Applied Numerical Mathematics 83 (2014) 12–21

Contents lists available at ScienceDirect

Applied Numerical Mathematics

www.elsevier.com/locate/apnum

A modified alternating projection based prediction–correction method for structured variational inequalities

Wenxing Zhang [a], Deren Han [b,*], Suoliang Jiang [c]

[a] School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu 611731, PR China
[b] School of Mathematical Sciences, Jiangsu Key Laboratory for NSLSCS, Nanjing Normal University, Nanjing 210023, PR China
[c] School of Computer Sciences, Nanjing Normal University, Nanjing 210097, PR China

A R T I C L E I N F O

Article history:
Received 6 August 2011
Received in revised form 25 March 2014
Accepted 6 April 2014
Available online 18 April 2014

Keywords:
Prediction–correction
Contraction
Separable
Self-adaptive
Alternating projection

A B S T R A C T

In this paper, we propose a novel alternating projection based prediction–correction method for solving the monotone variational inequalities with separable structures. At each iteration, we adopt the weak requirements for the step sizes to derive the predictors, which affords fewer trial and error steps to accomplish the prediction phase. Moreover, we design a new descent direction for the merit function in correction phase. Under some mild assumptions, we prove the global convergence of the modified method. Some preliminary computational results are reported to demonstrate the promising and attractive performance of the modified method compared to some state-of-the-art prediction–contraction methods.

1. Introduction

Let $\Omega \subseteq \mathcal{R}^n$ be a nonempty closed convex set and $F : \Omega \to \mathcal{R}^n$ be a continuous mapping. A classical variational inequality, denoted by VI(Ω, F), is to find a vector $u^* \in \Omega$ such that

$$\left(u' - u^*\right)^T F\left(u^*\right) \geq 0, \quad \forall u' \in \Omega. \tag{1}$$

Variational inequalities play fundamental roles in diversified applications, e.g., traffic assignment, game theory, economics, etc. (see e.g., [1,7,8,10,11,17] and references therein). Herein, we concentrate on the VI(Ω, F) (1) with the following special structures, i.e., the mapping F and the constraint Ω in (1) satisfy

$$u = \begin{pmatrix} x \\ y \end{pmatrix}, \qquad F(u) = \begin{pmatrix} f(x) \\ g(y) \end{pmatrix} \quad \text{and} \quad \Omega = \left\{(x, y) \,\middle|\, Ax + By = b,\ x \in \mathcal{X},\ y \in \mathcal{Y}\right\}, \tag{2}$$

where $\mathcal{X} \subseteq \mathcal{R}^{n_1}$ and $\mathcal{Y} \subseteq \mathcal{R}^{n_2}$ are nonempty closed convex sets with $n_1 + n_2 = n$; $A \in \mathcal{R}^{m\times n_1}$ and $B \in \mathcal{R}^{m\times n_2}$ are full column rank matrices; $b \in \mathcal{R}^m$ is a given vector; $f : \mathcal{X} \to \mathcal{R}^{n_1}$ and $g : \mathcal{Y} \to \mathcal{R}^{n_2}$ are monotone mappings (see Definition 1 in the next section). Note that although we focus on the variational inequalities in vector variables, the results can be tractably extended to the case of matrix variables. VI(Ω, F) in (1)–(2) has been successfully employed in the recent hot-investigated fields, e.g., signal and image processing, machine learning and statistics. Take the *background extraction of surveillance video* arising from the image processing for example, the problem can be modeled as (see e.g., [2,4,19])

* Corresponding author.
E-mail address: handeren@njnu.edu.cn (D. Han).

http://dx.doi.org/10.1016/j.apnum.2014.04.007
0168-9274/

$$\min\{\|X\|_* + \tau\|Y\|_1 \mid X + Y = D,\ X \geq 0,\ Y \geq 0\}, \tag{3}$$

where D represents the observed surveillance video; X and Y denote the background and foreground, respectively; $\|\cdot\|_*$ denotes the nuclear norm (the sum of all singular values) which can induce the low-rank component and $\|\cdot\|_1$ denotes the l_1 norm (the sum of absolute values of all entries) which can induce the sparse component; and $\tau > 0$ is the trade-off balancing the low-rank and sparsity. Theoretically, the model (3) can be cast as (1)–(2) with matrix variables (see e.g., [9] for the relations of convex minimizations with variational inequalities)

$$f(X) := \partial(\|X\|_*), \qquad g(Y) := \partial(\|Y\|_1) \quad \text{and} \quad \Omega := \{(X, Y) \mid X + Y = D,\ X \geq 0,\ Y \geq 0\},$$

where $\partial(\cdot): \mathcal{R}^n \to 2^{\mathcal{R}^n}$ is the subdifferential operator (see Definition 2 in the next section).

