On Convergence Rate of the Augmented Lagrangian Algorithm for Nonsymmetric Saddle Point Problems

G. M. Awanou [a,*], M.J. Lai [b]

[a] *Institute for Mathematics and its Applications, University of Minnesota*

[b] *Department of Mathematics, University of Georgia*

Abstract

We are interested in solving the system

$$\begin{pmatrix} A & L^T \\ L & 0 \end{pmatrix} \begin{pmatrix} c \\ \lambda \end{pmatrix} = \begin{bmatrix} F \\ G \end{bmatrix} \tag{1}$$

by a variant of the augmented lagrangian algorithm. This type of problem with nonsymmetric A typically arises in certain discretizations of the Navier-Stokes equations. Here A is a (n, n) matrix, c, $F \in \mathbb{R}^n$, L is a (m, n) matrix, and $\lambda, G \in \mathbb{R}^m$. We assume that A is invertible on the kernel of L. Convergence rates of the augmented lagrangian algorithm are known in the symmetric case but the proofs in [4] used spectral arguments and cannot be extended to the nonsymmetric case. The purpose of this paper is to give a rate of convergence of a variant of the algorithm in the nonsymmetric case. We illustrate the performance of this algorithm with numerical simulations of the lid-driven cavity flow problem for the 2D Navier-Stokes equations.

Key words: Augmented Lagrangian Algorithm, Saddle Point Problems
1991 MSC: 76-08, 65F10

1 Introduction

We will use the same notations $(.,.)$ and $\|\cdot\|$ for the inner products and norms in $\mathbb{R}^n$ and $\mathbb{R}^m$. The particular inner product will be identified by the types

* 400 Lind Hall, 207 Church Street S.E. Minneapolis, MN 55455-0436

of matrices appearing. The augmented lagrangian algorithm for symmetric problems can be derived by minimization arguments. We refer to [4] for details. It is described as follows: with $r > 0$ and $\rho_l > 0$ for all l as parameters, given $\lambda^{(0)} \in \mathbb{R}^m$ specified arbitrarily, with $\lambda^{(l)}$ known, compute $c^{(l)}$ then $\lambda^{(l+1)}$ by

$$\begin{cases} (A + rL^T L)c^{(l)} + L^T\lambda^{(l)} = F + rL^T G \\ \lambda^{(l+1)} = \lambda^{(l)} + \rho_l(Lc^{(l)} - G). \end{cases} \tag{2}$$

In [10], we were interested in a variant of this algorithm for $\rho_l = \rho = \frac{1}{\epsilon}$ for all l, $r = \frac{1}{\epsilon}$ where $\epsilon > 0$ is fixed. More precisely, for this choice of the parameters, the algorithm reads

$$\begin{cases} (A + \frac{1}{\epsilon}L^T L)c^{(l)} + L^T\lambda^{(l)} = F + \frac{1}{\epsilon}L^T G \\ \lambda^{(l+1)} = \lambda^{(l)} + \frac{1}{\epsilon}(Lc^{(l)} - G). \end{cases} \tag{3}$$

The variant we considered is the following algorithm

$$\begin{cases} (A + \frac{1}{\epsilon}L^T L)c^{(l+1)} + L^T\lambda^{(l)} = F + \frac{1}{\epsilon}L^T G \\ \lambda^{(l+1)} = \lambda^{(l)} + \frac{1}{\epsilon}(Lc^{(l+1)} - G), \end{cases} \tag{4}$$

which can be easily shown to be equivalent to the following sequence of problems

$$\begin{pmatrix} A & L^T \\ L & -\epsilon M \end{pmatrix} \begin{pmatrix} c^{(l+1)} \\ \lambda^{(l+1)} \end{pmatrix} = \begin{bmatrix} F \\ G - \epsilon M\lambda^{(l)} \end{bmatrix}, \tag{5}$$

for $M = I$ is the identity matrix of size $m \times m$. Here, M is a suitable chosen matrix. In [5], it was claimed that the later algorithm converges to the solution c of (1) and

$$\|c - c^{(l+1)}\| \leq C\epsilon\|c - c^{(l)}\|,$$

for a constant $C > 0$ independent of ϵ. We have not however been able to find a proof of this result in the literature. The main objective of this article is to prove the convergence of the following algorithm

$$\begin{cases} (A + rL^T M^{-1}L)c^{(l+1)} + L^T\lambda^{(l)} = F + rL^T M^{-1}G \\ \lambda^{(l+1)} = \lambda^{(l)} + \rho M^{-1}(Lc^{(l+1)} - G), \ \rho > 0, \end{cases} \tag{6}$$

which generalizes (5) and give a convergence rate similar to the one above. A fine study of the convergence rate still appears to be difficult (cf. remark 2.12 p. 64 [4]).

The algorithm (6) is the Uzawa algorithm applied to the augmented system

$$(A + rL^T M^{-1}L)c + L^T\lambda = F + rL^T M^{-1}G \tag{7}$$

and hence can have an improved convergence rate. One other advantage of the augmented lagrangian algorithm is to solve compared to (1), systems of smaller size.

However, following [3] p 15, $A_r = A + rL^T M^{-1} L$ has a condition number $\mathcal{K}(A_r)$ asymptotically proportional to r, that is

$$\mathcal{K}(A_r) \approx r \frac{||M^{-1}||\ ||L||^2}{\sigma} \quad \text{when } r \to \infty,$$

σ depending only on A and we again denote by $\|\cdot\|$ the matrix norm associated with $\|\cdot\|$ on $\mathbb{R}^n$. This prevents the choice of extremely large values of $r = \rho$ to get convergence in one step. Our convergence result and numerical evidence also suggest that one does not have a good convergence rate for r small and ρ large. We also see that a judicious choice of M can balance the deterioration of the condition number.

