BIT Numer Math (2016) 56:399–422
DOI 10.1007/s10543-015-0563-z

Rigorous convergence analysis of alternating variable minimization with multiplier methods for quadratic programming problems with equality constraints

Zhong-Zhi Bai[1] · **Min Tao**[2]

Received: 2 February 2015 / Accepted: 19 May 2015 / Published online: 24 July 2015

Abstract We discuss unique solvability of the equality-constraint quadratic programming problem, establish a class of *preconditioned alternating variable minimization with multiplier* (*PAVMM*) methods for iteratively computing its solution, and demonstrate asymptotic convergence property of these PAVMM methods. We also discuss an algebraic derivation of the PAVMM method by making use of matrix splitting, which reveals that the PAVMM method is actually a modified block Gauss–Seidel iteration method for solving the augmented Lagrangian linear system resulting from the weighted Lagrangian function with respect to the equality-constraint quadratic programming problem.

Keywords Equality-constraint quadratic programming problem · Solvability · Iteration method · Preconditioning · Asymptotic convergence

---

Communicated by Lothar Reichel.

The work of these authors is supported by The National Basic Research Program (No. 2011CB309703), The National Natural Science Foundation (Nos. 11301280, 91118001), The National Natural Science Foundation for Creative Research Groups (No. 11321061) and The Fundamental Research Funds for the Central Universities (No. 020314330019), People's Republic of China.

✉ Zhong-Zhi Bai
bzz@lsec.cc.ac.cn

Min Tao
taom@nju.edu.cn

[1] State Key Laboratory of Scientific/Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, P.O. Box 2719, Beijing 100190, People's Republic of China

[2] Department of Mathematics, Nanjing University, Nanjing 210008, People's Republic of China

Mathematics Subject Classification 65F08 · 65F10 · 65K05 · 90C20 · 90C25 · CR: G1.3

1 Introduction

Let $\mathbb{R}$ be the domain of all real numbers, $\mathbb{R}^n$ be the n-dimensional real linear space equipped with the Euclidean inner product, say $\langle \cdot, \cdot \rangle$, and $\mathbb{R}^{m \times n}$ be the m-by-n real matrix space. Denote by $(\cdot)^T$ and $\|\cdot\|$ the transpose and the Euclidean norm of either a vector or a matrix of suitable dimension, respectively. We consider numerical solutions of equality-constraint quadratic programming problems of the form

$$\begin{cases} \min & \phi(x) + \psi(y), \\ \text{s.t.} & Ax + By = b, \end{cases} \tag{1.1}$$

where $A \in \mathbb{R}^{p \times n}$ and $B \in \mathbb{R}^{p \times m}$ are two matrices, $b \in \mathbb{R}^p$ is a known vector, and $\phi: \mathbb{R}^n \to \mathbb{R}$ and $\psi: \mathbb{R}^m \to \mathbb{R}$ are two quadratic functions defined by

$$\begin{cases} \phi(x) = \frac{1}{2} x^T F x + x^T f, \\ \psi(y) = \frac{1}{2} y^T G y + y^T g, \end{cases} \tag{1.2}$$

with $F \in \mathbb{R}^{n \times n}$, $G \in \mathbb{R}^{m \times m}$ being symmetric positive semidefinite matrices and $f \in \mathbb{R}^n$, $g \in \mathbb{R}^m$ being given vectors. We assume that some standard assumptions are imposed on the matrices F, G and A, B as well as on the vectors f, g and b such that the solution set of the problem (1.1)–(1.2) is nonempty. This class of constraint programming problems occurs in many areas of computational science and engineering applications such as economics [1], electrical circuits and networks [7,13,40], electromagnetism [9,35], finance [30,31], image reconstruction [24], image registration [22,33] and optimal control [8]. It also captures a number of important applications arising in various areas such as the l_1-norm regularized least-squares problems, the total variation image restoration and the standard quadratic programming problems; see, e.g., [25,28] for more details.

In fact, the constraint quadratic programming problem (1.1) is mathematically equivalent to the unconstraint optimization problem

$$\max_{z} \min_{x,y} \mathscr{L}_a(x, y, z), \tag{1.3}$$

where $\mathscr{L}_a(x, y, z)$ is the augmented Lagrangian function defined by

$$\mathscr{L}_a(x, y, z) = \phi(x) + \psi(y) - \langle Ax + By - b, z \rangle + \frac{\beta}{2} \|Ax + By - b\|^2, \tag{1.4}$$

with $z \in \mathbb{R}^p$ being the Lagrange multiplier and β a penalty or a regularization parameter. That is to say, a point $(x_*, y_*) \in \mathbb{R}^n \times \mathbb{R}^m$ is a solution of the problem (1.1)–(1.2) if and only if there exists a $z_* \in \mathbb{R}^p$ such that the point $(x_*, y_*, z_*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ is a solution of the problem (1.3)–(1.4).

One of the most popular and effective iterative methods for solving the equality-constraint quadratic programming problem (1.1) is the so-called *alternating direction* method with *multipliers*, or in short, the *ADM* method. At each iteration step, it first alternatively minimizes the augmented Lagrangian function $\mathscr{L}_a(x, y, z)$ with respect to the variables x, y, and then update the Lagrange multiplier z according to the steepest ascent principle so that violation of the original constraint $Ax + By = b$ is penalized. More precisely, the ADM method for solving the problem (1.1) can be algorithmically described as follows.

Method 1.1 (*ADM method for the problem* (1.1)) Given initial guesses $y^{(0)} \in \mathbb{R}^m$ and $z^{(0)} \in \mathbb{R}^p$, for $k = 0, 1, 2, \ldots$ until the iteration sequences $\{x^{(k)}\}_{k=0}^{\infty} \subset \mathbb{R}^n$, $\{y^{(k)}\}_{k=0}^{\infty} \subset \mathbb{R}^m$ and $\{z^{(k)}\}_{k=0}^{\infty} \subset \mathbb{R}^p$ are convergent, compute $x^{(k+1)}$, $y^{(k+1)}$ and $z^{(k+1)}$ according to the following rule:

$$\begin{aligned}
x^{(k+1)} &= \arg\min_{x\in\mathbb{R}^n} \left\{\phi(x) - \langle Ax + By^{(k)} - b, z^{(k)}\rangle + \frac{\beta}{2}\|Ax + By^{(k)} - b\|^2\right\},\\
y^{(k+1)} &= \arg\min_{y\in\mathbb{R}^m} \left\{\psi(y) - \langle Ax^{(k+1)} + By - b, z^{(k)}\rangle + \frac{\beta}{2}\|Ax^{(k+1)} + By - b\|^2\right\},\\
z^{(k+1)} &= z^{(k)} - \beta(Ax^{(k+1)} + By^{(k+1)} - b).
\end{aligned} \tag{1.5}$$

Intuitively, Method 1.1 is an *alternating variable minimization with multiplier* (*AVMM*) method. The AVMM method is intended to blend the decomposability of dual ascent with the superior convergence properties of the method of multipliers [11]. It computes a saddle point of the augmented Lagrangian function $\mathscr{L}_a(x, y, z)$ by adopting the idea of block Gauss–Seidel iteration for solving block system of linear or nonlinear equations [21,34,45], in which a single block Gauss–Seidel pass over the variables x and y is used instead of the usual joint minimization. In [18] Gabay illustrated this iteration scheme as an application of the Douglas-Rachford splitting method [29] to the dual of the problem (1.1), and Eckstein and Bertsekas [15] showed in turn that Douglas-Rachford splitting is a special case of the proximal point method. Hence, AVMM is a special case of the proximal point method; see Eckstein and Ferris [16] for more discussions explaining this approach. On the other hand, it is also a natural generalization of the classical Uzawa method for solving the saddle-point problems; see [1,10,14].

Many papers have analyzed the AVMM method from the perspective of maximal monotone operators [15,36–39]. Its global convergence was proved under some mild conditions such as the solution set of the problem (1.1) is nonempty; see [17–19]. Also, it has been known that this method converges linearly, but an accurate estimate about the convergence rate is still in its infancy; see, e.g., [20,25,28,29,41].

In this paper, based on a weighted inner product and the corresponding weighted norm, by adopting matrix preconditioning strategy and utilizing parameter accelerating technique, we establish a class of *preconditioned alternating variable minimization with multiplier* (*PAVMM*) methods for iteratively solving the equality-constraint quadratic programming problem (1.1)–(1.2). This method includes the AVMM or the ADM method as a special case. By making use of blockwise matrix transformation, from null space relationships of the involved sub-matrices we discuss solvability of the problem (1.1)–(1.2) and give sufficient and necessary conditions for guaranteeing existence and uniqueness of its solution. By exploring an explicit formula about eigenvalues of the iteration matrix, we demonstrate asymptotic convergence property and analyze asymptotic convergence rate of the PAVMM method. By making use of matrix splitting, we also discuss an algebraic derivation of the PAVMM method, which shows that this method is actually a modified block Gauss–Seidel iteration method for solving the augmented Lagrangian linear system resulting from the weighted Lagrangian function with respect to the equality-constraint quadratic programming problem (1.1)–(1.2).

