Journal of Convex Analysis
Volume 25 (2018), No. 4, 1121–1138

The Dual Gap Function and Error Bounds for Strongly Monotone Variational Inequalities

D. Aussel
Laboratoire PROMES, Université de Perpignan, 66100 Perpignan, France
aussel@univ-perp.fr

J. Dutta
Dept. of Economic Sciences, Indian Inst. of Technology, 208016 Kanpur, India
jdutta@iitk.ac.in

A. C. Xu
Massachusetts Institute of Technology, Cambridge, MA 02139, U.S.A.
andrewxu@mit.edu

Received: January 27, 2017
Accepted: July 10, 2017

In the literature of variational inequalities, there has been a lot of studies about the role of gap functions in the development of error bounds specially for the case where the variational inequality is described by a strongly monotone mapping. However the role of the dual gap function in devising error bounds, to the best of our knowledge, has not been thoroughly investigated. In this article we focus on the dual gap function for monotone variational inequalities. We highlight some properties of the dual gap function which are not shared by the other gap functions and also show how it can be used to develop error bounds for strongly monotone variational inequalities with convex and compact feasible sets.

Keywords: Variational inequalities, gap functions, error bounds, strongly monotone maps.

2010 Mathematics Subject Classification: 90C33, 90C25.

1. Introduction and Motivation

Consider a continuous vector function $f\colon \mathbb{R}^n \to \mathbb{R}^n$ and a non-empty closed convex set C. The variational inequality problem $VI(f,C)$ consists of finding $\bar{x} \in C$ such that

$$\langle f(\bar{x}), x - \bar{x}\rangle \geq 0, \quad \forall x \in C.$$

If f is assumed to be strongly monotone over C then the $VI(f,C)$ has a unique solution. Let us recall f is called strongly monotone over the set C if there exists $\mu > 0$ such that for all $x, y \in C$,

$$\langle f(y) - f(x), y - x\rangle \geq \mu ||y - x||^2.$$

ISSN 0944-6532 / $ 2.50

In this article we focus on variational inequalities where f is strongly monotone over the set C. Our aim is to develop error bounds in terms of the so called *dual gap function* which we will soon describe below. Error bounds are important since they allow to estimate the distance of a feasible point from the solution by providing an upper bound. This becomes helpful in numerical computation since it can be used as a stopping criterion or can also be used to judge the quality of the approximate solution generated by an iterative process. Gap functions are the main vehicles of devising error bounds for variational inequalities. A *gap function* h is a numerical function on $\mathbb{R}^n$ such that $h(x) \geq 0$ for all $x \in C$ and

$$\left.\begin{array}{l} h(\bar{x}) = 0 \\ \bar{x} \in C \end{array}\right\} \iff \bar{x} \text{ solution of } VI(f, C).$$

There are numerous gap functions available in the literature. See for example Auslender [3], Peng [17], Fukushima [14], Auchmuty [1], Yamashita, Taji and Fukushima [22], Dutta [11], Aussel and Dutta [4], and the references therein. The notion of the gap function suggests an optimization reformulation of the variational inequality problem $VI(f, C)$. If we minimize the function h over C and the infimal value is zero and there exists a minimizer of h over C, then that minimizer is a solution of $VI(f, C)$. Consider the following function

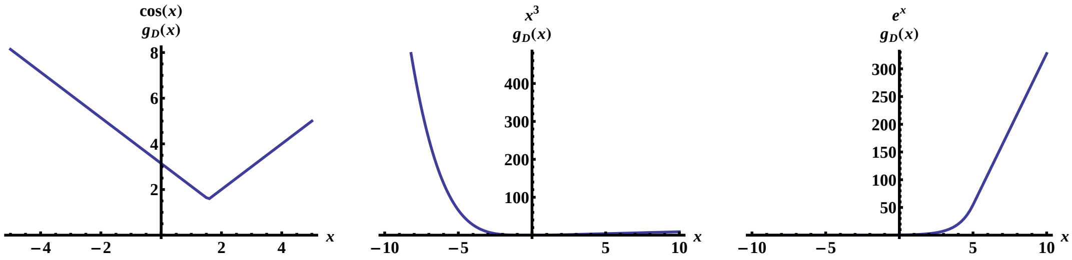
$$g_D(x) = \sup_{y \in C} \langle f(y), x - y \rangle.$$

It is not difficult to see that if f is monotone over C then g_D is a gap function for $VI(f, C)$. Further it is important to note that the dual gap function that we discuss in this paper was first suggested and studied in [24] and [25]. This dual gap function g_D has been also studied in [16] and used in [23] in order to derive the finite convergence of algorithms for variational inequalities. However it need not be finite for each x unless there is some addditional assumption on C or f. If C is compact or, as will be shown in the next section, f is strongly monotone on C, then g_D is finite-valued. This fact is important since we will focus on the strongly monotone case. The fact that g_D is finite in this case makes it an ideal vehicle for devising error bounds. Further observe that since C is non-empty, g_D never takes the value $-\infty$. Note also that g_D, being the supremum of a family of affine functions in x, is convex over $\mathbb{R}^n$. Thus when f is strongly monotone over C then g_D is also continuous. The figures below show g_D for $f(x) = \cos x,\ x^3,\ e^x$ with corresponding $C = [0, \pi], [-7, 1], [-4, 4]$, respectively.

In general g_D is a gap function for the Minty variational inequality, which seeks to find $x \in C$ such that

$$\langle f(y), y - x \rangle \geq 0, \quad \forall y \in C.$$

When f is monotone then the solution sets of the Minty variational inequality coincides with the solution set of $VI(f, C)$. At this point it is worthwhile to



mention that in the literature $VI(f, C)$ is usually known as the Stampacchia variational inequality.

It is interesting to notice that the variational inequality, as a separate problem, was considered after the pioneering works of Fichera [13] and Stampacchia [19]. while the set-valued formulation of the variational inequalities was first considered in [9]. On the other hand, the above called Minty variational inequality was first studied in Debrunner-Flor [10], where the first existence result for this problem has been obtained.