It is reasonable to expect that an inexact Uzawa algorithm applied to (7) might perform as well. Our numerical experiments did not yield convergence.

Other references for the augmented lagrangian algorithm are [1], [2], [3], [4] and [7].

The paper is organized as follows. We first give a sufficient condition for solvability of (1) which leads to a Ladyzhenskaya–Babuška–Brezzi (LBB) type condition. We then prove the convergence rate. Finally, we will give numerical experiments for the 2D Navier-Stokes equations.

2 Solvability

In this section, we derive a sufficient condition for the solvability of (1). Let $\mathrm{Ker}(X)$ and $\mathrm{Im}(X)$ denote the kernel and range of the operator X. We first give a few lemmas:

Lemma 1 $\mathbb{R}^n = \mathrm{Ker(L)} \oplus \mathrm{Im(L^T)}$ *and* $\mathbb{R}^m = \mathrm{Ker(L^T)} \oplus \mathrm{Im(L)}$.

Proof: It is enough to prove only one of the decompositions. Since $\mathrm{Im(L)} \subseteq \mathbb{R}^\mathrm{m}$, we have $\mathbb{R}^m = \mathrm{Im(L)} \oplus \mathrm{Im(L)}^\perp$, where $\mathrm{Im(L)}^\perp$ denotes the orthogonal of $\mathrm{Im(L)}$. We need to show that

$$\mathrm{Im(L)}^\perp = \mathrm{Ker(L^T)}.$$

Let $q \in \mathrm{Im(L)}^\perp$. For $w \in \mathbb{R}^n$, $Lw \in \mathrm{Im(L)}$, so $q^T(Lw) = 0$. Therefore $w^T L^T q = 0$ so that $L^T q$ is orthogonal to $\mathbb{R}^n$, that is $L^T q = 0$, i.e. $q \in \mathrm{Ker(L^T)}$. This argument also shows that $\mathrm{Ker(L^T)} \subset \mathrm{Im(L)}^\perp$ and the result follows. □

The following result can be found in [6], we give here a detailed proof for convenience.

Lemma 2 *Suppose A is an invertible linear operator with positive definite symmetric part $A_s = \frac{1}{2}(A + A^T)$ that satisfy*

$$(Ax, y) \leq \alpha (A_s x, x)^{\frac{1}{2}} (A_s y, y)^{\frac{1}{2}}, \quad \textit{for all } x, y \in \mathbb{R}^n \textit{ and } \alpha \geq 1, \tag{8}$$

then $(A^{-1})_s$ is positive definite and satisfy

$$((A^{-1})_s w, w) \leq ((A_s)^{-1} w, w) \leq \alpha^2 (A^{-1})_s w, w) \quad \textit{for all } w \in \mathbb{R}^n.$$

Moreover

$$(A^{-1}x, y) \leq \Big((A_s)^{-1} x, x\Big)^{\frac{1}{2}} \Big((A_s)^{-1} y, y\Big)^{\frac{1}{2}}.$$

Proof:

$$\Big((A_s)^{-1} w, w\Big) = \sup_{y \in \mathbf{R}^n} \frac{(w, y)^2}{(A_s y, y)},$$

and

$$(w, y)^2 = (w, A^{-1} A y)^2 = ((A^{-1})^T w, Ay)^2 \leq \alpha^2 (A_s y, y)(A_s (A^{-1})^T w, (A^{-1})^T w)$$

by (8). So

$$\begin{aligned}(w, y)^2 \leq \alpha^2 ||y||^2_{A_s} (A(A^{-1})^T w, (A^{-1})^T w) &= ||y||^2_{A_s} ((A^{-1})^T w, A^T (A^{-1})^T w) \\ &= ||y||^2_{A_s} ((A^{-1})^T w, w),\end{aligned}$$

where $||.||^2_{A_s} = (A_s ., .)$ and we used $(Aw, w) = (A_s w, w)$ for all $w \in \mathbb{R}^n$. It follows that

$$\Big((A_s)^{-1} w, w\Big) \leq \alpha^2 (A^{-1})_s w, w).$$

On the other hand (using fractional power of symmetric positive definite matrices which can be defined via singular values decomposition)

$$\begin{aligned}((A^{-1})_s w, w) = (A^{-1} w, w) &= (A_s^{\frac{1}{2}} A^{-1} w, (A_s)^{-\frac{1}{2}} w) \leq ||A_s^{\frac{1}{2}} A^{-1} w|| \;\; ||(A_s)^{-\frac{1}{2}} w|| \\ &= ||A^{-1} w||_{A_s} ||w||_{(A_s)^{-1}} = (A A^{-1} w, A^{-1} w)^{\frac{1}{2}} ||w||_{(A_s)^{-1}} \\ &= (A^{-1} w, w)^{\frac{1}{2}} ||w||_{(A_s)^{-1}}.\end{aligned}$$

It follows that

$$((A^{-1})_s w, w) \leq \Big((A_s)^{-1} w, w\Big). \tag{9}$$

In addition

$$
\begin{aligned}
(A^{-1}x, y) &= (A_s^{\frac{1}{2}} A^{-1}x, (A_s)^{-\frac{1}{2}} y) \\
&\leq (A^{-1}x, x)^{\frac{1}{2}} \Big((A_s)^{-1} y, y \Big)^{\frac{1}{2}} \quad \text{using the same arguments as above} \\
&\leq \Big((A_s)^{-1}x, x \Big)^{\frac{1}{2}} \Big((A_s)^{-1} y, y \Big)^{\frac{1}{2}} \quad \text{using (9) .} \qquad \square
\end{aligned}
$$