2 The PAVMM methods

For a symmetric positive definite matrix $H \in \mathbb{R}^{p\times p}$, let $\langle\cdot,\cdot\rangle_H$ be the weighted inner product with the weighting matrix H, or the H-inner product, in $\mathbb{R}^p$, and $\|\cdot\|_H$ be the corresponding weighted matrix norm, or the H-norm, in $\mathbb{R}^{p\times p}$. Note that for $u, v \in \mathbb{R}^p$ and $X \in \mathbb{R}^{p\times p}$, it holds that

$$\langle u, v\rangle_H = \langle Hu, v\rangle, \quad \|u\|_H = \|H^{\frac{1}{2}}u\| \quad \text{and} \quad \|X\|_H = \|H^{\frac{1}{2}}XH^{-\frac{1}{2}}\|.$$

We say that $u, v \in \mathbb{R}^p$ are H-orthogonal, denoted by $u \perp_H v$, if $\langle u, v\rangle_H = 0$. If, in particular, $H = I$ (the identity matrix), then the vectors u and v are said to be orthogonal, which is simply represented as $u \perp v$.

For a given symmetric positive definite matrix $W \in \mathbb{R}^{p\times p}$, we can prove that the constraint quadratic programming problem (1.1) is mathematically equivalent to the unconstraint optimization problem

$$\max_{z} \min_{x,y} \mathscr{L}_{\text{wa}}(x, y, z), \tag{2.1}$$

where $\mathscr{L}_{\text{wa}}(x, y, z)$ is the weighted augmented Lagrangian function defined by

$$\mathscr{L}_{\text{wa}}(x, y, z) = \phi(x) + \psi(y) - \langle Ax + By - b, z\rangle_{W^{-1}} + \frac{\beta}{2}\|Ax + By - b\|^2_{W^{-1}}, \tag{2.2}$$

with $W \in \mathbb{R}^{p\times p}$ being the weighting matrix, $z \in \mathbb{R}^p$ being the Lagrange multiplier, and β a penalty or a regularization parameter. To be more precise, a point $(x_*, y_*) \in \mathbb{R}^n \times \mathbb{R}^m$ is a solution of the problem (1.1) if and only if there exists a $z_* \in \mathbb{R}^p$ such that the point $(x_*, y_*, z_*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ is a solution of the problem (2.1)–(2.2).

Analogous to Method 1.1, we can establish the following *PAVMM* method for iteratively solving the problem (1.1)–(1.2).

Method 2.1 (*PAVMM method for the problem* (1.1)–(1.2)) Let $W \in \mathbb{R}^{p\times p}$ be a symmetric positive definite matrix, $Q \in \mathbb{R}^{p\times p}$ be a nonsingular matrix, and α be a positive constant. Given initial guesses $y^{(0)} \in \mathbb{R}^m$ and $z^{(0)} \in \mathbb{R}^p$, for $k = 0, 1, 2, \ldots$ until the iteration sequences $\{x^{(k)}\}_{k=0}^{\infty} \subset \mathbb{R}^n$, $\{y^{(k)}\}_{k=0}^{\infty} \subset \mathbb{R}^m$ and $\{z^{(k)}\}_{k=0}^{\infty} \subset \mathbb{R}^p$ are convergent, compute $x^{(k+1)}$, $y^{(k+1)}$ and $z^{(k+1)}$ according to the following rule:

$$
\begin{aligned}
x^{(k+1)} &= \arg\min_{x\in\mathbb{R}^n} \left\{\phi(x) - \langle Ax + By^{(k)} - b, z^{(k)}\rangle_{W^{-1}} + \frac{\beta}{2}\|Ax + By^{(k)} - b\|^2_{W^{-1}}\right\},\\
y^{(k+1)} &= \arg\min_{y\in\mathbb{R}^m} \left\{\psi(y) - \langle Ax^{(k+1)} + By - b, z^{(k)}\rangle_{W^{-1}} + \frac{\beta}{2}\|Ax^{(k+1)} + By - b\|^2_{W^{-1}}\right\},\\
z^{(k+1)} &= z^{(k)} - \alpha Q^{-1}W^{-1}(Ax^{(k+1)} + By^{(k+1)} - b). \qquad (2.3)
\end{aligned}
$$

Note that when $\alpha = \beta$ and $W = Q = I$ the PAVMM method reduces to the AVMM or the ADM method. The weighting matrix W can be used to balance the cost function and the equality constraint such that the conditioning of the weighted augmented Lagrangian function $\mathscr{L}_{\mathrm{wa}}(x, y, z)$ is greatly improved, and the preconditioning matrix Q and the relaxation parameter α can be chosen such that the convergence rate of the PAVMM method is further accelerated.

By straightforward calculations, we can rewrite the iteration scheme (2.3) into an explicit form as follows:

$$
\begin{cases}
\phi'(x^{(k+1)}) - A^T W^{-1} z^{(k)} + \beta A^T W^{-1}(Ax^{(k+1)} + By^{(k)} - b) = 0,\\
\psi'(y^{(k+1)}) - B^T W^{-1} z^{(k)} + \beta B^T W^{-1}(Ax^{(k+1)} + By^{(k+1)} - b) = 0,\\
z^{(k+1)} = z^{(k)} - \alpha Q^{-1}W^{-1}(Ax^{(k+1)} + By^{(k+1)} - b),
\end{cases}
$$

which can be equivalently reformulated as

$$
\begin{cases}
Fx^{(k+1)} + f - A^T W^{-1} z^{(k)} + \beta A^T W^{-1}(Ax^{(k+1)} + By^{(k)} - b) = 0,\\
Gy^{(k+1)} + g - B^T W^{-1} z^{(k)} + \beta B^T W^{-1}(Ax^{(k+1)} + By^{(k+1)} - b) = 0,\\
z^{(k+1)} = z^{(k)} - \alpha Q^{-1}W^{-1}(Ax^{(k+1)} + By^{(k+1)} - b)
\end{cases}
$$

due to the concrete expressions

$$
\phi'(x) = Fx + f \quad \text{and} \quad \psi'(y) = Gy + g
$$

of the derivatives of the quadratic functions $\phi(x)$ and $\psi(y)$. After rearrangement we obtain the matrix–vector form of the PAVMM method as follows:

$$
\begin{cases}
(F + \beta A^T W^{-1} A)x^{(k+1)} = A^T W^{-1}[\beta(b - By^{(k)}) + z^{(k)}] - f,\\
(G + \beta B^T W^{-1} B)y^{(k+1)} = B^T W^{-1}[\beta(b - Ax^{(k+1)}) + z^{(k)}] - g,\\
z^{(k+1)} = z^{(k)} - \alpha Q^{-1}W^{-1}(Ax^{(k+1)} + By^{(k+1)} - b).
\end{cases} \qquad (2.4)
$$

The main costs of the iteration scheme (2.4) are solving two systems of linear equations having the coefficient matrices $F + \beta A^T W^{-1} A$ and $G + \beta B^T W^{-1} B$, which are of the sizes $n \times n$ and $m \times m$, respectively. When

$$
\mathrm{null}(F) \cap \mathrm{null}(A) = \{0\} \quad \text{and} \quad \mathrm{null}(G) \cap \mathrm{null}(B) = \{0\},
$$

with $\mathrm{null}(\cdot)$ indicating the null space of the corresponding matrix, the matrices $F + \beta A^T W^{-1} A$ and $G + \beta B^T W^{-1} B$ are symmetric positive definite, so that the corresponding systems of linear equations can be solved effectively either by direct methods such as the Cholesky factorization or by iterative methods such as the preconditioned conjugate gradient method; see [21]. Of course, the weighting matrix W and the regularization parameter β should be chosen such that both augmented Lagrangian matrices $F + \beta A^T W^{-1} A$ and $G + \beta B^T W^{-1} B$ are much better conditioned than the original matrices F and G, respectively.

3 Preliminaries

We first introduce some notations necessitating for the subsequent statements. We use $\mathbb{C}$ to denote the domain of all complex numbers. For a $\zeta \in \mathbb{C}$, $\bar{\zeta}$ stands for its conjugate complex. $\mathbb{C}^n$ represents the n-dimensional complex linear space equipped with the Euclidean inner product, say, $\langle \cdot, \cdot \rangle$, and $\mathbb{C}^{m\times n}$ is the m-by-n complex matrix space. For a given matrix, we use range$(\cdot)$ to indicate the range space spanned by all of its columns. The symbol $\oplus$ is the direct sum of subspaces in $\mathbb{C}^n$. Denote by $(\cdot)^*$ and $\|\cdot\|$ the conjugate transpose and the Euclidean norm of either a vector in $\mathbb{C}^n$ or a matrix in $\mathbb{C}^{m\times n}$, respectively. For two symmetric matrices X and Y, we say $X \succ Y$ (or $X \succeq Y$) if $X - Y$ is positive definite (or semidefinite). Alternatively, $X \succ Y$ (or $X \succeq Y$) is also written as $Y \prec X$ (or $Y \preceq X$). We use $\lambda(\cdot)$ and $\rho(\cdot)$ to indicate an eigenvalue and the spectral radius of a square matrix, and $\lambda_{\min}(\cdot)$ and $\lambda_{\max}(\cdot)$ the smallest and the largest of its eigenvalues, respectively.

The following result presents necessary and sufficient conditions for the nonsingularity of block two-by-two matrices of the saddle-point form; see [2,3,5] and the references therein.

Lemma 3.1 *Let $H \in \mathbb{C}^{n\times n}$ be a Hermitian positive semidefinite matrix and $E \in \mathbb{C}^{m\times n}$. Then the saddle-point matrix*

$$K = \begin{pmatrix} H & E^* \\ E & 0 \end{pmatrix}$$

is nonsingular if and only if $\mathrm{null}(H) \cap \mathrm{null}(E) = \{0\}$ *and E is of full row rank.*

For the nonsingularity of block three-by-three matrices, we can demonstrate necessary and sufficient conditions as follows.