Now, in contrast, the Auslender gap function ([3] [24], [25]):

$$\theta(x) = \sup_{y \in C} \langle f(x), x - y \rangle$$

is a gap function for $VI(f, C)$ regardless of the nature of f. This gap function was also first suggested in [24], [25]. However it is important to know that unless we assume some conditions like compactness of C, it is not possible to guarantee the finiteness of the Auslender gap function even if f is strongly monotone. For example consider $C = \mathbb{R}^n_+$ which corresponds to a classical complementarity problem (see [12] for details). In this particular case we have $\theta(x) = \langle f(x), x \rangle$ if $f(x) \in \mathbb{R}^n_+$ and $\theta(x) = +\infty$ otherwise. This happens even if f is strongly monotone, the class of functions which are amenable to the development of error bounds.

There has been a lot of studies about the role of gap functions in the literature of variational inequalities and their role in the development of error bounds specially for the case where the variational inequality is described by a strongly monotone mapping. However the role of the dual gap function in devising error bounds, to the best of our knowledge, has not been thoroughly investigated. The only result involving the dual gap function in an error bound seems to have appeared in Marcotte and Zhu [16]. It is also to be noted that Marcotte and Zhu [16] develops the error bound for a weak-sharp solution of a variational inequality when the underlying map f is monotone plus (see [16] for details). Thus one of our main focus would be to develop error bounds for $VI(f, C)$ when f is strongly monotone without any additional assumption on the nature of the solution. Though this is one of our major goals we nevertheless start our article by discussing some basic properties of the dual gap function. We can conclusively show that g_D is a gap function $VI(f, C)$ when f is monotone. We also discuss some key properties of the g_D namely its finiteness and continuity which are the two crucial properties needed to develop error bounds. From a very practical point of view one might also feel that computing the value of g_D at a given x is indeed very difficult since one needs to solve a non-convex optimization problem. This disadvantage possibly may offset the fact that g_D is indeed a convex function while other gap functions known in the literature are not. For example the Auslender gap function mentioned above though not convex is more easy to compute as one needs to maximize an affine function over a convex set. But as we have mentioned above g_D is always finite when f is strongly monotone, while the Auslender gap function θ is not. Thus again from a theoretical point of view we believe that it would be interesting to investigate the role of g_D in devising error bounds for strongly monotone variational inequality problem.

The article is arranged as follows. In section 2 we mainly discuss the finiteness of g_D. As mentioned above we show that when f is strongly monotone g_D is finite. In this section we shall show that just having monotonicity of f need not give us finiteness over all the feasible set. We demonstrate this by giving two important examples of variational inequalities with monotone maps which are not strongly monotone. One of them is generated through the primal-dual pair of linear programming problems and the other one through the primal-dual pair of semidefinite programming problems. In fact for primal-dual pair of points in both these cases the function g_D can be explicitly computed and is nothing but the duality gap and is $+\infty$ everywhere else. These two examples are important examples because the associated variational inequalities cannot be represented as an optimality condition for any convex optimization problem. In Section 3 we present the error bounds for $VI(f, C)$ when f is a strongly monotone and the set C is convex and compact. We end Section 3 with some simple examples of $VI(f, C)$, where f is strongly monotone and g_D can be explicitly computed. Finally a concluding remark is given in Section 4 where, based on a very simple example, we suggest a possible extension of this work to non convex optimization problems.

2. Finiteness of the dual gap function

In this section we briefly comment on the relation between the Auslender's gap function and the dual gap function and show that when f is strongly monotone the dual gap function is always finite. To begin let us recall the standard definition of a monotonicity of a function $f\colon \mathbb{R}^n \to \mathbb{R}^n$. The function f is *monotone on C* if for any $x, y \in C$ we have

$$\langle f(y) - f(x), y - x\rangle \geq 0.$$

This simply shows that $g_D(x) \leq \theta(x)$ for all $x \in \mathbb{R}^n$. The fundamental difference between these two gap functions is the following: In order to compute the value at a given point for the function θ we need to maximize an affine function over the convex set C while for the dual gap function one needs to solve a non-convex minimization problem. We have however shown that for simple cases one can use a modern mathematical package and at least approximately compute the gap function g_D. Consider the graph of the function g_D for the one dimensional examples in Section 1. We admit however that it will be difficult to compute g_D in higher dimensional cases but one may be able to make some upper approximations of the value of g_D at each point. Fortunately one of the main strengths of g_D is that when f is strongly monotone, it is finite without any boundedness assumption on the set C. This is what we show below.

Proposition 2.1. *Let f be continuous and μ-strongly monotone on a non-empty closed and convex subset C. Then g_D is finite-valued on C.*

Proof. Since f is μ-strongly monotone we have for any x, y

$$\langle f(x) - f(y), x - y\rangle \geq \mu||y - x||^2.$$

This shows that

$$\sup_{y\in C}\{\langle f(x), x - y\rangle - \mu||y - x||^2\} \geq g_D(x).$$

Note that we have

$$\sup_{y\in C}\{\langle f(x), x - y\rangle - \mu||y - x||^2\} = -\inf_{y\in C}\{\langle f(x), y - x\rangle + \mu||y - x||^2\}.$$

But, for any fixed x, the function $y \mapsto \langle f(x),\, y - x\rangle + \mu||y - x||^2$ is strongly convex and thus has a unique minimizer over C since C is closed and convex. Thus $g_D(x)$ is finite for any $x \in C$. □

Example 2.2. We shall now give an example of a strongly monotone variational inequality with non-compact C and show that the Auslender gap function can fail to be finite but the dual gap function will be finite. Consider the $VI(f, C)$ with $f\colon \mathbb{R} \to \mathbb{R}$ given as $f(x) = -1 + x$ and $C = \mathbb{R}_+$. Then at $x = \dfrac{1}{2}$ we have $g_D(\frac{1}{2}) = \dfrac{1}{16}$ while $\theta(\frac{1}{2}) = +\infty$. □

The following example which is in Borwein and Dutta [7] highlights the fact that if f is just monotone the dual gap function g_D need not be finite everywhere. In fact the variational inequality in the following example is important since in this case f cannot be expressed as the gradient of any convex function.

