Convergence of a Class of Stationary Iterative Methods for Saddle Point Problems

Yin Zhang[*](张寅)

August, 2010

Abstract

A unified convergence result is derived for an entire class of stationary iterative methods for solving equality constrained quadratic programs or saddle point problems. This class is constructed from essentially all possible splittings of the $n \times n$ submatrix residing in the (1,1)-block of the $(n+m) \times (n+m)$ augmented matrix that would generate non-expansive iterations in $\mathbb{R}^n$. The classic multiplier method and augmented Lagrangian alternating direction method are two special members of this class.

Keywords. Saddle point problem, quadratic program, splitting, stationary iterations, alternating direction augmented Lagrangian method, Q-linear convergence.

1 Introduction

First consider the equality constrained quadratic program:

$$\min q(x) := \frac{1}{2} x^T A x - b^T x \ \text{ s.t. } \ B x = c. \tag{1}$$

where $A \in \mathbb{R}^{n \times n}$ is symmetric and $B \in \mathbb{R}^{m \times n}$ with $m < n$. The matrix A can be indefinite, but is assumed to be positive definite in the null space of B. Without loss of generality, we assume that B is of full rank m. The system of stationarity for the quadratic program (1) is

$$
\begin{aligned}
Ax + B^T y - b &= 0, \\
Bx - c &= 0,
\end{aligned}
$$

---
[*]Department of Computational and Applied Mathematics, Rice University, Houston, Texas 77005, U.S.A. (yzhang@rice.edu)

where x is the primal variable and y is the Lagrangian multiplier (or dual variable). In matrix form, the system is

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b \\ c \end{pmatrix}, \tag{2}$$

which is commonly called the augmented system or saddle point system – a problem with a wide range of applications in various areas of computational science and engineering. Numerical solutions of this problem have been extensively studied in the literature; see the survey paper [1] for a comprehensive review and a thorough list of references up to 2004.

The augmented lagrangian technique has been used to make the (1,1)-block of the saddle point system positive definite. In this approach, an equivalent system is solved,

$$\begin{aligned} Ax + B^T y - b + \gamma B^T (Bx - c) &= 0, \\ Bx - c &= 0, \end{aligned}$$

with a parameter $\gamma > 0$, which has the matrix form

$$\begin{pmatrix} A + \gamma B^T B & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} b + \gamma B^T c \\ c \end{pmatrix}. \tag{3}$$

The following result is a well-known fact.

Proposition 1. *Let A be symmetric positive definite in the null space of B. If $A \succeq 0$, then $A + \gamma B^T B \succ 0$ for $\gamma \in (0, +\infty)$; otherwise, there exists some $\hat{\gamma} > 0$ such that*

$$\gamma \in (\hat{\gamma}, +\infty) \implies A + \gamma B^T B \succ 0. \tag{4}$$

1.1 Notation

For matrix $M \in \mathbb{R}^{n \times n}$, $\sigma(M)$ denotes the spectrum of M and $\rho(M)$ the spectral radius of M. For symmetric M, $\lambda_{\max}(M)$ $(\lambda_{\min}(M))$ is the maximum (minimum) eigenvalue of M. By $M \succ 0$ $(M \succeq 0)$, we mean that M is symmetric positive definite (semi-definite). For a complex number $z \in \mathbb{C}$, $\Re(z)$ denotes the real part of z and $\Im(z)$ the imaginary part.

2 A Class of Stationary Iterative Methods

In this section, we describe a class of stationary iterative methods for solving the saddle point problem (3) where the (1,1)-block has been made positive definite. However, we will re-parameterize

the first equation and introducing another parameter into the the second. The equivalent system under consideration is

$$\begin{pmatrix} H(\alpha) & -B^T \\ \tau B & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha b + B^T c \\ \tau c \end{pmatrix}, \tag{5}$$

where $\alpha, \tau > 0$ and

$$H(\alpha) = \alpha A + B^T B \succ 0.$$

Comparing (5) to (3), we see that $\alpha = 1/\gamma$ and the multiplier y has been rescaled along with a sign change. These changes are cosmetic except that one more parameter τ is introduced into (5).