Lemma 3.2 *Let $H_a \in \mathbb{C}^{n\times n}$ and $H_b \in \mathbb{C}^{m\times m}$ be Hermitian positive semidefinite matrices, and $E_a \in \mathbb{C}^{p\times n}$ and $E_b \in \mathbb{C}^{p\times m}$. Then the block three-by-three matrix*

$$K = \begin{pmatrix} H_a & 0 & E_a^* \\ 0 & H_b & E_b^* \\ E_a & E_b & 0 \end{pmatrix}$$

is nonsingular if and only if

$$(\mathrm{null}(H_a) \oplus \mathrm{null}(H_b)) \cap \mathrm{null}((E_a \;\; E_b)) = \{0\} \quad \textit{and} \quad \mathrm{null}(E_a^*) \cap \mathrm{null}(E_b^*) = \{0\}.$$

Proof Denote

$$H = \begin{pmatrix} H_a & 0 \\ 0 & H_b \end{pmatrix} \quad \text{and} \quad E = \begin{pmatrix} E_a & E_b \end{pmatrix}.$$

Then we can rewrite the block three-by-three matrix K as the block two-by-two one

$$K = \begin{pmatrix} H & E^* \\ E & 0 \end{pmatrix}.$$

From Lemma 3.1 we know that K is nonsingular if and only if $\text{null}(H) \cap \text{null}(E) = \{0\}$ and E^* is of full column rank. By straightforward operations we can obtain the following statements:

(i) $\text{null}(H) \cap \text{null}(E) = \{0\}$ if and only if

$$(\text{null}(H_a) \oplus \text{null}(H_b)) \cap \text{null}((E_a \ \ E_b)) = \{0\};$$

(ii) E^* is of full column rank if and only if $\text{null}(E_a^*) \cap \text{null}(E_b^*) = \{0\}$.

□

As

$$(\text{null}(H_a) \oplus \text{null}(H_b)) \cap \text{null}((E_a \ \ E_b)) = \{0\}$$

implies

$$\text{null}(H_a) \cap \text{null}(E_a) = \{0\} \quad \text{and} \quad \text{null}(H_b) \cap \text{null}(E_b) = \{0\},$$

by Lemma 3.2 we can directly obtain a necessary condition for the nonsingularity of the block three-by-three matrix K.

Lemma 3.3 *Let $H_a \in \mathbb{C}^{n\times n}$ and $H_b \in \mathbb{C}^{m\times m}$ be Hermitian positive semidefinite matrices, and $E_a \in \mathbb{C}^{p\times n}$ and $E_b \in \mathbb{C}^{p\times m}$. If*

$$K = \begin{pmatrix} H_a & 0 & E_a^* \\ 0 & H_b & E_b^* \\ E_a & E_b & 0 \end{pmatrix}$$

is nonsingular, then

$$\text{null}(H_a) \cap \text{null}(E_a) = \{0\} \quad \textit{and} \quad \text{null}(H_b) \cap \text{null}(E_b) = \{0\}.$$

Below we demonstrate some characteristics of a saddle point of the weighted augmented Lagrangian function $\mathscr{L}_{\text{wa}}(x, y, z)$ defined in (2.2).

Theorem 3.1 *Let $F \in \mathbb{R}^{n\times n}$, $G \in \mathbb{R}^{m\times m}$ be symmetric positive semidefinite matrices, and $A \in \mathbb{R}^{p\times n}$, $B \in \mathbb{R}^{p\times m}$ be two matrices. Let*

$$\mathbf{A}(\beta) = \begin{pmatrix} F + \beta A^T W^{-1} A & \beta A^T W^{-1} B & -A^T W^{-1} \\ \beta B^T W^{-1} A & G + \beta B^T W^{-1} B & -B^T W^{-1} \\ W^{-1} A & W^{-1} B & 0 \end{pmatrix} \tag{3.1}$$

and

$$\mathbf{b}(\beta) = \begin{pmatrix} \beta A^T W^{-1} b - f \\ \beta B^T W^{-1} b - g \\ W^{-1} b \end{pmatrix}, \quad \mathbf{x} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{3.2}$$

Then the following statements hold true:

(i) *For* $x_* \in \mathbb{R}^n$, $y_* \in \mathbb{R}^m$ *and* $z_* \in \mathbb{R}^p$, (x_*, y_*, z_*) *is a saddle point of the weighted augmented Lagrangian function* $\mathscr{L}_{\rm wa}(x, y, z)$ *defined in* (2.2) *if and only if* $\mathbf{x}_* = (x_*^T, y_*^T, z_*^T)^T \in \mathbb{R}^{n+m+p}$ *is a solution of the linear system* $\mathbf{A}(\beta)\,\mathbf{x} = \mathbf{b}(\beta)$;

(ii) *The matrix* $\mathbf{A}(\beta) \in \mathbb{R}^{(n+m+p)\times(n+m+p)}$ *is nonsingular if and only if*
   (a) $(\mathrm{null}(F) \oplus \mathrm{null}(G)) \cap \mathrm{null}((A\ \ B)) = \{0\}$, *and*
   (b) $\mathrm{null}(A^T) \cap \mathrm{null}(B^T) = \{0\}$.

Proof We first demonstrate (i). By straightforward calculations we obtain the derivatives of the weighted augmented Lagrangian function $\mathscr{L}_{\rm wa}(x, y, z)$ with respect to x, y and z as follows:

$$\begin{cases} \frac{\partial \mathscr{L}_{\rm wa}(x,y,z)}{\partial x} = (F + \beta A^T W^{-1} A)x + \beta A^T W^{-1} B y - A^T W^{-1} z + f - \beta A^T W^{-1} b, \\ \frac{\partial \mathscr{L}_{\rm wa}(x,y,z)}{\partial y} = \beta B^T W^{-1} A x + (G + \beta B^T W^{-1} B) y - B^T W^{-1} z + g - \beta B^T W^{-1} b, \\ \frac{\partial \mathscr{L}_{\rm wa}(x,y,z)}{\partial z} = -W^{-1}(Ax + By - b). \end{cases}$$

Note that (x_*, y_*, z_*) is a saddle point of $\mathscr{L}_{\rm wa}(x, y, z)$ if and only if

$$\frac{\partial \mathscr{L}_{\rm wa}(x_*, y_*, z_*)}{\partial x} = 0, \quad \frac{\partial \mathscr{L}_{\rm wa}(x_*, y_*, z_*)}{\partial y} = 0 \quad \text{and} \quad \frac{\partial \mathscr{L}_{\rm wa}(x_*, y_*, z_*)}{\partial z} = 0,$$

which can be rewritten in matrix–vector form into $\mathbf{A}(\beta)\,\mathbf{x} = \mathbf{b}(\beta)$.

To verify the validity of (ii), we denote by

$$\mathbf{P}(\beta) = \begin{pmatrix} I & 0 & -\beta A^T \\ 0 & I & -\beta B^T \\ 0 & 0 & I \end{pmatrix},$$

which is a three-by-three block unit upper-triangular matrix. Then it holds that

$$\widetilde{\mathbf{A}}(\beta) = \mathbf{P}(\beta)\,\mathbf{A}(\beta) = \begin{pmatrix} F & 0 & -A^T W^{-1} \\ 0 & G & -B^T W^{-1} \\ W^{-1} A & W^{-1} B & 0 \end{pmatrix},$$

which implies that $\mathbf{A}(\beta)$ is nonsingular if and only if $\widetilde{\mathbf{A}}(\beta)$ is nonsingular. From Lemma 3.2 we know that $\widetilde{\mathbf{A}}(\beta)$ is nonsingular if and only if

$$(\text{null}(F) \oplus \text{null}(G)) \cap \text{null}((W^{-1}A \;\; W^{-1}B)) = \{0\}$$

and

$$\text{null}(A^T W^{-1}) \cap \text{null}(B^T W^{-1}) = \{0\}.$$

The result what we were proving then follows straightforwardly from

$$\text{null}((W^{-1}A \;\; W^{-1}B)) = \text{null}((A \;\; B)),$$

and

$$\text{null}(A^T W^{-1}) \cap \text{null}(B^T W^{-1}) = \{0\}$$

if and only if

$$\text{null}(A^T) \cap \text{null}(B^T) = \{0\}.$$

□

In the sequel, for notational convenience we define matrices

$$\widehat{A} = W^{-1/2}A, \quad \widehat{B} = W^{-1/2}B, \quad \widehat{Q} = W^{1/2}QW^{1/2} \tag{3.3}$$

and

$$\widehat{F} = F + \beta A^T W^{-1} A, \quad \widehat{G} = G + \beta B^T W^{-1} B. \tag{3.4}$$

The positive definiteness and the spectral radius of the symmetric matrix

$$\widehat{H} = \widehat{A}\widehat{F}^{-1}\widehat{A}^T + \widehat{B}\widehat{G}^{-1}\widehat{B}^T$$

are indispensable in discussion of the convergence property and the convergence rate of the PAVMM method, which are precisely stated in the following lemma.

