A Numerical Minimization Algorithm for Convex Unconstrained Problems [1]

I. Raykov [2] and M.Z. Nashed [3]

Virtual Meeting of Oklahoma-Arkansas Section of the MAA

April 9, 2021

Abstract We introduce a numerical algorithm for solving convex smooth minimization problems by reducing the number of independent variables. We present a comparative study with numerical examples.

Keywords: Maximal monotone operator, lower semicontinuous (l.s.c.) convex maps, differential inclusions (d.i.), optimization problems.

1. Introduction

We consider minimization problems for a class of convex functions on $H \times H$, where H is a real Hilbert space. We define and study a monotone operator on H using a convex function on $H \times H$, presented in [2].

Definition 1.1.(see [2]) For each convex function f on $H \times H$ we define the operator

[1] We thank to Professor Bruce Lundberg from CSU-Pueblo and Professor Hans D. Mittelmann from ASU for the cooperation and support.

[2] University of Arkansas at Pine Bluff, AR, USA

[3] University of Central Florida, FL, USA

$T_f x := \{u \in H | (u, x) \in \partial f(x, u)\}$ for each $x \in H$.

Fitzpatrick proved that T_f is a monotone operator (see [2] Proposition 2.2).

We consider next optimization problem

$$\inf_{(x,u) \in H \times H} f(x, u) = f^{inf} \tag{1}$$

supposing its solution exists, i.e. we suppose that there exists at least one point $(x^*, u^*) \in H \times H$ for which the solution of (1) exists. Therefore we can solve (1) by finding the zero(s) of $T_f x$.

2. Application

By using the above and the convexity we create a simplified algorithm for solving unconstrained convex smooth minimization problems by reducing the number of independent variables. We present two iterative procedures of one independent variable for solving convex minimization problems of two independent variables. We show with a comparative study that the new iterative procedures solve the convex minimization problems faster with the same precision than the existing procedures.

3. Comparison algorithms

We compare the results of our algorithm with the results of four other algorithms. Two of them are well-known variants of Newton method from MATLAB Optimization Tool: FMINUNC and FMINSEARCH. The third one is The Proximal-point algorithm presented by Bauschke and Combetts (see Theorem 23.41 from [1]). Similar algorithm is presented also by Pavel an Raykov (see [3]). The fourth one is The Regularized Minimization Algorithm introduced also by Bauschke and Combetts (see Theorem 27.23 from [1]).

We present in short the last two algorithms and the statements of the theorems of their covergence here.

We introduce next Theorem 3.1 (Theorem 23.41 from [1]) proving the convergence of the Proximal-point algorithm.

Theorem 3.1. (Theorem 23.41 (Proximal-point algorithm) from [1]) *Let* $A : H \to 2^H$ *be a maximally monotone operator such that* $zerA \neq \emptyset$*, let* $(\gamma)_{n\in N}$ *be a sequence in* R_{++} *such that* $\Sigma_{n\in N}\gamma_n^2 = +\infty$*, and let* $x_0 \in H$*. Set* $(\forall n \in N)x_{n+1}J_{\gamma_n A}x_n$ *Then the following hold: (i)* $(x_n)_{n\in N}$ *converges weakly to a point in* $zerA$*.*

(ii) Suppose that A *is uniformly monotone on every bounded subset of* H*. Then* $(x_n)_{n\in N}$ *coverges strongly to the unique point in* $zerA$*.*

In the above theorem $J_{\gamma_n A}$ is the *resolvent* of A: $J_{\gamma_n A} = (Id + \gamma_n A)^{-1}$. For the proof see [1].