We can now prove the following theorem

Theorem 3 *Let A be a matrix which satisfies the condition (8) and has a symmetric part A_s positive definite with respect to L in the sense that $x^T A_s x \geq 0$ and $x^T A_s x = 0$ with $Lx = 0$ implies $x = 0$. In addition, assume that $G \in \text{Im}(L)$. Then (1) is solvable and moreover*

$$
sup_{u \in \mathbf{R}^n} \frac{(y, Lu)^2}{(A_r u, u)} \geq c_1 ||y||^2 \quad for \;\; all \; y \in Im(L), \tag{10}
$$

where $c_1 > 0$ is a positive constant which depends on r and $A_r = A + rL^T M^{-1} L$.

Proof: We have

$$
\begin{cases} Ac + L^T \lambda &= F \\ Lc = G, \end{cases} \tag{11}
$$

so $rL^T M^{-1} Lc = rL^T M^{-1} G$ which gives

$$
(A + rL^T M^{-1} L)c + L^T \lambda = F + rL^T M^{-1} G.
$$

We first show that $A + rL^T M^{-1} L$ is invertible. Since A is a square matrix, it is enough to show that

$$
(A + rL^T M^{-1} L)x = 0 \Rightarrow x = 0.
$$

We have

$$
x^T (A + rL^T M^{-1} L)x = x^T (A_s + rL^T M^{-1} L)x = x^T A_s x + r(Lx)^T M^{-1} (Lx),
$$

so by the assumptions on A,

$$
x^T (A + rL^T M^{-1} L)x = 0 \Rightarrow x^T A_s x = 0 \text{ and } (Lx)^T M^{-1} (Lx) = 0.
$$

It follows that $x^T A_s x = 0$ and $Lx = 0$. Since A_s is assumed to be symmetric positive definite with respect to L, we get $x = 0$. We can therefore write

$$
c = (A + rL^T M^{-1} L)^{-1} (F + rL^T M^{-1} G) - (A + rL^T M^{-1} L)^{-1} L^T \lambda.
$$

Since $Lc = G$, we see that the solvability of (1) is equivalent to solving

$$
L(A + rL^T M^{-1} L)^{-1} L^T \lambda = L(A + rL^T M^{-1} L)^{-1} (F + rL^T M^{-1} G) - G
$$

for λ. By Lemma 1, we have $\lambda = \lambda_0 + \overline{\lambda}$ with $\lambda_0 \in \mathrm{Ker}(L^T)$ and $\overline{\lambda} \in \mathrm{Im}(L)$. Clearly it is enough to find $\overline{\lambda}$. We show that there is $c_1 > 0$ such that

$$y^T L(A + rL^T M^{-1} L)^{-1} L^T y \geq c_1 ||y||^2 \quad \text{for all } y \in \mathrm{Im}(L).$$

This will imply that $L(A + rL^T M^{-1} L)^{-1} L^T$ is invertible on $\mathrm{Im}(L)$ and show that (1) is solvable.

Since A satisfies (8) and because

$$(A_s x, x) \leq ((A_s + rL^T M^{-1} L)x, x) \quad \text{and} \quad (rL^T M^{-1} Lx, x) \leq ((A_s + rL^T M^{-1} L)x, x),$$

we have

$$((A + rL^T M^{-1} L)x, y) \leq \alpha ((A_s + rL^T M^{-1} L)x, x)^{\frac{1}{2}} (A_s + rL^T M^{-1} L)y, y)^{\frac{1}{2}}, \quad \text{for all } x, y \in \mathbb{R}^n,$$

and $\alpha = 1$. It follows from Lemma 2 that

$$([(A + rL^T M^{-1} L)^{-1}]_s w, w) \geq \frac{1}{\alpha^2} ((A_s + rL^T M^{-1} L)^{-1} w, w) \quad \text{for all } w \in \mathbb{R}^n.$$

Because A_s is positive definite with respect to L, $A_s + rL^T M^{-1} L$ is symmetric positive definite and so $L(A_s + rL^T M^{-1} L)^{-1} L^T$ is symmetric positive definite on $\mathrm{Im}(L)$ ($L^T z = 0$ and $z \in \mathrm{Im}(L)$ implies $z = 0$) so that we have

$$y^T L(A_s + rL^T M^{-1} L)^{-1} L^T y \geq c_0 ||y||^2, \quad \text{for all } y \in \mathrm{Im}(L), \tag{12}$$

with $c_0 > 0$ depending on r. This gives

$$\begin{aligned} y^T L(A + rL^T M^{-1} L)^{-1} L^T y &= (L^T y)^T (A + rL^T M^{-1} L)^{-1} (L^T y) \\ &= (L^T y)^T [(A + rL^T M^{-1} L)^{-1}]_s (L^T y) \\ &\geq \frac{1}{\alpha^2} (L^T y)^T (A_s + rL^T M^{-1} L)^{-1} (L^T y) \\ &= \frac{1}{\alpha^2} y^T L(A_s + rL^T M^{-1} L)^{-1} L^T y \\ &\geq \frac{c_0}{\alpha^2} ||y||^2, \quad \text{for all } y \in \ \mathrm{Im}(L). \end{aligned}$$

