A variant of projected gradient method for quasi-convex optimization problems with a competitive search strategy

J. Y. Bello Cruz* L. R. Lucambio Pérez†

April 14, 2010

Abstract

We present the projected gradient method for solving constrained quasi-convex minimization problem with a competitive search strategy, i.e., an appropriate stepsize rule through an Armijo-search along feasible direction obtaining global convergence properties. Differently from other similar stepsize rule, we perform only one projection onto the feasible set per iteration, rather than one projection for each tentative step during the search of the stepsize, which represents a considerable saving when the projection is computationally expensive.

Keywords: Quasi-convex optimization, Armijo search, Projected gradient method.
Mathematical Subject Classification (2008): 90C47, 49J35.

1 Introduction

Consider the following constrained minimization problem:

$$\min \quad f(x) \quad \text{s.t.} \quad x \in C, \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable and $C \subset \mathbb{R}^n$ is closed and convex. We denote by S^* the solution set and by $\bar{S}$ the set of stationary points of this problem. We remind that $\overline{x} \in C$ is stationary for problem (1) if and only if $\langle \nabla f(\overline{x}), x - \overline{x} \rangle \geq 0$ for all $x \in C$.

*IME, Universidade Federal de Goiás, Goiânia, GO, CEP 74001-970, Brazil, e-mail: yunier@impa.br
†IME, Universidade Federal de Goiás, Brazil, e-mail: lrlp@impa.br

Many methods have been proposed to solve the problem (1). The simplest is the projected gradient method. The projected gradient method with constants stepsize was proposed by Goldstein in [11]. It was also studied by Levitin and Polyak in [14]. The use of the Armijo-rule for the search along projection arcs was proposed by Bertsekas in [5] and its use for the search along feasible directions was studied by Iusem in [12].

The projected gradient method has some advantages. Firstly, it is easy to implement (specially, for optimization problems with relative simple constrains), uses little storage and readily exploits any sparsity or separable structure of ∇f or C. Secondly, it is able to drop from or add to the active set many constraints at each iteration. Hence, it has been developed for solving various cases of problem (1).

Each iteration of the projected gradient method, which we describe formally later, consists, basically, of two stages: starting from the iterate x^k, we move us on the direction of $-\nabla f(x^k)$, the resulting vector is projected onto C, and on the direction from this projection to x^k we find the next iterate, namely x^{k+1}, such that a sufficient decrease of the function value is guaranteed.

1.1 The projected gradient method

We describe the main iterative scheme of the projected gradient method and the possible chooses for the stepsize. A formal description of the algorithm is the following:

Projected gradient method

Initialization: Take $x^0 \in C$.

Iterative step: Given x^k, compute

$$z^k = x^k - \beta_k \nabla f(x^k), \tag{2}$$

If $x^k = P_C(z^k)$ then stop. Otherwise, let

$$x^{k+1} = \alpha_k P_C(z^k) + (1 - \alpha_k)x^k, \tag{3}$$

where β_k, α_k are positive.

The coefficients β_k and α_k are called stepsizes and $P_C : \mathcal{H} \to C$ is the orthogonal projection onto C, i.e., $P_C(x) = \text{argmin}_{y \in C} \|x - y\|$. Several choices are possible for the stepsizes. Before discussing them, we make mention that in the unconstrained case, i.e. $C = \mathbb{R}^n$, the method given by (2)-(3) with $\alpha_k = 1$, for all k, reduces to the iteration $x^{k+1} = x^k - \beta_k \nabla f(x^k)$, called the steepest descent method (see [6]).

Following [4] and [12], we will focus our comments in four strategies for the stepsizes:

a) Constant stepsize: $\beta_k = \beta$ and $\alpha_k = 1$, for all k, where $\beta > 0$ is a fixed number.

b) Armijo search along the boundary of C: $\alpha_k = 1$, for all k, and β_k is given by

$$\beta_k = \beta 2^{-j(k)}$$

with

$$j(k) = \min\left\{ j \geq 0 : f(z^{k,j}) - f(x^k) \leq -\delta \langle \nabla f(x^k), x^k - P_C(z^{k,j}) \rangle \right\}$$

and

$$z^{k,j} = x^k - \beta 2^{-j} \nabla f(x^k),$$

for some $\beta > 0$, and $\delta \in (0,1)$.