Since the equation $Bx = c$ is equivalent to $QBx = Qc$ for any nonsingular $Q \in \mathbb{R}^{m \times m}$, B and c in (5) can obviously be replaced by QB and Qc, respectively.

2.1 Splitting of the (1,1)-block

In our framework, the (1,1)-block submatrix $H(\alpha)$ in (5) is split into a "left part" L and a "right part" R; that is,

$$H := \alpha A + B^T B = L - R. \tag{6}$$

We drop the α-dependence from H, as well as from L and R, since α will always be fixed in our analysis as long as $H \succ 0$ is maintained, even though it can also be varied to improve convergence performance.

In this report, unless otherwise noted, splittings refer to those for the (1,1)-block submatrix H rather than for the entire (2×2)-block augmented matrix of the saddle point problem. Moreover, we will associate a splitting with a left-right pair (L, R). Simplest examples of splittings include

$$L = H, \quad R = 0;$$

or after partitioning H into 2 by 2 blocks,

$$L = \begin{pmatrix} H_{11} & 0 \\ 0 & H_{22} \end{pmatrix}, \quad R = - \begin{pmatrix} 0 & H_{12} \\ H_{21} & 0 \end{pmatrix},$$

which is of block Jacobi type; or

$$L = \begin{pmatrix} H_{11} & 0 \\ H_{21} & H_{22} \end{pmatrix}, \quad R = - \begin{pmatrix} 0 & H_{12} \\ 0 & 0 \end{pmatrix}, \tag{7}$$

which is of block Gauss-Seidel type. We note that when $H \succ 0$ and (L, R) is a Gauss-Seidel splitting, either element-wise or block-wise, it is known that $\rho(L^{-1}R) < 1$.

3

In general, one can first partition H into p-by-p blocks for any $p \in \{1, 2, \cdots, n\}$, then perform a block splitting. In addition, splittings can be of SOR type involving an extra relaxation parameter. To keep notation simple, however, we will not carry such a parameters in a splitting (L, R) since it does not affect our analysis.

2.2 A stationary iteration class

We consider a class of stationary iterations consisting of all possible splittings (L, R) for which the spectral radius of $L^{-1}R$ does not exceed the unity (plus an additional technical condition to be specified soon). This class of stationary iterative methods, that we call the $\{\texttt{L,R}\}$-class for lack of a more descriptive term, iterates as follows:

$$x^{k+1} = L^{-1}\left(Rx^k + B^T(y^k + c) + \alpha b\right), \tag{8a}$$

$$y^{k+1} = y^k - \tau\left(Bx^{k+1} - c\right), \tag{8b}$$

where (L, R) is any admissable splitting and τ represents a step length in multiplier updates.

It is easy to see that the $\{\texttt{L,R}\}$-class iterations (8) correspond to the following splitting of the (2×2)-block augmented matrix in system (5):

$$\begin{pmatrix} H & -B^T \\ \tau B & 0 \end{pmatrix} = \begin{pmatrix} L & 0 \\ \tau B & I \end{pmatrix} - \begin{pmatrix} R & B^T \\ 0 & I \end{pmatrix}. \tag{9}$$

Therefore, the resulting iteration matrix is

$$M(\tau) := \begin{pmatrix} L & 0 \\ \tau B & I \end{pmatrix}^{-1} \begin{pmatrix} R & B^T \\ 0 & I \end{pmatrix} = \begin{pmatrix} L^{-1}R & L^{-1}B^T \\ -\tau B L^{-1}R & I - \tau B L^{-1}B^T \end{pmatrix}. \tag{10}$$

It is worth observing that the results of the present paper still hold if in the right-hand side of (9) the identity matrix in the (2,2)-blocks is replaced by any symmetric positive definite matrix [7].

