Journal of Physics:
Conference Series

PAPER • OPEN ACCESS

A proximal bundle algorithm for solving generalized variational inequalities with inexact data

To cite this article: Siqi Zhang *et al* 2024 *J. Phys.: Conf. Ser.* **2905** 012020

A proximal bundle algorithm for solving generalized variational inequalities with inexact data

Siqi Zhang[1], Ming Huang[1,2,*], Yongxiu Feng[1], Sida Lin[1] and Yajing Zhang[1]

[1] School of Science, Dalian Maritime University, Dalian 116026, China
[2] Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong, China

[*]Corresponding author's e-mail: huangming0224@163.com.

Abstract. This paper proposes a proximal bundle algorithm based on the proximal point method for solving generalized variational inequalities with inexact data. First, the explanation of subgradient values and function inexactness is given. The algorithm's fundamental steps are then presented. Ultimately, the algorithm's convergence is confirmed under specific circumstances.

1. Introduction

Variational inequality is a generalization and development of classical variational problems. It originated in the 19th century and was developed to some extent under the influence of the Lagrange multiplier method. Its theme can be simplified to construct a variational type so that it has the best solution in meeting relevant conditions, constraints, and other main formulas that reflect the essence of the problem. Therefore, it is widely used to solve optimization problems, such as equilibrium problems, operational research problems, and urban traffic network modeling problems, so it is a very useful analysis algorithm. As an important research direction in the field of mathematics, it has a wide application background and rich historical development. Therefore, the study of variational inequalities not only has theoretical value but also has extensive application prospects.

We are concerned with the general variational inequality problem[1] $GVI(K,\Xi)$: $x^* \in K$ and $\varpi(x^*) \in \Xi(x^*)$ are found, and subject for every $x \in K$ is:

$$\left\langle \varpi(x^*), x - x^* \right\rangle + f(x) - f(x^*) \geq 0, \tag{1}$$

where Ξ is a monotone multivalued operator which is defined on a real Hilbert space H equipped with the inner product $\langle \cdot,\cdot \rangle$, $f: \mathrm{H} \to R \cup \{+\infty\}$ is a lower semi-continuous (l.s.c.) proper convex function and the set $K \subseteq \mathrm{H}$ is a nonempty closed subset such that $K \subseteq \mathrm{int}(dom f)$. In addition, this problem has been shown to have at least one solution.

The Generalized Variational Inequality Problem $GVI(K,\Xi)$ is closely connected to many problems, and is even equivalent to some of them. Later, we will introduce two classic examples[2].

First, we show that the general convex optimization can be regarded as a general variational inequality problem. The following questions are considered:

$$\min_{x \in K} f(x) \tag{2}$$

It is worth noting that $f: K \to R$ is a differentiable convex function defined on K. Assuming that the problem has an optimal solution x^*, then using the optimal condition of convex optimization, we know that the optimal solution of the problem is equivalent to that all feasible directions at x^* are not gradient descent directions. Then the feasible direction set and gradient descent set at point $x \in K$ are defined as $S_f := \{s \in R_n \mid s = x' - x,\ x' \in K\}$ and $S_d := \{s \in R_n \mid s^T \nabla f(x) < 0\}$. So we know that the optimal solution $x^* = argmin_{x \in C} f(x)$ for a convex optimization problem is equivalent to $x^* \in K,\ S_f \cap S_d = \varnothing$, that is for any $x' \in K$.

$$\nabla f(x^*)^T (x' - x^*) \geq 0. \tag{3}$$

Next, the variational inequalities are also equivalent to a more general class of nonlinear complementarity problems (NCP)[3]. Specifically, the nonlinear complementarity problem is described as:

$$x \geq 0,\ f(x) \geq 0,\ x^T f(x) = 0, \tag{4}$$

where $f = (f_1, f_2, ..., f_n)^T : R^n \to R^n$ is a continuously differentiable function. It is worth noting that if taking $K = R_+^n$, then $GVI(K,\Xi)$ is equivalent to (NCP).

