NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS

Numer. Linear Algebra Appl. 2000; **00**:1–6 *Prepared using* nlaauth.cls *[Version: 2002/09/18 v1.02]*

Optimal multigrid preconditioned semi–monotonic augmented Lagrangians applied to the Stokes problem

D. Lukáš[1]*and Z. Dostál[1]

[1] *Department of Applied Mathematics, VŠB–Technical University, 17. listopadu 15, 70833 Ostrava–Poruba, Czech Republic*

SUMMARY

We propose an optimal computational complexity algorithm for the solution of quadratic programming problems with equality constraints arising from partial differential equations. The algorithm combines a variant of the semi–monotonic augmented Lagrangian (SMALE) method with adaptive precision control and a multigrid preconditioning for the Hessian of the cost function and for the inner product on the space of Lagrange variables. In our approach there is no need for preconditioning of the constraints. The optimality of the algorithm is theoretically proven and confirmed by numerical experiments for the 2–dimensional Stokes problem. Copyright © 2000 John Wiley & Sons, Ltd.

KEY WORDS: multigrid, augmented Lagrangians, Stokes problem

*Correspondence to: dalibor.lukas@vsb.cz

Received 3 May 2004

1. Introduction

Let V, Q be Hilbert spaces. Denote their respective inner products by $(.,.)_V$, $(.,.)_Q$, their dual spaces by V', Q' and the duality pairings by $\langle .,. \rangle_V$, $\langle .,. \rangle_Q$. Let I_V, I_Q denote the inner product (Riesz isomorphism) operators on V and Q, respectively, i.e. $\langle I_V .,. \rangle_V := (.,.)_V$ and $\langle I_Q .,. \rangle_Q := (.,.)_Q$. Let further $A : V \to V'$ be a symmetric positive definite (spd) linear bounded operator, let $B : V \to Q'$ be another linear bounded operator, let $f \in V'$, and let $g \in \mathrm{Range}(B)$. Denote by $B^T : Q \to V'$ the adjoint operator of B. We consider the following equality constrained quadratic programming problem:

$$\min_{u \in V} h(u) \quad \text{s.t.} \quad Bu = g \text{ on } Q', \tag{1}$$

where $h(u) := (1/2)\langle Au, u \rangle_V - \langle f, u \rangle_V$. By introducing a Lagrange multiplier $p \in Q$, the problem (1) is equivalent to the saddle–point problem

$$\min_{u \in V} \max_{p \in Q} \left\{ h(u) + \langle Bu - g, p \rangle_Q \right\}, \tag{2}$$

which is also equivalent to the mixed linear system

$$\begin{pmatrix} A & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} u^* \\ p^* \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix} \text{ on } V' \times Q'. \tag{3}$$

The latter is known as the Karush–Kuhn–Tucker (KKT) system for (1). It is well–known, cf. [16, 17], that there exists a unique primal solution $u^* \in V$. Moreover, if $\mathrm{Kernel}(B^T) = \{0\}$, then there is also a unique Lagrange multiplier $p^* \in Q$.

This type of problems arises in mixed variational formulations of the Stokes problem [23], in elliptic partial differential equations (pde) with periodic boundary conditions [7], in applications of the domain decomposition methods to parallel solution of three–dimensional elasticity problems [20], in modeling of laminated composites [19], in development of scalable algorithms for parallel solution of variational inequalities [13], or in large–scale optimization [15]. The results involving the solution of (1) are also useful in solution of problems with bound and inequality constraints [9].

Basically, there are two approaches to solve (1), cf. [3] for an overview. The first class of methods, called also black–box or nested methods, is based on an elimination of u^* or p^* from the mixed system (3). It involves the Schur complement (range space) methods, which are based on the solution of the system with the operator $S := BA^{-1}B^T$ in terms of the dual variable p, and the null space method, which attempts to exploit a knowledge of Kernel(B) and solves the auxiliary system with the operator Z^TAZ, where $BZ = 0$. Note also that this class involves such pde–constrained optimization methods where the operator B is nonsingular and the pde–equality constraint is eliminated for each design u.