Method of Regularization of Minimization Problems (see [1])

Let $f \in \Gamma_0(H)$ and suppose that Argmin $f \neq \emptyset$, i.e., the minimization problem $\min_{x\in H} f(x)$ has at least one solution. In order to obtain a specific minimizer, one can introduce, for every $\varepsilon \in]0, 1[$, the regularized problem $\min_{x\in H} f(x) + \varepsilon g(x)$, where $g \in \Gamma_0(H)$. The objective is to choose the regularization function g such that the last minimization problem admits a unique solution x_ε and such that the net $(x_\varepsilon)_{\varepsilon\in]0,1[}$ converges to a specific point of Argmin f. We choose $g = (1/2)\|\cdot\|^2$(*Tykhonov Regu larization* framework.)

In this case the regularized is equivalent to find $x_\varepsilon \in H$ such that $0 \in \partial f(x_\varepsilon) + \varepsilon x_\varepsilon$. If we denote by x_0 the minimum norm minimizer of f, Theorem 23.44(i) ([1]) proves that $x_\varepsilon \to x_0$ as $\varepsilon \downarrow 0$.

The next Theorem 3.2 (Theorem 27.23 from [1]) explores the asymptotic behavior of the curve $(x_\varepsilon)_{\varepsilon\in]0,1[}$

for more general choices of the regularization function g.

Theorem 3.2. (Theorem 27.23 from [1]) *Let f and g be in $\Gamma_0(H)$. Suppose that* Argmin $f \cap$ dom $g \neq \emptyset$ *and that g coercive and strictly convex. Then g admits a unique minimizer x_0 over* Argmin$f$ *and, for every $\varepsilon \in]0,1[$, the regularized problem $\min_{x \in H} f(x) + \varepsilon g(x)$ admits a unique solution x_ε. Moreover, following hold:*

(i) $(x_\varepsilon) \rightharpoonup x_0$ as $\varepsilon \downarrow 0$. (ii) $g(x_\varepsilon) \to g(x_0)$ as $\varepsilon \downarrow 0$.

(iii) Suppose that g is uniformly convex on every closed ball in H. Then $(x_\varepsilon) \to x_0$ as $\varepsilon \downarrow 0$.

For the proof see [1].

4. Examples

We present now next two examples:

Example 1. Let us solve the unconstrained convex optimization problem : min $f_p(x,y)$, where $f_p(x,y) = 4(x+p)^2 + 4y^2 - 4(x+p)y - 12y$, p is a parameter of disturbances. The partial derivatives should be equal to zero at the min $f_p(x,y)$:

$f_{p_x}(x,y) = 8(x+p) - 4y = 0$ and $f_{p_y}(x,y) = 8y - 4(x+p) - 12 = 0.$

We have to find the conditions for the parameter of distubances such that to have a monotone convergence algorithm. Because both partial derivatives should be equal to zero at the solution point, then we have

$$y_p = 2(x+p) \quad \text{and} \quad y_p = \frac{x+p+3}{2} \tag{2}$$

They are monotone, because $f_p(x,y)$ is convex. These two lines intersect at different solution points for a given value of the parameter p and we use a sequence $\{p_k\}$ to create an iterative procedure depending only on x_{p_k} and the parameter of disturbances p_{k+1} on each step. As the solution is a unique one when $p = 0$ we can find sufficient conditions for the parameter of disturbances p to ensure convergence. We create an iterative procedute for $x_{p_{k+1}}$ for finding the minimum of $f_p(x, y)$:

$$x_{p_{k+1}} = x_{p_k} - \frac{\frac{9}{4}(x_{p_k} + p_{k+1} - 1)^2}{\frac{9}{2}(x_{p_k} + p_{k+1} - 1)} = x_{p_k} - \frac{x_{p_k} + p_{k+1} - 1}{2}$$

To ensure convergence we have to have $\max \|(x_{p_k}, y_{p_k}) - (x_{p_{k+1}}, y_{p_{k+1}})\| < \varepsilon_k$, where the sequence $\{\varepsilon_k\}$ is convergent to zero, i.e. the $\liminf_{k\to\infty} \varepsilon_k = 0^+$, $\varepsilon_k > 0$ for every positive integer k. Because