Recall that $A_r = (A + rL^T M^{-1} L)^{-1}$ and notice that

$$\begin{aligned} y^T L A_r^{-1} L^T y &= (A_r^{-1} L^T y, L^T y) \\ &= \sup_{u \in \mathbf{R}^n} \frac{(L^T y, u)^2}{(A_r u, u)} \\ &= \sup_{u \in \mathbf{R}^n} \frac{(y, Lu)^2}{(A_r u, u)}. \end{aligned}$$

We have therefore proved that

$$\sup_{u\in \mathbb{R}^n} \frac{(y,Lu)^2}{(A_r u,u)} \geq \frac{c_0}{\alpha^2}||y||^2, \quad \text{for all } y \in \text{Im}(L),$$

which is a LBB type condition. □

It would be desirable to have more information on the dependence on r of the constant c_1 in Theorem (3). Put

$$E = A_s + rL^T M^{-1} L,$$

and recall from (12) that $LE^{-1}L^T$ is symmetric positive definite on Im(L). We let $D = LE^{-1}L^T$ considered as a mapping from Im(L) to Im(L) and seek a lower bound of

$$R(y) = \frac{y^T D y}{y^T y} > 0, \ y \in \text{Im}(L), \ y \neq 0,$$

where $R(y)$ is the Raleigh quotient. We have

$$\begin{aligned} R(y) &= \frac{y^T LE^{-1}L^T y}{y^T y} = \frac{y^T LE^{-1}(EE^{-1})L^T y}{y^T y} \\ &= \frac{(y^T LE^{-1})E(E^{-1}L^T y)}{y^T y}. \end{aligned}$$

So

$$R(y) = \frac{||E^{-1}L^T y||_E^2}{y^T y}, \tag{13}$$

where we have defined a norm $||\ ||_E$ associated with the symmetric positive definite matrix E;

$$||u||_E^2 = b(u,u) \text{ with } b(u,v) = v^T E u.$$

We have

$$||E^{-1}L^T y||_E = \sup_{\substack{v\in \mathbb{R}^n \\ v\neq 0}} \frac{b(E^{-1}L^T y, v)}{||v||_E},$$

with

$$\frac{b(E^{-1}L^T y, v)}{||v||_E} = \frac{v^T E E^{-1} L^T y}{||v||_E} = \frac{v^T L^T y}{||v||_E} = \frac{y^T L v}{||v||_E}, \quad \forall\ v \neq 0 \ \in \ \mathbb{R}^n.$$

So

$$||E^{-1}L^T y||_E = \sup_{\substack{v\in \mathbb{R}^n \\ v\neq 0}} \frac{y^T L v}{||v||_E}. \tag{14}$$

We claim that L is an invertible mapping from $\mathrm{Im}(L^T)$ to $\mathrm{Im}(L)$, [8]. Since $y \in \mathrm{Im}(L)$, there is v in $\mathrm{Im}(L^T)$ such that $y = Lv$. We write $||v||^2 = v^T v$. Then,

$$R(y) \geq \frac{1}{||Lv||^2}\left(\frac{(Lv)^T(Lv)}{||v||_E}\right)^2 = \frac{||Lv||^2}{||v||_E^2}$$
$$= \frac{||Lv||^2}{v^T A_s v + r v^T L^T M^{-1} L v} = \frac{||Lv||^2}{v^T A_s v + r||Lv||_{M^{-1}}^2}.$$

For an operator X, $\mu_{X\max}$ denotes the greatest eigenvalue of X. We have, using Raleigh's principle,

$$v^T A_s v \leq \mu_{A_s\max}||v||^2.$$

On the other hand, since $y = Lv \neq 0$, $||Lv|| \geq \frac{||v||}{||L^{-1}||}$. We therefore have

$$R(y) \geq \frac{\frac{1}{||L^{-1}||^2}}{\mu_{A_s\max} + r||L||^2} > 0,$$

We have

$$c_1 = \frac{c_0}{\alpha^2} \geq \frac{1}{\alpha^2}\frac{\frac{1}{||L^{-1}||^2}}{\mu_{A_s\max} + r||L||^2}$$

3 Convergence

In this section, we prove the convergence of the iterative algorithm (6). In the next section we give the convergence rate.

Theorem 4 *Suppose that the linear system (1) has a unique solution c and that A_s the symmetric part of A is positive definite with respect to L. Moreover, assume that M is symmetric positive definite. Then, the sequence $(c^{(l)})$ defined in (6) converges to the solution c of (1) for $r \geq \frac{\varrho}{2}$.*

Proof: Clearly (6) is solvable since A_r is invertible. With $\lambda^{(l)}$ given one computes successively $c^{(l+1)}$ and $\lambda^{(l+1)}$.

The original problem (1),

$$\begin{cases} Ac + L^T\lambda = F \\ Lc = G, \end{cases}$$

can be rewritten as

$$\begin{cases} (A + rL^T M^{-1} L)c + L^T \lambda = F + rL^T M^{-1} G \\ \lambda = \lambda + \rho M^{-1}(Lc - G). \end{cases} \tag{15}$$