Example 2.3. In this example we shall look into a very special type of affine variational inequality. Thus we consider $VI(f, C)$ with $f(x) = Mx+q$ where M is a $n \times n$ skew-symmetric matrix. One of the major source of such variational inequalities comes from the primal-dual pair of linear programming problems. This was discussed in the recent work of Borwein and Dutta [7] and also in Borwein and Lewis [8]. Borwein and Dutta [7] has also calculated the dual gap function for such a variational inequality explicitly. Consider now the following pair of primal-dual linear programs. The primal problem denoted as (LP) is given as

$$\min\langle c, x\rangle \quad \text{subject to} \quad Ax \geq b, \quad x \geq 0,$$

and the dual problem denoted as (DP) is given as

$$\max\langle b, y\rangle \quad \text{subject to} \quad A^T y \leq c, \quad y \geq 0.$$

Consider the following affine variational inequality $VI(\hat{M}z + q,\ \mathbb{R}^n_+ \times \mathbb{R}^m_+)$, where

$$\hat{M} = \begin{bmatrix} 0 & -A^T \\ A & 0 \end{bmatrix}$$

and $z = (x, y)$. It was proved in [7] that the solution set of the above mentioned affine variational inequality coincides with the primal-dual solution set of the primal-dual pair of linear programming problem when both of them are non-empty. Let us denote the primal-dual feasible set of the pair of linear programming problem (LP)-(DP) by S and the primal-dual solution set as $\overline{S}$. In fact it has been shown in [7] that when $(x, y) \in S$ then

$$g_D(x, y) = \langle c, x\rangle - \langle b, y\rangle,$$

and $g_D(x, y) = +\infty$ when $(x, y) \notin S$. □

We shall now show that a similar thing can happen if we consider a variational inequality constructed using the a primal-dual pair of semidefinite programming problems.

Example 2.4. We begin by considering the space $\mathbb{S}^n$ of all $n \times n$ real symmetric matrices, $\mathbb{S}^n_+$ the cone of all positive semidefinite matrices and $\mathbb{S}^n_{++}$ be the cone of positive definite matrices. The inner product of two matrices X and Y in $\mathbb{S}^n$ is defined by $\langle X, Y\rangle_{\mathbb{S}^n} = trace(XY)$ An important feature of the cone $\mathbb{S}^n_+$ is that it is not polyhedral and it is self-dual, which means that it coincides with its dual cone. The dual cone of $\mathbb{S}^n_+$ is denoted as $(\mathbb{S}^n_+)^*$ and is given as

$$(\mathbb{S}^n_+)^* = \{W \in S^n : \langle W, V\rangle \geq 0, \quad \forall V \in \mathbb{S}^n_+\}.$$

The semidefinite programming problem (SDP) is given as follows:

$$\min\langle C, X\rangle_{\mathbb{S}^n_+}, \quad \text{subject to} \quad , \langle A_i, X\rangle_{\mathbb{S}^n} \geq b_i, i = 1, \ldots, m, \quad X \in \mathbb{S}^n_+,$$

where $C, A_1, \ldots A_m \in \mathbb{S}^n$. The dual problem of (SDP) which we denote by (DSDP) is given as

$$\max\langle b, y\rangle, \quad \text{subject to} \quad C - \sum_{i=1}^{m} y_i A_i \in \mathbb{S}^n_+, \quad y \in \mathbb{R}^m_+.$$

Given the matrices $A_1, \ldots A_m \in \mathbb{S}^n$ we can define the linear operator $\mathcal{A}\colon \mathbb{S}^n \to \mathbb{R}^m$ as

$$\mathcal{A}X = (\langle A_1, X\rangle_{\mathbb{S}^n}, \ldots, \langle A_m, X\rangle_{\mathbb{S}^n}).$$

We can also define the adjoint of the operator $\mathcal{A}$. The adjoint of $\mathcal{A}$ is denoted as $\mathcal{A}^*\colon \mathbb{R}^m \to \mathbb{S}^n$, is given as

$$\mathcal{A}^* y = y_1 A_1 + \cdots y_m A_m.$$

We shall now define an inner product in the space $\mathbb{S}^n \times \mathbb{R}^m$. Consider any (X, y) and (Z, w) in $\mathbb{S}^n \times \mathbb{R}^m$ then

$$\langle (X, y), (Z, w)\rangle = \langle X, Z\rangle_{\mathbb{S}^n} + \langle y, w\rangle.$$

Once we are equipped with this definition it is easy to argue that $\mathcal{A}^*$ is indeed an adjoint operator of the linear operator $\mathcal{A}$. Note that

$$\langle \mathcal{A}X, y\rangle = \sum_{i=1} y_i \langle A_i, X\rangle_{\mathbb{S}^n} = \langle \sum_{i=1}^{m} y_i A_i, X\rangle_{\mathbb{S}^n} = \langle \mathcal{A}^* y, X\rangle_{\mathbb{S}^n}.$$