$\max |x_{p_k} - x_{p_{k+1}}| \leq \max \|(x_{p_k}, y_{p_k}) - (x_{p_{k+1}}, y_{p_{k+1}})\| < \varepsilon_k$, means that $|x_{p_k} - x_{p_{k+1}}| = |x_{p_k} - 1 + p_{k+1}| < 2\varepsilon_k$. We receive $|x_{p_k} - 1 + p_{k+1}| \leq |x_{p_k} - 1| + |p_{k+1}| < 2\varepsilon_k^1$ where the sequence $\{\varepsilon_k^1\}$ is also convergent to zero, i.e. the $\liminf_{k\to\infty} \varepsilon_k^1 = 0^+$, $\varepsilon_k^1 \geq \varepsilon_k > 0$ for every positive integer k, such that when $|x_{p_k} - 1| + |p_{k+1}| < 2\varepsilon_k^1$ then $|x_{p_k} - 1 + p_{k+1}| < 2\varepsilon_k$, or $0 \leq |p_{k+1}| < 2\varepsilon_k^1 - |x_{p_k} - 1|$, or $|x_{p_k} - 1| - 2\varepsilon_k^1 \leq p_{k+1} \leq 2\varepsilon_k^1 - |x_{p_k} - 1|$. From equations (2) it follows that when p_k approaches 0 and x_{p_k} approaches $x^* = 1$ then y_{p_k} approaches $y^* = 2$. We can simplify the iterative procedure by presenting the disturbances on each step k for any $x_k \in \{x_{p_k}\}$ with addition a one parameter p_{k+1}: $x_{p_{k+1}} = x_k - \frac{x_k - 1}{2} + p_{k+1}$. To have for any $x_{k+1} \in \{x_{p_{k+1}}\}$ $\lim\limits_{k\to\infty} |x_k - x_{k+1}| \to 0$ p_{k+1} should satisfy next inequalities: $-\left|\frac{x_k - 1}{2}\right| < p_{k+1} < \left|\frac{x_k - 1}{2}\right|$ for every $x_k \in \{x_{p_k}\}$.

Example 2. We consider the following unconstrained convex optimization problem in which the surface is flatter close to the solution point: min $f_p(x, y)$, where

$f_p(x, y) = (x + p)^6 + e^{(x+p)} + (x + p)^2 + 2(x + p)y + y^2 + y^4$, p is again a parameter of disturbances.

The partial derivatives should be equal to zero at the min $f_p(x,y)$:

$$f_{p_x}(x,y) = 6(x+p)^5 + e^{(x+p)} + 2(x+p) + 2y = 0 \qquad \text{and} \qquad f_{p_y}(x,y) = 4y^3 + 2y + 2(x+p) = 0.$$

Because both partial derivatives should be equal to zero at the solution point, then we have

$$y_p = -\frac{6(x+p)^5 + e^{(x+p)} + 2(x+p)}{2} \quad \text{and} \quad y_p = \Big(\frac{6(x+p)^5 + e^{(x+p)}}{4}\Big)^{1/3} \tag{3}$$

These two lines intersect at different solution points for a given value of the parameter p and we use a sequence $\{p_k\}$, $(|p_k| \downarrow 0)$ to create an iterative procedure depending only on x_{p_k} and the parameter of disturbances p_{k+1} on each step. As the solution is a unique one when $p = 0$ we can find sufficient conditions for the parameter of disturbances p to ensure convergence. We create an iterative procedute for $x_{p_{k+1}}$ for finding the minimum of $f_p(x,y)$: Let

$$A_{p_{k+1}} = \Big(\frac{6(x_{p_k}+p_{k+1})^5 + e^{(x_{p_k}+p_{k+1})}}{4}\Big)^{1/3}, \quad B_{p_{k+1}} = \frac{6(x_{p_k}+p_{k+1})^5 + 2(x_{p_k}+p_{k+1}) + e^{(x_{p_k}+p_{k+1})}}{2},$$

