Hindawi
Mathematical Problems in Engineering
Volume 2020, Article ID 6872810, 10 pages
https://doi.org/10.1155/2020/6872810

Research Article

A Parallel Splitting Augmented Lagrangian Method for Two-Block Separable Convex Programming with Application in Image Processing

Jing Liu[1,2] **Yongrui Duan,**[1] **and Tonghui Wang**[3]

[1]*School of Economics and Management, Tongji University, Shanghai, China*
[2]*School of Date Scineces, Zhejiang University of Finance and Economics, Hangzhou, China*
[3]*Department of Mathematical Science, New Mexico State University, Las Cruces, NM, USA*

Correspondence should be addressed to Jing Liu; ljlj8899@163.com

Received 30 August 2019; Revised 30 December 2019; Accepted 13 January 2020; Published 21 February 2020

Guest Editor: Marco Perez-Cisneros

The augmented Lagrangian method (ALM) is one of the most successful first-order methods for convex programming with linear equality constraints. To solve the two-block separable convex minimization problem, we always use the parallel splitting ALM method. In this paper, we will show that no matter how small the step size and the penalty parameter are, the convergence of the parallel splitting ALM is not guaranteed. We propose a new convergent parallel splitting ALM (PSALM), which is the regularizing ALM's minimization subproblem by some simple proximal terms. In application this new PSALM is used to solve video background extraction problems and our numerical results indicate that this new PSALM is efficient.

1. Introduction

Many problems arising from machine learning, such as compressive sensing [1, 2], the video background extraction problem [3–5], batch images alignment [6, 7], and transform invariant low-rank textures [8, 9] can be formulated as separable convex programming with linear constraints. In this paper, we consider the following two-block separable convex programming:

$$\min\left\{\theta_1(x_1)+\theta_2(x_2)\mid A_1x_1+A_2x_2=b,\quad x_1\in\mathcal{X}_1,\ x_2\in\mathcal{X}_2\right\},\tag{1}$$

where $i=1,2$, $A_i\in\mathfrak{R}^{\ell\times n_i}$, the $\ell\times n_i$ matrix on $\mathfrak{R}$; $b\in\mathfrak{R}^{\ell}$, the ℓ-dimensional vector on $\mathfrak{R}$; $\mathcal{X}_i$ is the subset of $\mathfrak{R}^{n_i}$. Here, θ_i: $\mathcal{X}_i\longmapsto\mathfrak{R}\ (i=1,2)$ are the proper lower semicontinuous convex functions, i.e.,

$$\begin{aligned}&\liminf_{x\longrightarrow\tilde{x}}\theta_i(x)=\theta_i(\tilde{x}),\quad\text{for all }\tilde{x}\in\operatorname{dom}(\theta_i)\neq\varnothing,\\&\theta_i(\lambda x+(1-\lambda)y)\leq\lambda\theta_i(x)+(1-\lambda)\theta_i(y),\\&\qquad\qquad\text{for all }x,y\in\mathcal{X}_i,\lambda\in[0,1].\end{aligned}\tag{2}$$

The solution set of problem (1), denoted by $\mathcal{X}^*$, is assumed to be nonempty. Note that theoretical results to problem (1) can easily be extended to matrix variables.

Let $\beta>0$ be a penalty parameter, and the augmented Lagrangian functions of problem (1) can be written as

$$\begin{aligned}\mathcal{L}_\beta(x_1,x_2,\lambda)=\theta_1(x_2)+\theta_2(x_2)-\lambda^\top(A_1x_1+A_2x_2-b)\\+\frac{\beta}{2}\|A_1x_1+A_2x_2-b\|^2,\end{aligned}\tag{3}$$

where $\lambda\in\mathfrak{R}^{\ell}$ is the Lagrangian multiplier. It is well-known that the augmented Lagrangian method (ALM) [10] is one of the most successful first-order methods for convex programming with linear constraints. Applying it to problem (1), we can obtain the following procedure:

$$\begin{cases}(x_1^{k+1},x_2^{k+1})=\arg\min\left\{\mathcal{L}_\beta(x_1,x_2,\lambda^k)\mid x_1\in\mathcal{X}_1,\ x_2\in\mathcal{X}_2\right\},\\\lambda^{k+1}=\lambda^k-\alpha\beta(A_1x_1^{k+1}+A_2x_2^{k+1}-b),\end{cases}\tag{4}$$

where $\alpha \in (0,2)$ is the step size. It has been proved that the sequence generated by the iterative scheme (4) is Fejér monotone and has the global convergence. In many practical problems, the proximal mapping $\mathrm{prox}_{\nu\theta_i}$ of θ_i, defined as

$$\mathrm{prox}_{\nu\theta_i}(x_i) := \arg\min_{y\in\mathscr{R}^{n_i}}\left\{\theta_i(x_i) + \frac{1}{2\nu}\|y - x_i\|^2\right\}, \tag{5}$$

has closed-form expression. For example, consider the video background extraction problem given by

$$\begin{aligned} \min_{L,S} \quad & \{\|L\|_* + \tau\|S\|_1\}, \\ \text{subject to} \quad & L + S = D, \end{aligned} \tag{6}$$

where

$$\|L\|_* = \sum_{i=1}^{\min\{m,n\}} \sqrt{\sigma_i(L)}, \quad \|S\|_1 = \sum_{i=1}^{m}\sum_{j=1}^{n}\left|s_{ij}\right|. \tag{7}$$

$\sigma_i(L)$ are the singular values of the $m \times n$ matrix L, $U\sum V^\top$ is the singular value decomposition of L, and s_{ij} are elements of the $m \times n$ matrix S. Then, by [11], the proximal mapping (5) is rewritten as

$$\mathrm{prox}_{\nu\|\cdot\|_*}(L) = U\mathrm{diag}(\mathcal{S}_c(\Sigma))V^\top. \tag{8}$$

Similarly, by [12], the proximal mapping (5) is given as

$$\mathrm{prox}_{\nu\|\cdot\|_1}(S) = \mathcal{S}_\nu(S), \tag{9}$$

where $\mathcal{S}_\nu(E)$ is defined by

$$(\mathcal{S}_\nu(E))_{ij} := \mathrm{sign}\left(E_{ij}\right)\cdot\max\left\{\left|E_{ij}\right| - \nu, 0\right\}, \quad 1 \le i \le m,\ 1 \le j \le n, \tag{10}$$

and $\mathrm{sign}(\cdot)$ is the sign function. However, the proximal mapping of $\|L\|_* + \tau\|S\|_1$ does not admit closed-form expression so that the iterative scheme (4) for problem (5) has to resort some inner solver to compute its minimization problem in each iteration, which inevitably creates a chain of problems, such as an efficient solver to solve the minimization problem and stopping criterion to ensure the global convergence of inexact version of the iterative scheme (5). Therefore, we should not ignore the separable structure of the objective function of problem (1). That is, we'd better split the minimization subproblem in (5) into some smaller scale subproblems to fully utilize the individual property of θ_i.