By attaching a Lagrange multiplier $\lambda \in \mathcal{R}^m$ to the constraint $Ax + By = b$, (1)–(2) can be reformulated as the following variational inequality: Find $w^* \in \mathcal{W}$, such that

$$(w' - w^*)^T Q(w^*) \geq 0, \quad \forall w' \in \mathcal{W}, \tag{4}$$

where

$$w = \begin{pmatrix} x \\ y \\ \lambda \end{pmatrix}, \quad Q(w) = \begin{pmatrix} f(x) - A^T\lambda \\ g(y) - B^T\lambda \\ Ax + By - b \end{pmatrix} \quad \text{and} \quad \mathcal{W} = \mathcal{X} \times \mathcal{Y} \times \mathcal{R}^m. \tag{5}$$

The exploitable structure of VI$(\mathcal{W}, Q)$ in (4)–(5) provides us the opportunities to design methods with separable algorithmic framework, i.e., decompose the original problem as a series of small scale subproblems involving x or y only. Actually, some classical decomposition methods have been developed and investigated in, e.g., [6,8,11–13]. The alternating direction method of multiplier (ADMM) is the benchmark among those decomposition methods. Specifically, given $w^k = (x^k, y^k, \lambda^k) \in \mathcal{W}$, ADMM produces the next iterate $w^{k+1} = (x^{k+1}, y^{k+1}, \lambda^{k+1})$ by solving the following variational inequalities

$$(x' - x)^T\{f(x) - A^T[\lambda^k - H(Ax + By^k - b)]\} \geq 0, \quad \forall x' \in \mathcal{X}, \tag{6}$$

$$(y' - y)^T\{g(y) - B^T[\lambda^k - H(Ax^{k+1} + By - b)]\} \geq 0, \quad \forall y' \in \mathcal{Y}, \tag{7}$$

$$\lambda^{k+1} = \lambda^k - H(Ax^{k+1} + By^{k+1} - b), \tag{8}$$

where $H \in \mathcal{R}^{m\times m}$ is a positive definite matrix which performs as a penalty parameter associated with the linear constraint.

Due to adequately exploiting the separable structure of VI$(\mathcal{W}, Q)$ in (4)–(5), ADMM exhibits great superiority to some state-of-the-art methods in many practical applications, e.g., signal processing, image restoration and matrix completion (see e.g., [5,18,20,21]). The success of ADMM in these applications is mainly attributed to its resulting subproblems (6)–(7) possessing closed-form solutions (see [5,18,20,21] for details). However, for the generic applications, if (6)–(7) have no exploitable structures to render closed-form solutions, it may be numerically intensive to be solved. To make (6)–(7) easier, He et al. [14] developed an alternating projection method for solving VI$(\mathcal{W}, Q)$ in (4)–(5). Instead of tackling (6)–(7) as variational inequalities, they utilized a projection–contraction algorithmic framework. Specifically, given $w^k = (x^k, y^k, \lambda^k) \in \mathcal{W}$, the algorithm in [14] (denoted by "APP") generates the predictor $\tilde{w}^k = (\tilde{x}^k, \tilde{y}^k, \tilde{\lambda}^k)$ by

$$\tilde{x}^k = P_{\mathcal{X}}\left\{x^k - \frac{1}{r_k}\left(f(x^k) - A^T[\lambda^k - H(Ax^k + By^k - b)]\right)\right\},$$

$$\tilde{y}^k = P_{\mathcal{Y}}\left\{y^k - \frac{1}{s_k}\left(g(y^k) - B^T[\lambda^k - H(A\tilde{x}^k + By^k - b)]\right)\right\},$$

$$\tilde{\lambda}^k = \lambda^k - H(A\tilde{x}^k + B\tilde{y}^k - b),$$

where the step sizes r_k and s_k are selected *self-adaptively* (see [14] for details) to satisfy the inequalities

$$\|f(x^k) - f(\tilde{x}^k) + A^T HA(x^k - \tilde{x}^k)\| \leq \nu r_k\|x^k - \tilde{x}^k\| \tag{9}$$

and

$$\|g(y^k) - g(\tilde{y}^k) + B^T HB(y^k - \tilde{y}^k)\| \leq \nu s_k\|y^k - \tilde{y}^k\|, \tag{10}$$

respectively; where ν is a parameter in $(0, 1)$.

The main computational effort of APP is the parameter selections of the step sizes r_k and s_k, i.e., selecting r_k (resp. s_k) such that (9) (resp. (10)) is satisfied at each iteration. Computationally, the smaller the left-hand sides of (9)–(10), the more opportunities for selecting the optimal step sizes for APP. Recently, Hu [16] proposed a modified alternating projection-based prediction–correction method (denoted by "MAPP") with the step sizes r_k and s_k being selected self-adaptively to satisfy

$$\|f(x^k) - f(\tilde{x}^k) + (1/2)A^T HA(x^k - \tilde{x}^k)\| \leq \nu r_k\|x^k - \tilde{x}^k\|, \tag{11}$$