Lemma 3.4 *Let $F \in \mathbb{R}^{n\times n}, G \in \mathbb{R}^{m\times m}$ and $W \in \mathbb{R}^{p\times p}$ be symmetric positive definite matrices, and $A \in \mathbb{R}^{p\times n}$ and $B \in \mathbb{R}^{p\times m}$ be two arbitrary matrices. Define matrices $\widehat{A}$, $\widehat{B}$ and $\widehat{F}, \widehat{G}$ as in (3.3) and (3.4), respectively. Then the following two statements hold true*:

(i) $0 \preceq \widehat{A}\widehat{F}^{-1}\widehat{A}^T \prec \frac{1}{\beta}I, 0 \preceq \widehat{B}\widehat{G}^{-1}\widehat{B}^T \prec \frac{1}{\beta}I$ *and* $0 \preceq \widehat{H} \prec \frac{2}{\beta}I$;

(ii) $\widehat{H} \succ 0$ *if A and B further satisfy*

$$\text{null}(A^T) \cap \text{null}(B^T) = \{0\}.$$

Proof By denoting

$$\widehat{\Omega}_\beta = \beta \widehat{A}^T \widehat{A} \quad \text{and} \quad \widehat{\Gamma}_\beta = \beta \widehat{B}^T \widehat{B},$$

we have

$$\widehat{F} = F + \widehat{\Omega}_\beta \quad \text{and} \quad \widehat{G} = G + \widehat{\Gamma}_\beta.$$

So

$$\lambda(\beta \widehat{A}\widehat{F}^{-1}\widehat{A}^T) = \lambda(\widehat{F}^{-1/2}\widehat{\Omega}_\beta \widehat{F}^{-1/2}) < 1$$

and

$$\lambda(\beta \widehat{B}\widehat{G}^{-1}\widehat{B}^T) = \lambda(\widehat{G}^{-1/2}\widehat{\Gamma}_\beta \widehat{G}^{-1/2}) < 1.$$

It follows that

$$0 \preceq \beta \widehat{A}\widehat{F}^{-1}\widehat{A}^T \prec I \quad \text{and} \quad 0 \preceq \beta \widehat{B}\widehat{G}^{-1}\widehat{B}^T \prec I,$$

or equivalently,

$$0 \preceq \widehat{A}\widehat{F}^{-1}\widehat{A}^T \prec \frac{1}{\beta} I \quad \text{and} \quad 0 \preceq \widehat{B}\widehat{G}^{-1}\widehat{B}^T \prec \frac{1}{\beta} I.$$

As a result, we straightforwardly have

$$0 \preceq \widehat{H} \prec \frac{2}{\beta} I.$$

So the statement (i) is valid.

In addition, we can assert $\widehat{H} \succ 0$. Otherwise, if there exists a nonzero vector $\widehat{z} \in \mathbb{R}^p$ such that $\widehat{z}^T \widehat{H}\widehat{z} = 0$, it must hold

$$\widehat{z}^T \widehat{A}\widehat{F}^{-1}\widehat{A}^T \widehat{z} = 0 \quad \text{and} \quad \widehat{z}^T \widehat{B}\widehat{G}^{-1}\widehat{B}^T \widehat{z} = 0$$

due to

$$\widehat{A}\widehat{F}^{-1}\widehat{A}^T \succeq 0 \quad \text{and} \quad \widehat{B}\widehat{G}^{-1}\widehat{B}^T \succeq 0.$$

Because $\widehat{F} \succ 0$ and $\widehat{G} \succ 0$, we have

$$\widehat{A}^T \widehat{z} = 0 \quad \text{and} \quad \widehat{B}^T \widehat{z} = 0,$$

or in other words,

$$\widehat{z} \in \text{null}(\widehat{A}^T) \cap \text{null}(\widehat{B}^T).$$

In addition, as

$$\text{null}(\widehat{A}^T) = \text{null}(A^T W^{-1/2}) \quad \text{and} \quad \text{null}(\widehat{B}^T) = \text{null}(B^T W^{-1/2}),$$

from

$$\text{null}(A^T) \cap \text{null}(B^T) = \{0\}$$

we see that $\widehat{z} = 0$, which is a contradiction. So $\widehat{H} \succ 0$ and the statement (ii) holds true. □

In addition, the following determinant criterion for locations of the two roots of a complex quadratic polynomial equation is fundamental for us to demonstrate the asymptotic convergence of the PAVMM method.

Lemma 3.5 *Let η and ζ be two complex constants. Then both roots of the complex quadratic polynomial equation*

$$\lambda^2 + \zeta\lambda + \eta = 0$$

have modulus less than one if and only if

$$|\zeta - \bar{\zeta}\eta| + |\eta|^2 < 1; \tag{3.5}$$

see [6,32]. *In particular, if both η and ζ are real constants, then the condition* (3.5) *reduces to*

$$|\eta| < 1 \quad \textit{and} \quad |\zeta| < 1 + \eta;$$

see [6,45].

4 The asymptotic convergence

In this section, we demonstrate the asymptotic convergence and estimate the corresponding asymptotic convergence rate for the PAVMM method.

To this end, we introduce block matrices

$$\mathbf{M}(\alpha, \beta) = \begin{pmatrix} F + \beta A^T W^{-1} A & 0 & 0 \\ \beta B^T W^{-1} A & G + \beta B^T W^{-1} B & 0 \\ W^{-1} A & W^{-1} B & \frac{1}{\alpha} Q \end{pmatrix}$$

and

$$\mathbf{N}(\alpha, \beta) = \begin{pmatrix} 0 & -\beta A^T W^{-1} B & A^T W^{-1} \\ 0 & 0 & B^T W^{-1} \\ 0 & 0 & \frac{1}{\alpha} Q \end{pmatrix},$$

which satisfy

$$\mathbf{A}(\beta) = \mathbf{M}(\alpha, \beta) - \mathbf{N}(\alpha, \beta),$$

with $\mathbf{A}(\beta)$ being defined in (3.1). And we also introduce block vectors

$$\mathbf{x}^{(k+1)} = \begin{pmatrix} x^{(k+1)} \\ y^{(k+1)} \\ z^{(k+1)} \end{pmatrix} \quad \text{and} \quad \mathbf{x}^{(k)} = \begin{pmatrix} x^{(k)} \\ y^{(k)} \\ z^{(k)} \end{pmatrix}.$$

Then the iteration scheme (2.4) can be briefly rewritten into the standard stationary iteration scheme

$$\mathbf{M}(\alpha, \beta)\,\mathbf{x}^{(k+1)} = \mathbf{N}(\alpha, \beta)\,\mathbf{x}^{(k)} + \mathbf{b}(\beta), \tag{4.1}$$

with $\mathbf{b}(\beta)$ being defined in (3.2). Using these notations we can alternatively translate the PAVMM method as a stationary matrix splitting iteration method for solving the block three-by-three system of linear equations

$$\mathbf{A}(\beta)\,\mathbf{x} = \mathbf{b}(\beta),$$

which is induced by the splitting

$$\mathbf{A}(\beta) = \mathbf{M}(\alpha, \beta) - \mathbf{N}(\alpha, \beta)$$

of the matrix $\mathbf{A}(\beta) \in \mathbb{R}^{(n+m+p)\times(n+m+p)}$. As a result, the PAVMM method is convergent if and only if the spectral radius $\rho(\mathbf{L}(\alpha, \beta))$ of its iteration matrix $\mathbf{L}(\alpha, \beta) = \mathbf{M}(\alpha, \beta)^{-1}\mathbf{N}(\alpha, \beta)$ is less than one, i.e., $\rho(\mathbf{L}(\alpha, \beta)) < 1$, and, in the convergence situation, its asymptotic convergence factor is given by $\rho(\mathbf{L}(\alpha, \beta))$; see [42].

Using the matrices

$$\mathbf{P}_l(\beta) = \begin{pmatrix} I & 0 & -\beta A^T \\ 0 & I & -\beta B^T \\ 0 & 0 & W^{1/2} \end{pmatrix} \quad \text{and} \quad \mathbf{P}_r(\beta) = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & W^{1/2} \end{pmatrix},$$

we have

$$\begin{aligned} \widehat{\mathbf{M}}(\alpha, \beta) &= \mathbf{P}_l(\beta)\,\mathbf{M}(\alpha, \beta)\,\mathbf{P}_r(\beta) \\ &= \begin{pmatrix} F & -\beta A^T W^{-1} B & -\frac{\beta}{\alpha} A^T Q W^{1/2} \\ 0 & G & -\frac{\beta}{\alpha} B^T Q W^{1/2} \\ W^{-1/2} A & W^{-1/2} B & \frac{1}{\alpha} W^{1/2} Q W^{1/2} \end{pmatrix} \\ &= \begin{pmatrix} F & -\beta \widehat{A}^T \widehat{B} & -\frac{\beta}{\alpha} \widehat{A}^T \widehat{Q} \\ 0 & G & -\frac{\beta}{\alpha} \widehat{B}^T \widehat{Q} \\ \widehat{A} & \widehat{B} & \frac{1}{\alpha} \widehat{Q} \end{pmatrix} \end{aligned}$$

and

$$
\begin{aligned}
\widehat{\mathbf{N}}(\alpha,\beta) &= \mathbf{P}_l(\beta)\,\mathbf{N}(\alpha,\beta)\,\mathbf{P}_r(\beta)\\
&= \begin{pmatrix} 0 & -\beta A^T W^{-1} B & A^T\left(W^{-1}-\frac{\beta}{\alpha}Q\right)W^{1/2}\\ 0 & 0 & B^T\left(W^{-1}-\frac{\beta}{\alpha}Q\right)W^{1/2}\\ 0 & 0 & \frac{1}{\alpha}W^{1/2}QW^{1/2}\end{pmatrix}\\
&= \begin{pmatrix} 0 & -\beta\widehat{A}^T\widehat{B} & \widehat{A}^T\left(I-\frac{\beta}{\alpha}\widehat{Q}\right)\\ 0 & 0 & \widehat{B}^T\left(I-\frac{\beta}{\alpha}\widehat{Q}\right)\\ 0 & 0 & \frac{1}{\alpha}\widehat{Q}\end{pmatrix},
\end{aligned}
$$

where the matrices $\widehat{A}$, $\widehat{B}$ and $\widehat{Q}$ are defined as in (3.3). Hence, it holds that

$$
\begin{aligned}
\widehat{\mathbf{L}}(\alpha,\beta) &= \widehat{\mathbf{M}}(\alpha,\beta)^{-1}\widehat{\mathbf{N}}(\alpha,\beta) = \mathbf{P}_r(\beta)^{-1}\,\mathbf{M}(\alpha,\beta)^{-1}\mathbf{N}(\alpha,\beta)\,\mathbf{P}_r(\beta)\\
&= \mathbf{P}_r(\beta)^{-1}\,\mathbf{L}(\alpha,\beta)\,\mathbf{P}_r(\beta),
\end{aligned}
$$

which implies that $\mathbf{L}(\alpha,\beta)$ and $\widehat{\mathbf{L}}(\alpha,\beta)$ are similar matrices so that they have the same eigenvalue set.