The alternating direction method of multipliers (ADMM) can decompose a complicated problem into some small-scale subproblems and solve these subproblems in a sequential manner or a parallel manner. Focusing on the way of splitting the augmented Lagrangian function in the spirit of the well-known alternating direction method (ADM), Tao and Yuan [4] have proposed a variant of the alternating splitting augmented Lagrangian method (ASALM) with convergent property, which can solve three-block separable convex programming. He et al. [13] propose a splitting method for solving a separable convex minimization problem with linear constraints, where the objective function is expressed as the sum of m individual functions without coupled variables. Aybat and Iyenga [14] have designed a variant of ADMM with increasing penalty for stable principal component pursuit (ADMIP) to solve the SPCP problem. Its preliminary computational tests show that ADMIP works very well in practice and outperforms ASALM which is a state-of-the-art ADMM algorithm for the SPCP problem with a constant penalty parameter.

In this paper, following the abovementioned procedure, we split the minimization problem of ALM in the Jacobian manner and get the following iterative scheme:

$$\begin{cases} x_1^{k+1} = \arg\min\left\{\mathscr{L}_\beta\left(x_1, x_2^k, \lambda^k\right) \mid x_1 \in \mathscr{X}_1\right\}, \\ x_2^{k+1} = \arg\min\left\{\mathscr{L}_\beta\left(x_1^k, x_2, \lambda^k\right) \mid x_2 \in \mathscr{X}_2\right\}, \\ \lambda^{k+1} = \lambda^k - \alpha\beta\left(A_1x_1^{k+1} + A_2x_2^{k+1} - b\right). \end{cases} \tag{11}$$

The iterative scheme (11) takes advantages of the separable structure and parallel architectures. Unfortunately, it suffers from the weakness that its convergence cannot be guaranteed under the convex assumption of θ_i and $\mathscr{X}^* \neq \varnothing$. To verify this, we shall apply the iterative scheme (11) to solve problem (1) given in Section 2 and show that the generated sequence diverges even $\beta \longrightarrow 0^+$. So, in Section 2, we propose a new parallel splitting ALM for problem (1), which is not only convergent but also keeps the good properties of the original ALM. Our numerical results in Section 3 indicate that the new PSALM is efficient.

The remainder of this paper is organized as follows. In Section 2, an example is given showing that the iterative scheme (11) maybe diverge, and a new parallel splitting ALM for problem (1) is proposed, which has the global convergence in both ergodic and nonergodic senses. For illustration of our main results, two numerical examples are given in Section 3 and a brief conclusion is drawn in Section 4.

2. Algorithm and Convergence Results

In this section, we will show that the iterative scheme (11) with any small positive α and β can be diverge so that a new parallel splitting ALM for problem (1) is proposed and the corresponding convergence results are established.

In what follows, $\mathfrak{R}^n$ will stand for the n-dimensional Euclidean space, and set

$$\begin{aligned} x &= (x_1, x_2), \\ w &= (x_1, x_2, \lambda), \\ \theta(x) &= \theta_1(x_1) + \theta_2(x_2), \\ \mathscr{A} &= (A_1, A_2), \\ \mathscr{X} &= \mathscr{X}_1 \times \mathscr{X}_2, \\ \mathscr{W} &= \mathscr{X} \times \mathfrak{R}^l. \end{aligned} \tag{12}$$

By the characterization of optimality condition for convex optimization, problem (1) can be readily characterized as the following mixed variational inequalities denoted by MVI $(\mathscr{W}, F, \theta)$:

$$\theta(x) - \theta(x^*) + (w - w^*)^\top F(w^*) \ge 0, \quad \forall w \in \mathscr{W}, \tag{13}$$

where

$$F(w)=\begin{pmatrix}-A_1^\top\lambda\\ -A_2^\top\lambda\\ \mathscr{A}x-b\end{pmatrix}. \tag{14}$$

Because $F(w)$ is a linear mapping with the skew-symmetric coefficient matrix, it satisfies the following property:

$$(w'-w)^\top F(w')=(w'-w)^\top F(w),\quad \forall w',\, w\in\mathscr{W}. \tag{15}$$

The solution set of MVI $(\mathscr{W},F,\theta)$ is denoted by $\mathscr{W}^*$, which is nonempty under the assumption $\mathscr{X}^*\neq\varnothing$.

Based on the prototype algorithm proposed by He and Yuan in [15], we can introduce the following prototype algorithm to solve MVI $(\mathscr{W},F,\theta)$:

A prototype algorithm for MVI $(\mathscr{W},F,\theta)$ is denoted by ProAlo

[Prediction] For given w^k, find $\widehat{w}^k\in\mathscr{W}$ and Q satisfying

$$\theta(x)-\theta(\widehat{x}^k)+(w-\widehat{w}^k)^\top F(\widehat{w}^k)\geq(w-\widehat{w}^k)^\top Q(w^k-\widehat{w}^k),\quad \forall w\in\mathscr{W}, \tag{16}$$

where the matrix Q has the property $(Q+Q^\top)$ which is positive definite

[Correction] Determine a nonsingular matrix M, a scalar $\alpha>0$, and generate the new iterate w^{k+1} via

$$w^{k+1}=w^k-\alpha M(w^k-\widehat{w}^k). \tag{17}$$

Similar to [15], we have the following convergence condition, which ensures that the ProAlo is globally convergent and has worst-case $\mathcal{O}(1/t)$ convergence rate in ergodic and nonergodic senses.

Convergence Condition

The matrices $Q+Q^\top, H:=QM^{-1}, G(\alpha):=Q+Q^\top-\alpha M^\top HM$ are positive definite.

Now we give an example to show that the iterative scheme (11) maybe diverge for problem (1).