c) Armijo search along the feasible direction: the sequence $\{\beta_k\}$ is contained in $[\hat{\beta}, \tilde{\beta}]$, where $0 < \hat{\beta} \leq \tilde{\beta}$, and α_k is determined by an Armijo rule, namely

$$\alpha_k = 2^{-j(k)}$$

with

$$j(k) = \min\left\{ j \in \mathbb{N} \colon f\left(2^{-j} P_C(z^k) + (1 - 2^{-j}) x^k\right) - f(x^k) \leq -\delta 2^{-j} \langle \nabla f(x^k), x^k - P_C(z^k) \rangle \right\},$$

for some $\delta \in (0,1)$.

d) Exogenous stepsize before projecting: The α_k's are constant equal to 1 and the elements of the sequence $\{\beta_k\}$ are given by

$$\beta_k = \frac{\delta_k}{\|\nabla f(x^k)\|}$$

with

$$\sum_{k=0}^{\infty} \delta_k = \infty \qquad \text{and} \qquad \sum_{k=0}^{\infty} \delta_k^2 < \infty. \tag{4}$$

Several comments are in order.

In Strategy (a) it is necessary to assume Lipschitz continuity of ∇f and to choose $\beta \in (0, \frac{2}{L})$, where L is the Lipschitz constant, in order to ensure that the cluster points of $\{x^k\}$ are stationary, see [4].

Note that Strategy (b) requires one projection onto C for each step of the inner loop resulting from the Armijo search, i.e. possibly many projections for each k, while Strategy (c) demands only one projection for each outer step, i.e. for each k. Thus, Strategy (b) is competitive only when P_C is very easy to compute (e.g. when C is a halfspace, or a box, or a ball, or a subspace).

We make mention that Strategy (d) fails to be a descent method, also in the unconstrained case. Finally, it is easy to show that $\|x^{k+1} - x^k\| \leq \delta_k$ for all k, with δ_k exogenous and satisfying (4). In view of (4), all stepsizes in Strategy (d) are small, while Strategies (c) and (b) allow, occasionally, long steps. More important, Strategy (d) does not take into account the information available at iteration k for determining the stepsizes, which, in general, entails worse computational performance. Its redeeming feature is that its convergence properties hold also in the nonsmooth case, in which the Armijo searches given by (b) and (c) may be unsuccessful.

Without assuming convexity of f, the convergence results for these methods closely mirror the ones for the steepest descent method in the unconstrained case: cluster points may fail to exist, even when (1) has solutions, but if they exist, they are stationary and feasible. These results can be found in Section 2.3.2 of [4].

The convergence results for Strategy (b) can be found in [10]. In order to ensure existence of cluster points, it is necessary to assume that the starting iterate x^0 belongs to a bounded level set of f.

On the other hand, when f is convex, it is proved for Strategies (b) and (c) the convergence of the whole sequence to a minimizer of f under the sole assumption of existence of minimizers, i.e., without any additional assumption on boundedness of level sets. These results can be found in [12].

The projected gradient method under Strategy (d) keeps its good convergence properties in an arbitrary Hilbert space also when f is convex but nonsmooth, after replacing $\nabla f(x^k)$ by a subgradient u^k of f at x^k, i.e., $u^k \in \partial f(x^k)$. See [2] and [1] for convergence properties in this setting. It is proved in these references that the whole sequence $\{x^k\}$ converges weakly to a solution of problem (1) under the sole assumption of existence of solutions. This strategy has been used to solve non-smooth variational inequalities problems in [3].

Another stepsize rule was proposed in [7] by Calamai and Moré. They studied convergence properties of the projected gradient method with a general steplength rule given by:

e) $\alpha_k = 1$ for all k, and β_k is given by

$$\beta_k = \beta 2^{-j(k)}$$

with

$$j(k) = \min\left\{j \geq 0 : f(z^{k,j}) - f(x^k) \leq -\delta_1 \langle \nabla f(x^k), x^k - P_C(z^{k,j})\rangle\right\} \tag{5}$$

and

$$z^{k,j} = x^k - \bar{\beta} 2^{-j} \nabla f(x^k),$$

for some $\beta > 0$, $\delta_1 \in (0, 1)$ and

$$\beta_k \geq \gamma_1 \quad \text{or} \quad \beta_k \geq \gamma_2 \bar{\beta}_k,$$

where $\bar{\beta}_k$ satisfies

$$f(\bar{z}^k) - f(x^k) > -\delta_2 \langle \nabla f(x^k), x^k - P_C(\bar{z}^k)\rangle \tag{6}$$

and

$$\bar{z}^k = x^k - \bar{\beta}_k \nabla f(x^k),$$

with γ_1 and γ_2 are positive constants and $\delta_2 \in (0, 1)$.