The second class of methods, called also simultaneous, one–shot, primal–dual or all–at–once methods, solves for both u^* and p^* simultaneously. A classical prototype is the Uzawa algorithm [2], which relies on exact solver for the operator A and Jacobi–like iteration with the Schur complement S, which may be interpreted as the Hessian of the dual function. Here we are especially interested in the inexact Uzawa methods, replacing the action of A^{-1} by an approximation $\widehat{A}^{-1}$, whose convergence typically depends on the preconditioning of both A and $\widehat{S} := B\widehat{A}^{-1}B^T$; see [5] for the original result and [25] for a generalization. Recently, in [21, 22] the algorithm has been modified and proven to converge efficiently even with a relatively poor preconditioning of $\widehat{S}$. Another approach, which is based on multigrid methods as an outer iteration combined with appropriate smoothers as a sort of inner iteration, was proposed in [24]. In [23], a multigrid method with an additive Schwarz smoother is proposed and analyzed with an application to the Stokes problem. It is proven that the method is equivalent to the symmetric Uzawa algorithm. Another possibility is the choice of a multiplicative smoother, which is numerically shown to work optimally, however, it has not been proven theoretically yet.

Our development is based on the well–known augmented Lagrangian algorithm, which may be interpreted as the Uzawa algorithm applied to the penalized problem. It generates approximations of the Lagrange multipliers in the outer loop while the following unconstrained auxiliary subproblems are solved in the inner loop of the k-th iteration:

$$\min_{u\in V}\left\{h(u)+\langle Bu-g,p^{(k)}\rangle_Q+\frac{\rho^{(k)}}{2}\|Bu-g\|^2_{Q'}\right\}. \tag{4}$$

The latter is equivalent to solution to the following KKT linear system:

$$\left[A+\rho^{(k)}B^TI_Q^{-1}B\right]u=f-B^Tp^{(k)}+\rho^{(k)}B^TI_Q^{-1}g. \tag{5}$$

We use a variant of the augmented Lagrangian method proposed by Dostál, see [10, 11], which controls the precision of the solution of the auxiliary unconstrained problem in the inner loop by a norm of the feasibility error, and relates the update of the penalization parameter to the increase of the Lagrangian so that it is possible to give upper bounds on both the $\{\rho^{(k)}\}$ and the number of the outer iterations which is independent of the form of the constraints and depends on the ellipticity constant of A only. Therefore, a proper preconditioning of A results in an optimal algorithm. This algorithm also turns out to be very robust with respect to additional algorithmic parameters and it does not require independent constraints, i.e. $\mathrm{Kernel}(B^T)=\{0\}$ is not necessary.

The remainder of the paper is organized as follows: In Section 2, we describe the semi–monotonic augmented Lagrangian method without preconditioning and we present the analysis. In Section 3 we introduce efficient preconditioners and apply the previous analysis to prove the optimal, i.e. linear assymptotical computational complexity. Finally. in Section 4, we will present the performance of the algorithm for an application to the 2–dimensional Stokes problem.

2. Semi–monotonic augmented Lagrangian method

Let us denote the augmented Lagrangian by

$$L(u, p, \rho) := h(u) + \langle Bu - g, p\rangle_Q + \frac{\rho}{2}\|Bu - g\|^2_{Q'}$$

and its Fréchet derivative by

$$F(u, p, \rho) := \nabla_u L(u, p, \rho) = Au - f + B^T p + \rho B^T I_Q^{-1}(Bu - g).$$

Note that evaluations of the dual norms are due to the Riesz theorem as follows:

$$\|\varphi\|_{V'} = \sqrt{\langle \varphi, I_V^{-1}\varphi\rangle_V}, \quad \|\xi\|_{Q'} = \sqrt{\langle \xi, I_Q^{-1}\xi\rangle_Q}. \tag{6}$$

Algorithm 1, see also [10], is a modification of the classical augmented Lagrangian method for the solution of strictly convex quadratic programing problems with equality constraints that enables adaptive precision control of the solution of the auxiliary subproblems (4).

2.1. The analysis

This section is a straightforward extension of the analysis presented in [10], which was done in $\mathbb{R}^n$. The difference here is that we incorporate into the proofs the proper scaling (6) of the dual norms.