From the well-known theory for stationary iterative methods for linear systems, we have

Proposition 2. *A member of the $\{\texttt{L,R}\}$-class converges Q-linearly from any initial point if and only if the corresponding iteration matrix $M(\tau)$ satisfies, for some $\tau > 0$,*

$$\rho(M(\tau)) < 1. \tag{11}$$

In this report, we establish that, under two mild assumptions, condition (11) holds for the entire $\{\texttt{L,R}\}$-class.

4

2.3 Classic multiplier and ADM methods

The trivial splitting $(L, R) = (H, 0)$ gives the classic multiplier method [2, 3], and is also equivalent to Uzawa's method [6] applied to (3). In this case,

$$M(\tau) = \begin{pmatrix} 0 & H^{-1}B^T \\ 0 & I - \tau B H^{-1} B^T \end{pmatrix},$$

and

$$\rho(M(\tau)) = \rho\left(I - \tau B H^{-1} B^T\right), \tag{12}$$

leading to the well-known convergence result for the multiplier method.

Proposition 3. *The Lagrangian multiplier method applied to the quadratic program (1) converges Q-linearly from any initial point for $\tau \in \left(0, 2/\lambda_{\max}(BH^{-1}B^T)\right)$ where $H = \alpha A + B^T B \succ 0$. Moreover, when $A \succeq 0$, $\tau \in (0, 2)$ suffices for convergence.*

The classic multiplier method, or Uzawa's method applied to (3), is a special and unique member of the {L,R}-class, which requires solving systems involving the entire (1,1)-block submatrix H with a different right-hand side from iteration to iteration. On the other hand, all other {L,R}-class members only require solving systems involving the left part L which can be much less expensive if L are chosen to exploit problem structures.

When the splitting of H is of the (2×2)-block Gauss-Seidel type as is defined in (7), the associated {L,R}-class member reduces to the classic augmented Lagrangian alternating direction method (AMD) [5, 4], for which convergence has been established for general convex functions not restricted to quadratics. However, such general theory requires objective functions to be separable with respect to the two block variables and convex in the entire space. Apparently no convergence results are available, to the best of our knowledge, when the objective is inseparable or convex only in a subspace, or the number of block variables exceeds two.

3 Convergence of the Entire Class

We present a unified convergence result for the entire {L,R}-class under two assumptions:

A1. $H := \alpha A + B^T B \succ 0$.

A2. $H = L - R$ satisfies $\rho(L^{-1}R) \leq 1$ and condition (13) below.

We know that Assumption A1 holds for appropriate α values if $A \in \mathbb{R}^{n \times n}$ is positive definite in the null space of B, see Proposition 1. We further require that $L^{-1}R$ have no eigenvalue of unit modulus or greater except possibly the unity itself being an eigenvalue; that is,

$$\max \left\{ |\mu| : \mu \in \sigma(L^{-1}R) \setminus \{1\} \right\} < 1. \tag{13}$$

Now we present a unified convergence theorem for the entire $\{\texttt{L,R}\}$-class.

Theorem 1. *Let $\{(x^k, y^k)\}$ be generated from any initial point by a member of the $\{\texttt{L,R}\}$-class defined by (8). Under Assumptions A1-A2, there exists $\eta > 0$ such that for all $\tau \in (0, 2\eta)$ the sequence $\{(x^k, y^k)\}$ converges Q-linearly to the solution of (1).*

The proof is left to the next section after we develop some technical results. We note that the convergence interval $(0, 2\eta)$ is member-dependent. It can also depend on the value of parameter $\alpha > 0$ in $H(\alpha) = \alpha A + B^T B \succ 0$.

It is worth noting that the theorem only requires $L^{-1}R$, as a linear mapping in $\mathbb{R}^n$, to be non-expansive (plus a technical condition) rather than contractive. Convergence would not necessarily happen if one kept iterating on the primal variable x only. However, timely updating the multiplier y helps the pair converge together.

