International Mathematical Forum, 5, 2010, no. 44, 2199 - 2204

A New Descent Method With Optimal Step Size for Structured Co-coercive Variational Inequalities

Min Sun

Department of Mathematics and Information Science
Zaozhuang University, Shandong 277160, China
sunmin_2008@yahoo.com.cn

Abstract

This paper presents a new descent method with optimal step size for structured co-coercive variational inequalities. Without carrying out any line search technique, we prove the global convergence of the new method. At each iteration, the new method needs only to perform some orthogonal projections and some function evaluations, so its computational load is very tiny. Preliminary numerical experiments for some practical problems illustrate the effectiveness of the proposed method.

Mathematics Subject Classification: 90C25, 90C30, 90C90

Keywords: variational inequalities, co-coercive function, global convergence

1 Introduction

Consider the following structured monotone variational inequality problem with linear constraint: Find $u^* \in \Omega$, such that

$$(u - u^*)^\top T(u^*) \geq 0, \quad u \in \Omega \tag{1}$$

where

$$u = \begin{pmatrix} x \\ y \end{pmatrix}, \ T(u) = \begin{pmatrix} f(x) \\ g(y) \end{pmatrix}, \ \Omega = \{(x, y)|x \in X, y \in Y, Ax + By = b\},$$

$X \subseteq R^n$ and $Y \subseteq R^m$ are given nonempty closed convex sets; $f : X \to R^n$ and $g : Y \to R^m$ are given continuous monotone operators; $A \in R^{r\times n}$

and $B \in R^{r\times m}$ are given matrices; $b \in R^r$ is a given vector. This problem has several important applications in many fields, such as network economics, traffic assignment, game theoretic problems, etc.

By attaching a Lagrange multiplier vector $\lambda \in R^r$ to the linear constraints $Ax + By = b$, (1) can be equivalently transformed into the following compact form, denoted by VI: Find $w^* \in W$, such that

$$(w - w^*)^\top Q(w^*) \geq 0, \quad \forall w \in W \tag{2}$$

where

$$w = \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix}, Q(w) = \begin{pmatrix} f(x) - A^\top \lambda \\ g(y) - B^\top \lambda \\ Ax + By - b \end{pmatrix}, W = X \times Y \times R^r.$$

The alternating direction method (ADM) is a powerful method for solving the structured problem (2), since it decompose the original problems into a series subproblesm with lower scale. It was originally proposed by Gabay and Mercier[1] and Gabay[2]. Recently, Ye and Yuan [3] proposed a new descent method for VI by adding an additional projection step to the above ADM. Han [4] proposed a modified alternating direction method for variational inequalities with linear constraints. At each iteration, the method only makes an orthogonal projection to simple set and some function evaluations.

Motivated by [3-4], we give a new descent method for VI, and the computational load of the new method is very tiny.

2 Preliminaries

Definition 2.1 *A mapping $f : R^n \to R^n$ is said to be co-coercive with modulus $\mu > 0$ is*

$$(x - y)^\top (f(x) - f(y)) \geq \mu \|f(x) - f(y)\|^2, \quad \forall x, y \in R^n.$$

In the following, we always assume that the underlying function f, g are co-coercive with modulus μ_1, μ_2, respectively. Let P_W denote the orthogonal projection mapping from R^{n+m+r} onto W.

It is well known that VI is equivalent to the projection equation

$$e(w, \beta) = w - P_W[w - \beta Q(w)](\beta > 0).$$

Thus, VI is also equivalent to the projection equation

$$r(w,\beta)=\begin{pmatrix} r_1(w,\beta)\\ r_2(w,\beta)\\ r_3(w,\beta)\end{pmatrix}=$$
$$\begin{pmatrix} x-P_X[x-\beta(f(x)-A^\top(\lambda-\beta(Ax+By-b)))]\\ y-P_Y[y-\beta(g(y)-B^\top(\lambda-\beta(Ax+By-b)))]\\ \beta(Ax+By-b)\end{pmatrix}.$$

For a given $w\in W$, $\forall v\in R^{n+m+r}$ satisfies the inequality

$$(w-P_W(w))^\top(v-P_W(w))\le 0. \tag{3}$$

Similar to the monotonicity of $\|e(w,\beta)\|$, we have

$$\|r(w,\beta_1)\|\le\|r_2(w,\beta_2)\|, \tag{4}$$

where $0<\beta_1\le\beta_2$.

