^{-4}
HS-78	5	3	5	46	5×10^{-5}	5	46	5×10^{-5}
HS-79	5	3	5	19	5×10^{-5}	4	18	5×10^{-4}
S-219	4	2	6	32	5×10^{-6}	6	32	5×10^{-6}
S-394	20	1	5	76	5×10^{-5}	5	76	5×10^{-5}
S-395	50	1	5	84	5×10^{-5}	5	84	5×10^{-5}

5.1. The Hard-Spheres Problem as an Optimization Problem

The *Hard-Spheres* problem belongs to a family of well-known mathematical problems dated from the century XVII. Some of them were solved but many still remain as open problems. They are related to practical problems in chemistry, biology and physics. For more details see [3].

The problem consists in maximizing the minimum distance among p points on a sphere in \mathbb{R}^n . This problem can be formulated as a nonlinear optimization problem with a large number of KKT points which are not optimal. A whole family of problems can be obtained that can be used as test problems if the parameters n and p are allowed to vary.

On the other hand, the *Hard-Spheres Problem* is associated to other well-known problem called the *Kissing Number Problem*. It consists in determining the maximum number of spheres of given radius such that, without overlapping, can touch

simultaneously a centered sphere with the same radius. Given n and p , if the distance obtained, when the *Hard-Spheres Problem* is solved, is greater than or equal to the radius of the sphere where the points are located, we conclude that the *Kissing number* is $K_n \geq p$.

The formulation of the *Hard-Spheres problem* leads to a problem known as “min-max” where $2r$ is the radius of the sphere centered in the origin. On this sphere one can locate the p points. That is

$$\begin{aligned} \max \quad & \min_{i \neq j} \|y^i - y^j\|_2 \\ \text{s.t} \quad & \|y^k\|_2 = 2r, \quad k = 1, \dots, p. \end{aligned} \quad (5.1)$$

The vectors $y^k \in \mathbb{R}^n$. Since the answer to the problem is invariant under the choice of r , it can be chosen as $r = \frac{1}{2}$.

The problem is equivalent to

$$\begin{aligned} \min \quad & \max_{i \neq j} \langle y^i, y^j \rangle \\ \text{s.t} \quad & \|y^k\|_2 = 1, \quad k = 1, \dots, p, \end{aligned} \quad (5.2)$$

where \langle, \rangle is the inner product defined in \mathbb{R}^n . Furthermore, by using a standard transformation, the “min-max” problem (5.2) can be stated as a constrained minimization problem,

$$\begin{aligned} \min \quad & z \\ \text{s.t} \quad & z \geq \langle y^i, y^j \rangle, \quad \forall i \neq j \\ & \|y^k\|_2 = 1, \quad k = 1, \dots, p. \end{aligned} \quad (5.3)$$

Adding a squared slack variable to each inequality constraint of (5.3), it becomes a nonlinear programming problem with equality constraints:

$$\begin{aligned} \min \quad & z \\ \text{s.t} \quad & z - \langle y^i, y^j \rangle - s_{ij}^2 = 0, \quad \forall i \neq j \\ & \|y^k\|_2^2 = 1, \quad k = 1, \dots, p. \end{aligned}$$

In order to put the problem under the assumptions of differentiability, observe that the set of equality constraints of (5.3) are squared.

5.2. Numerical Results

We have solved a small set of problems of the *Hard-Spheres Problem*. Each problem was solved 50 times with starting points chosen at random so that each variable belongs to the interval $[-1, 1]$. The Lagrange multiplier estimate was updated by using the Hestenes and Powell formula. The table II shows the sizes of the problems, the minimum, the maximum and the average amount of the optimal distances found, the outer iterations, the functions evaluations and the CPU-time in seconds required to obtain the solution.

Table II: Numerical Results of the *Hard-Spheres Problem*

Size of the Prob.		Min. distance	NOI	Func.	CPU-time
n, p	V, C				
2, 4	15, 10	1.414213,1.414213,1.414213	3,7,3.3	74,201,112.4	0.16,055,0.24
2, 6	28, 21	0.891448,1.000000,0.995432	4,8,4.9	123,408,231.2	1.34,7.21,3.08
3, 10	76, 55	1.046976,1.154700,1.096688	4,9,6.2	312,627,289.7	9.34, 92.6, 31.0
3, 12	103, 78	0.946381,1.051462,1.027298	4,17,11.05	301,729,345.1	12.35,89.03,21.39

6. Final Remarks

An algorithm based on penalization combined with the conjugate-gradient method and the trust-region technique has been proposed. It was proved that the method is well defined and convergent under suitable conditions on the problem.