It was showed that any cluster point of the generated sequence by the projected gradient method under Strategy (e) is a stationary point of (1). These stepsize rule is fairly mild and is satisfied by several well-known stepsize rules for the gradient projection method, including Strategies (a) and (b). We remark that the condition (6) guarantees that β_k is not too small, thus, such algorithm allows long steps.

Later, Wang and Xiu studied in [15] convergence properties of the gradient projection algorithm with the stepsize rule given by Strategy (e) when f is quasi-convex or pseudo-convex function. They showed that there are only two possibilities: either the sequence has cluster points, in this case such points are minimizer, when f is quasi-convex, or stationary, when f is pseudo-convex, or the sequence has not cluster points, the solution set is empty and the sequence of function values converges to the infimum of f over C. This result improves upon or extend the one by Cheng in [8], and by Kiwiel and Murty in [13], and are also related to the results in [9], [16], [17], [18], [19] and [20]. We recall that $f\colon D \subset \mathbb{R}^n \to \mathbb{R}$, where D is convex, is called quasi-convex if $f(\lambda x + (1-\lambda)y) \le \max\{f(x)\,,\, f(y)\}$, for all x, $y \in D$, and $\lambda \in [0,1]$, and is called pseudo-convex when it is of class $\mathcal{C}^1$ and $\langle \nabla f(y)\,,\, x-y\rangle \ge 0$ implies that $f(x) \ge f(y)$. Other equivalent definitions of quasi-convexity are: all level sets of f are convex, and for differentiable f, $f(x) \le f(y)$ implies that $\langle x-y\,,\, \nabla f(y)\rangle \le 0$. Finally, observe that Strategy (e), as Strategy (b), requires one projection onto C for each tentative step of the Armijo-search in (5), i.e., possibly, many projections for each k.

Reduction of the number of required projections onto the feasible set is quite significant in cases in which the evaluation of this projection is considerably harder than the evaluation of ∇f. This is a rather frequent situation. While many operators of interest in the applications are given by easily computable closed formula, this is exceptional for orthogonal projections onto convex sets. Outside the cases of boxes, balls or hyperplanes, the computation of a projection is in itself an optimization problem which must be approximately solved with the help of some auxiliary numerical procedure.

2 Preliminary materials

In this Section we present two facts that are used in the convergence analysis. Projection has been extensively studied and we recall, briefly, some of its properties.

Lemma 1. *Take $x \in C$ and $\alpha > 0$, and define $x(\alpha) := P_C\left(x - \alpha \nabla f(x)\right)$, then*

i) $\langle x(\alpha) - x + \alpha\nabla f(x), y - x(\alpha)\rangle \ge 0$, *for all* $y \in C$;

ii) $\langle \nabla f(x), x - x(\alpha)\rangle \ge \dfrac{\|x(\alpha) - x\|^2}{\alpha}$;

iii) $x = x(\alpha)$ *if and only if* x *is stationary point for* (1).

Proof. See [7]. □

An immediate fact on descent directions is the following.

Proposition 1. *Take $\delta \in (0,1)$, $x \in C$ and $v \in \mathbb{R}^n$ such that $\langle \nabla f(x), v\rangle < 0$. Then there exists $\bar\zeta$ such that $f(x + \zeta\nabla f(x)) < f(x)$ for all $\zeta \in (0, \bar\zeta]$.*

Proof. The result follows from the differentiability of f. □

3 Statement of Algorithm A

We present the projected gradient method with Strategy (c) for solving constrained quasi-convex minimization problems. The algorithm requires the following exogenous parameters: $\delta \in (0,1)$, $\hat{\beta}$ and $\tilde{\beta}$ satisfying $0 < \hat{\beta} \leq \tilde{\beta}$ and a sequence $\{\beta_k\} \subset [\hat{\beta}, \tilde{\beta}]$.

Algorithm A

Initialization: Take $x^0 \in C$.