Example 1. Consider the linear equation [16]

$$x_1+x_2=0. \tag{18}$$

Obviously, the linear equation (18) is a special case of problem (1) with the specifications: $\theta_1=\theta_2=0, A_1=A_2=1, b=0$, and $\mathscr{X}_1=\mathscr{X}_2=\mathfrak{R}$. The augmented Lagrangian function of problem (18) is given by

$$\mathscr{L}_\beta(x_1,x_2,\lambda)=-\lambda^\top(x_1+x_2)+\frac{\beta}{2}\|x_1+x_2\|^2. \tag{19}$$

Then, applying the iterative scheme (11) to solve problem (18) gives the following procedure:

$$\begin{pmatrix}x_1^{k+1}\\ x_2^{k+1}\\ \lambda^{k+1}\end{pmatrix}=P\begin{pmatrix}x_1^k\\ x_2^k\\ \lambda^k\end{pmatrix}, \tag{20}$$

where

$$P=\begin{pmatrix}0 & -1 & \dfrac{1}{\beta}\\ -1 & 0 & \dfrac{1}{\beta}\\ \alpha\beta & \alpha\beta & 1-2\alpha\end{pmatrix}. \tag{21}$$

Three eigenvalues of the matrix P are

$$\begin{aligned}\lambda_1&=1,\\ \lambda_2&=-\alpha+\sqrt{\alpha^2+1},\\ \lambda_3&=-\alpha-\sqrt{\alpha^2+1}.\end{aligned} \tag{22}$$

For any $\alpha,\beta>0$, we have $\rho(P)=\alpha+\sqrt{\alpha^2+1}>1$, let $\rho(P)$ is the spectral radius of P. Hence, the iterative scheme (20) with $\alpha,\beta>0$ is divergent, which also indicates that the iterative scheme (11) diverge for problem (1) under the traditional assumption. In the following, we shall design a new parallel splitting ALM for problem (1), which maximally inherits the structure of the iterative scheme (11).

Algorithm 1. new PSALM

Step 1. Given an initial point $w^0=(x_1^0,x_2^0,\lambda^0)\in\mathscr{W}$; let parameters $\nu>0, \alpha\in(0,2\nu/1+\nu)$, $\beta>0$, and the tolerance $\varepsilon>0$. Set $k=0$.

Step 2. Compute the new iterate $w^{k+1}=(x_1^{k+1},x_2^{k+1},\lambda^{k+1})$ via

$$\begin{cases}x_1^{k+1}=\arg\min\left\{\mathscr{L}_\beta(x_1,x_2^k,\lambda^k)+\dfrac{\nu\beta}{2}\|A_1x_1-A_1x_1^k\|^2 \mid x_1\in\mathscr{X}_1\right\},\\[2ex] x_2^{k+1}=\arg\min\left\{\mathscr{L}_\beta(x_1^k,x_2,\lambda^k)+\dfrac{\nu\beta}{2}\|A_2x_2-A_2x_2^k\|^2 \mid x_2\in\mathscr{X}_2\right\},\\[2ex] \lambda^{k+1}=\lambda^k-\alpha\beta(A_1x_1^{k+1}+A_2x_2^{k+1}-b).\end{cases} \tag{23}$$

Step 3. If $\max\{\|A_1x_1^k - A_1x_1^{k+1}\|, \|A_2x_2^k - A_2x_2^{k+1}\|, \|\lambda^k - \lambda^{k+1}\|\} \le \varepsilon$, then stop; otherwise, replace $k+1$ by k, and go to Step 2.

Remark 1. The new PSALM iterative scheme (23) not only preserves the main advantage of the ALM iterative scheme (11) in treating the objective functions θ_1 and θ_2 individually, but also numerically accelerates (11) with some values of the step size α, e.g., $\alpha \in (1,2)$, which can be achieved when $\nu > 1$.

In the following, we further make the following assumption about problem (1).

Assumption 1. The matrices A_1 and A_2 in problem (1) are both full column ranks.

To establish the convergence results of new PSALM, we first cast it to a special case of the ProAlo. To accomplish this, let us define an auxiliary sequence $\{\widehat{w}^k = (\widehat{x}_1^k, \widehat{x}_2^k, \widehat{\lambda}^k)\}$ as follows:

$$\begin{aligned} \widehat{x}_i^k &= x_i^{k+1}, \quad i = 1,2, \\ \widehat{\lambda}^k &= \lambda^k - \beta\left(A_1x_1^k + A_2x_2^k - b\right). \end{aligned} \tag{24}$$

This together with the updated formula of λ in (23) gives

$$\lambda^{k+1} = \lambda^k - \alpha\left[-\beta A_1\left(x_1^k - \widehat{x}_1^k\right) - \beta A_2\left(x_2^k - \widehat{x}_2^k\right) + \left(\lambda^k - \widehat{\lambda}^k\right)\right]. \tag{25}$$

Thus,

$$w^{k+1} = w^k - \alpha M\left(w^k - \widehat{w}^k\right), \tag{26}$$

where the matrix M is defined as

$$M = \begin{pmatrix} \frac{I_{n_1}}{\alpha} & 0 & 0 \\ 0 & \frac{I_{n_2}}{\alpha} & 0 \\ -\beta A_1 & -\beta A_2 & I_l \end{pmatrix}. \tag{27}$$

Lemma 1. *Let $\{w^k\}$ be the sequence generated by the new PSALM and $\{\widehat{w}^k\}$ be defined as in (24), and it holds that*

$$\theta(x) - \theta\left(\widehat{x}^k\right) + \left(w - \widehat{w}^k\right)^\top F\left(\widehat{w}^k\right) \ge \left(w - \widehat{w}^k\right)^\top Q\left(w^k - \widehat{w}^k\right), \quad \forall w \in \mathcal{W}, \tag{28}$$

where the matrix Q is defined by

$$Q = \begin{pmatrix} \beta(1+\nu)A_1^\top A_1 & 0 & 0 \\ 0 & \beta(1+\nu)A_2^\top A_2 & 0 \\ -A_1 & -A_2 & \frac{I_l}{\beta} \end{pmatrix}. \tag{29}$$