4 Technical Results and Proof of Convergence

We first derive some useful technical lemmas. Let $\lambda(\tau)$ be an eigenvalue of $M(\tau)$, i.e.,

$$\lambda(\tau) \in \sigma(M(\tau)). \tag{14}$$

The eigenvalue system corresponding to λ is

$$\begin{pmatrix} L^{-1}R & L^{-1}B^T \\ -\tau B L^{-1}R & I - \tau B L^{-1}B^T \end{pmatrix} \begin{pmatrix} u(\tau) \\ v(\tau) \end{pmatrix} = \lambda(\tau) \begin{pmatrix} u(\tau) \\ v(\tau) \end{pmatrix}, \tag{15}$$

where $(u, v) \in \mathbb{C}^n \times \mathbb{C}^m$ is nonzero. For simplicity, we will often skip the τ-dependence of the eigen-pair if no confusion arises.

Lemma 1. *If $\rho(L^{-1}R) \leq 1$, then*

$$\rho(M(0)) = 1.$$

Under condition (13) the maximum eigenvalue of $M(\tau)$ in modulus, $\lambda(\tau)$, satisfies

$$\lim_{\tau \to 0} \frac{\lambda(\tau) - 1}{\lambda(\tau)} = 0. \tag{16}$$

6

Proof. From the definition of $M(\tau)$ in (10),

$$M(0) = \begin{pmatrix} L^{-1}R & L^{-1}B^T \\ 0 & I \end{pmatrix}.$$

Hence by our assumption $\rho(M(0)) = \max(1, \rho(L^{-1}R)) = 1 \in \sigma(M(0))$. The second part follows from the continuity of eigenvalues as functions of matrix elements and condition (13). $\square$

Lemma 2. *Let A be positive definite in the null space of B and $\alpha > 0$, or $H = \alpha A + B^T B \succ 0$. For any $\tau > 0$,*

$$1 \notin \sigma(M(\tau))$$

where $M(\tau)$ is given in (10).

Proof. We examine the eigen-system (15). Rearranging the first equation of (15), we have

$$(\lambda L - R)u = B^T v. \tag{17}$$

Multiplying the first equation by τB and adding to the second of (15), after rearranging we obtain

$$(1 - \lambda)v = \lambda \tau Bu. \tag{18}$$

Suppose that $\lambda = 1$. Then (18) implies $Bu = 0$. By the definition (6), equation (17) reduces to

$$(L - R)u \equiv (\alpha A + B^T B)u = B^T v.$$

Multiplying the above equation by u^* and invoking $Bu = 0$, we arrive at $u^* H u = u^* A u = 0$, contradicting to the assumption of the lemma. $\square$

Lemma 3. *Let $(\lambda, (u, v))$ be an eigen-pair of $M(\tau)$ as is given in (15) where $\lambda \notin \{0, 1\}$ and $Bu \neq 0$, then*

$$\lambda = 1 - \tau \left(\frac{u^* H u}{u^* B^T Bu} + \frac{\lambda - 1}{\lambda} \frac{u^* R u}{u^* B^T Bu} \right)^{-1}. \tag{19}$$

Proof. It follows readily from (18) that

$$v = \frac{\lambda \tau}{1 - \lambda} Bu. \tag{20}$$

Substituting the above into (17) and in view of (6), we have

$$(\lambda H + (\lambda - 1)R)\, u = \frac{\lambda \tau}{1 - \lambda} B^T Bu,$$

7

or after a rearrangement,

$$\left(H - \frac{\tau}{1-\lambda}B^TB\right)u = \frac{1-\lambda}{\lambda}Ru. \tag{21}$$

Multiplying both sides of (21) by u^*, we have

$$u^*Hu - \frac{\tau}{1-\lambda}u^*B^TBu = \frac{1-\lambda}{\lambda}u^*Ru.$$

Since $u^*B^TBu \neq 0$, the above equation can be rewritten into

$$\frac{\tau}{1-\lambda} = \frac{u^*Hu}{u^*B^TBu} + \frac{\lambda-1}{\lambda}\frac{u^*Ru}{u^*B^TBu}. \tag{22}$$