Let us consider the following operator $\mathcal{F}\colon \mathbb{S}^n \times \mathbb{R}^m \to \mathbb{S}^n \times \mathbb{R}^m$ described as

$$\mathcal{F}(X, y) = (-\mathcal{A}^* y, \mathcal{A}X) + (C, -b).$$

It is important to note that for any $(X, y) \in \mathbb{S}^n \times \mathbb{R}^m$ we have

$$\langle (-\mathcal{A}^* y, \mathcal{A}X), (X, y)\rangle = 0. \tag{1}$$

We shall consider the variational inequality $VI(\mathcal{F}, \mathbb{S}^n_+ \times \mathbb{R}^m_+)$. It is not difficult to show that $\mathcal{F}$ is a monotone operator. Our first aim is to compute g_D for $VI(\mathcal{F}, \mathbb{S}^n_+ \times \mathbb{R}^m_+)$. Let V denote the primal-dual feasible set to the primal-dual pair (SDP) and (DSDP) which is naturally given as

$$V = \{(X, y) \in \mathbb{S}^m_+ \times \mathbb{R}^m_+ : \mathcal{A}X \geq b, \quad C - \mathcal{A}^* y \in \mathbb{S}^n_+\},$$

where $\geq$ denotes component-wise ordering. The primal-dual solution set will be denoted by $\overline{V}$. Thus we have

$$g_D(X, y) = \sup_{(X', y') \in \mathbb{S}^n_+ \times \mathbb{R}^m_+} \langle \mathcal{F}(X', y'), (X, y) - (X', y')\rangle.$$

A simple calculation and keeping in mind (1) we have

$$g_D(X,y) = \sup_{(X',y')\in\mathbb{S}^n_+\times\mathbb{R}^m_+} \langle C,X\rangle_{\mathbb{S}^n} - \langle b,y\rangle + \langle -\mathcal{A}^*y', X\rangle_{\mathbb{S}^n} + \\ + \langle \mathcal{A}X', y\rangle - \langle C, X'\rangle_{\mathbb{S}^n} + \langle b, y'\rangle.$$

Further simplifying allows us to conclude that

$$g_D(X,y) = \langle C,X\rangle_{\mathbb{S}^n} - \langle b,y\rangle + \sup_{(X',y')\in\mathbb{S}^n_+\times\mathbb{R}^m_+} \langle (\mathcal{A}^*y - C, b - \mathcal{A}X), (X', y')\rangle).$$

If $(X,y)\in V$ that is primal-dual feasible then it is simple to see that

$$g_D(X,y) = \langle C,X\rangle_{\mathbb{S}^n} - \langle b,y\rangle.$$

Further if $(X,y)\notin V$ then we have $g_D(X,y) = +\infty$. $\square$

As a byproduct of the computation of the dual gap function in Example 2.4 we get the following result which is of independent interest.

Proposition 2.5. *Let $SVI(\mathcal{F},\mathbb{S}^m_+\times\mathbb{R}^m_+)$ denote the solution set of $VI(\mathcal{F},\mathbb{S}^m_+\times\mathbb{R}^m_+)$. Let as before $\overline{V}$ denote the primal-dual solution set of the primal-dual pair of semi-definite programs (SDP)–(DSDP). Assume that both solution sets are non-empty. Further assume that the Slater condition holds for (SDP), which means that there exists $\hat{X}\in\mathbb{S}^n_{++}$ such that $\langle A_i,\hat{X}\rangle_{\mathbb{S}^n} > b_i$ for all $i = 1,\dots,m$. Then*

$$\overline{V} = SVI(\mathcal{F},\mathbb{S}^m_+\times\mathbb{R}^m_+)$$

Proof. Let $(X,y)\in SVI(\mathcal{F},\mathbb{S}^m_+\times\mathbb{R}^m_+)$. Then we have $g_D(X,y) = 0$. Note that one must have $(X,y)\in V$ or else $g_D(X,y) = +\infty$. This implies that $\langle C,X\rangle_{\mathbb{S}^m} = \langle b,y\rangle$. Hence the strong duality holds for (SDP)-(DSDP) and thus it is simple to show that X solves (SDP) and y solves (DSDP). Hence $(X,y)\in\overline{V}$. Conversely assume that X solves (SDP) and y solves (DSDP) and thus $(X,y)\in\overline{V}$. Since Slater condition holds for (SDP) then strong duality holds (see for example Tuncel [20]) and we have $\langle C,X\rangle_{\mathbb{S}^m} = \langle b,y\rangle$. Thus $g_D(X,y) = 0$ showing that $(X,y)\in SVI(\mathcal{F},\mathbb{S}^m_+\times\mathbb{R}^m_+))$. Hence the result. $\square$

In fact in order to develop an error bound or even if we try to minimize the gap function it is important that the gap functions are continuous. If C is compact then both θ and g_D are continuous. If C is compact the continuity of g_D is immediate since g_D is finite and convex. On the other hand we need a bit more work to show that θ is continuous when C is compact. However if C is not compact we cannot say anything about the continuity of θ even if we have f to be strongly monotone. Nevertheless from Proposition 2.1 we see that when f is strongly monotone, without any compactness assumption on the set C, g_D is continuous since it is convex and finite.

3. Error bounds

Though the dual gap function is convex it has hardly been studied in the literature on the gap function for variational inequalities. In this section we will try to develop an error bound in terms of g_D where C is a compact set. We will have two different assumptions on f. In the first case we shall consider f to be strongly monotone over the set C and in the second case f is additionally Lipschitz on the set C.

Theorem 3.1. *Let us consider the problem $VI(f, C)$, where C is a nonempty compact convex set. Further assume that f is μ-strongly monotone ($\mu > 0$) over the set C. Let $\bar{x}$ be the unique solution of $VI(f, C)$. Then for any $x \in C$ we have*

$$||x - \bar{x}|| \leq 2 \sqrt{\frac{||f(x)||}{\mu} \sqrt{\frac{g_D(x)}{\mu}} - \frac{g_D(x)}{\mu}}. \tag{2}$$

Proof. Since C is a compact set and f is continuous it is clear that for any x the set $Y(x) = \text{argmax}_{y \in C} \langle f(y), x - y \rangle \neq \emptyset$. Consider $y \in Y(x)$. Let us consider $x \neq \bar{x}$ since if $x = \bar{x}$ the error bound given in (2) holds trivially. Note that $x \notin Y(x)$ since if $x \in Y(x)$ then $g_D(x) = 0$ and therefore $x = \bar{x}$. This also shows that $y \neq x$. Consider any $\lambda \in (0, 1)$ and construct the point $y' = x + \lambda(y - x)$. By convexity of C, $y' \in C$ and since $y \in Y(x)$ we have

$$\langle f(y), x - y \rangle \geq \langle f(y'), x - y' \rangle.$$

Note that $x - y' = \lambda(x - y)$. Thus from the previous expression we deduce that

$$\frac{1}{\lambda} \langle f(y), x - y \rangle \geq \langle f(y'), x - y \rangle. \tag{3}$$