Proof. By the first-order optimality condition of the two minimization problems in (23), for any $x_1 \in \mathcal{X}_1$ and $x_2 \in \mathcal{X}_2$, we have

$$\begin{aligned} &\theta_1(x_1) - \theta_1\left(\widehat{x}_1^k\right) + \left(x_1 - \widehat{x}_1^k\right)^\top \Big(-A_1^\top\widehat{\lambda}^k - \beta(1+\nu)A_1^\top A_1 \\ &\quad \cdot \left(x_1^k - \widehat{x}_1^k\right)\Big) \ge 0, \\ &\theta_2(x_2) - \theta_2\left(\widehat{x}_2^k\right) + \left(x_2 - \widehat{x}_2^k\right)^\top \Big(-A_2^\top\widehat{\lambda}^k - \beta(1+\nu)A_2^\top A_2 \\ &\quad \cdot \left(x_2^k - \widehat{x}_2^k\right)\Big) \ge 0. \end{aligned} \tag{30}$$

The definition of the variable $\widehat{\lambda}^k$ in (24) gives

$$\begin{aligned} &\left(\lambda - \widehat{\lambda}^k\right)^\top \Big(A_1\widehat{x}_1^k + A_2\widehat{x}_2^k - b + A_1\left(x_1^k - \widehat{x}_1^k\right) + A_2\left(x_2^k - \widehat{x}_2^k\right) \\ &\quad + \frac{1}{\beta}\left(\widehat{\lambda}^k - \lambda^k\right)\Big) \ge 0, \quad \forall \lambda \in \mathfrak{R}^l. \end{aligned} \tag{31}$$

Adding the above three inequalities, using the notation of $\theta(x)$ and $F(w)$, we can get the compact form (28) and this completes the proof. □

Remark 2. If $\max\{\|A_1x_1^k - A_1x_1^{k+1}\|, \|A_2x_2^k - A_2x_2^{k+1}\|, \|\lambda^k - \lambda^{k+1}\|\} = 0$, then by (24), we have $A_1x_1^k = A_1\widehat{x}_1^k, A_2x_2^k = A_2\widehat{x}_2^k, \lambda^k = \widehat{\lambda}^k$. So, $Q(w^k - \widehat{w}^k) = 0$. Substituting it into the right-hand side of (28), we have

$$\theta(x) - \theta\left(\widehat{x}^k\right) + \left(w - \widehat{w}^k\right)^\top F\left(\widehat{w}^k\right) \ge 0, \quad \forall w \in \mathcal{W}. \tag{32}$$

So, $w^k \in \mathcal{W}^*$, and the stopping criterion of the PSALM is reasonable.

With inequalities (26) and (28) in hand, to establish the convergence results of the PSALM, we only need to ensure the matrices M and Q defined as in (27) and (29) satisfy the convergence condition.

Lemma 2. *Let the matrices Q and M be defined as in (27) and (29). If $\alpha \in (0, 2\nu/1+\nu)$ and Assumption 1 holds, then*

(i) The matrix $Q + Q^\top$ is positive definite

(ii) The matrix $H = QM^{-1}$ is symmetric and positive definite

(iii) The matrix $G(\alpha) = Q + Q^\top - \alpha M^\top HM$ is symmetric and positive definite

Proof

(i) From the definition of the matrix Q, we have

$$Q + Q^\top = \begin{pmatrix} 2\beta(1+\nu)A_1^\top A_1 & 0 & -A_1^\top \\ 0 & 2\beta(1+\nu)A_2^\top A_2 & -A_2^\top \\ -A_1 & -A_2 & \frac{2I_l}{\beta} \end{pmatrix}, \tag{33}$$

which is positive definite from $\nu > 0$ and Assumption 1.

(ii) From the definition of the matrices Q and M, we have

$$H = \begin{pmatrix} \alpha\beta(1+\nu)A_1^\top A_1 & 0 & 0 \\ 0 & \alpha\beta(1+\nu)A_2^\top A_2 & 0 \\ 0 & 0 & \frac{I_l}{\beta} \end{pmatrix}, \tag{34}$$

which is symmetric and positive definite by Assumption 1.

(iii) Similarly, from the definition of Q, M, we have

$$\begin{aligned} G(\alpha) &= \begin{pmatrix} \beta(1+\nu-\alpha)A_1^\top A_1 & -\alpha\beta A_1^\top A_2 & -(1-\alpha)A_1^\top \\ -\alpha\beta A_2^\top A_1 & \beta(1+\nu-\alpha)A_2^\top A_2 & -(1-\alpha)A_2^\top \\ -(1-\alpha)A_1 & -(1-\alpha)A_2 & \frac{(2-\alpha)I_l}{\beta} \end{pmatrix} \\ &= L^\top R(\alpha) L, \end{aligned} \tag{35}$$

where

$$L = \begin{pmatrix} \sqrt{\beta}A_1 & 0 & 0 \\ 0 & \sqrt{\beta}A_2 & 0 \\ 0 & 0 & \frac{I_l}{\sqrt{\beta}} \end{pmatrix}, \tag{36}$$

$$R(\alpha) = \begin{pmatrix} (1+\nu-\alpha)I_l & -\alpha I_l & -(1-\alpha)I_l \\ -\alpha I_l & (1+\nu-\alpha)I_l & -(1-\alpha)I_l \\ -(1-\alpha)I_l & -(1-\alpha)I_l & (2-\alpha)I_l \end{pmatrix}.$$

From Assumption 1, we only need to prove the matrix $R(\alpha)$ is positive definite. In fact, it can be written as

$$\begin{pmatrix} 1+\nu-\alpha & -\alpha & -(1-\alpha) \\ -\alpha & 1+\nu-\alpha & -(1-\alpha) \\ -(1-\alpha) & -(1-\alpha) & 2-\alpha \end{pmatrix} \otimes I_l, \tag{37}$$

where $\otimes$ denotes the matrix Kronecker product. Thus, we only need to prove the 3 order matrix

$$\begin{pmatrix} 1+\nu-\alpha & -\alpha & -(1-\alpha) \\ -\alpha & 1+\nu-\alpha & -(1-\alpha) \\ -(1-\alpha) & -(1-\alpha) & 2-\alpha \end{pmatrix}, \tag{38}$$

is positive definite. Its three eigenvalues are $\lambda_1 = 1+\nu$, $\lambda_2 = (3+\nu-3\alpha-\sqrt{9\alpha^2-2\alpha\nu-14\alpha+\nu^2-2\nu+9})/2$, $\lambda_3 = (3+\nu-3\alpha+\sqrt{9\alpha^2-2\alpha\nu-14\alpha+\nu^2-2\nu+9})/2$. Then, solving the following three inequalities

$$\begin{cases} 1+\nu > 0, \\ 3+\nu-3\alpha-\sqrt{9\alpha^2-2\alpha\nu-14\alpha+\nu^2-2\nu+9} > 0, \\ 3+\nu-3\alpha+\sqrt{9\alpha^2-2\alpha\nu-14\alpha+\nu^2-2\nu+9} > 0, \end{cases} \tag{39}$$

we get $0 < \alpha < 2\nu/1+\nu$. Therefore, the matrix G is positive definite for any $\alpha \in (0, 2\nu/1+\nu)$. In addition, from its expression, the matrix $G(\alpha)$ is obviously symmetric. This completes the proof.