3 Algorithm and convergence

Set $r_i=r_i(w,\beta), i=1,2,3,\ F=f(x)-A^\top(\lambda-\beta(Ax+By-b)),\ G=g(y)-B^\top(\lambda-\beta(Ax+By-b))$.

Lemma 3.1 *Let $w^*=(x^*,y^*,z^*)\in W$ be an arbitrary solution of VI,*

$$d(w,\beta):=\begin{pmatrix} r_1+\beta^2A^\top Ar_1+\beta^2A^\top Br_2\\ r_2+\beta^2B^\top Ar_1+\beta^2B^\top Br_2\\ r_3-\beta Ar_1-\beta Br_2\end{pmatrix},$$

then for any $w=(x,y,z)\in R^{n+m+r}$, we have

$$(w-w^*)^\top d(w,\beta)\ge c\|r(w,\beta)\|^2,$$

where $c=\min\{1-\beta/(4\mu_1),1-\beta/(4\mu_2)\}$.

Proof. From (3), we have

$$\{x-\beta F-P_X[x-\beta F]\}^\top\{P_X[x-\beta F]-x^*\}\ge 0.$$

So

$$r_1^\top(x-x^*)\ge\|r_1\|^2+\beta F^\top(P_X[x-\beta F]-x^*). \tag{5}$$

Similarly, we have

$$r_2^\top(y-y^*)\ge\|r_2\|^2+\beta G^\top(P_Y[y-\beta G]-y^*). \tag{6}$$

As w^* is a solution of VI, we have

$$\beta(P_X[x-\beta F]-x^*)^\top(f(x^*)-A^\top\lambda^*)+\beta(P_Y[y-\beta G]-y^*)^\top(g(y^*)-B^\top\lambda^*)\geq 0, \tag{7}$$

$$Ax^*+By^*=b. \tag{8}$$

Adding (5)-(8), from the co-coercivity of f,g, we get the desired result.

From Lemma 3.1, we have that $d(w,\beta)$ is a descent direction of the function $\|w-w^*\|^2/2$. This motivates us to design the following algorithm.

Algorithm 3.1

Step 0: Given $\varepsilon>0$. Choose $w^0\in W$ and $\gamma\in(0,2)$, $\mu\in(0,1)$, $\delta\in(0,1)$, $0<\beta_L<\beta_U<\min\{4\mu_1,4\mu_2\}$, and $\beta_0\in(\beta_L,\beta_U)$. Set $k:=0$;

Step 1: If $\|e(w^k,\beta_k)\|<\varepsilon$, then stop; else calculate $d(w^k,\beta_k)$ by the expression of $d(w,\beta)$ in Lemma 3.1 and the optimal step size

$$\rho(w^k,\beta_k)=[1-\beta_k/(4\mu)]\|r(w^k,\beta_k)\|^2/\|d(w^k,\beta_k)\|^2,$$

where $\mu=\min\{\mu_1,\mu_2\}$.

Step 2: Calculate the new iterate

$$w^{k+1}=P_W[w^k-\gamma\rho(w^k,\beta_k)d(w^k,\beta_k)].$$

Choose $\beta_{k+1}\in(\beta_L,\beta_U)$ and set $k:=k+1$, and go to Step 1.

Theorem 3.2 *Suppose that the operator f,g are co-coercive monotone, the solution set W^* of VI is nonempty. Then the sequence of $\{w^k\}$ generated by Algorithm 3.1 is bounded. More specifically, we have*

$$\|w^{k+1}-w^*\|^2\leq\|w^k-w^*\|^2-c\rho(w^k,\beta_k)\|r(w^k,\beta_k)\|^2.$$

Proof. It is proof is similar to that of Theorem 4.1. of Han[4].