In recent years, the development of variational inequalities has attracted the interest of many scholars, so both theoretical research and algorithm research have made great progress. This paper focuses on the algorithm aspect, so there are many excellent research results only from the aspect of the algorithm, for instance, the Projection method, Proximal point method, Splitting method, Auxiliary Principle, Bundle method, and Alternating Direction Method of Multipliers (ADMM)[6]. All of them are important and representative solutions for $GVI(K,\Xi)$. However, the research results of domestic research on $GVI(K,\Xi)$ are still relatively weak compared with other fields. Therefore, this paper hopes to explore some new research directions of $GVI(K,\Xi)$ algorithm on the basis of existing research.

This paper is organized as follows: Section 2 introduces novel inexact techniques and basic bundle methods; Section 3 suggests an executable inexact proximal bundle method; Section 4 shows how the algorithm converges; and Section 4 summarizes this paper and suggests further work.

2. Preliminary results

2.1. Inexact oracle

Throughout the following content, we make the following assumption[4][5]: it is supposed that the function f is a continuously differentiable and strongly convex function, a convex set K is the domain of the function f, and there exists a parameter $\rho > 0$ and for each $x, y \in K$.

$$f(x) + g(x)^T (x - y) + (\rho / 2)\|x - y\|^2 \leq f(y). \tag{5}$$

It is assumed that $f = \tau(\sigma(x))$, where $\sigma(x)$: $R^n \to R^m$ is continuously differentiable and τ: $R^m \to R$ convex, then we have $\partial f := \{\sum_{i=1}^{m} \delta_i \nabla \sigma_i \mid \delta \in \partial(\tau(\sigma(x)))\}$. Obviously, computing a subgradient $g(x) \in \partial f(x)$ is great costly great, then we try to find an inexact oracle to replace the exact value. Therefore, we assume that we have an approximation $\nabla \sigma'$ of $\nabla \sigma$ satisfying $\|\nabla \sigma' - \nabla \sigma\| \leq \kappa(\alpha)$ (typically $\kappa(\alpha) \to 0$ when $\alpha \to 0$). Then we define the approximate subgradients as $\hat{g}(x) := \sum_{i=1}^{m} \delta_i \cdot \nabla \sigma_i'(x)$, $\delta_i \in \partial \tau(\sigma(x))$, hence,

$$\begin{aligned} f(x)+\hat{g}(x)^T(y-x) &= f(x)+g(x)^T(y-x)+\|\xi\|\cdot\|\nabla\sigma'(x)-\nabla\sigma(x)\|\cdot\|y-x\| \\ &= f(y)-\frac{\rho}{2}\|y-x\|^2+\|\xi\|\cdot\|\kappa(\alpha)\|\cdot\|y-x\| \\ &\le f(y)+\frac{1}{2\rho}\|\xi\|^2\|\kappa(\alpha)\|^2. \end{aligned} \tag{6}$$

From the above derivation, we obtain $\hat{g}(x)\in\partial_\varepsilon f(x)$ when $\varepsilon=\frac{1}{2\rho}\|\xi\|^2\|\kappa(\alpha)\|^2$, namely, $\hat{g}(x)$ represents an ε -subgradients of f at x . At the same time, from the local boundedness of $\partial(\tau(\sigma(x)))$ (which implies that $\|\xi\|$ is bounded), we infer that ε is locally bounded. x remains in K , which is a bounded subset of R^n , and a uniform bound $\delta>0$ exists such that $\varepsilon\le\delta$.

And then, for the given positive number η_i , at each given point $x\in K$, we can find some $\tilde{f}\in R$ such that $\tilde{f}(x)=f(x)-\eta_i$, where $\eta_{i+1}=\gamma_i\eta_i$, $\{\gamma_i\}_{i\in N_+}$ is a non-increasing sequence and satisfies $0<\gamma_i<1$.