Lemma 2 indicates that the matrices Q and M defined as in (27) and (29) satisfy convergence condition, and according to the result in [15], we get the following convergence results of new PSALM. □

Theorem 1 (global convergence). *Let $\{w^k\}$ be the sequence generated by new PSALM. Then, it converges to some w^∞, which belongs to $\mathcal{W}^*$.*

Theorem 2 (ergodic convergence rate). *Let $\{w^k\}$ be the sequence generated by new PSALM, $\{\widehat{w}^k\}$ be the corresponding sequence defined as in (24). Set*

$$\overline{w}^t = \frac{1}{t}\sum_{k=0}^{t-1} \widehat{w}^k. \tag{40}$$

Then, for any integer $t \geq 0$, we have

$$\theta(\overline{x}^t) - \theta(x) + (\overline{w}^t - w)^\top F(w) \leq \frac{1}{2\alpha t}\|w - w^0\|_H^2, \quad \forall w \in \mathcal{W}. \tag{41}$$

Theorem 3 (nonergodic convergence rate). *Let $\{w^k\}$ be the sequence generated by new PSALM. Then, for any $w^* \in \mathcal{W}^*$ and integer $t \geq 0$, we have*

$$\|M(w^t - \widehat{w}^t)\|_H^2 \leq \frac{1}{c_0 t}\|w^0 - w^*\|_H^2, \tag{42}$$

where $c_0 > 0$ is a constant.

Note that the right-hand side of (41) is independent of the distance between the initial iterate w^0 and the solution set $\mathcal{W}^*$. Therefore, (41) is not a reasonable criterion to measure the nonergodic convergence rate. In the following, we give a refined result to measure the nonergodic convergence rate of the sequence $\{w^k\}$ generated by new PSALM.

Lemma 4. *Let $\{w^k\}$ be the sequence generated by new PSALM. Then, for any $w \in \mathcal{W}$, we have*

$$\begin{aligned} &\alpha\left(\theta(x) - \theta(\widehat{x}^k) + (w - \widehat{w}^k)^\top F(\widehat{w}^k)\right) \\ &\geq \frac{1}{2}\left(\|w - w^{k+1}\|_H^2 - \|w - w^k\|_H^2\right) + \frac{\alpha}{2} w^k - \widehat{w}_H^{k2}. \end{aligned} \tag{43}$$

Proof. The proof is similar to that of Lemma 3.1 in [15] and is omitted for brevity of this paper.

Theorem 4 (refined ergodic convergence rate). *Let $\{w^k\}$ be the sequence generated by the PSALM, and set*

$$\overline{w}^t = \frac{1}{t}\sum_{k=0}^{t-1}\widehat{w}^k. \tag{44}$$

Then, for any integer $t \ge 0$, there exists a constant $c > 0$ such that

$$\begin{cases} \left|\theta(\overline{x}^t) - \theta(x^*)\right| \le \dfrac{c}{2\alpha t}, \\ \|\mathscr{A}\overline{x}^t - b\| \le \dfrac{c}{2\alpha t}. \end{cases} \tag{45}$$

Proof. Choose $w^* = (x^*, \lambda^*) \in \mathscr{W}^*$. Then, for any $\lambda \in \mathfrak{R}^l$, we have $\widetilde{w}^* := (x^*, \lambda) \in \mathscr{W}$. From the notion of $F(w)$ and (15), we have

$$\begin{aligned} &\left(\widetilde{w}^* - \widehat{w}^k\right)^\top F\left(\widehat{w}^k\right) \\ &= \left(\widetilde{w}^* - \widehat{w}^k\right)^\top F\left(\widetilde{w}^*\right) \\ &= \begin{pmatrix} x^* - \widehat{x}^k \\ \lambda - \widehat{\lambda}^k \end{pmatrix}^\top \begin{pmatrix} -\mathscr{A}^\top\lambda \\ \mathscr{A}x^* - b \end{pmatrix} \\ &= -\lambda^\top\left(\mathscr{A}x^* - \mathscr{A}\widehat{x}^k\right) \\ &= \lambda^\top\left(\mathscr{A}\widehat{x}^k - b\right). \end{aligned} \tag{46}$$

Set $w = \widetilde{w}^*$ in (43), and we obtain

$$\begin{aligned} &\alpha\left(\theta\left(\widehat{x}^k\right) - \theta(x^*) - \left(\widetilde{w}^* - \widehat{w}^k\right)^\top F\left(\widehat{w}^k\right)\right) \\ &\quad \le \frac{1}{2}\left(\left\|\widetilde{w}^* - w^k\right\|_H^2 - \left\|\widetilde{w}^* - w^{k+1}\right\|_H^2\right) - \frac{\alpha}{2}\left\|w^k - \widehat{w}^k\right\|_H^2. \end{aligned} \tag{47}$$

Combining the above two inequalities gives

$$\begin{aligned} &\alpha\left(\theta\left(\widehat{x}^k\right) - \theta(x^*) - \lambda^\top\left(\mathscr{A}\widehat{x}^k - b\right)\right) \\ &\quad \le \frac{1}{2}\left(\left\|\widetilde{w}^* - w^k\right\|_H^2 - \left\|\widetilde{w}^* - w^{k+1}\right\|_H^2\right) - \frac{\alpha}{2}\left\|w^k - \widehat{w}^k\right\|_H^2. \end{aligned} \tag{48}$$