Based on this fact, below we describe analytic formulas for the eigenvalues of the matrix $\widehat{\mathbf{L}}(\alpha,\beta)$ when the matrices F and G are symmetric positive definite.

Theorem 4.1 *Let $F\in\mathbb{R}^{n\times n}$, $G\in\mathbb{R}^{m\times m}$ be symmetric positive definite matrices, and $A\in\mathbb{R}^{p\times n}$, $B\in\mathbb{R}^{p\times m}$ be two matrices such that*

$$
\text{null}(A^T)\cap\text{null}(B^T)=\{0\}.
$$

Denote

$$
\widetilde{R}=\widehat{A}\widehat{F}^{-1}\widehat{A}^T,\quad \widetilde{S}=\widehat{B}\widehat{G}^{-1}\widehat{B}^T \quad \textit{and} \quad \widetilde{Q}=(I-\beta\widetilde{S})^{-1}\widehat{Q}(I-\beta\widetilde{S}),
$$

where $\widehat{A}$, $\widehat{B}$, $\widehat{Q}$ and $\widehat{F}$, $\widehat{G}$ are defined as in (3.3)–(3.4). Then the iteration sequence $\{\mathbf{x}^{(k)}\}_{k=0}^{\infty}$, generated by the PAVMM iteration scheme (2.4) or (4.1), is convergent to the exact solution of the equality-constraint quadratic programming problem (1.1)–(1.2), provided the modulus of any eigenvalue λ of the quadratic eigenvalue problem

$$
\left[\lambda^2\widetilde{Q}+\lambda\left(\alpha(\widetilde{R}+\widetilde{S}-\beta\widetilde{R}\widetilde{S})-(I+\beta^2\widetilde{R}\widetilde{S})\widetilde{Q}\right)-\beta\widetilde{R}\widetilde{S}(\alpha I-\beta\widetilde{Q})\right]\widetilde{w}=0 \quad (4.2)
$$

is less than 1*. Moreover, the convergence factor of the PAVMM iteration method is given by* $\max\{|\lambda|\,|\,\lambda$ *is an eigenvalue of the problem* (4.2)$\}$.

Proof As $\widehat{\mathbf{L}}(\alpha,\beta)$ and $\mathbf{L}(\alpha,\beta)$ are similar matrices, we only need to analyze the eigenvalues of the matrix $\widehat{\mathbf{L}}(\alpha,\beta)$.

Let λ be a nonzero eigenvalue of the matrix $\widehat{\mathbf{L}}(\alpha, \beta)$ and $\mathbf{u} = (u^T, v^T, w^T)^T \in \mathbb{C}^{n+m+p}$, with $u \in \mathbb{C}^n$, $v \in \mathbb{C}^m$ and $w \in \mathbb{C}^p$, be the corresponding eigenvector, i.e., $\widehat{\mathbf{L}}(\alpha, \beta)\, \mathbf{u} = \lambda \mathbf{u}$. Then it holds that

$$\widehat{\mathbf{N}}(\alpha, \beta)\, \mathbf{u} = \lambda\, \widehat{\mathbf{M}}(\alpha, \beta)\, \mathbf{u},$$

or equivalently,

$$\begin{cases} -\beta \widehat{A}^T \widehat{B} v + \widehat{A}^T \left(I - \frac{\beta}{\alpha} \widehat{Q} \right) w = \lambda \left(F u - \beta \widehat{A}^T \widehat{B} v - \frac{\beta}{\alpha} \widehat{A}^T \widehat{Q} w \right), \\ \widehat{B}^T \left(I - \frac{\beta}{\alpha} \widehat{Q} \right) w = \lambda \left(G v - \frac{\beta}{\alpha} \widehat{B}^T \widehat{Q} w \right), \\ \frac{1}{\alpha} \widehat{Q}\, w = \lambda \left(\widehat{A} u + \widehat{B} v + \frac{1}{\alpha} \widehat{Q} w \right). \end{cases}$$

After some manipulations, we can rewrite the above generalized eigenvalue problem as

$$\begin{cases} \lambda F u + (1-\lambda) \beta \widehat{A}^T \widehat{B} v = \widehat{A}^T \left(I + \frac{(\lambda-1)\beta}{\alpha} \widehat{Q} \right) w, \\ \lambda G v = \widehat{B}^T \left(I + \frac{(\lambda-1)\beta}{\alpha} \widehat{Q} \right) w, \\ \lambda (\widehat{A} u + \widehat{B} v) = \frac{1-\lambda}{\alpha} \widehat{Q} w. \end{cases} \tag{4.3}$$

From the second equation of (4.3) we have

$$v = \frac{1}{\lambda} G^{-1} \widehat{B}^T \left(I + \frac{(\lambda-1)\beta}{\alpha} \widehat{Q} \right) w. \tag{4.4}$$

By substituting this expression into the first equation of (4.3) we obtain

$$u = \frac{1}{\lambda} F^{-1} \widehat{A}^T \left(I + \frac{(\lambda-1)\beta}{\lambda} \widehat{B} G^{-1} \widehat{B}^T \right) \left(I + \frac{(\lambda-1)\beta}{\alpha} \widehat{Q} \right) w. \tag{4.5}$$

Substitutions of (4.4) and (4.5) into the third equation of (4.3) result in

$$\left[\widehat{A} F^{-1} \widehat{A}^T \left(I + \frac{(\lambda-1)\beta}{\lambda} \widehat{B} G^{-1} \widehat{B}^T \right) + \widehat{B} G^{-1} \widehat{B}^T \right] \cdot \left(I + \frac{(\lambda-1)\beta}{\alpha} \widehat{Q} \right) w = \frac{1-\lambda}{\alpha} \widehat{Q} w. \tag{4.6}$$

Recalling

$$\widehat{F} = F + \beta A^T W^{-1} A = F + \beta \widehat{A}^T \widehat{A}$$

and

$$\widehat{G} = G + \beta B^T W^{-1} B = G + \beta \widehat{B}^T \widehat{B},$$

we know that both $\widehat{F}$ and $\widehat{G}$ are symmetric positive definite matrices. Also, it holds that

$$F=\widehat{F}-\beta\widehat{A}^T\widehat{A}\quad\text{and}\quad G=\widehat{G}-\beta\widehat{B}^T\widehat{B}.$$

From Lemma 3.4 (i) we know that

$$0\preceq\beta\widehat{A}\widehat{F}^{-1}\widehat{A}^T\prec I\quad\text{and}\quad 0\preceq\beta\widehat{B}\widehat{G}^{-1}\widehat{B}^T\prec I,$$

which straightforwardly imply that

$$I-\beta\widehat{A}\widehat{F}^{-1}\widehat{A}^T\quad\text{and}\quad I-\beta\widehat{B}\widehat{G}^{-1}\widehat{B}^T$$

are symmetric positive definite, or in other words, $I-\beta\widetilde{R}$ and $I-\beta\widetilde{S}$ are symmetric positive definite. By making use of the Sherman-Morrison-Woodbury formula [21] we obtain

$$F^{-1}=\widehat{F}^{-1}+\beta\widehat{F}^{-1}\widehat{A}^T(I-\beta\widehat{A}\widehat{F}^{-1}\widehat{A}^T)^{-1}\widehat{A}\widehat{F}^{-1}$$

and

$$G^{-1}=\widehat{G}^{-1}+\beta\widehat{G}^{-1}\widehat{B}^T(I-\beta\widehat{B}\widehat{G}^{-1}\widehat{B}^T)^{-1}\widehat{B}\widehat{G}^{-1}.$$

Therefore,

$$\widehat{A}F^{-1}\widehat{A}^T=\widetilde{R}+\beta\widetilde{R}(I-\beta\widetilde{R})^{-1}\widetilde{R}=\widetilde{R}(I-\beta\widetilde{R})^{-1} \tag{4.7}$$

and

$$\widehat{B}G^{-1}\widehat{B}^T=\widetilde{S}+\beta\widetilde{S}(I-\beta\widetilde{S})^{-1}\widetilde{S}=\widetilde{S}(I-\beta\widetilde{S})^{-1}. \tag{4.8}$$

By substituting (4.7) and (4.8) into (4.6), after some manipulations we have

$$(I-\beta\widetilde{R})^{-1}\left(\widetilde{R}+\widetilde{S}-\frac{(\lambda+1)\beta}{\lambda}\widetilde{R}\widetilde{S}\right)(I-\beta\widetilde{S})^{-1}\left(I+\frac{(\lambda-1)\beta}{\alpha}\widehat{Q}\right)w=\frac{1-\lambda}{\alpha}\widehat{Q}w.$$

Further rearrangements straightforwardly lead to the equation

$$\alpha[\lambda(\widetilde{R}+\widetilde{S})-(\lambda+1)\beta\widetilde{R}\widetilde{S}](I-\beta\widetilde{S})^{-1}w=(1-\lambda)(\lambda I-\beta^2\widetilde{R}\widetilde{S})(I-\beta\widetilde{S})^{-1}\widehat{Q}w.$$

By noticing

$$\widetilde{Q}=(I-\beta\widetilde{S})^{-1}\widehat{Q}(I-\beta\widetilde{S})$$

and letting

$$\widetilde{w}=(I-\beta\widetilde{S})^{-1}w,$$

after combining terms of the same kind with respect to the power exponents of λ we finally obtain

$$[\lambda^2\widetilde{Q} + \lambda(\alpha(\widetilde{R} + \widetilde{S} - \beta\widetilde{R}\widetilde{S}) - (I + \beta^2\widetilde{R}\widetilde{S})\widetilde{Q}) - \beta\widetilde{R}\widetilde{S}(\alpha I - \beta\widetilde{Q})]\widetilde{w} = 0,$$

which is exactly the quadratic eigenvalue problem (4.2). Here we should assert that $\widetilde{w} \neq 0$, as, otherwise, if $\widetilde{w} = 0$ then $w = 0$ and, from (4.4) and (4.5), we see that $v = 0$ and $u = 0$. This contradicts with the assumption that $\mathbf{u} = (u^T, v^T, w^T)^T$ is an eigenvector. □

We remark that the derivation of the quadratic eigenvalue problem (4.2) does not require the condition $\mathrm{null}(A^T) \cap \mathrm{null}(B^T) = \{0\}$. This condition is only imposed to guarantee the unique solvability of the equality-constraint quadratic programming problem (1.1)–(1.2).