Let $u^{(l+1)} = c^{(l+1)} - c$ and $p^{(l+1)} = \lambda^{(l+1)} - \lambda$. We have, using (15) and (6),

$$(A + rL^T M^{-1} L)u^{(l+1)} + L^T p^{(l)} = 0 \tag{16}$$

and

$$p^{(l+1)} = p^{(l)} + \rho M^{-1} L u^{(l+1)}. \tag{17}$$

We deduce from (17) that

$$\|p^{(l+1)}\|_M^2 = (Mp^{(l+1)}, p^{(l+1)}) = (Mp^{(l)} + \rho L u^{(l+1)}, p^{(l)} + \rho M^{-1} L u^{(l+1)})$$

which gives

$$\begin{aligned} \|p^{(l+1)}\|_M^2 &= \|p^{(l)}\|_M^2 + (Mp^{(l)}, \rho M^{-1} L u^{(l+1)}) \\ &\quad + (\rho L u^{(l+1)}, p^{(l)}) + \rho^2 (L u^{(l+1)}, M^{-1} L u^{(l+1)}) \\ &= \|p^{(l)}\|_M^2 + 2\rho(p^{(l)}, L u^{(l+1)}) + \rho^2 (L u^{(l+1)}, M^{-1} L u^{(l+1)}), \end{aligned}$$

since M^{-1} is symmetric and hence

$$\|p^{(l)}\|_M^2 - \|p^{(l+1)}\|_M^2 = -2\rho(p^{(l)}, L u^{(l+1)}) - \rho^2 (L u^{(l+1)}, M^{-1} L u^{(l+1)}). \tag{18}$$

It follows from (16) that

$$(Au^{(l+1)}, u^{(l+1)}) + r(L^T M^{-1} L u^{(l+1)}, u^{(l+1)}) = -(p^{(l)}, L u^{(l+1)}),$$

and hence by substituing this into (18), we get

$$\begin{aligned} \|p^{(l)}\|_M^2 - \|p^{(l+1)}\|_M^2 &= 2\rho(Au^{(l+1)}, u^{(l+1)}) + (2r\rho - \rho^2)(L u^{(l+1)}, M^{-1} L u^{(l+1)}) \\ &= 2\rho(Au^{(l+1)}, u^{(l+1)}) + (2r\rho - \rho^2)\|L u^{(l+1)}\|_{M^{-1}}^2 \end{aligned}$$

Therefore for $r \geq \frac{\rho}{2}$, since A_s is nonnegative,

$$\|p^{(l)}\|_M^2 - \|p^{(l+1)}\|_M^2 \geq 0,$$

i.e.

$$\|p^{(l+1)}\|_M \leq \|p^{(l)}\|_M, \quad \text{for all } l \tag{19}$$

and the sequence $\{\|p^{(l)}\|_M\}$ is seen to be decreasing. If for some l, $\|p^{(l)}\|_M - \|p^{(l+1)}\|_M = 0$, then $(Au^{(l+1)}, u^{(l+1)}) = Lu^{(l+1)} = 0$ and we have $u^{(l+1)} = 0$, since A_s is symmetric positive definite with respect to L, and hence convergence. Otherwise, the sequence being bounded below by 0 converges; hence $\|p^{(l)}\|^2 - \|p^{(l+1)}\|^2$ converges to 0 which implies that $(A_s u^{(l)}, u^{(l)})$ and $\|Lu^{(l)}\|^2$ converge to 0. Since $A_s + rL^T M^{-1} L$ is positive definite, it follows that $u^{(l)}$ converges to 0 and finally $c^{(l)}$ converges to c. $\square$

4 Convergence Rate

To show the convergence rate, we will need the following lemma.

Lemma 5 . *The mappings* $L : \mathrm{Im}(\mathrm{L^T}) \to \mathrm{Im}(\mathrm{L})$ *and* $L^T : \mathrm{Im}(\mathrm{L}) \to \mathrm{Im}(\mathrm{L^T})$ *are bijections with bounded inverses.*

Proof: We show that L is one-one on $\mathrm{Im}(\mathrm{L^T})$. This is immediate since $Lx = 0$ with $x \in \mathrm{Im}(\mathrm{L^T})$ implies $x \in \mathrm{Im}(\mathrm{L^T}) \cap \mathrm{Ker}(\mathrm{L}) = \{0\}$ by Lemma 1. As a linear mapping between finite dimensional spaces, L has a bounded inverse on $\mathrm{Im}(\mathrm{L})$ and there exists $k_0 > 0$ such that for any $g \in \mathrm{Im}(\mathrm{L})$, there exists $v_g \in \mathrm{Im}(\mathrm{L^T})$ such that $Lv_g = g$ with

$$\|v_g\| \leq \frac{1}{k_0}\|g\|. \tag{20}$$

A similar proof applies to L^T. This completes the proof. $\square$

We would like to elaborate on this last inequality. Let $v \in \mathbb{R}^n$ and $g = Lv \in \mathrm{Im}(\mathrm{L})$ so there is $v_g \in \mathrm{Im}(\mathrm{L^T})$ for which $g = Lv_g$. That is, $Lv = Lv_g$. It follows that $v_g = v + v_0$ for some $v_0 \in \mathrm{Ker}(\mathrm{L})$. We therefore have

$$\|v + v_0\| \leq \frac{1}{k_0}\|Lv\| = \frac{1}{k_0}\|L(v + v_0)\|$$

for any $v + v_0 \in \mathrm{Im}(\mathrm{L^T})$. It follows that

$$\|v\| \leq \frac{1}{k_0}\|Lv\| = \frac{1}{k_0}\sup_{q \in \mathbf{R}^n} \frac{(q, Lv)}{\|q\|}, \quad \text{for all } v \in \mathrm{Im}(\mathrm{L^T}). \tag{21}$$

The same arguments applied to L^T show that

$$\|q\| \leq \frac{1}{k_0}\|L^T q\| = \frac{1}{k_0}\sup_{v \in \mathbf{R}^n} \frac{(v, L^T q)}{\|v\|} \quad \text{for all } q \in \mathrm{Im}(\mathrm{L}). \tag{22}$$

We have the following theorem

Theorem 6 *Suppose that the linear system (1) has a unique solution c and that A_s the symmetric part of A is positive definite with respect to L. Moreover,* *assume that M is symmetric positive definite. Then,*

$$\|c - c^{(l+1)}\| \leq C\|c - c^{(l)}\|,$$

for a positive constant C which depends on r and ρ but independent of l. Moreover for $r = \rho = \frac{1}{\epsilon}$,

$$\|c - c^{(l+1)}\| \leq C\epsilon\|c - c^{(l)}\|,$$

for a positive constant C independent of l and ϵ.