$$F_{p_{k+1}} = \Big\{A_{p_{k+1}} + B_{p_{k+1}}\Big\}^2, \quad G_{p_{k+1}} = 2\Big\{A_{p_{k+1}} + B_{p_{k+1}}\Big\}, \quad \text{and}$$

$$H_{p_{k+1}} = \Big\{\Big(\frac{1}{12}\Big)\Big[\frac{4}{6(x_{p_k}+p_{k+1})^5 + e^{(x_{p_k}+p_{k+1})}}\Big]^{2/3}(30(x_{p_k}+p_{k+1})^4 + e^{(x_{p_k}+p_{k+1})}) + 0.5(30(x_{p_k}+p_{k+1})^4 + e^{(x_{p_k}+p_{k+1})} + 2)\Big\}$$

Then $x_{p_{k+1}} = x_{p_k} - \frac{F_{p_{k+1}}}{G_{p_{k+1}} H_{p_{k+1}}}$. To ensure convergence we should have $\max\|(x_{p_k}, y_{p_k}) - (x_{p_{k+1}}, y_{p_{k+1}})\| < \varepsilon_k$, where the sequence $\{\varepsilon_k\}$ is convergent to zero, i.e. the $\liminf_{k\to\infty} \varepsilon_k = 0^+$, $\varepsilon_k > 0$ for every positive integer k. Because $\max|x_{p_k} - x_{p_{k+1}}| \le \max\|(x_{p_k}, y_{p_k}) - (x_{p_{k+1}}, y_{p_{k+1}})\| < \varepsilon_k$, means that $|x_{p_k} - x_{p_{k+1}}| = \frac{F_{p_{k+1}}}{|G_{p_{k+1}}| H_{p_{k+1}}} < \varepsilon_k$, since $F_{p_{k+1}} \ge 0$ and $H_{p_{k+1}} \ge 0$ for every p_{k+1} satisfying the convergnce conditions. Therefore we have to restrict p_{k+1} such that, when $|p_{k+1}| < \varepsilon^1_{k+1}$, then $\frac{F_{p_k}}{|G_{p_k}| H_{p_k}} < \varepsilon_k$. Here $\{\varepsilon^1_k\}$ is a decreasing sequence with positive elemets that ensures these requirements. From equations (3) it follows that when p_k approaches 0 and x_{p_k} approaches $x^* = -0.5493$ then y_{p_k} approaches $y^* = 0.4107$.

We can simplify the iterative procedure by presenting the disturbances on each step k for any $x_k \in \{x_{p_k}\}$ with addition a one parameter p_{k+1}. We use the following simplified analogical notarions: $F_k = \left\{\left(\frac{6x_k^5+e^{x_k}}{4}\right)^{1/3} + \frac{6x_k^5+2x_k+e^{x_k}}{2}\right\}^2$, $\quad G_k = 2\left\{\left(\frac{6x_k^5+e^{x_k}}{4}\right)^{1/3} + \frac{6x_k^5+2x_k+e^{x_k}}{2}\right\}$, and $H_k = \left\{\left(\frac{1}{12}\right)\left[\frac{4}{6x_k^5+e^{x_k}}\right]^{2/3}(30x_k^4 + e^{x_k}) + 0.5(30x_k^4 + e^{x_k} + 2)\right\}$. Then we receive the iterative procedure: $x_{p_{k+1}} = x_k - \frac{F_k}{G_k H_k} + p_{k+1}$. To have for any $x_k \in \{x_{p_k}\}$ and $x_{k+1} \in \{x_{p_{k+1}}\}$ $\lim\limits_{k\to\infty} |x_k - x_{k+1}| \to 0$ where $|x_k - x_{k+1}| = \left|\frac{F_k}{|G_k|H_k} - p_{k+1}\right|$, p_{k+1} should satisfy next inequalities:

$$-\frac{F_k}{|G_k|H_k} < p_{k+1} < \frac{F_k}{|G_k|H_k}$$

5. Numerical Results

We present in this Section a comparative study of results with our algorithm with the results from four other algorithms for Examples considered at the end of Section 5 using MATLAB. Our results are for both Examples with initial point $(x_0, y_0) = (5, 5)$ and with precision $\varepsilon = 0.000001$.