Important and useful choices of the preconditioning matrix $Q \in \mathbb{R}^{p\times p}$ are symmetric positive definite matrix and symmetric negative definite matrix. Correspondingly, in the light of Theorem 4.1 and Lemma 3.5 we can obtain a deepgoing convergence result for the PAVMM iteration method as follows.

Theorem 4.2 *Let $F \in \mathbb{R}^{n\times n}$, $G \in \mathbb{R}^{m\times m}$ be symmetric positive definite matrices, and $A \in \mathbb{R}^{p\times n}$, $B \in \mathbb{R}^{p\times m}$ be two matrices such that*

$$\mathrm{null}(A^T) \cap \mathrm{null}(B^T) = \{0\}.$$

Assume that $Q \in \mathbb{R}^{p\times p}$ is either a symmetric positive definite matrix or a symmetric negative definite matrix. Denote

$$\widetilde{R} = \widehat{A}\widehat{F}^{-1}\widehat{A}^T, \quad \widetilde{S} = \widehat{B}\widehat{G}^{-1}\widehat{B}^T \quad \text{and} \quad \widehat{R} = (I - \beta\widetilde{S})\widetilde{R}(I - \beta\widetilde{S})^{-1},$$

where $\widehat{A}$, $\widehat{B}$ and $\widehat{F}$, $\widehat{G}$ are defined as in (3.3)–(3.4). And for any nonzero vector $w \in \mathbb{C}^p$, define

$$\nu = \frac{w^*\widehat{Q}w}{w^*w}, \quad \varkappa = \frac{w^*(\widehat{R} + \widetilde{S})w}{w^*w} \tag{4.9}$$

and

$$\chi_- = \frac{w^*\widehat{R}\widetilde{S}(\alpha I - \beta\widehat{Q})w}{w^*w}, \quad \chi = \frac{w^*\widehat{R}\widetilde{S}w}{w^*w}, \tag{4.10}$$

where $\widehat{Q}$ is defined as in (3.3). Then the iteration sequence generated by the PAVMM iteration scheme (2.4) or (4.1) is convergent to the exact solution of the equality-constraint quadratic programming problem (1.1)–(1.2), provided the parameters α and β satisfy the condition

$$|\nu\,\delta(\alpha,\beta) + \beta\chi_-\,\bar{\delta}(\alpha,\beta)| + \beta^2|\chi_-|^2 < \nu^2, \tag{4.11}$$

where

$$\delta(\alpha, \beta) = \alpha \varkappa - \nu - \beta(2\alpha\chi - \chi_-).$$

Moreover, the convergence factor of the PAVMM iteration method is given by

$$\sigma(\alpha, \beta) = \max\{\lambda_+^{(\max)}, \lambda_-^{(\max)}\}, \tag{4.12}$$

where

$$\lambda_\pm^{(\max)} = \max_{w \in \mathbb{C}^p \setminus \{0\}} \left\{ \frac{|-\delta(\alpha, \beta) \pm \sqrt{[\delta(\alpha, \beta)]^2 + 4\beta\nu\chi_-}|}{2\nu} \right\}.$$

Proof With the variable replacement

$$w = (I - \beta\widetilde{S})\widetilde{w},$$

after straightforward operations we can equivalently rewrite the quadratic eigenvalue problem (4.2) as

$$[\lambda^2 \widehat{Q} + \lambda(\alpha(\widehat{R} + \widetilde{S}) - \widehat{Q} - \beta\widehat{R}\widetilde{S}(\alpha I + \beta\widehat{Q})) - \beta\widehat{R}\widetilde{S}(\alpha I - \beta\widehat{Q})]w = 0. \tag{4.13}$$

Hence, with the notations in (4.9) and (4.10) we know that λ is an eigenvalue of the quadratic eigenvalue problem (4.13) if it is a root of the quadratic polynomial equation

$$\nu\lambda^2 + (\alpha\varkappa - \nu - \beta\chi_+)\lambda - \beta\chi_- = 0, \tag{4.14}$$

where

$$\chi_+ = \frac{w^* \widehat{R}\widetilde{S}(\alpha I + \beta\widehat{Q})w}{w^* w}.$$

Note that

$$\chi_- + \chi_+ = 2\alpha\chi.$$

As the matrix Q is either symmetric positive definite or symmetric negative definite, so is $\widehat{Q}$. This shows that either $\nu > 0$ or, correspondingly, $\nu < 0$. In accordance with Lemma 3.5, a necessary and sufficient condition for guaranteeing that both roots of the quadratic polynomial equation (4.14) have modulus less than one is that

$$|\nu(\alpha\varkappa - \nu - \beta\chi_+) + \beta(\alpha\bar{\varkappa} - \nu - \beta\bar{\chi}_+)\chi_-| + \beta^2|\chi_-|^2 < \nu^2.$$

After simplifying this inequality, with appropriate arrangements we then obtain the convergence condition (4.11) imposed on the parameters α and β.

In addition, because the two roots of the quadratic polynomial equation (4.14) are

$$\lambda_{\pm} = \frac{\nu - \alpha\varkappa + \beta(2\alpha\chi - \chi_{-}) \pm \sqrt{[\nu - \alpha\varkappa + \beta(2\alpha\chi - \chi_{-})]^2 + 4\beta\nu\chi_{-}}}{2\nu},$$

the convergence factor of the PAVMM iteration method is given by $\sigma(\alpha, \beta)$ in (4.12). □

Furthermore, in actual applications we can choose the preconditioning matrix $Q \in \mathbb{R}^{p\times p}$ depending on the weighting matrix $W \in \mathbb{R}^{p\times p}$. For example, for a prescribed constant $\theta \neq 0$ we can take $Q = \theta W^{-1}$. It then follows that $\widehat{Q} = \theta I$; see (3.3). Correspondingly, we can obtain a convergence result for the PAVMM iteration method as follows.

Theorem 4.3 *Let $F \in \mathbb{R}^{n\times n}$, $G \in \mathbb{R}^{m\times m}$ be symmetric positive definite matrices, and $A \in \mathbb{R}^{p\times n}$, $B \in \mathbb{R}^{p\times m}$ be two matrices such that*

$$\mathrm{null}(A^T) \cap \mathrm{null}(B^T) = \{0\}.$$

Assume that $Q = \theta W^{-1}$, with $\theta \neq 0$ being a prescribed constant. Denote

$$\widetilde{R} = \widehat{A}\widehat{F}^{-1}\widehat{A}^T \quad \textit{and} \quad \widetilde{S} = \widehat{B}\widehat{G}^{-1}\widehat{B}^T,$$

where $\widehat{A}$, $\widehat{B}$ and $\widehat{F}$, $\widehat{G}$ are defined as in (3.3)–(3.4). And for any nonzero vector $\widetilde{w} \in \mathbb{C}^p$, define

$$\varkappa = \frac{\widetilde{w}^*(\widetilde{R} + \widetilde{S})\widetilde{w}}{\widetilde{w}^*\widetilde{w}} \quad \textit{and} \quad \chi = \frac{\widetilde{w}^*\widetilde{R}\widetilde{S}\widetilde{w}}{\widetilde{w}^*\widetilde{w}}. \tag{4.15}$$

Then the iteration sequence generated by the PAVMM iteration scheme (2.4) or (4.1) is convergent to the exact solution of the equality-constraint quadratic programming problem (1.1)–(1.2), provided the parameters α and β satisfy the condition

$$|\theta\,\delta(\alpha, \beta) + \beta(\alpha - \beta\theta)\chi\,\bar{\delta}(\alpha, \beta)| + \beta^2(\alpha - \beta\theta)^2|\chi|^2 < \theta^2, \tag{4.16}$$

where

$$\delta(\alpha, \beta) = \alpha\varkappa - \theta - \beta(\alpha + \beta\theta)\chi.$$

Moreover, the convergence factor of the PAVMM iteration method is given by

$$\sigma(\alpha, \beta) = \max\{\lambda_{+}^{(\max)}, \lambda_{-}^{(\max)}\}, \tag{4.17}$$

where

$$\lambda_{\pm}^{(\max)} = \max_{\widetilde{w}\in\mathbb{C}^p\setminus\{0\}} \left\{ \frac{|-\delta(\alpha, \beta) \pm \sqrt{[\delta(\alpha, \beta)]^2 + 4\beta(\alpha - \beta\theta)\theta\chi}|}{2\theta} \right\}.$$