Iterative step: Given x^k, compute

$$z^k = x^k - \beta_k \nabla f(x^k).$$

If $x^k = P_C(z^k)$, then stop. Otherwise, set

$$x^{k+1} = \alpha_k P_C(z^k) + (1-\alpha_k)x^k, \tag{7}$$

where α_k is computed by the following Armijo search along the feasible direction through x^k and $P_C(z^k)$. Set

$$z^{k,j} = 2^{-j} P_C(z^k) + \left(1 - 2^{-j}\right) x^k, \quad \text{for } j = 0, 1, \dots,$$

compute

$$j(k) = \min \left\{ j \in \mathbb{N} : \ \delta 2^{-j} \langle \nabla f(x^k) \, , \, x^k - P_C(z^k) \rangle \leq f(x^k) - f(z^{k,j}) \right\},$$

and define

$$\alpha_k = 2^{-j(k)}.$$

Next we establish that Algorithm A is well defined and it is descent one.

Proposition 2. *Let $\{x^k\}$ and $\{z^k\}$ be the sequences generated by Algorithm A.*

i) x^k belongs to C for all k.

ii) If x^k is not stationary point, then $\langle \nabla f(x^k), P_C(z^k) - x^k \rangle < 0$.

iii) $j(k)$ is well defined.

Proof. For (i), using induction in k, (ii) follows from Lemma 1(ii)-(iii), and (iii) is the classical result consequence of (ii). □

4 Convergence analysis

Stationarity of cluster point of the sequence is a classical result.

Proposition 3. *Every cluster point of* $\{x^k\}$ *is stationary.*

Proof. See Proposition 4 in [12]. □

Two comment are in order. First, no result proved up to this point requires any assumption over f, only f must be continuously differentiable. Second, all these results are rather standard and well known. The novelty of this paper occurs in that follows.

Next, we present a vital lemma.

Lemma 2. *For all* $x \in C$ *and all* k*, we have*

$$\|x^{k+1} - x\|^2 \le \|x^k - x\|^2 + m\left\{f(x^k) - f(x^{k+1})\right\} + 2\alpha_k\beta_k\langle \nabla f(x^k), x - x^k\rangle,$$

where $m = \dfrac{2\,\tilde{\beta}}{\delta}$.

Proof. Using (7) and elementary algebra, we get

$$\|x^{k+1} - x^k\|^2 + \|x^k - x\|^2 - \|x^{k+1} - x\|^2 = 2\langle x^k - x^{k+1}, x^k - x\rangle = 2\alpha_k\langle P_C(z^k) - x^k, x - x^k\rangle. \quad (8)$$

Lemma 1(i) and (3) imply that

$$\begin{aligned} 0 &\le \langle P_C(z^k) - x^k + \beta_k\nabla f(x^k), x - P_C(z^k)\rangle \\ &= \langle P_C(z^k) - x^k + \beta_k\nabla f(x^k), x - x^k\rangle + \langle P_C(z^k) - x^k + \beta_k\nabla f(x^k), x^k - P_C(z^k)\rangle. \end{aligned}$$

Then,

$$\begin{aligned} \langle P_C(z^k) - x^k, x - x^k\rangle &\ge \beta_k\langle \nabla f(x^k), x^k - x\rangle - \langle P_C(z^k) - x^k + \beta_k\nabla f(x^k), x^k - P_C(z^k)\rangle \\ &= \beta_k\langle \nabla f(x^k), P_C(z^k) - x\rangle + \|P_C(z^k) - x^k\|^2. \end{aligned} \quad (9)$$

Combining now (8) and (9), and taking into account (7), we obtain that

$$\begin{aligned} \|x^{k+1} - x\|^2 &\le \|x^k - x\|^2 + (1 - \frac{2}{\alpha_k})\|x^{k+1} - x^k\|^2 + 2\alpha_k\beta_k\langle \nabla f(x^k), x - P_C(z^k)\rangle \\ &\le \|x^k - x\|^2 + 2\alpha_k\beta_k\langle \nabla f(x^k), x - P_C(z^k)\rangle \\ &= \|x^k - x\|^2 + 2\alpha_k\beta_k\langle \nabla f(x^k), x^k - P_C(z^k)\rangle + 2\alpha_k\beta_k\langle \nabla f(x^k), x - x^k\rangle. \end{aligned}$$

Then by (3)

$$\|x^{k+1}-x\|^2 \leq \|x^k-x\|^2 + \frac{2\beta_k}{\delta}\left\{f(x^k)-f(x^{k+1})\right\} + 2\alpha_k\beta_k\langle\nabla f(x^k), x-x^k\rangle.$$

□

The aim of this paper is to analyze the behavior of the projected gradient method with Armijo search along the feasible direction as strategy for the stepsize computation, i.e., Algorithm A, under the hypothesis of quasi-convexity of f or of pseudo-convexity of f. We present now the main convergence result in the quasi-convex case.