Summing the abovementioned inequality from $k = 0$ to $t-1$ yields

$$\sum_{k=0}^{t-1}\theta\left(\widehat{x}^k\right) - t\theta(x^*) - \lambda^\top\left(\mathscr{A}\sum_{k=0}^{t-1}\widehat{x}^k - tb\right) \le \frac{1}{2\alpha}\left\|\widetilde{w}^* - w^0\right\|_H^2. \tag{49}$$

Dividing the both sides of the above inequality by t, we obtain

$$\frac{1}{t}\sum_{k=0}^{t-1}\theta\left(\widehat{x}^k\right) - \theta(x^*) - \lambda^\top\left(\mathscr{A}\overline{x}^t - b\right) \le \frac{1}{2\alpha t}\left\|\widetilde{w}^* - w^0\right\|_H^2. \tag{50}$$

Then, it follows from the convexity of θ_1 and θ_2 that

$$\theta\left(\overline{x}^t\right) - \theta(x^*) - \lambda^\top\left(\mathscr{A}\overline{x}^t - b\right) \le \frac{1}{2\alpha t}\left\|\begin{pmatrix} x^0 - x^* \\ \lambda^0 - \lambda \end{pmatrix}\right\|_H^2. \tag{51}$$

Since (51) holds for any λ, we set

$$\lambda = -\frac{\mathscr{A}\overline{x}^t - b}{\|\mathscr{A}\overline{x}^t - b\|}, \tag{52}$$

and obtain

$$\theta\left(\overline{x}^t\right) - \theta(x^*) + \left\|\mathscr{A}\overline{x}^t - b\right\| \le \frac{1}{2\alpha t}\sup_{\|\lambda\|\le 1}\left\|\begin{pmatrix} x^0 - x^* \\ \lambda^0 - \lambda \end{pmatrix}\right\|_H^2. \tag{53}$$

Set

$$c = \sup_{\|\lambda\|\le 1}\left\|\begin{pmatrix} x^0 - x^* \\ \lambda^0 - \lambda \end{pmatrix}\right\|_H^2, \tag{54}$$

and we thus obtain

$$\theta\left(\overline{x}^t\right) - \theta(x^*) + \left\|\mathscr{A}\overline{x}^t - b\right\| \le \frac{c}{2\alpha t}. \tag{55}$$

Since $x^* \in \mathscr{X}^*$, we have

$$\theta\left(\overline{x}^t\right) - \theta(x^*) \ge 0. \tag{56}$$

Combining the above two inequalities gives

$$\begin{cases} \left|\theta(\overline{x}^t) - \theta(x^*)\right| \le \dfrac{c}{2\alpha t}, \\ \|\mathscr{A}\overline{x}^t - b\| \le \dfrac{c}{2\alpha t}, \end{cases} \tag{57}$$

which completes the proof. □

The following example is to test the influence of the parameters ν and α on this new PSALM.

Example 2 (go on with Example 1). For this new PSALM, we set $\beta = 1$, and the resulting iterative scheme is

$$\begin{pmatrix} x_1^{k+1} \\ x_2^{k+1} \\ \lambda^{k+1} \end{pmatrix} = \frac{1}{1+\nu}\begin{pmatrix} \nu & -1 & 1 \\ -1 & \nu & 1 \\ \alpha(1-\nu) & \alpha(1-\nu) & 1+\nu-2\alpha \end{pmatrix}\begin{pmatrix} x_1^k \\ x_2^k \\ \lambda^k \end{pmatrix}. \tag{58}$$

Furthermore, we set the initial point $x_1^0 = x_2^0 = 0$ and $\lambda^0 = 1$, and the stopping criterion is set as

$$\max\left\{\left|x_1^k + x_2^k\right|, \left|\lambda^k\right|\right\} \le 10^{-5}, \tag{59}$$

or the number of iterations exceeds 10000.

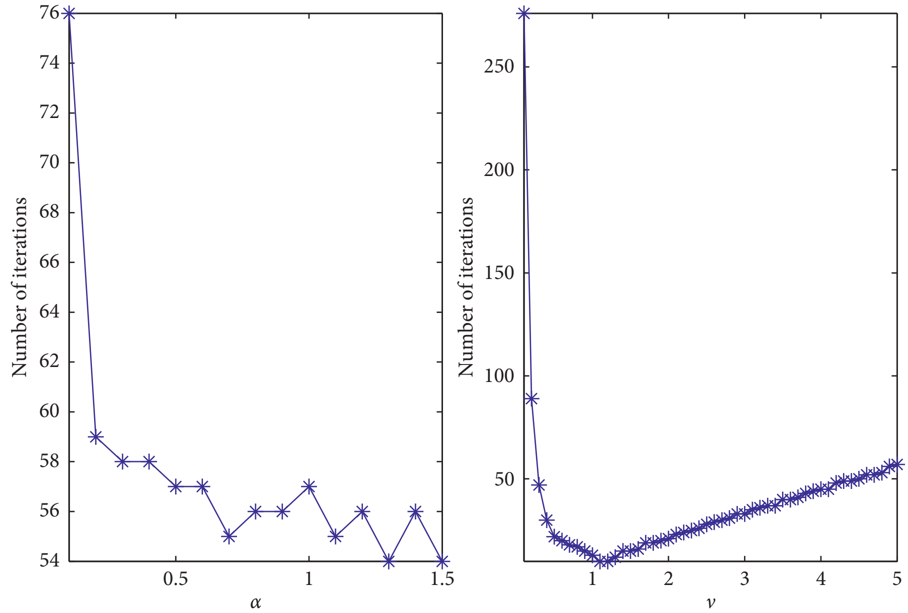


Figure 1: Sensitivity test on the parameters ν and α.

The left subplot of Figure 1 graphically depicts the numerical results when $\nu = 5$ and $\alpha = 0.1\colon 0.1\colon 1.5$. It reveals that when ν is fixed, the larger α is, the better of the new PSALM's performance is. The right subplot of Figure 1 graphically depicts the numerical results when $\nu = 0.1\colon 0.1\colon 5$ and $\alpha = (1.9\nu/1+\nu)$. It reveals that when ν takes some values near 1, the new PSALM always performs best. This is in conformity with the intuition because the large value of this parameter makes the proximal term $\nu\beta/2\|A_1x_1 - A_1x_1^k\|^2$ taking a too heavy weight in the objective function $\mathscr{L}_\beta(x_1, x_2^k, \lambda^k) + \nu\beta/2\|A_1x_1 - A_1x_1^k\|^2$, leads to a very small increment of the variable x_1, and finally decreases the convergence speed of the new PSALM. On the contrary, the small value of this parameter also makes the step size $\alpha = 1.9\nu/1+\nu$ to be very small. Therefore, how to choose a suitable combination of α and ν is a challenging issue.