Proof Recalling the definition of $\widehat{Q}$ given in (3.3), when $Q = \theta W^{-1}$ we have $\widehat{Q} = \theta I$. Hence, it holds that

$$\widetilde{Q} = (I - \beta\widetilde{S})^{-1}\widehat{Q}(I - \beta\widetilde{S}) = \theta I.$$

Accordingly, in Theorem 4.1 the quadratic eigenvalue problem (4.2) becomes

$$[\theta\lambda^2 I + \lambda(\alpha(\widetilde{R} + \widetilde{S} - \beta\widetilde{R}\widetilde{S}) - \theta(I + \beta^2\widetilde{R}\widetilde{S})) - \beta(\alpha - \beta\theta)\widetilde{R}\widetilde{S}]\widetilde{w} = 0.$$

Hence, with the notations in (4.15) we know that λ is an eigenvalue of the quadratic eigenvalue problem (4.2) if it is a root of the quadratic polynomial equation

$$\theta\lambda^2 + [\alpha\varkappa - \theta - \beta(\alpha + \beta\theta)\chi]\lambda - \beta(\alpha - \beta\theta)\chi = 0. \tag{4.18}$$

Note that Lemma 3.4 (i) shows that $0 < \varkappa < \frac{2}{\beta}$. Again, in accordance with Lemma 3.5, based on the quadratic polynomial equation (4.18) we can obtain the convergence condition (4.16) imposed on the parameters α and β, as well as the convergence factor $\sigma(\alpha, \beta)$ in (4.17), for the PAVMM iteration method in an analogous fashion to the demonstration of Theorem 4.2. □

In particular, if $\alpha = \beta$ and $Q = W^{-1}$, the result in Theorem 4.3 reduces to the following one, which straightforwardly gives a convergence result for the AVMM iteration method defined by the specific choice $Q = W = I$.

Corollary 4.1 *Let $F \in \mathbb{R}^{n\times n}$, $G \in \mathbb{R}^{m\times m}$ be symmetric positive definite matrices, and $A \in \mathbb{R}^{p\times n}$, $B \in \mathbb{R}^{p\times m}$ be two matrices such that*

$$\text{null}(A^T) \cap \text{null}(B^T) = \{0\}.$$

Assume that $Q = W^{-1}$. Denote

$$\widetilde{R} = \widehat{A}\widehat{F}^{-1}\widehat{A}^T \quad \textit{and} \quad \widetilde{S} = \widehat{B}\widehat{G}^{-1}\widehat{B}^T,$$

where $\widehat{A}$, $\widehat{B}$ and $\widehat{F}$, $\widehat{G}$ are defined as in (3.3)–(3.4). And for any nonzero vector $\widetilde{w} \in \mathbb{C}^p$, define

$$\varkappa = \frac{\widetilde{w}^*(\widetilde{R} + \widetilde{S})\widetilde{w}}{\widetilde{w}^*\widetilde{w}} \quad \textit{and} \quad \chi = \frac{\widetilde{w}^*\widetilde{R}\widetilde{S}\widetilde{w}}{\widetilde{w}^*\widetilde{w}}. \tag{4.19}$$

Then the iteration sequence generated by the AVMM iteration scheme (1.5) is convergent to the exact solution of the equality-constraint quadratic programming problem (1.1)–(1.2), provided the parameter β satisfies the condition

$$|2\beta^2\chi - \beta\varkappa + 1| < 1. \tag{4.20}$$

Moreover, the convergence factor of the AVMM iteration method is given by

$$\sigma(\beta) = \max_{\widetilde{w} \in \mathbb{C}^p \setminus \{0\}} \{|2\beta^2 \chi - \beta \varkappa + 1|\}. \tag{4.21}$$

Proof Using the notation in Theorem 4.3, when $Q = W^{-1}$ we have $\theta = 1$. Moreover, as $\alpha = \beta$, from the proof of Theorem 4.3 we see that the quadratic polynomial equation (4.18) reduces to

$$\lambda^2 + (\beta\varkappa - 2\beta^2\chi - 1)\lambda = 0, \tag{4.22}$$

where $\varkappa$ and χ are defined as in (4.19). Again, note that Lemma 3.4 (i) shows that $0 < \varkappa < \frac{2}{\beta}$.

The two roots of this quadratic polynomial equation are

$$\lambda_- = 0 \quad \text{and} \quad \lambda_+ = 2\beta^2\chi - \beta\varkappa + 1.$$

Hence, we can obtain the convergence condition (4.20) imposed on the regularization parameter β, as well as the convergence factor $\sigma(\beta)$ in (4.21), for the AVMM iteration method. □

In general, the matrices $\widehat{R}$, $\widetilde{S}$ in Theorem 4.2, and $\widetilde{R}$, $\widetilde{S}$ in Theorem 4.3 and Corollary 4.1 may be symmetric positive semidefinite. It turns out that the constant terms and the coefficients of the first-order terms of the quadratic polynomial equations in (4.14), (4.18) and (4.22) are possibly complex. Hence, we are not capable to derive simpler sufficient conditions than those given in Theorems 4.2, 4.3 and Corollary 4.1 for guaranteeing the asymptotic convergence of the PAVMM method. However, this could become possible if more specific restrictions are imposed on the equality-constraint quadratic programming problem (1.1)–(1.2), as well as on the weighting matrix W and the preconditioning matrix Q. In fact, deriving easily checkable convergence conditions for the PAVMM method is a theoretically meaningful and practically useful topic that deserves further study.

5 An algebraic derivation of PAVMM

Given a symmetric positive definite matrix $W \in \mathbb{R}^{p \times p}$, the equality-constraint programming problem (1.1) is mathematically equivalent to the unconstraint optimization problem

$$\max_{z} \min_{x,y} \mathscr{L}_w(x, y, z), \tag{5.1}$$

where $\mathscr{L}_w(x, y, z)$ is the weighted Lagrangian function defined by

$$\mathscr{L}_w(x, y, z) = \phi(x) + \psi(y) - \langle Ax + By - b, z \rangle_{W^{-1}}, \tag{5.2}$$

with $z \in \mathbb{R}^p$ being the Lagrange multiplier and $W \in \mathbb{R}^{p\times p}$ being the weighting matrix. This implies that a point $(x_*, y_*) \in \mathbb{R}^n \times \mathbb{R}^m$ is a solution of the problem (1.1) if and only if there exists a $z_* \in \mathbb{R}^p$ such that the point $(x_*, y_*, z_*) \in \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p$ is a solution of the problem (5.1)–(5.2). It follows that the first-order necessary condition corresponding to (5.1)–(5.2), with the functions $\phi(x)$ and $\psi(y)$ being the quadratic ones given in (1.2), reads as

$$\begin{cases} Fx - A^T W^{-1} z = -f, \\ Gy - B^T W^{-1} z = -g, \\ -W^{-1} Ax - W^{-1} By = -W^{-1} b. \end{cases} \tag{5.3}$$

This condition is also sufficient as the functions $\phi(x)$ and $\psi(y)$ defined in (1.2) are convex.

After pre-multiplying -1 on both sides of the third equation, we can equivalently write the linear system (5.3) as

$$\begin{cases} Fx - A^T W^{-1} z = -f, \\ Gy - B^T W^{-1} z = -g, \\ W^{-1} Ax + W^{-1} By = W^{-1} b, \end{cases} \tag{5.4}$$

or in matrix–vector form as

$$\begin{pmatrix} F & 0 & -A^T W^{-1} \\ 0 & G & -B^T W^{-1} \\ W^{-1} A & W^{-1} B & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -f \\ -g \\ W^{-1} b \end{pmatrix}. \tag{5.5}$$

With an additional regularization parameter $\beta > 0$, first multiplying both sides of the third equation in (5.4) from left by, respectively, βA^T and βB^T and, then, adding the resulting two equations successively to the first and the second equations in (5.4), we can obtain the augmented Lagrangian linear system corresponding to (5.4) as follows:

$$\begin{cases} (F + \beta A^T W^{-1} A)x + \beta A^T W^{-1} By - A^T W^{-1} z = \beta A^T W^{-1} b - f, \\ \beta B^T W^{-1} Ax + (G + \beta B^T W^{-1} B)y - B^T W^{-1} z = \beta B^T W^{-1} b - g, \\ W^{-1} Ax + W^{-1} By = W^{-1} b. \end{cases}$$

Again, in matrix–vector form it can be written as the block three-by-three linear system

$$\mathbf{A}(\beta)\, \mathbf{x} = \mathbf{b}(\beta),$$

where the coefficient matrix $\mathbf{A}(\beta)$, the right-hand side vector $\mathbf{b}(\beta)$ and the unknown vector $\mathbf{x}$ are defined as in (3.1) and (3.2), respectively.

Now the PAVMM method (see (2.4)) for solving the equality-constraint quadratic programming problem (1.1)–(1.2) is exactly the modified block Gauss–Seidel iteration method used to solve the augmented Lagrangian linear system $\mathbf{A}(\beta)\, \mathbf{x} = \mathbf{b}(\beta)$ defined in (3.1)–(3.2) corresponding to the matrix splitting

$$\mathbf{A}(\beta) = \mathbf{D}(\alpha, \beta) - \mathbf{L}(\alpha, \beta) - \mathbf{U}(\alpha, \beta),$$

with

$$\mathbf{D}(\alpha,\beta) = \begin{pmatrix} F + \beta A^T W^{-1} A & 0 & 0 \\ 0 & G + \beta B^T W^{-1} B & 0 \\ 0 & 0 & \frac{1}{\alpha} Q \end{pmatrix}$$

being a block diagonal matrix,

$$\mathbf{L}(\alpha,\beta) = \begin{pmatrix} 0 & 0 & 0 \\ -\beta B^T W^{-1} A & 0 & 0 \\ -W^{-1} A & -W^{-1} B & 0 \end{pmatrix}$$

being a strictly block lower-triangular matrix, and

$$\mathbf{U}(\alpha,\beta) = \begin{pmatrix} 0 & -\beta A^T W^{-1} B & A^T W^{-1} \\ 0 & 0 & B^T W^{-1} \\ 0 & 0 & \frac{1}{\alpha} Q \end{pmatrix}$$

being a block upper-triangular matrix. Here α is a nonzero constant and $Q \in \mathbb{R}^{p \times p}$ is a nonsingular matrix.