2.2. Bundle strategy

Correa and Lemarechal introduced the bundle method in the 20th century[10]. The method's approach is to use a piecewise linear convex function f that is generated step by step to approximate the function at the most recent iteration. Iterations are only advanced when the approximation is sufficient.

The method yields two sequences: the first, denoted as a series of candidate points and represented by $\{y_i\}_{i\in I_k}$, where *i* is a non-empty index set, is made up of sample points that are used to define the model. The second sequence, known as the stable center sequence and defined by $\{x^k\}_{k\in N_+}$, is made up of sample locations that is considerably lower the goal function f .

Assuming that x^k is the stable center, we approximate f below by a polyhedral model $\hat{f}_i^k$ defined by:

$$\hat{f}_i^k := \tilde{f}(y_i)+\max_{i\in I_l}\{\langle\hat{g}(y_i),x-y_i\rangle-\varepsilon\}, \tag{7}$$

where $I_k=\{0,1,...,k\}$ represents the iterated internal indicator set. Then due to the instability of the cutting-planes model, the model usually becomes coarser when farther away from x^k . For this, the proximal term t^k is introduced, and appropriate adjustments will confine the model to the region where it is a reliable approximation. Then (1.1) becomes:

$$GVI-s \qquad \min_{x\in K}\{\hat{f}_i^k(x)+\langle\varpi(x),x-x^k\rangle+t^{-k}[s(x)-s(x^k)-\langle\nabla s(x^k),x-x^k\rangle]\}, \tag{8}$$

with strongly convex functions $s(x)$ and $\{t^k\}_{k\in N_+}$ being positive sequences.

Here we offer several helpful components to ensure inexact optimality and limit the amount of information in the bundle. It is n necessary to notice that there are multipliers $\{\alpha_i\}_{i\in I_k}$, also known as Lagrange multipliers, such that:

$$G_f=\sum_{i\in I_k}\alpha_i^k\hat{g}_i^k, \tag{9}$$

$$\sum_{i\in I_k}\alpha_i^k=1,\ \alpha_i^k\ge 0, \tag{10}$$

$$\alpha_i^k[\hat{f}_i^k(y_i)-\tilde{f}(y_i)]=0. \tag{11}$$

The subgradient relation above gives us the aggregation linearization function $\overline{f}_i^k$ of $\hat{f}_i^k$, which is denoted as:

$$\overline{f}_i^k(x^k):=\hat{f}_i^k+\langle G_f(x^k),\ x^k-y_i\rangle, \tag{12}$$

with the linearization error $e^k:=f(x^k)-\overline{f}(x^k)$.

Hence, the stability is usually updated according to the following form[7]:

$$\upsilon^k := \tilde{f}(x^k) - \hat{f}_i^k(y_i), \tag{13}$$

We let $\beta \in (0,1)$ be a parameter. If $f(y_i) \le f(x^k) - \beta\upsilon^k$, then $x^{k+1} = y_i$; otherwise, $x^{k+1} = x^k$.

First of all, it is worth noting that the descent test is worthless because of the inexact oracle settings in Sect. 2.1, which makes ν^k non-positive. As a result, the proximal parameter t^k must be increased. This is motivated by the following: Increasing t^k reduces the weight of the quadratic terms, raises the relative weight of $\hat{f}$, and lowers the model value of the candidate points until υ^k is positive. (If υ^k is not positive, then y_i does not adequately minimize the model). After that, iteration continues in the same manner as the basic bundle method. We refer to this procedure as **Noise Attenuation[8].**

3. Main algorithm

Below we give the general steps of the Proximal Bundle Algorithm.

Proximal Bundle Algorithm with inexact data (PBAI).