From the μ-strong monotonicity of f and since $y - y' = (1 - \lambda)(y - x)$, we have

$$\langle f(y) - f(y'), y - y' \rangle \geq \mu(1 - \lambda)^2 ||y - x||^2.$$

Now replacing $y - y'$ by $(1 - \lambda)(y - x)$ in the left hand side of the above expression we get

$$\langle f(y), y - x \rangle - \langle f(y'), y - x \rangle \geq \mu(1 - \lambda)||y - x||^2$$

and thus combining the previous inequality with (3) and since $y \in Y(x)$, we get

$$(\frac{1}{\lambda} - 1)\, g_D(x) \geq (\frac{1}{\lambda} - 1)\langle f(y), x - y \rangle \geq \mu(1 - \lambda)||y - x||^2.$$

This shows that

$$||y - x||^2 \leq \frac{g_D(x)}{\lambda\mu}.$$

As $\lambda \uparrow 1$ we have

$$||y - x|| \leq \sqrt{\frac{g_D(x)}{\mu}}. \tag{4}$$

Now by strong monotonicity of f and Cauchy-Schwarz inequality we have

$$||f(x)||.||x - y|| - \mu||x - y||^2 \geq \langle f(y), x - y\rangle. \tag{5}$$

For $k \in (0, 1)$ consider the point $x' = x + k(\bar{x} - x)$ of C. Since $y \in Y(x)$ from (5) we have

$$||f(x)||.||x - y|| - \mu||x - y||^2 \geq \langle f(x'), x - x'\rangle = k\langle f(x'), x - \bar{x}\rangle. \tag{6}$$

Again using the strong monotonicity of f we have

$$\langle f(x'), x' - \bar{x}\rangle \geq \mu||x' - \bar{x}||^2 + \langle f(\bar{x}), x' - \bar{x}\rangle.$$

Now by definition of $\bar{x}$, we have $\langle f(\bar{x}), x' - \bar{x}\rangle \geq 0$ and thus using that $x' - \bar{x} = (1 - k)(x - \bar{x})$ we get

$$\langle f(x'), x - \bar{x}\rangle \geq \mu(1 - k)||x - \bar{x}||^2.$$

Therefore combining with (6) we have for any $k \in (0, 1)$

$$||f(x)||.||x - y|| - \mu||x - y||^2 \geq \mu k(1 - k)||x - \bar{x}||^2. \tag{7}$$

Note that the function $\psi(k) = k(1 - k)$ attains its maximum over $(0, 1)$ at $k = \frac{1}{2}$. Thus from (7) we have

$$||f(x)||.||x - y|| - \mu||x - y||^2 \geq \frac{\mu}{4}||x - \bar{x}||^2.$$

A simple manipulation proves that

$$\left(\frac{||f(x)||}{2\mu} - ||x - y||\right)^2 \leq \frac{||f(x)||^2}{4\mu^2} - \frac{1}{4}||x - \bar{x}||^2.$$

This shows that

$$\frac{||f(x)||}{2\mu} \leq ||y - x|| + \sqrt{\frac{||f(x)||^2}{4\mu^2} - \frac{1}{4}||x - \bar{x}||^2}.$$

Now using (4) we have

$$\frac{||f(x)||}{2\mu} - \sqrt{\frac{g_D(x)}{\mu}} \leq \sqrt{\frac{||f(x)||^2}{4\mu^2} - \frac{1}{4}||x - \bar{x}||^2}. \tag{8}$$

We claim that

$$\frac{||f(x)||}{2\mu} - \sqrt{\frac{g_D(x)}{\mu}} \geq 0.$$

Note that using the strong monotonicity of f and the fact that $y \in Y(x)$ shows that

$$\langle f(x), x - y \rangle \geq \mu ||y - x||^2 + g_D(x).$$

Since we have already shown that $y \neq x$, we have using the Cauchy-Schwarz inequality that

$$||f(x)|| \geq \mu ||y - x|| + \frac{g_D(x)}{||x - y||}.$$

Since both sides of the above inequality are non-negative we have by squaring both sides

$$||f(x)||^2 \geq \mu^2 ||x - y||^2 + 2\mu g_D(x) + \frac{g_D(x)^2}{||x - y||^2}. \tag{9}$$

But one clearly has

$$\frac{g_D(x)^2}{||x - y||^2} + \mu^2 ||x - y||^2 \geq 2\mu g_D(x).$$

Therefore from (9), $||f(x)||^2 \geq 4\mu g_D(x)$ and the claim is proved. Since both sides of (8) are non-negative, by squaring both sides and from simple manipulations we obtain

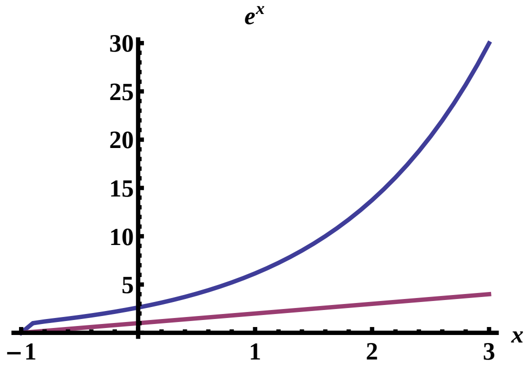
$$\frac{1}{4}||x - \bar{x}||^2 \leq \frac{||f(x)||}{\mu}\sqrt{\frac{g_D(x)}{\mu}} - \frac{g_D(x)}{\mu}$$

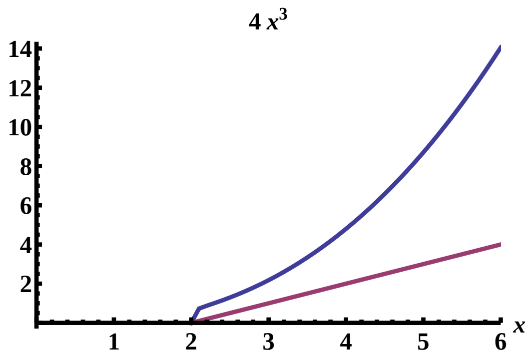
and the proof is complete. □

Through the following example we shall illustrate the error bound in Theorem 3.1 in two simple cases.