It has been tested with two group of problems. A qualitative analysis of the numerical results shows that the behavior of the algorithm is acceptable.

The algorithm is numerically stable. The values of the penalty parameter for which convergence has been obtained show that they belong to a range of acceptable values in the sense that they do not cause ill-conditioning troubles, in particular when the Hessian matrix of the augmented Lagrangian function is evaluated to construct the quadratic model formulation.

The choice of the Lagrange multiplier estimate shows a little better performance when the least squares approximation formula is used instead of the Hestenes and Powell formula. However this difference is not representative to justify the computational cost of using it.

Among the possible lines for continuation of this work, a more general result, similar to Theorem 2, without specifying the formula of the Lagrange multiplier estimate, but with the assumption that the sequence of $\{\lambda_k\}$ be uniformly bounded should be investigated.

Resumo. Neste trabalho é apresentado um método iterativo do tipo Lagrangeano aumentado para resolver o problema de minimização com restrições de igualdade. Para resolver cada subproblema de minimização irrestrita é usado o método de gradientes conjugados combinado com uma estratégia de região de confiança. Esta estratégia de globalização é muito eficiente para resolver problemas de grande porte. Propriedades teóricas do algoritmo e detalhes da implementação são discutidos. O algoritmo é testado usando um conjunto de problemas-teste clássicos e com uma formulação min-max dos problemas conhecidos como “Hard-Spheres Problems”.

References

- [1] D.P. Bertsekas, “Constrained Optimization and Lagrange Multiplier Methods”, Athenas Scientific, Belmont, MA, 1996.

- [2] A.R. Conn, N.I.M. Gould and Ph.L. Toint, A globally convergent augmented Lagrangian algorithm for optimization with general constraints and simple bounds, *SIAM Journal on Numer. Analysis*, **28**, No.3 (1991), 545-572.
- [3] J.H. Conway and N.J.C. Sloane, “Sphere Packing, Lattices and Groups”, Springer-Verlag, New York, 1988.
- [4] J.P. Dussault, Augmented penalty algorithms, *IMA Journal of Numerical Analysis*, **18** (1998), 355-372.
- [5] N.I.M. Gould, On the convergence of a sequential penalty function method for constrained minimization, *SIAM Journal on Numer. Analysis*, **26** (1989), 107-128.
- [6] M.R. Hestenes, Multiplier and gradient methods, *Journal of Optimization Theory and Applications*, **4** (1969), 303-320.
- [7] W. Hock and K. Schittkowski, Test examples for nonlinear programming codes, in “Lecture Notes in Economics and Mathematical Systems” (M. Beckmann and H.P. Künzi, eds.), **187**, Springer-Verlag, Berlin, 1981.
- [8] N. Krejić, J.M. Martínez, M. Mello and E.A. Pilotta, Validation of an augmented Lagrangian algorithm with a Gauss-Newton Hessian approximation using a set of Hard-Spheres problems, *Computational Optimization and Applications*, **16** (2000), 247-263.
- [9] M.J.D. Powell, A method for nonlinear constraints in minimization problems, in “Optimization” (R. Fletcher, ed.), pp. 283-298, Academic Press, London, 1969.
- [10] K. Schittkowski, More test examples for nonlinear programming codes, in “Lectures Notes in Economics and Mathematical Systems” (M. Beckmann and W. Krelle, eds.), Springer-Verlag, Berlin, 1987.
- [11] G.N. Sottosanto, “Un algoritmo de penalización aumentada para el problema de minimización con restricciones”, Master’s thesis, Departamento de Matemática, Universidad Nacional del Sur, Bahía Blanca, Argentina, 2001.
- [12] T. Steihaug, The conjugate-gradient method and trust-regions in large-scale optimization, *SIAM Journal on Numer. Analysis*, **20(3)** (1983), 626-637.