Numerical Results for Example 1.

a) Ther results with our algorithm are: $x = 1.0000$ and $y = 2.0000$ for time $t = 1.8896 \times 10^{-4}$ sec.

b) The results with FMINUNC (MATLAB) are: $x = 1.0000$ and $y = 2.0000$ for time $t = 0.0089$ sec.

c) The results with FMINSEARCH (MATLAB) are: $x = 1.0000$ and $y = 2.0000$ for time $t = 0.0025$ sec.

d) The results with Proximal-point algorithm are: $x = 1.0000$ and $y = 2.0000$ for time $t = 2.0426 \times 10^{-3}$ sec.

e) The results with Regularization algorithm are: $x = 1.0000$ and $y = 2.0000$ for time $t = 0.0143$ sec.

These results are summarized in the next Table 1.

Table 1.

Algorithm	x - results	y - results	time
Our Algorithm	1.0000	2.0000	1.8896 $\times 10^{-4}$ sec.
FMINUNC	1.0000	2.0000	0.0089 sec.
FMINSEARCH	1.0000	2.0000	0.0026 sec.
Proximal-point	1.0000	2.0000	2.0426$\times 10^{-3}$ sec.
Regularization	1.0000	2.0000	0.0143 sec.

The advantages of our algorithm and the Proximal-point algorithm are obvious. See next Figure 1.

$$\mathbf{f(x,y)} = 4x^2 + 4y^2 - 4xy - 12y$$

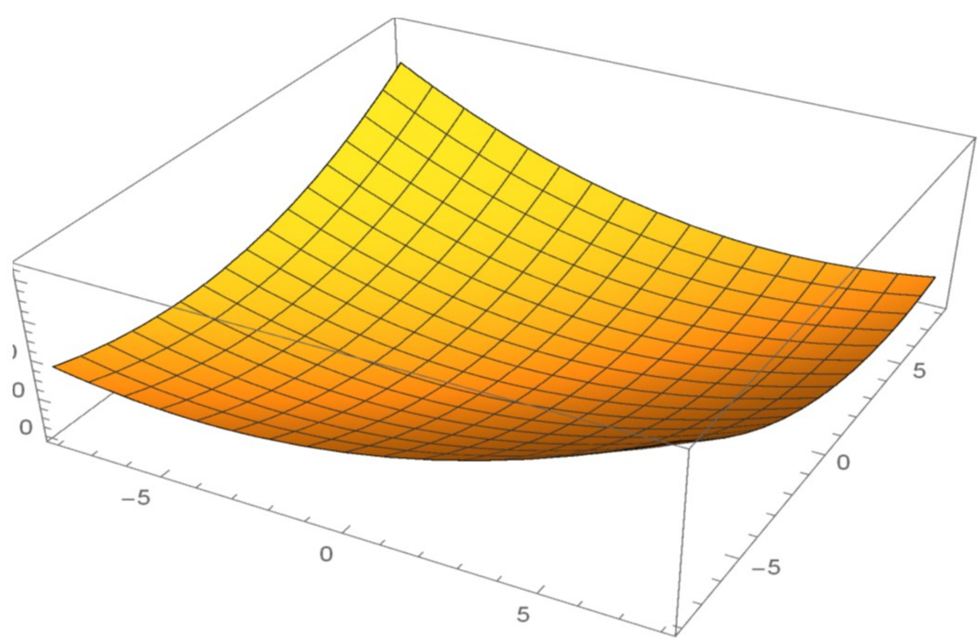

Figure 1.