Step 0. (*Initialization*) Stopping tolerances $\mu_1, \mu_2 \ge 0$, a descent parameter $m \in (0,1)$, a stepsize $t_0 \ge t > 0$, and $\alpha \in [1/2,\ 1)$; Choosing a starting point $x^0 \in K$, and calling the oracle to compute $f(x^0),\ g(x^0)$; Setting $I_0 := \{0\},\ k = 0$ and $\text{na} := 0$.

Step 1. (*Trial point calculation*) Selecting a piecewise linear convex function $\hat{f} \le f$ and computing $GVI - s$ to get the unique solution y_i; Computing $\mathrm{T}^k := (x^k - y_i)/t^k$, $\mathrm{P}^k := f(x^k) - \hat{f}(y_i)$, and $e^k := \mathrm{P}^k - t^k \left\|\mathrm{T}^k\right\|^2$.

Step 2. (*Stopping criterion*) If $\left\|\mathrm{T}^k\right\|^2 \le \mu_1$ and $e^k \le \mu_2$, stopping.

Step 3. (*Inaccuracy detection*) Setting $t^k := 10t^k$, na:=k and proceeding to Step 1 if $e^k \le -\alpha t^k \left\|\mathrm{T}^k\right\|^2$.

Step 4. (*Descent test*) Calling the inexact oracle to calculate $(f(x^k),\ \hat{g}(x^k))$; If the descent (13) is met, then declaring the iterate to be serious, and setting $x^{k+1} = y_i$, $\text{na} := 0$; Otherwise, declaring the iterate to be null and setting $x^{k+1} = x^k$.

Step 5. (*Bundle Management*) Choosing $I_{k+1} \supseteq \{i \in I_k\} \cup \{k+1\}$.

Step 6. (*Stepsize updating and loop*) Choosing $t^{k+1} \ge t_k$ if the situation is deemed serious; Otherwise, if the iterate is deemed null, either setting $t^{k+1} := t_k$, or choosing $t^{k+1} \in [0.1t_k,\ t_k]$ such that $t^{k+1} \ge t^*$ if $\text{na} = 0$; Increasing k by 1 and going to Step 1.

Remark: In Step 1, the assumption of strong convexity of the function $s(x)$ ensures that $GVI - s$ has one and only one solution.

In Step 4, if the algorithm takes a null step, the simplest and most effective method is to add iteration information at the candidate point y_i to the model. The new model $\hat{f}_{i+1}^k$ obtained has a better approximation effect than $\hat{f}_i^k$, because it contains more information about the past iteration; otherwise, the algorithm takes a serious step. We only update the stable center x^{k+1} and the piecewise linear model $\hat{f}_i^k$ [8].

4. Convergence Results

We go over three cases of algorithm convergence in this section.

- An infinite degree of noise attenuation is produced by the algorithm.
- An infinite number of null steps are produced by the algorithm.

- An infinite number of null steps are produced by the algorithm.

The ensuing conclusions can be made in light of the findings and discussion above.

Theorem 3.1. (Infinite Loop of Noise Attenuation) It is suppose that the PBAI for $GVI(K,\Xi)$ yields x^k in its final significant iteration, the algorithm proceeds indefinitely between Step 1 and Step 3, then x^k represents an approximate optimal solution to $GVI(K,\Xi)$.

Theorem 3.2. (Infinite Loop of Null Steps) It is supposed that the PBAI for $GVI(K,\Xi)$ yields iteration x^k and k stays constant thereafter, namely, the algorithm only creates null steps, and then x^k represents an approximate optimal solution to $GVI(K,\Xi)$.

Theorem 3.3. (Infinite Loop of Serious Steps) It is supposed that the PBAI for $GVI(K,\Xi)$ produces an infinite series of serious steps $\{x^k\}_{k\in N_+}$, then each accumulation point of $\{x^k\}_{k\in N_+}$ represents an approximately optimal solution to $GVI(K,\Xi)$.