Solving for the λ on the left-hand side of (22) while fixing the ones on the right, we obtain the desired result where the denominator term must be nonzero. □

Lemma 4. *Let $\tau, \kappa \in \mathbb{R}$ and $z = \Re(z) + i\Im(z) \in \mathbb{C}$ such that $\kappa + \Re(z) > 0$. Then*

$$\tau \in (0, 2(\kappa + \Re(z))) \iff \left|1 - \frac{\tau}{\kappa + z}\right| < 1. \tag{23}$$

Moreover, $\tau = \kappa + \Re(z)$ minimizes the above modulus so that

$$\min_{\tau}\left|1 - \frac{\tau}{\kappa + z}\right| = \left|1 - \frac{\kappa + \Re(z)}{\kappa + z}\right| = \frac{|\Im(z)|}{|\kappa + z|}. \tag{24}$$

Proof. By direct calculation,

$$\left|1 - \frac{\tau}{\kappa + z}\right|^2 = 1 - \tau\frac{2(\kappa + \Re(z)) - \tau}{|\kappa + z|^2} = \frac{(\kappa + \Re(z) - \tau)^2 + \Im(z)^2}{(\kappa + \Re(z))^2 + \Im(z)^2}, \tag{25}$$

from which both (23) and (24) follow. □

Now we are ready to prove Theorem 1.

Proof. The proof follows from Lemmas 1, 3 and 4, while Lemma 2 is implicitly used.

Let $(\lambda(\tau), (u(\tau), v(\tau)))$ be an eigen-pair of $M(\tau)$ corresponding to an eigenvalue of maximum modulus. Clearly, $\lambda(\tau) \notin \{0, 1\}$. We need to prove that $|\lambda(\tau)| < 1$ for some values of $\tau > 0$. In the rest of the proof, we often skip the dependence on τ.

We consider two cases: $Bu = 0$ and $Bu \neq 0$. If $Bu = 0$, then (18) implies $v = 0$, and (17) implies that (λ, u) is an eigen-pair of $L^{-1}R$. Therefore, $|\lambda| < 1$ by Assumption A2. Now we assume that $Bu \neq 0$. By Lemmas 3 and 4, $|\lambda(\tau)| < 1$ if and only if the following inclusion is feasible,

$$\tau \in (0, 2\Theta(\tau)), \tag{26}$$

8

where

$$\Theta(\tau) \quad := \quad \frac{u^* H u}{u^* B^T B u} + \Re\left(\frac{\lambda - 1}{\lambda}\frac{u^* R u}{u^* B^T B u}\right)$$

$$= \quad \frac{u^* u}{u^* B^T B u}\left(\frac{u^* H u}{u^* u} + \Re\left(\frac{\lambda - 1}{\lambda}\frac{u^* R u}{u^* u}\right)\right) \tag{27}$$

Under Assumption A2, we know from (16) in Lemma 1 that $1 - 1/\lambda(\tau) \to 0$ as $\tau \to 0$. Hence, in view of the boundedness of $u^* R u / u^* u$, for any $\delta \in (0,1)$ there exists $\xi_\delta > 0$ such that

$$\Re\left(\frac{\lambda - 1}{\lambda}\frac{u^* R u}{u^* u}\right) \geq -\left|\frac{\lambda - 1}{\lambda}\right|\frac{|u^* R u|}{u^* u} \geq -\delta\lambda_{\min}(H), \quad \forall\,\tau \in (0, 2\xi_\delta). \tag{28}$$