Numerical Results for Example 2.

We compare our results with the results with the four other algorithms next:

a) The results with our algorithm are: $x = -0.5493$ and $y = 0.4107$ for time $t = 3.6133 \times 10^{-4}$ sec.

b) The results with FMINUNC (MATLAB) are: $x = -0.5487$ and $y = 0.4117$ for time $t = 0.0110$ sec.

c) The results with FMINSEARCH (MATLAB) are: $x = -0.5493$ and $y = 0.4107$ for time $t = 0.0034$ sec.

d) The results with Proximal-point algorithm are: $x = -0.5493$ and $y = 0.4107$ for time $t = 3.2726 \times 10^{-3}$ sec.

e) The results with Regularization algorithm are: $x = -0.5469$ and $y = 0.4120$ for time $t = 0.0175$ sec.

These results are summarized in the next Table 2.

Table 2.

Algorithm	x - results	y - results	time
Our Algorithm	-0.5493	0.4107	3.6133 $\times 10^{-4}$ sec.
FMINUNC	-0.5487	0.4117	0.0110 sec.
FMINSEARCH	-0.5493	0.4107	0.0034 sec.
Proximal-point	-0.5493	0.4107	3.2726$\times 10^{-3}$ sec.
Regularization	-0.5469	0.4120	0.0175 sec.

The advantages of our algorithm and of the Proximal-point algorithm are clear again. See next Figure 2.

$$\mathbf{f(x,y)} = x^6 + e^x + x^2 + 2xy + y^2 + y^4$$

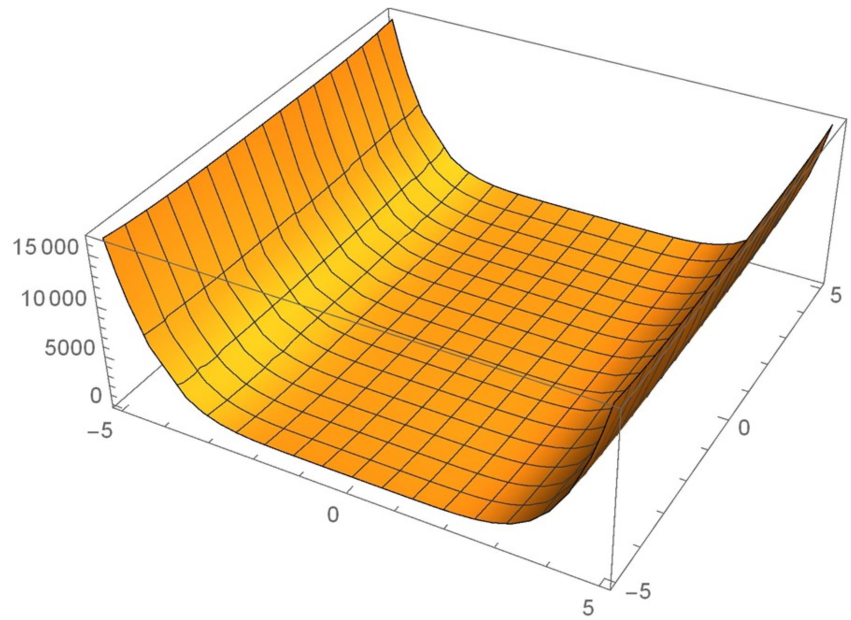

Figure 2.

References

[1] Bauschke, H. H., Combetts, P. L. (2017). *Convex Analysis and Monotone Operator Theory in Hilbert Spaces*. Springer.

[2] Fitzpatrick, S. (1988). Representing monotone operators by convex functions. *Workshop/Miniconference on Functional Analysis and Optimization (Canberra, 1988), Proc.Centre for Mathematical Analysis, Australian National University, Canberra*. 20: pp.59-65.

[3] Pavel, N., Raykov, I. (2008). Parametric proximal-point algorithms. *JOTA*. 139(1): pp.85-107.