Theorem 3.3 *Suppose that the assumptions in Theorem 3.1 hold. Then the whole sequence $\{w^k\}$ converges to a solution of VI.*

Proof. From the definition of $d(w^k,\beta_k)$, it is easy to deduce that the step size is larger than a positive constant. That is, there exists a constant $\tau>0$ such that

$$\rho(w^k,\beta_k)\geq\tau.$$

This with Theorem 3.1 lead to

$$\lim_{k\to\infty}\|r(w^k,\beta_k)\|=0.$$

From (4) and $\beta_k>\beta_L>0$, we have

$$\lim_{k\to\infty}\|r(w^k,\beta_L)\|=0.$$

This shows that every cluster point of $\{w^k\}$ is a solution of VI. The following proof is similar to that of Theorem 4.2 in [4], so is omitted.

4 Preliminary Computational Results

In this section, we implement Algorithm in MATLAB and test it on a PC. The examples used here are the test problem of Han[4]. The constraint set S and the mapping f are taken respectively as

$$S = \{x \in R^5_+ | \sum_{i=1}^{5} x_i = 10\}, \quad f(x) = Mx + \rho C(x) + q,$$

where M is an $R^{5\times 5}$ asymmetric positive matrix and $C_i(x) = \arctan(x_i - 2), i = 1, 2, \cdots, 5$. The parameter ρ is used to vary the degree of asymmetry and nonlinearity, and the data of example are illustrate as follows:

$$M = \begin{pmatrix} 0.726 & -0.949 & 0.266 & -1.193 & -0.504 \\ 1.645 & 0.678 & 0.333 & -0.217 & -1.443 \\ -1.016 & -0.225 & 0.769 & 0.943 & 1.007 \\ 1.063 & 0.587 & -1.144 & 0.550 & -0.548 \\ -0.256 & 1.453 & -1.073 & 0.509 & 1.026 \end{pmatrix}$$

and

$$q = (5.308,\ 0.008\ , -0.938,\ 1.024,\ \ -1.312)^{\top}.$$

Thus, $g = 0, B = 0$ in this example. The stopping criterion is $\varepsilon = 10^{-6}$. Table 1 report the computational results for $\rho = 10$ and $\rho = 20$. All programs are coded in Matlab 7.1. 'IN' denotes the number of iterations and 'CPU' denotes the CPU time in seconds. The results in tables 1 indicate the algorithm is

Table 1: Numerical results

Starting point	ρ	IN	CPU
	$\rho = 10$	7	0.01
(0 2.5 2.5 2.5 2.5)	$\rho = 20$	21	0.01
	$\rho = 10$	12	0.01
(25 0 0 0 0)	$\rho = 20$	20	0.01
	$\rho = 10$	10	0.01
(10 0 0 0 0)	$\rho = 20$	24	0.01
	$\rho = 10$	12	0.01
(10 0 10 0 10)	$\rho = 20$	22	0.01

efficient for the given problem.

References

[1] Gabay, D..Applications of the method of multipliers to variational inequalities.In:M.Fortin and R.Glowinski(Eds) Augmented Lagrange Methods: Applications to the Solution of Boundary valued Problems(Amsterdam: North Holland),1983: 299-331.

[2] Gabay, D. and Mercier, B., A dual algorithm for the solution of nonlinear variational problem svia finite-element approximations. Computers and Mathematics with Applications, 1976, 2:17-40.

[3] Ye, C. H. and Yuan, X. M. A descent method for structured monotone variational inequalities.Optimization Methods and Software, 2007, 22(2): 329-338.

[4] Han, D.R. A modified alternating direction method for variational inequality problems. Applied Mathematics and Optimization, 2002, 45: 63-74.

[5] Sun, M. A new alternating direction method for co-coercive variational inequality problems with linear equality and inequality constraints. AMO - Advanced Modeling and Optimization, 2010, 12(2):161-176.

Received: March, 2010