We now estimate $\Theta(\tau)$ for $\tau \in (0, 2\xi_\delta)$ from (27) and (28),

$$\Theta(\tau) \geq \frac{\lambda_{\min}(H) - \delta\lambda_{\min}(H)}{\lambda_{\max}(B^T B)} = (1 - \delta)\frac{\lambda_{\min}(H)}{\lambda_{\max}(B^T B)} := \theta_\delta > 0, \quad \forall\,\tau \in (0, 2\xi_\delta). \tag{29}$$

It follows from (29) that inclusion (26) indeed holds for all $\tau \in (0, 2\eta)$ where

$$\eta := \min(\xi_\delta, \theta_\delta). \tag{30}$$

This completes the proof. $\qquad\qquad\square$

In view of the second part of Lemma 4, if τ_o solves the equation $\tau = \Theta(\tau)$, then the best achievable rate of convergence (for a given α) is no greater than

$$\frac{|\Im(z(\tau_o))|}{|u(\tau_o)^* H u(\tau_o) + z(\tau_o)|} = \left(1 + \frac{(u(\tau_o)^* H u(\tau_o) + \Re(z(\tau_o))^2}{\Im(z(\tau_o))^2}\right)^{-\frac{1}{2}} < 1, \tag{31}$$

where $z(\tau) := \frac{\lambda(\tau) - 1}{\lambda(\tau)} u(\tau)^* R u(\tau)$. The formula is of course impractical to compute, but might serve as an estimation tool.

5 Remarks

The $\{\texttt{L},\texttt{R}\}$-class, defined by (8), is constructed from splitting the (1,1)-block of the saddle point system matrix that includes, but is not limited to, all known convergent splittings for positive definite matrices, offering adaptivity to problem structures with guaranteed convergence.

Those $\{\texttt{L},\texttt{R}\}$-class members associated with block Gauss-Seidel splittings are natural extensions to the classic ADM specialized to quadratics. In contrast to the existing general convergence theory for ADM, Theorem 1 does not require separability, nor convexity in the entire space, imposes no restriction on the number of block variables, and gives a Q-linear rate of convergence. It

should be of great interest to extend these properties beyond quadratic functions, which is a topic to be addressed in another report.

The convergence of certain members of the {L,R}-class has been studied in [9] under the assumption that L is symmetric positive definite. In [8], a special case corresponding to the SOR-splitting has been analyzed.

Acknowledgments

References

[1] M. Benzi, G. H. Golub and J. Liesen. Numerical solution of saddle point problems. *Acta Numerica* (2005), pp. 1–137. Cambridge University Press, 2005. United Kingdom.

[2] M. R. Hestenes. Multiplier and gradient methods. J. Optimiz. Theory App., vol. 4, pp. 303–320, 1969.

[3] M. J. D. Powell. A method for nonlinear constraints in minimization problems. Optimization, R. Fletcher, ed., Academic Press, New York, NY, pp. 283–298, 1969.

[4] D. Gabay, and B. Mercier. A dual algorithm for the solution of nonlinear variational problems via finite-element approximations. Comp. Math. Appl., vol. 2, pp. 17–40, 1976.

[5] R. Glowinski, and A. Marrocco. Sur lapproximation par elements finis dordre un, et la resolution par penalisation-dualite dune classe de problemes de Dirichlet nonlineaires. Rev. Francaise dAut. Inf. Rech. Oper., R-2, pp. 41–76, 1975.

[6] H. Uzawa. Iterative methods for concave programming. in *Studies in Linear and Nonlinear Programming*. (K. J. Arrow, L. Hurwicz and H. Uzawa, eds), Stanford University Press, Stanford, CA, pp. 154–165. (1958).

[7] Junfeng Yang. Private communications. August, 2010.

[8] J. Zhang and J. Shang. A class of Uzawa-SOR methods for saddle point problems. Applied Mathematics and Computation. 216, pp. 2163-2168 (2010).

[9] W. Zulehner. Analysis of Iterative Methods for Saddle Point Problems: A Unified Approach. Mathematics of Computation. Vol. 71, No. 238, pp. 479-505 (2001).