Proof: We write $u^{(l+1)} = \hat{u}^{(l+1)} + \overline{u}^{(l+1)}$ with $\hat{u}^{(l+1)} \in \mathrm{Ker}(L)$ and $\overline{u}^{(l+1)} \in \mathrm{Im}(L^T)$. Using (21), we have

$$k_0\|\overline{u}^{(l+1)}\| \leq \|Lu^{(l+1)}\|,$$

where we used the inner product $(M.,.)$ instead of the canonical one. Using (17), we have

$$Lu^{(l+1)} = \frac{1}{\rho}Mp^{(l+1)} - \frac{1}{\rho}Mp^{(l)},$$

so

$$k_0\|\overline{u}^{(l+1)}\| \leq \frac{||M||}{\rho}\Big(\|p^{(l+1)}\|_M + \|p^{(l)}\|_M\Big),$$

where we used the equivalence of norms on $\mathbb{R}^m$. By (19) we have $\|p^{(l+1)}\|_M \leq \|p^{(l)}\|_M$ which gives

$$\|\overline{u}^{(l+1)}\| \leq \frac{2||M||}{\rho k_0}\|p^{(l)}\|_M.$$

We next give a bound on $\hat{u}^{(l+1)}$. Since $A+rL^TM^{-1}L$ is invertible, A is invertible on $\mathrm{Ker}(L)$. Indeed if $Ax = 0$ and $Lx = 0$, then $(A + rL^TM^{-1}L)x = 0$ which implies that $x = 0$. Therefore there is $\alpha_0 > 0$ such that

$$\begin{aligned} \alpha_0\|\hat{u}^{(l+1)}\| &\leq \sup_{v_0\in \mathrm{Ker}(L)} \frac{(v_0, A\hat{u}^{(l+1)})}{\|v_0\|} \\ &= \sup_{v_0\in \mathrm{Ker}(L)} \frac{v_0^TAu^{(l+1)} - v_0^TA\overline{u}^{(l+1)}}{\|v_0\|} \end{aligned}$$

However, from (16), we have $Au^{(l+1)} = -L^Tp^{(l)} - rL^TM^{-1}Lu^{(l+1)}$ which implies

$$\begin{aligned} v_0^TAu^{(l+1)} &= -v_0^TL^Tp^{(l)} - rv_0^TL^TM^{-1}Lu^{(l+1)} \\ &= -(Lv_0)^Tp^{(l)} - r(Lv_0)^TM^{-1}Lu^{(l+1)} = 0 \end{aligned}$$

for $v_0 \in \mathrm{Ker}(L)$. Thus,

$$\begin{aligned}
\alpha_0 \|\hat{u}^{(l+1)}\| &\leq \sup_{v_0 \in \mathrm{Ker}(L)} \frac{-v_0^T A \overline{u}^{(l+1)}}{\|v_0\|} \\
&\leq \|A\| \, \|\overline{u}^{(l+1)}\| \\
&\leq \frac{2\|A\|\|M\|}{\rho k_0} \, \|p^{(l)}\|_M.
\end{aligned}$$

We therefore have

$$\begin{aligned}
\|u^{(l+1)}\| &\leq \|\overline{u}^{(l+1)}\| + \|\hat{u}^{(l+1)}\| \\
&\leq \frac{2}{\rho k_0}\Big(\frac{||A||}{\alpha_0} + 1\Big)||M||\|p^{(l)}\|_M.
\end{aligned}$$

We now give a bound on $\|p^{(l)}\|_M$ in terms of $\|u^{(l)}\|$. First we establish that $Mp^{(l)} \in \mathrm{Im}(L)$. Solving for $u^{(l+1)}$ in (16) and substituting in (17), we get

$$p^{(l+1)} = (I - \rho M^{-1} L (A + r L^T M^{-1} L)^{-1} L^T) p^{(l)}.$$

It follows that $M(p^{(l+1)} - p^{(l)})$ is in the range of L. Since

$$p^{(l+1)} = \sum_{j=1}^{k+1} p^{(j)} - p^{(j-1)} + p^{(0)},$$

we have $Mp^{(l)} \in \mathrm{Im}(L)$ provided $Mp^{(0)} = M\lambda - M\lambda_0 \in \mathrm{Im}(L)$ which is possible by a suitable choice of λ_0. One way to do this is to first notice that by (11), we may assume that $\lambda \in \mathrm{Im}(L)$. So if $M = I$, where I is the identity matrix of $\mathbb{R}^m$, we may choose λ_0 in $\mathrm{Im}(L)$. Otherwise, we can choose M in such a way that M maps $\mathrm{Im}(L)$ into $\mathrm{Im}(L)$.