Theorem 1. *Assume that f is a quasi-convex function. Then, either*

i) $\{x^k\}$ converges to a stationary point, or

ii) $\lim_{k\to\infty}\|x^k\| = \infty$*, the problem 1 has not solution and* $\lim_{k\to\infty} f(x^k) = \inf\{f(x)\colon x \in C\}$.

Proof. There exist three different cases to analyze.

First: the sequence $\{x^k\}$ converges to $\bar{x}$. By Proposition 3, $\bar{x}$ is stationary point.

Second: the sequence $\{x^k\}$ is divergent and has at least one cluster point. Let $\bar{x}$ be cluster point of $\{x^k\}$. Since the sequence $\{f(x^k)\}$ is decreasing and f is continuos, we have that $\lim_{k\to\infty} f(x^k) = f(\bar{x})$ and $f(\overline{x}) \leq f(x^k)$, for all k. Using the quasi-convexity of f, we get $\langle\nabla f(x^k)\,,\,\overline{x}-x^k\rangle \leq 0$, for all k. It follows from Lemma 2 that

$$mf(\overline{x}) \leq \|x^{k+1}-\overline{x}\|^2 + mf(x^{k+1}) \leq \|x^k-\overline{x}\|^2 + mf(x^k), \tag{10}$$

for all k. By (10), the sequence $\left\{\|x^k-\overline{x}\|^2 + mf(x^k)\right\}$ is convergent. Using that one cluster point of this sequence is $mf(\bar{x})$, we obtain that $\|x^k-\bar{x}\|$ goes to zero. This fact is a contradiction with what we have assumed. Therefore this option is not possible.

Third: the sequence $\{x^k\}$ diverges and has not cluster points, i.e. $\lim_{k\to\infty}\|x^k\| = \infty$. Suppose $\bar{x}$ is solution of Problem (1). Then $\bar{x}$ is also stationary point. Using Lemma 2, we obtain

$$\|x^k-\bar{x}\|^2 + mf(x^k) \leq \|x^0-\bar{x}\|^2 + mf(x^0).$$

We conclude that $\{\|x^k-\bar{x}\|\}$ is bounded, contradicting $\lim_{k\to\infty}\|x^k\| = \infty$. Henceforth, Problem (1) can not has solution.

We claim that $\lim_{k\to\infty} f(x^k) = \inf\{x \in C : f(x)\}$. It is clear that $\lim_{k\to\infty} f(x^k) = \tilde{f} \geq \inf\{f(x) \colon x \in C\}$. Suppose that $\tilde{f} > \inf\{f(x) \colon x \in C\}$. By continuity of f, there exists $\tilde{x} \in C$ such that $f(\tilde{x}) = \tilde{f}$. Then $f(\tilde{x}) \leq f(x^k)$, for all k, because the sequence $\{f(x^k)\}$ is decreasing. Using the quasi-convexity of f and Lemma 2 with $x = \tilde{x}$, we get

$$\|x^k - \tilde{x}\|^2 + mf(x^k) \leq \|x^0 - \tilde{x}\|^2 + mf(x^0).$$

Again, we conclude that $\{\|x^k - \tilde{x}\|\}$ is bounded, contradicting $\lim_{k\to\infty} \|x^k\| = \infty$. □

Define $T := \{x \in C : f(x) \leq f(x^k), \ \ \forall k\}$. Note that if the problem (1) has solutions or $\{x^k\}$ has cluster points, then $T \neq \emptyset$.