Remark 3. The step size α is at the denominator of the right-hand side terms of (45). Therefore, the bigger α is, the smaller the right-hand side terms of (45) are, and the new PSALM may need less iteration to generate an ε-approximate solution of problem (1).

3. Numerical Results

In this section, we demonstrate the potential efficiency of this new PSALM by solving the video background extraction problem (6).

3.1. Simulation Study. Firstly, we generate a low-rank matrix $L^* \in \mathfrak{R}^{m\times m}$ and a sparse matrix $S^* \in \mathfrak{R}^{m\times m}$. Secondly, we get an observation matrix $D = L^* + S^*$ be an observation matrix. Thirdly, we want to recover the low-rank matrix L^* and the sparse matrix S^* by problem (6) with D as the input matrix and L and S as the unknown matrices. In the following, let r and spr represent the low-rank ratio of L^* and the ratios of the number of nonzeros entries of S^* (i.e., $\|S^*\|_0/(pq)$), respectively. We use the following Matlab script to generate matrix D:

(i) $L = \text{rand}n(m, r) * \text{rand}n(r, n)$

(ii) $S = \text{zeros}(m, n)$; $q = \text{randperm}(m * n)$; $K = \text{round}(\text{spr} * m * n)$; $S(q(1\colon K)) = \text{rand}n(K, 1)$

(iii) $D = L + S$

We set $m = n = 100$ and tested

$$(r, \text{spr}) = (1.0, 0.05), (5, 0.05), (10, 0.05), (15, 0.05), \\ (1.0, 0.1), (5, 0.1), (10, 0.1), (15, 0.1). \tag{60}$$

We empirically set $\tau = 0.1$ in problem (6), and $\alpha = 1.3, \beta = 0.05/\text{mean}(\text{abs}(D(:))), \nu = 2$ in the PSALM. The initial iterate $(L^0, S^0, \Lambda^0) = (0, 0, 0)$, and use the stopping criterion

$$\text{RelChg} := \frac{\|(L^{k+1}, S^{k+1}) - (L^k, S^k)\|}{1 + \|(L^k, S^k)\|} < 10^{-8}, \tag{61}$$

or the number of iterations exceeds 1000. Let $\widehat{L}$ and $\widehat{S}$ be the numerical solution of problem (6) obtained by the PSALM, and we measure the quality of recovery by the relative error to (L^*, S^*), which is defined by

$$\text{RelErr} := \frac{\|(\widehat{L}, \widehat{S}) - (L^*, S^*)\|}{1 + \|(L^*, S^*)\|}. \tag{62}$$

In Table 1, we report the recovery results on simulated matrices with size $m = n = 100$. Note that for the recovered sparse matrix $\widehat{S}$, to eliminate the roundabout error, we have reset the component whose the absolute value is less than 0.01 to be 0, i.e., $S(\text{abs}(S) < 0.01) = 0$. From this table, it can be seen that (i) the rank of the low-rank matrix L^* can be

Table 1: Numerical results on simulated matrices with size $m = n = 100$.

(r, spr)	RelErr	Rank (L^*)	Rank $(\widehat{L})$	$\|S^*\|_0$	$\|\widehat{S}\|_0$
(1, 0.05)	$3.7215e-05$	1	1	500	499
(1, 0.1)	$2.6670e-05$	1	1	1000	999
(5, 0.05)	$3.7402e-05$	5	5	500	499
(5, 0.1)	$2.7067e-05$	5	5	1000	1001
(10, 0.05)	$4.2738e$-06	10	10	500	500
(10, 0.1)	$4.9249e-06$	10	10	1000	1000
(15, 0.05)	$3.8210e-05$	15	15	500	504
(15, 0.1)	$3.1827e-05$	15	15	1000	1028

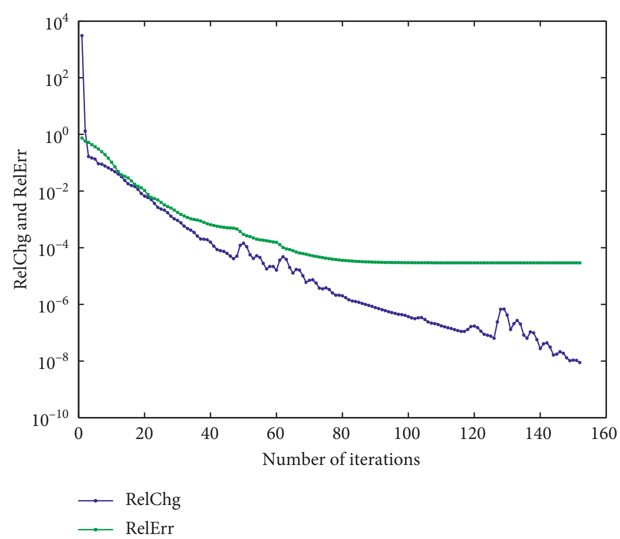
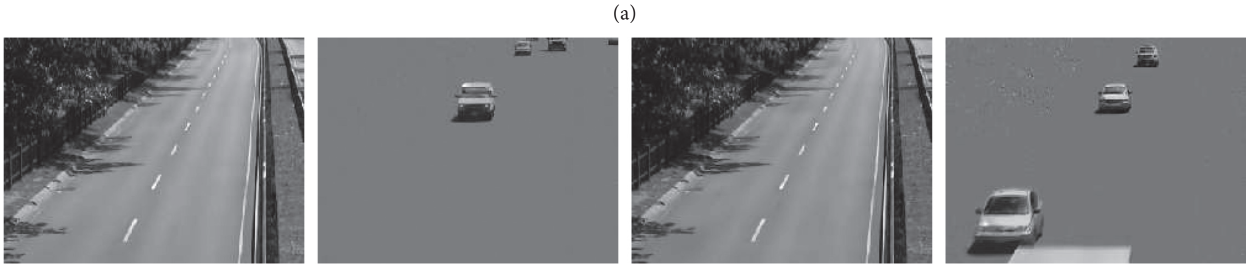


Figure 2: Convergence results for $r = 5$ and spr $= 0.1$.