Theorem 3.3's result can be broadly broken down into three steps: first, it establishes the boundedness of the infinite sequence that the algorithm generates; second, it establishes the existence of some convergence points in the series; and third, it verifies that these convergence points are the solutions of $GVI(K,\Xi)$.

5. Conclusions and future work

For the general variational inequality problem with non-smooth and lower semi-continuous convex function, an executable Proximal Bundle Algorithm with inexact data is proposed, which approximates the original function by using a series of piecewise linear convex functions.

It is noted that in each iteration, the approximate model needs to meet certain conditions in order to be considered suitable. At the same time, due to inexact data, the predicted descent is not positive, triggering noise attenuation, so the parameters and errors should be updated according to specific criteria until the predicted descent value is not negative. Finally, under appropriate conditions, it is theoretically proved that the algorithm has good convergence.

To put it briefly, we present a new proximal bundle method that can be used to solve more non-smooth general variational inequality issues. As a result, there is greater universality in the suggested algorithm.

The convergence study demonstrates that the technique is capable of efficiently resolving the non-smooth convex general variational inequality problem. However, there are still certain topics that require more research:

- Can other classical problems be solved by using the proximal bundle algorithm of inexact data?
- How much of the algorithm's efficiency can be increased by selecting the appropriate parameters?
- Does this algorithm still hold true if the objective function is non-convex?

These issues still merit more investigation and study.

Acknowledgments

This work is financially supported by the National Key Research and Development Program of China under No.2022YFB3304600; the International Postdoctoral Exchange Fellowship Program 2020 by the Office of China Postdoctoral Council under No.2020003; the National Natural Science Foundation of China under No.11701063, 11901075, and 12071342; the Project by China Postdoctoral Science Foundation under No.2019M651091 and 2019M661073; the Fundamental Research Funds for the Central Universities of DMU under No. 3132023205, and 3132023206; Dalian Youth Science and Technology Star under No.2020RQ047; the Project by Natural Science Foundation of Hebei Province under No. A2020202030.

References

[1] Rockafellar R T, Wets R J B. Variational analysis [M]. Springer Science & Business Media, 2009.
[2] Rockafellar R T, Wets R J B. Variational analysis [M]. Springer Science & Business Media, 2009.

[3] Salmon G, Strodiot J J, Nguyen V H. A bundle method for solving variational inequalities [J]. SIAM Journal on Optimization, 2004, 14(3): 869-893.
[4] Shen J, Pang L P. An approximate bundle method for solving variational inequalities [J]. Communications in Optimization Theory, 2012, 1: 1-18.
[5] He B, Yuan X. On the direct extension of ADMM for multi-block separable convex programming and beyond: from variational inequality perspective [J]. Optimization-online, 2014, 4293.
[6] Huang M, He Y, Qiao P P, et al. Proximal decomposition of convex optimization via an alternating linearization algorithm with inexact oracles [J]. Journal of Industrial and Management Optimization, 2024, doi:10.3934/jimo.2024057.
[7] Huang M, Niu H M, Lin S D, et al. A redistributed proximal bundle method for nonsmooth nonconvex functions with inexact information [J]. Journal of Industrial and Management Optimization, 2023, 19(12): 8691-8708.
[8] Huang M, Yuan J, Lin S, et al. On Solving the Convex Semi-Infinite Minimax Problems via Superlinear Incremental Bundle Technique with Partial Inexact Oracle [J]. Asia-Pacific Journal of Operational Research, 2021, 38(5).
[9] Tasora A, Mangoni D, Benatti S, et al. Solving variational inequalities and cone complementarity problems in nonsmooth dynamics using the alternating direction method of multipliers [J]. International Journal for Numerical Methods in Engineering, 2021, 122(16): 4093-4113.
[10] Alakoya T O, Uzor V A, Mewomo O T. A new projection and contraction method for solving split monotone variational inclusion, pseudo monotone variational inequality, and common fixed point problems [J]. Computational and Applied Mathematics, 2023, 42(1): 3.