Example 3.2. In this example we plot the left-hand side (LHS) and the right-hand side (RHS) of Theorem 3.1 for $e^x, 4x^3$ which is strongly monotone over $C = [-1, 2], [2, 5]$ with $\mu = \frac{1}{e}$, 48, respectively. Observe that there are two curves in each diagram with one curve always staying above the other. The top curve represents the values of the RHS of the error bound and lower curve represents the LHS, which is the distance to the solution.

It is interesting to observe that near the solution, the RHS value drops towards zero faster than the LHS. □





Remark 3.3. Let us observe that from Theorem 3.1 we immediately obtain that for any $x \in C$,

$$g_D(x) \leq \frac{||f(x)||^2}{\mu}. \qquad \square$$

If, in addition to strong monotonicity, f also satisfies a Lipschitz property, then an alternative error bound can be proved in terms of the dual gap function g_D. That is the aim of the following theorem.

Theorem 3.4. *Let us consider the problem $VI(f, C)$ where C is a convex compact set and f is μ-strongly monotone on C. Further assume that f is Lipschitz on C with Lipschitz rank L. Let $\bar{x}$ be the unique solution of $VI(f, C)$. Then for any $x \in C$ we have*

$$||x - \bar{x}|| \leq (1 + \frac{L}{\mu}) \sqrt{\frac{g_D(x)}{\mu}}. \tag{10}$$

Proof. As in the proof of Theorem 3.1, by considering $x \in C$, $x \neq \bar{x}$, we immediately get that $x \notin Y(x)$, this latter set being nonempty by compactness of C. Consider $y \in Y(x)$. Our main step is to show that

$$L||x - y|| \geq \mu||y - \bar{x}||. \tag{11}$$

Consider $k \in (0, 1)$ and let $y' = \bar{x} + (1 - k)(y - \bar{x})$. Clearly $y' \in C$ and since $y \in Y(x)$ we have $\langle f(y), x - y\rangle \geq \langle f(y'), x - y'\rangle$ and thus

$$\langle f(y) - f(y'), x - y\rangle \geq \langle f(y'), y - y'\rangle.$$

A simple calculation shows that $y - y' = k(y - \bar{x}) = \dfrac{k}{1-k}(y' - \bar{x})$ and thus

$$\langle f(y) - f(y'), x - y\rangle \geq \frac{k}{1-k}\langle f(y'), y' - \bar{x}\rangle.$$

Now by using the strong monotonicity of f on C we have

$$\langle f(y) - f(y'), x - y\rangle \geq \frac{k}{1-k}\left(\langle f(\bar{x}), y' - \bar{x}\rangle + \mu||y' - \bar{x}||^2\right).$$

But since $\bar{x}$ is the solution of $VI(f, C)$ we can deduce, together with the Cauchy-Schwarz inequality, that

$$||f(y) - f(y')||.||y - x|| \geq \langle f(y) - f(y'), x - y\rangle \geq \frac{\mu k}{1-k}||y' - \bar{x}||^2.$$

The fact that f is Lipschitz on C with rank L allows us to conclude that

$$L||y - y'||.||x - y|| \geq \frac{\mu k}{1-k}||y' - \bar{x}||^2$$

or in other words

$$Lk||y - \bar{x}||.||x - y|| \geq \frac{\mu k}{(1-k)}(1-k)^2||y - \bar{x}||^2.$$

This leads to the expression $L||x - y|| \geq (1-k)\mu||y - \bar{x}||$ and as $k \downarrow 0$ we have $L||x - y|| \geq \mu||y - \bar{x}||$ thus establishing the claimed inequality (11). Now the proof is complete since, by (11), we have

$$||x - \bar{x}|| \leq ||x - y|| + ||y - \bar{x}|| \leq (1 + \frac{L}{\mu})||y - x||.$$

The result is now a direct consequence of inequality (4) of the first part of Theorem 3.1. $\square$

It will be important to note that though for a monotone f, $g_D \leq \theta$, a straightforward comparison between the error bounds obtained through these two gap functions as they appear cannot be made. Further the next important question would be how to develop an error bound when C is not compact. Surprisingly we found that it could be done for the strongly convex case by regularizing the dual gap function in a particular way. However we also found that Yamashita and Fukushima [21] had already published a similar result in 1997. Consider the following function

$$h_\beta(x) = \sup_{y \in C}\{\langle f(y), x - y\rangle + \beta\|y - x\|^2\}.$$

If f is strongly monotone function with $\mu > 0$ as the modulus of strong monotonicity, then it was shown in [21] that $h_\beta(x)$ is gap function for $VI(F, C)$ if $0 < \beta \leq \mu$. Further they have also showed that when f is strongly monotone and $< \beta \leq \mu$, one has

$$\|x - \bar{x}\| \leq \sqrt{\frac{h_\beta(x)}{\beta}},$$

where $x \in C$ and $\bar{x}$ is the unique solution of $VI(f, C)$ Surprisingly no error bounds were developed for g_D in [21] even with C compact. However it is important to note that in Marcotte and Zhu [16] (see Theorem 4.1) a very simple error bound has been developed in the case C is compact f is pseudomonotone plus and the solution set is weak-sharp. In fact the existence of such an error bound was shown to be the necessary condition for the variational inequality with pseudomonotone plus map. A map $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is said to be pseudomonotone plus on C if f is pseudonomotone and for $x, y \in C$ we have

$$\langle f(x), y-x\rangle \geq 0 \quad \text{and} \quad \langle f(y), y-x\rangle = 0, \implies f(y) = f(x).$$

In fact it is simple to observe that if f is strongly monotone then f is pseudomonotone plus. If f is strongly monotone then it is easily seen to be pseudomonotone plus. Further noting the definition of weak-sharp solution for a variational inequality from [16] we can deduce that when f is strongly monotone then its unique solution x^* is weak-sharp if

$$-f(x^*) \in \operatorname{int} N_C(x^*).$$

Thus we can use the Theorem 4.1 in Marcotte and Zhu [16] in order to get the following result.