(a)



(b)

Figure 3: Continued.

(c)

Figure 3: The decomposed results by TY_VASALM (a), HTY_IMA (b), and NEW_PSALM (c).

Table 2: Numerical results generated by TY_VASALM, HTY_IMA, and NEW_PSALM.

Algorithm	Iteration	CPU	Rank	Sparse
TY_VASALM	26	290.5518	58	5941722
HTY_IMA	20	389.5898	79	8146712
NEW_PSALM	17	51.8672	54	4683149

exactly recovered, while the sparse matrix S^* is recovered approximately; (ii) though S^* is not recovered completely, the relative error RelErr indicates that the precision of the recovered $\widehat{S}$ is high.

In Figure 2, we present the convergence result for problem (6) with $r = 5$ and spr $= 0.1$. From this figure, it can be observed that when the number of iterations $k = 80$, the relative error RelErr achieves at a stable state. It decreases almost linearly with k before $k = 80$, while it almost does not decrease any more after $k = 80$.

3.2. An Application Example. In this section, we apply the new PSALM(NEW_PSALM) to solve a concrete problem of model (6): the video background extraction problem [17], which aims to subtract the background from (the low-rank matrix L) a video clip (stacking as the matrix D) and meanwhile detect the moving objects (the sparse matrix S). We set $\tau = 1/(\sqrt{\max\{m,n\}})$ in problem (6). We have downloaded the video clip: Highway from the Internet. For comparison, we also applied the iteration method in [4] (denoted by TY_VASALM) and the iteration method in [13] (denoted by HTY_IMA) to solve this problem.

Figure 3 presents the extracted backgrounds and foregrounds images by TY_VASALM, HTY_IMA, and NEW_PSALM. Clearly, all methods can extract satisfactory results

Table 2 lists the detail experimental results, where the number of iterations (Iteration), CPU time in seconds (CPU), the rank of the recovered final low-rank matrix (Rank), and the number of nonzeros of the recovered final sparse matrix (Sparse) are reported. From the results given in Table 2, we see that NEW_PSALM is faster than the other two methods in terms of both iteration numbers and CPU time. In particular, NEW_PSALM outperforms all the other two methods in recovering low rank and sparse components. These results clearly illustrate the efficiency and robustness of NEW_PSALM.

4. Conclusion

In this paper, we have proposed a new parallel splitting augmented Lagrangian method(NEW_PSALM) for two-block separable convex programming and have established its various convergence results, including global convergence, ergodic, and nonerogdic convergence rate. Numerical results indicate that it is efficient for the video background extraction problem.

Practical experiments indicate that the ALM and its various variants with the step size $\alpha \in (1,2)$ can often accelerate the convergence speed, while in the new PSALM, to ensure the step size $\alpha \in (1,2)$, the parameter ν must larger than 1, such as $\alpha \in (0,1.9)$ only if $\nu \geq 19$. However, as we have mentioned in Example 2, large ν often decreases the convergence speed of the new PSALM. Therefore, the ALM type methods with less restrictions imposed on ν and α deserve further researching, and we leave this problem as a future research work.

Data Availability

The data used to support the findings of this study are available from http://www.SceneBackgroundModeling.net.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (71532015, 71528007, and 71771179) and Foundation of First Class Discipline of Zhejiang-A (Zhejiang University of Finance and Economics-Statistics).

References

[1] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.

[2] J. F. Yang and Y. Zhang, "Alternating direction algorithms for ℓ_1-problems in compressive sensing," *SIAM Journal on Scientific Computing*, vol. 33, no. 1, pp. 250–278, 2011.

[3] E. J. Candés, X. D. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of ACM*, vol. 58, no. 1, pp. 1–37, 2011.

[4] M. Tao and X. Yuan, "Recovering low-rank and sparse components of matrices from incomplete and noisy observations," *SIAM Journal on Optimization*, vol. 21, no. 1, pp. 57–81, 2011.

[5] L. Hou, H. He, and J. Yang, "A partially parallel splitting method for multiple-block separable convex programming with applications to robust PCA," *Computational Optimization and Applications*, vol. 63, no. 1, pp. 273–303, 2016.

[6] Y. Peng, A. Ganesh, J. Wright, W. Xu, and Y. Ma, "RASL: robust alignment by sparse and low-rank decomposition for linearly correlated images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2233–2246, 2012.

[7] S. Wang, Y. Xiao, and Z. Jin, "An efficient algorithm for batch images alignment with adaptive rank-correction term," *Journal of Computational and Applied Mathematics*, vol. 346, pp. 171–183, 2019.

[8] Z. Zhang, A. Ganesh, X. Liang, and Y. Ma, "TILT: transform invariant low-rank textures," *International Journal of Computer Vision*, vol. 99, no. 1, pp. 1–24, 2012.

[9] X. Ren and Z. Lin, "Linearized alternating direction method with adaptive penalty and warm starts for fast solving transform invariant low-rank textures," *International Journal of Computer Vision*, vol. 104, no. 1, pp. 1–14, 2013.

[10] M. R. Hestenes, "Multiplier and gradient methods," *Journal of Optimization Theory and Applications*, vol. 4, no. 5, pp. 303–320, 1969.

[11] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[12] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Review*, vol. 51, no. 1, pp. 34–81, 2009.

[13] B. He, M. Tao, and X. Yuan, "A splitting method for separable convex programming," *IMA Journal of Numerical Analysis*, vol. 35, no. 1, pp. 394–426, 2014.

[14] N. S. Aybat and G. Iyengar, "An alternating direction method with increasing penalty for stable principal component pursuit," *Computational Optimization and Applications*, vol. 61, no. 3, pp. 635–668, 2015.

[15] B. S. He and X. M. Yuan, *On the Direct Extension of ADMM for Multi-Block Separable Convex Programming and beyond: From Variational Inequality Perspective*, 2014, http://www.optimization-online.org/DB_FILE/2014/03/4293.pdf.

[16] B. He, L. Hou, and X. Yuan, "On full Jacobian decomposition of the augmented Lagrangian method for separable convex programming," *SIAM Journal on Optimization*, vol. 25, no. 4, pp. 2274–2312, 2015.

[17] T. Bouwmans, "Traditional and recent approaches in background modeling for foreground detection: an overview," *Computer Science Review*, vol. 11-12, pp. 31–66, 2014.