Admittedly, the *Hermitian and skew-Hermitian splitting* iteration method [4] should be an effective solver for the block three-by-three linear system (5.5), which straightforwardly results in another class of iteration methods for solving the equality-constraint quadratic programming problem (1.1)–(1.2), too; see, e.g., [12,23,44] for more details.

6 Concluding remarks

The so-called alternating direction method of multipliers is a practically effective solver for separable programming problems with proper constraints. When the cost function is quadratic and the constraints are linear, by generalizing and modifying this method we have constructed a class of PAVMM methods. Using matrix analysis we have established its asymptotic convergence theorem and derived its asymptotic convergence rate. This particularly results in rigorous convergence theory for the alternating direction method of multipliers.

For the convex quadratic programming problems with equality constraints, the basic principle for constructing the PAVMM iteration method is deeply revealed, which specifically shows that the two typical approaches: augment-then-optimize and optimize-then-augment, lead to exactly the same iteration method. As a result, the PAVMM method can be algorithmically formulated and theoretically analyzed from viewpoints of both numerical optimization and matrix computation.

Moreover, the algebraic methodology advocated in this paper is applicable to developing iteration methods of this kind and analyzing their convergence properties for more general separable programming problems with more complicated constraints; see, e.g., [26,27,43].

Acknowledgments The authors are very much indebted to Profs. Hua Dai, An-Ping Liao and Rui-Ping Wen for useful discussions. They also thank Prof. Zeng-Qi Wang for constructive suggestions and valuable discussions. The referees provided very useful comments and suggestions, which greatly improved the original manuscript of this paper.

References

1. Arrow, K., Hurwicz, L., Uzawa, H.: Studies in Nonlinear Programming. Stanford University Press, Stanford (1958)
2. Bai, Z.-J., Bai, Z.-Z.: On nonsingularity of block two-by-two matrices. Linear Algebra Appl. **439**, 2388–2404 (2013)
3. Bai, Z.-Z.: Eigenvalue estimates for saddle point matrices of Hermitian and indefinite leading blocks. J. Comput. Appl. Math. **237**, 295–306 (2013)
4. Bai, Z.-Z., Golub, G.H., Ng, M.K.: Hermitian and skew-Hermitian splitting methods for non-Hermitian positive definite linear systems. SIAM J. Matrix Anal. Appl. **24**, 603–626 (2003)
5. Bai, Z.-Z., Ng, M.K., Wang, Z.-Q.: Constraint preconditioners for symmetric indefinite matrices. SIAM J. Matrix Anal. Appl. **31**, 410–433 (2009)
6. Bai, Z.-Z., Parlett, B.N., Wang, Z.-Q.: On generalized successive overrelaxation methods for augmented linear systems. Numer. Math. **102**, 1–38 (2005)
7. Bergen, A.R.: Power Systems Analysis. Prentice Hall, Englewood Cliffs (1986)
8. Betts, J.T.: Practical Methods for Optimal Control Using Nonlinear Programming. SIAM, Philadelphia (2001)
9. Bossavit, A.: “Mixed” systems of algebraic equations in computational electromagnetics. COMPEL **17**, 59–63 (1998)
10. Brezzi, F., Fortin, M.: Mixed and Hybrid Finite Element Methods. Springer, New York (1991)
11. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends Mach. Learn. **3**, 1–122 (2010)
12. Chan, L.-C., Ng, M.K., Tsing, N.-K.: Spectral analysis for HSS preconditioners. Numer. Math.-Theory Methods Appl. **1**, 57–77 (2008)
13. Chua, L.O., Desoer, C.A., Kuh, E.S.: Linear and Nonlinear Circuits. McGraw-Hill, New York (1987)
14. Elman, H.C., Golub, G.H.: Inexact and preconditioned Uzawa algorithms for saddle point problems. SIAM J. Numer. Anal. **31**, 1645–1661 (1994)
15. Eckstein, J., Bertsekas, D.P.: On the Douglas-Rachford splitting method and the proximal point algorithm for maximal monotone operators. Math. Program. **55**, 293–318 (1992)
16. Eckstein, J., Ferris, M.C.: Operator-splitting methods for monotone affine variational inequalities, with a parallel application to optimal control. INFORMS J. Comput. **10**, 218–235 (1998)
17. Fortin, M., Glowinski, R.: Augmented Lagrangian Methods, Applications to the Numerical Solution of Boundary Value Problems. North-Holland, Amsterdam (1983)
18. Gabay, D.: Applications of the method of multipliers to variational inequalities. In: Fortin, M., Glowinski, R. (eds.) Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems, pp. 299–331. North-Holland, Amsterdam (1983)
19. Gabay, D., Mercier, B.: A dual algorithm for the solution of nonlinear variational problems via finite element approximations. Comput. Math. Appl. **2**, 17–40 (1976)
20. Glowinski, R., Le Tallec, P.: Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics. SIAM, Philadelphia (1989)
21. Golub, G.H., Van Loan, C.F.: Matrix Computations, 3rd edn. The Johns Hopkins University Press, Baltimore (1996)
22. Haber, E., Modersitzki, J.: Numerical methods for volume-preserving image registration. Inverse Probl. **20**, 1621–1638 (2004)
23. Hadjidimos, A., Lapidakis, M.: Optimal alternating direction implicit preconditioners for conjugate gradient methods. Appl. Math. Comput. **183**, 559–574 (2006)
24. Hall, E.L.: Computer Image Processing and Recognition. Academic Press, New York (1979)
25. Han, D.-R., Yuan, X.-M.: Local linear convergence of the alternating direction method of multipliers for quadratic programs. SIAM J. Numer. Anal. **51**, 3446–3457 (2013)
26. Han, D.-R., Yuan, X.-M.: A note on the alternating direction method of multipliers. J. Optim. Theory Appl. **155**, 227–238 (2012)

27. Han, D.-R., Yuan, X.-M., Zhang, W.-X.: An augmented Lagrangian based parallel splitting method for separable convex minimization with applications to image processing. Math. Comput. **83**, 2263–2291 (2014)
28. He, B.-S., Yuan, X.-M.: On the $\mathcal{O}(1/n)$ convergence rate of the Douglas-Rachford alternating direction method. SIAM J. Numer. Anal. **50**, 700–709 (2012)
29. Lions, P.L., Mercier, B.: Splitting algorithms for the sum of two nonlinear operators. SIAM J. Numer. Anal. **16**, 964–979 (1979)
30. Markowitz, H.M.: Portfolio Selection: Efficient Diversification of Investments. Wiley, New York (1959)
31. Markowitz, H.M., Perold, A.F.: Sparsity and piecewise linearity in large portfolio optimization problems. In: Duff, I.S. (ed.) Sparse Matrices and Their Uses, pp. 89–108. Academic Press, London (1981)
32. Miller, J.J.H.: On the location of zeros of certain classes of polynomials with applications to numerical analysis. J. Inst. Math. Appl. **8**, 397–406 (1971)
33. Modersitzki, J.: Numerical Methods for Image Registration. Oxford University Press, Oxford (2003)
34. Ortega, J.M., Rheinboldt, W.C.: Iterative Solution of Nonlinear Equations in Several Variables. SIAM, Philadelphia (2000)
35. Perugia, I.: A field-based mixed formulation for the two-dimensional magnetostatic problem. SIAM J. Numer. Anal. **34**, 2382–2391 (1997)
36. Rockafellar, R.T.: Augmented Lagrangians and applications of the proximal point algorithm in convex programming. Math. Oper. Res. **1**, 97–116 (1976)
37. Rockafellar, R.T.: Monotone operators and the proximal point algorithm. SIAM J. Control Optim. **14**, 877–898 (1976)
38. Rockafellar, R.T., Wets, R.J.-B.: Variational Analysis. Springer, Berlin (1998)
39. Shefi, R., Teboulle, M.: Rate of convergence analysis of decomposition methods based on the proximal method of multipliers for convex minimization. SIAM J. Optim. **24**, 269–297 (2014)
40. Strang, G.: Introduction to Applied Mathematics. Wellesley-Cambridge Press, Wellesley (1986)
41. Tao, M., Yuan, X.-M.: On the $\mathcal{O}(1/t)$ convergence rate of alternating direction method with logarithmic-quadratic proximal regularization. SIAM J. Optim. **22**, 1431–1448 (2012)
42. Varga, R.S.: Matrix Iterative Analysis. Prentice Hall, Englewood Cliffs (1962)
43. Wang, K., Han, D.-R., Xu, L.-L.: A parallel splitting method for separable convex programs. J. Optim. Theory Appl. **159**, 138–158 (2013)
44. Yang, A.-L., An, J., Wu, Y.-J.: A generalized preconditioned HSS method for non-Hermitian positive definite linear systems. Appl. Math. Comput. **216**, 1715–1722 (2010)
45. Young, D.M.: Iterative Solution of Large Linear Systems. Academic Press, New York (1971)