Proposition 3.5. *Consider the variational inequality $VI(f,C)$ where f is strongly monotone and C is compact. Let $x^* \in C$ be the unique solution of $VI(f,C)$. Then x^* is weak-sharp if and only if there exists $\alpha > 0$ such that*

$$d(x, SVI(f,C)) \leq \frac{g_D(x)}{\alpha},$$

where $SVI(f,C)$ is the solution set of $VI(f,C)$ and d is the distance function.

Further when f is Lipschitz on C and C is compact and C is written as the Cartesian product of several compact convex sets then if f is a unuiformly P-function then one can develop an error bound for $VI(f,C)$ using the natural residual map. See Proposition 6.3.1 in [12]. Note that a strongly monotone map is also uniformly-P function.

We would like to end this section by showing two simple examples where the dual gap function g_D can be explicitly computed.

Example 3.6. *The dual gap function in the strongly monotone case.*
In this example we will compute the dual gap function explicitly in two cases where the function f would be strongly monotone over C. We will consider the set C to be compact as required by our main results on error bounds in terms of g_D.

Our first problem is one dimensional. Consider the $VI(f,C)$ where $f\colon \mathbb{R} \to \mathbb{R}$, is given as $f(x) = e^x$ and $C = [0,1] \subset \mathbb{R}$. In fact this $VI(f,C)$ is in fact equivalent to the problem minimizing the function $\varphi(x) = e^x$ over the set $[0,1]$. We shall begin by showing that e^x is strongly convex on $C = [0,1]$ with $\rho = \frac{1}{e}$ showing that $f(x) = \varphi'(x) = e^x$ is strongly monotone with $\rho = \frac{1}{e}$.

In order to show that $\varphi(x) = e^x$ is strongly monotone with $\rho = \frac{1}{e}$ it is enough to show that the function $\psi(x) = e^x - \frac{1}{e}\|x\|^2$ is convex on $[0, 1]$. Note that

$$\psi''(x) = e^2 - \frac{2}{e}.$$

Note that as $x \in [0, 1]$ we have $e^x \geq 1$ and as $e \geq 2$ we have that $\psi''(x) \geq 0$ and thus ψ is convex on $[0, 1]$ showing that $\varphi(x) = e^x$ is strongly convex in $[0, 1]$ and hence $f(x)$ is strongly monotone with $\rho = \frac{1}{e}$. Further note that e^x is not strongly convex over the whole real line $\mathbb{R}$ and thus f is not strongly monotone over the whole real line $\mathbb{R}$. We shall now calculate g_D for the above mentioned variational inequalities. We gave

$$g_D(x) = \sup_{y \in [0,1]} e^y(x - y).$$

We shall show that for all $x \in [0, 1]$ we can explicitly compute x. Let us set for each $x \in [0, 1]$ the function

$$p_x(y) = xe^y - ye^y.$$

Thus we have

$$p'_x(y) = (x - 1)e^y - ye^y.$$

Hence, for $x = 0$, we get $p'_0(y) = -e^y - ye^y < 0$ for all $y \in [0, 1]$. Thus $p_0(y)$ is a strictly decreasing function in y and thus the function achieves it maximum at $y = 0$. Thus $g_D(0) = 0$. Hence $x = 0$ is the unique solution of the variational inequality. For $0 < x < 1$ we have

$$p'_y(x) = (x - 1)e^y - ye^y.$$

Thus, when $0 < x < 1$, we see that $p'(y) < 0$ for all $y \in [0, 1]$. Hence $p_x(y)$ is strictly decreasing and thus $y = 0$ is the maximizer and hence we have $g_D(x) = x$.

Now we shall consider the case $x = 1$. In this case $p_1(y) = (1 - y)e^y$. For $y = 0$ we have $p_1(0) = 1$ and $p_1(1) = 0$. The function is strictly decreasing for $y \in (0, 1)$ and thus noting the continuity of the function we conclude that $y = 0$ is the maximizer of $p_1(y)$ over $[0, 1]$. Thus $g_D(1) = 1$.

The previous discussion allows us to conclude that $g_D(x) = x$ for all $x \in [0, 1]$.

We shall now turn our attention to the two dimensional case. Let us consider $f : \mathbb{R}^2 \to \mathbb{R}^2$, given as $f(x_1, x_2) = (x_1, x_2)$ and $C = [0, 1] \times [0, 1]$. In this case we have

$$\begin{aligned}
g_D(x_1, x_2) &= \sup_{(y_1,y_2) \in C} \langle (y_1, y_2), (x_1, x_2) - (y_1, y_2) \rangle \\
&= \sup_{(y_1,y_2) \in C} y_1(x_1 - y_1) + y_2(x_2 - y_2) \\
&= - \inf_{(y_1,y_2) \in C} \{(y_1 - \frac{x_1}{2})^2 + (y_2 - \frac{x_2}{2})^2 - \frac{x_1^2}{4} - \frac{x_2^2}{4}\}.
\end{aligned}$$