Corollary 1. *Assume that f is a quasi-convex function. If $T \neq \emptyset$ then, $\{x^k\}$ converges to a stationary point.*

Proof. Take $\tilde{x} \in T$. Using the quasi-convexity of f and Lemma 2 with $x = \tilde{x}$, we get

$$\|x^k - \tilde{x}\|^2 + mf(x^k) \leq \|x^0 - \tilde{x}\|^2 + mf(x^0).$$

As in the proof of Theorem 1, we conclude that $\{\|x^k - \bar{x}\|\}$ and $\{x^k\}$ are bounded. Then, by Theorem 1, $\{x^k\}$ converges to a stationary point. □

In the pseudo-convex case we obtain the following result.

Corollary 2. *Assume that f is a pseudo-convex function. Then, $S^* \neq \emptyset$ if and only if there exists at least one cluster point of $\{x^k\}$. In this case $\{x^k\}$ converges to a stationary point.*

Proof. Follows from Theorem 1 and Corollary 1. □

Acknowledgment

The first auhor was partially supported by PROCAD-nf UFG/UnB/IMPA. The second author was partially supported by PRONEX-Faperj/CNPq.

References

[1] Ya.I. Alber, A. N. Iusem, Extension of subgradient techniques for nonsmooth optimization in Banach spaces, Set-Valued Analysis 9 (2001) 315-335.

[2] Ya. I. Alber, A.N. Iusem, M.V. Solodov, On the projected subgradient method for nonsmooth convex optimization in a Hilbert space, Mathematical Programming 81 (1998) 23-37.

[3] J. Y. Bello Cruz, A. N. Iusem, Convergence of direct methods for paramonotone variational inequalities, (to be published in Computational Optimization and Applications).

[4] D. Bertsekas, Nonlinear Programming, Athena Scientific, Belmont (1995).

[5] D. Bertsekas, On the Goldstein-Levitin-Polyak gradient projection method, IEEE Transactions on Automatic Control 21 (1976) 174-184.

[6] R. Burachik, L. M. Graña Drummond, A. N. Iusem, B. F. Svaiter, Full convergence of the steepest descent method with inexact line searches, Optimization 32 (1995) 137-146.

[7] P. H. Calamai, J. J. Moré, Projected gradient methods for linearly constrained problems. Mathematical Programming 39 (1987) 93-116.

[8] Y. C. Cheng, On the gradient-projection method for solving the nonsymmeytric linear complementary problem, Journal of Optimization Theory and Applications 43 (1984) 527-541.

[9] C. C. Chou, K. F. Ng, J. S. Pang, Minimizing and stationary sequences of optimization problems. Thecnical Reports, The Johns Hopkins University, Baltimore, Maryland, USA (1996).

[10] E. N. Gafni, D. Bertsekas, Convergence of a gradient projection method, Technical Report LIDS-P-1201, Laboratory for Information and Decision Systems, M.I.T. (1982).

[11] A. A. Goldstein, Convex programming in Hilbert space, Bulletin of the American Mathematical Society 70 (1964) 709-710.

[12] A. N. Iusem, On the convergence properties of the projected gradient method for convex optimization, Computational and Applied Mathematics 22 (2003) 37-52.

[13] K. C. Kiwiel, K. Murty, Convergence of the steepest descent method for minimizing quasiconvex functions, Journal of Optimization Theory and Applications 89 (1996) 221-226.

[14] E. S. Levitin, B. T. Polyak, Constrained minimization methods, Zhurnal Vychisliteľno ĭ Matematiki i Matematichesko ĭ Fiziki 6 (1966) 787-823.

[15] C. Wang, N. Xiu, Convergence of the gradient projection method for generalized convex minimization, Computational Optimization and Applications 16 (2000) 111-120.

[16] C. Y. Wang, Convergence characterization of both new pivote method and simplified Levitin-Polyak gradient projection method, Acta Mathematicae Applicatae Sinica 4 (1981) 37-52.

[17] C. Y. Wang, On convergence properties of an improved reduced gradient method, Ke Xue Tongbao 17 1030-1033.

[18] F. Wu, S. Wu, A modified Frank-Wolfe algorithm and its convergence properties, Acta Mathematicae Applicatae Sinica 11 (1995) 286-291.

[19] G.L. Xue, A family of gradient projection algorithm and their convergence properties, Acta Mathematicae Applicatae Sinica 10 (1987) 396-404.

[20] G. L. Xue, On convergence properties of a Least-Distance programming procedure for minimization problems under linear constraints, Journal of Optimization Theory and Applications 50 (1986) 365-370.