It follows from (22) that

$$k_0 \|Mp^{(l)}\|_{M^{-1}} = k_0 \|p^{(l)}\|_M \leq \sup_{v \in \mathbb{R}^n} \frac{v^T L^T p^{(l)}}{\|v\|}.$$

Combining the equations in (16) and (17), we get

$$(A + r L^T M^{-1} L - \rho L^T M^{-1} L) u^{(l)} + L^T p^{(l)} = 0,$$

thus $v^T L^T p^{(l)} = -v^T (A + r L^T M^{-1} L - \rho L^T M^{-1} L) u^{(l)}$. Therefore,

$$\|p^{(l)}\|_M \leq \frac{||A_{r,\rho}||}{k_0} \|u^{(l)}\|,$$

where $A_{r,\rho} = A + (r - \rho) L^T M^{-1} L$. It follows that

$$\|u^{(l+1)}\| \leq \frac{2}{\rho k_0}\Big(\frac{||A||}{\alpha_0} + 1\Big)||M||\frac{||A_{r,\rho}||}{k_0}\|u^{(l)}\|.$$

For $r = \rho$, $A_{r,\rho} = A$ and for $r = \rho = \frac{1}{\epsilon}$, we get

$$\|c - c^{(l+1)}\| \leq C\epsilon \|c - c^{(l)}\|,$$

for a positive constant C independent of l and ϵ. □

5 Numerical experiments

In this section, we first present a spline discretization of the 2D Navier-Stokes equations in velocity-pressure formulation. A simple iterative algorithm is used to linearize the nonlinear equations. We have applied the algorithm described here to the solutions of the linear systems which arise and display the maximum number of the iterations that was necessary to fully solve the problem for various choices of the parameters r and ρ.

Let $\Omega = \cup_{t \in \triangle} t$ be a polygonal domain in $\mathbb{R}^2$. Given two integers $d \geq 0$ and $0 \leq r < d$, we consider the spline space of degree d and smoothness r

$$S^r_d(\triangle) := \{s \in C^r(\Omega) : s|_t \in \mathbb{P}_d, \forall t \in \triangle\},$$

where $\mathbb{P}_d$ denotes the space of polynomials of degree less than or equal d. It is possible to represent in a unique fashion such a spline by a vector of coefficients, the B-net of the spline.

We refer to [10] for the 3D case and [11] in the 2D case for additional details. The weak form of the steady state Navier-Stokes equations is: Find $\mathbf{u} \in H^1(\Omega)^2$ such that

$$\begin{aligned} \nu \int_\Omega \nabla \mathbf{u} \cdot \nabla \mathbf{v} + \sum_{j=1}^{3} \int_\Omega u_j \frac{\partial \mathbf{u}}{\partial x_j} \cdot \mathbf{v} = \int_\Omega \mathbf{f} \cdot \mathbf{v} \qquad & \forall \mathbf{v} \in V_0 \\ \text{div } \mathbf{u} = 0 \qquad & \text{in } \Omega \\ \mathbf{u} = \mathbf{g} \qquad & \text{on } \partial\Omega, \end{aligned}$$

where

$$V_0 = \{\mathbf{v} \in H^1_0(\Omega)^2, \ \text{div } \mathbf{v} = 0\}$$

and $\partial\Omega$ is the boundary of Ω. We approximate elements of V_0 by vectors $\mathbf{d}$ in $\mathbf{R}^{2N}$, (each component of $\mathbf{u}$ is approximated by a spline with coefficient vector in $\mathbf{R}^N$), which represent smooth splines, $\overline{H}\mathbf{d} = 0$, are zero on the boundary, $\overline{R}\mathbf{d} = 0$ and satisfy the divergence-free condition, $D\mathbf{d} = 0$. Let us describe these elements as vectors $\mathbf{d}$ in $\mathbf{R}^{2N}$ satisfying $L\mathbf{d} = 0$ for some matrix L. If we let $\mathbf{c}$ encode the coefficients of the approximant of the velocity field, the discrete problem is:

Find $\mathbf{c}$ in $\mathbf{R}^{2N}$ satisfying $L\mathbf{c} = \overline{\mathbf{G}}$ with $\overline{\mathbf{G}}$ encoding the boundary conditions and

$$\nu \mathbf{c}^T \overline{K} \mathbf{d} + (\overline{B}(\mathbf{c})\mathbf{c})^T \mathbf{d} = \mathbf{d}^T \overline{M} \mathbf{F}$$

for all $\mathbf{d}$ in $\mathbf{R}^{2N}$ with constraint $L\mathbf{d} = 0$. Here, $\overline{K}$ and $\overline{M}$ are the stiffness and mass matrices respectively; $(\overline{B}(\mathbf{c})\mathbf{c})^T \mathbf{d}$ encodes the nonlinear term.

Using functional arguments, it can be shown that there exists a Lagrange multiplier λ such that:

$$\begin{aligned} \nu \overline{K} \mathbf{c} + \overline{B}(\mathbf{c})\mathbf{c} + L^T \lambda &= \overline{M} \mathbf{F} \\ L\mathbf{c} &= \overline{\mathbf{G}}. \end{aligned}$$

If we put $A = \nu \overline{K} + \overline{B}(\mathbf{c})$, we showed in [9] that the previous problem is solvable and that A_s is symmetric positive definite with respect to L for ν sufficiently large. The system of nonlinear equations is linearized as follows:

Let $(\mathbf{c}^{(0)}, \lambda^{(0)})$ be the solution of the linear problem (i.e. the associated Stokes equations) and for $n = 0, 1, \ldots,$ define $(\mathbf{c}^{(n+1)}, \lambda^{(n+1)})$ as the solution of

$$\begin{aligned} \nu \overline{K} \mathbf{c}^{(n+1)} + \overline{B}(\mathbf{c}^{(n)})\mathbf{c}^{(n+1)} + L^T \lambda^{(n+1)} &= \overline{M} \mathbf{F} \\ L\mathbf{c}^{(n+1)} &= \overline{\mathbf{G}}. \end{aligned}$$