Thus for any $(x_1, x_2) \in C = [0,1] \times [0,1]$ it is simple to see that

$$g_D(x_1, x_2) = \frac{x_1^2}{4} + \frac{x_2^2}{4}.$$

We would like to note that we have shown that for very simple cases we can explicitly compute the value of g_D over C. However for more complex and higher dimensional cases it might not be always simple to compute g_D. The computation of g_D and using it solve a monotone and strongly monotone variational inequalities will be one of our future research goals. $\square$

4. A final remark

As illustrated in the previous sections, the dual gap function enjoys some very interesting properties like convexity, finiteness and porwerful error bounds if the involved function f is strongly-monotone. thus wherever considering an optimization problem with differentiable strongly convex objective function h, the dual gap function can also be used since the associated variational inequality is defined by the monotone map $f(x) = \nabla h(x)$. However, by the following example, we would like to suggest that the dual gap function $g_D(f)$ can also be used for non convex optimization, while other gap functions can fail to be efficient. Let simply consider the problem of minimizing the differentiable function $h\colon x \mapsto x^3$ over the interval $C = [-1, 1]$. The associated variational inequality is of course defined by the map $f(x) = \nabla h(x) = 3x^2$ and the subset C. Then the Auslender function $\theta(x) = 3x^2(x+1)$ cannot act as a gap function for the minimization problem because actually $\theta(x) = 0$ admits two solutions, $x = -1$ and $x = 0$ while the minimization problem has only $x = -1$ as a solution. But since $g_D(x) = 3x + 3$ this point is the only one verifying $g_D(x) = 0$ thus showing that g_D perfectly plays the role of a gap function for this problem. An important characteristic of this simple example is that the objective function h is clearly not convex, it is only quasiconvex. Note that the function $f = \nabla h$ in this case is also not monotone. So here the dual gap function succeeded in selecting, between the different "stationary points" the one that is a solution of the initial optimization problem. A open question is on which class of non convex optimization problems the dual gap function can make such a selection.

References

[1] G. Auchmuty: Variational principles for variational inequalities, Num. Func. Anal. Optim. 10 (1989) 863–874.

[2] D. Aussel: New developments in quasiconvex optimization, in: Recents Advances in Fixed Point Theory and Optimization, Taylor and Francis (2013) 173–208.

[3] A. Auslender: Optimisation: Méthodes Numériques, Masson, Paris (1976).

[4] D. Aussel, J. Dutta: On gap functions for multivalued Stampacchia variational inequalities, J. Optim. Theory Appl. 149 (2011) 513-527.

[5] D. Aussel, N. Hadjisavvas: Adjusted sublevel sets, normal operator and quasiconvex programming, SIAM J. Optim. 16 (2005) 358–367.

[6] D. Aussel, J. Ye: Quasiconvex minimization on locally finite union of convex sets, J. Optim. Theory Appl. 139(1) (2008) 1–16.

[7] J. M. Borwein, J. Dutta: Maximal monotone inclusions and Fitzpatrick functions, J. Optim. Theory Appl. 171(3) (2016) 757–784.

[8] J. M. Borwein, A. S. Lewis: Convex analysis and nonlinear optimization. Theory and examples (2nd edition), CMS Books in Mathematics 3, Springer, New York (2006).

[9] F. E. Browder: Multivalued monotone nonlinear mappings and duality mappings in Banach spaces, Trans. Amer. Math. Soc. 71 (1965) 780–785.

[10] H. Debrunner, P. Flor: Ein Erweiterungssatz fur monotone Mengen, Arch. Math. 15 (1964) 445–447.

[11] J. Dutta: Gap functions and error bounds for variational and generalized variational inequalities, Vietnam J. Math. 40(2-3) (2012) 231–253.

[12] F. Facchinei, J.-S. Pang: Finite-Dimensional Variational Inequalities and Complementarity Problems I, Springer Series in Operations Research, Springer, New York (2003).

[13] G. Fichera: Problemi elastostatici con vincoli unilaterali; il problema di Signorini con ambigue al contorno, Atti Acad. Naz. Lincei. Memorie, Cl. Sci. Fis. Mat. Nat. Sez. I 8 (1964) 91–140.

[14] M. Fukushima: A class of gap functions for quasi-variational inequality problems, J. Ind. Manag. Optim. 3(2) (2007) 165–171.

[15] I. V. Konnov: On convergence properties of a subgradient method, Optim. Methods Softw. 18(1) (2003) 53–62.

[16] P. Marcotte, D. Zhu: Weak sharp solutions of variational inequalities, SIAM J. Optim. 9(1) (1999) 179–189.

[17] J. Peng: Equivalence of variational inequality problem to unconstrained minimization, Math. Prog. 78(3) (1997) 347–355.

[18] R. T. Rockafellar, R. J.-B. Wets: Variational Analysis, Fundamental Principles of Mathematical Sciences 317, Springer, Berlin (1998).

[19] G. Stampacchia: Formes bilinéaires coercitives sur le ensembles convexes, C. R. Acad. Sci. Paris 258 (1964) 4413–4416.

[20] L. Tunel: Polyhedral and Semidefinite Programming Methods in Combinatorial Optimization, Fields Institute Monographs 27, American Mathematical Society, Providence (2010).

[21] N. Yamashita, M. Fukushima: Equivalent unconstrained minimization and global error bounds for variational inequality problems, SIAM J. Control Optim. 35(1) (1997) 273–284.

[22] N. Yamashita, K. Taji, M. Fukushima: Unconstrained optimization reformulations of variational inequality problems, J. Optim. Theory Appl. 92(3) (1997) 439–456.

[23] J. Zhang, C. Wan, N. Xiu: The dual gap function for variational inequalities, Appl. Math. Optim. 48(2) (2003) 129–148.

[24] S. I. Zukhovitskii, R. A. Polyak, M. E. Primak: Two methods of search for equilibrium points of n-person concave games, Soviet Math. Doklady 10 (1969) 279–282.

[25] S. I. Zukhovitskii, R. A. Polyak, M. E. Primak: Concave multiperson games, Matekon 9 (1973) 10–30.