Because we are not using a basis to represent the discrete solution, the stiffness matrix $\overline{K}$ is singular. We have used the algorithm described in this paper to solve at each step the previous system of equations. The termination criterion for these steps is to require the maximum norm of $\mathbf{c}^{(n+1)} - \mathbf{c}^{(n)}$ to be less than 10^{-10} or the maximum norm of $\mathbf{c}^{(n+2)} - \mathbf{c}^{(n+1)}$. For the augmented Lagrangian iterations we use the same criteria over successive iterations. We display below the maximum number of iterations (over all steps) when this algorithm is applied to the lid driven cavity flow problem at Reynolds number 400. We used $d = 7$ and $r = 0$ in which case A has size 9216×9216 and L has size 6908×9216. The actual mesh size consists of 128 triangles obtained by refining 3 times a triangulation of the square into two triangles. Each triangle was subdivided into 4 triangles by connecting the midpoint of the edges.

r	10	10^2	10^3	10^4	10^5	10^6
ρ	10	10^2	10^3	10^4	10^5	10^6
Iterations	7	4	3	3	4	2

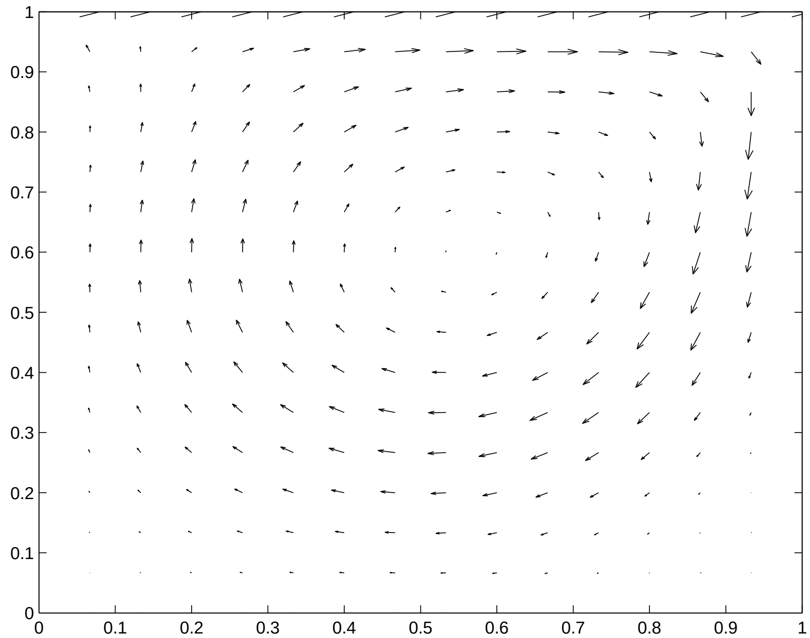

Fig. 1. 2D Cavity Flow Velocity Profile

r	10	10^2	10^3	10^4	10^5	10^6
$\rho = 2 \times r$	2×10	2×10^2	2×10^3	2×10^4	2×10^5	2×10^6
Iterations	N/A	N/A	3	3	4	4

N/A stands for not available. We did not get fast convergence for these values.

r	10^6	10^6	10^6	10^6	10^6
ρ	10^5	10^4	10^3	10^2	10
Iterations	1	4	3	4	3

These numerical results suggest that for this specific problem, the choice $r = 10^6$ and $\rho = 10^5$ is optimal. We finally display the velocity profile when $r = \rho = 10^3$ was used [cf. Fig 1].

6 Conclusion

In this paper, we have mainly given a convergence rate of a variant of the augmented lagrangian algorithm. We intend to undertake a study of the optimal choice of the parameters r and ρ when this algorithm is applied to the incompressible Navier-Stokes equations filling a gap left by earlier researchers.

References

[1] Brenner S.C. and Scott L.R., The Mathematical Theory of Finite Element Methods, Springer-Verlag, New-York, 1994.

[2] Brezzi, F. and M. Fortin, *Mixed and Hybrid Finite Element Methods*, Springer Verlag, New York, 1991.

[3] Fortin M. and Glowinski R., Augmented Lagrangian Methods: Applications to the Numerical Solution of Boundary-Value Problems, Studies in Mathematics and Its Applications 15, North-Holland, 1983.

[4] Glowinski R., LeTallec P., *Augmented Lagrangian and operator splitting methods in nonlinear mechanics*, Society for Industrial and Applied Mathematics (SIAM), 1989.

[5] Gunzburger M., *Finite Element Methods for Viscous Incompressible Flows*, Academic Press, 1989.

[6] Bramble, J. H., J. E. Pasciak, and A. T. Vassilev, Uzawa type algorithms for nonsymmetric saddle point problems, Math. Comp., 69 (1999), pp. 667–689.

[7] R. Bramley, X. Wang, D. Pelletier, Orthogonalization based iterative methods for generalized stokes problems, in Solution techniques for large scale CFD problems, W.G. Habashi, ed., Centre de Recherce en Calcul Applique, (1995), pp. 217-238.

[8] Segal, A., 1979, On the numerical solution of the Stokes equations using the finite element method, Computer Methods in Applied Mechanics and Engineering, **19** (1979), pp. 165–185.

[9] Awanou, G.M., Energy Methods in 3D Spline Approximations of the Navier-Stokes Equations, Ph.d. Dissertation, 2003, University of Georgia, Athens, Ga.

[10] Awanou, G.M. and Lai, M.J. Trivariate Spline approximations of 3D Navier-Stokes equations, Submitted, 2002.

[11] Lai, M.J. and Wenston, P. Bivariate splines for fluid flows, Submitted 2003.