Lemma 2.1. *Let $\nu > 0$, $p \in Q$, $\rho \geq 0$ be given and let $\{v^{(k)}\}$ denote any sequence that converges to the unique solution v^* of the problem*

$$\min_{v\in V} \widetilde{L}(v, p, \rho). \tag{7}$$

Algorithm 1 Semi–monotonic augmented Lagrangians with adaptive precision control

Given $\eta > 0$, $\beta > 1$, $\nu > 0$, $\rho^{(0)} > 0$, $p^{(0)} \in Q$, precision $\varepsilon > 0$, feasibility precision $\varepsilon_{\text{feas}} > 0$

for $k := 0, 1, 2, \dots$ **do**

  Find $u^{(k)} : \|F(u^{(k)}, p^{(k)}, \rho^{(k)})\|_{V'} \leq \min\left\{\nu\|Bu^{(k)} - g\|_{Q'}, \eta\right\}$

  if $\|F(u^{(k)}, p^{(k)}, \rho^{(k)})\|_{V'} \leq \varepsilon$ and $\|Bu^{(k)} - g\|_{Q'} \leq \varepsilon_{\text{feas}}$ **then**

    break

  end if

  $p^{(k+1)} := p^{(k)} + \rho^{(k)} I_Q^{-1}(Bu^{(k)} - g)$

  if $k > 0$ and $L(u^{(k)}, p^{(k)}, \rho^{(k)}) < L(u^{(k-1)}, p^{(k-1)}, \rho^{(k-1)}) + \frac{\rho^{(k)}}{2}\|Bu^{(k)} - g\|^2_{Q'}$ **then**

    $\rho^{(k+1)} := \beta\rho^{(k)}$

  else

    $\rho^{(k+1)} := \rho^{(k)}$

  end if

end for

$u^{(k)}$, $p^{(k)}$ is the solution.

Then $\{v^{(k)}\}$ *either converges to the solution* u^* *of the problem* (1) *or there is an index* k *such that*

$$\|F(v^{(k)}, p, \rho)\|_{V'} \leq \min\{\nu\|Bv^{(k)} - g\|_{Q'}, \eta\}. \tag{8}$$

Proof We strictly follow the proof of Lemma 2.2 in [10].

The assumptions $v^{(k)} \to v^*$ and v^* solves (7) imply $\|F(v^{(k)}, p, \rho)\|_{V'} \to 0$, therefore, there is an index k_0 such that $\|F(v^{(k)}, p, \rho)\|_{V'} \leq \eta$ for $k \geq k_0$. Hence, if (8) does not hold for any k, then $\|Bv^{(k)} - g\|_{Q'} \to 0$ and we must have $Bv^* = g$, which together with $F(v^*, p, \rho) = 0$ are the sufficient conditions for v^* to be the solution of (1), thus $v^* = u^*$. ■

Assumption 2.1. Let λ denote the ellipticity constant of A, i.e.

$$\forall v \in V : \lambda \|v\|_V^2 \leq \langle Av, v \rangle_V .$$

The following lemma will be the key ingredient in the proof of convergence of Algorithm 1.

Lemma 2.2. *Let $u, v \in V$, $p \in Q$, $\rho > 0$, $\eta > 0$, $\nu > 0$ and let $q = p + \rho I_Q^{-1}(Bu - g)$.*

(i) If

$$\|F(u,p,\rho)\|_{V'} \leq \nu \|Bu - g\|_{Q'}, \tag{9}$$

then

$$L(v,q,\rho) \geq L(u,p,\rho) + \frac{1}{2}\left(\rho - \frac{\nu^2}{\lambda}\right)\|Bu - g\|_{Q'}^2 + \frac{\rho}{2}\|Bv - g\|_{Q'}^2. \tag{10}$$

(ii) If

$$\|F(u,p,\rho)\|_{V'} \leq \eta, \tag{11}$$

then

$$L(v,q,\rho) \geq L(u,p,\rho) + \frac{\rho}{2}\|Bu - g\|_{Q'}^2 + \frac{\rho}{2}\|Bv - g\|_{Q'}^2 - \frac{\eta^2}{2\lambda}. \tag{12}$$

(iii) If (11) *holds and $w_0 \in V$ be such that $Bw_0 = g$ ($w_0 := u^*$ is the best choice), then*

$$L(u,p,\rho) \leq h(w_0) + \frac{\eta^2}{2\lambda}. \tag{13}$$

Proof We will strictly follow the proof of Lemma 3.1 in [10].

Denote $\delta := v - u$, $A_\rho := A + \rho B^T I_Q^{-1} B$. Using $L(u,q,\rho) = L(u,p,\rho) + \rho\|Bu-g\|^2_{Q'}$, and $F(u,q,\rho) = F(u,p,\rho) + \rho B^T I_Q^{-1}(Bu-g)$, we get

$$
\begin{aligned}
L(v,q,\rho) = L(u+\delta,q,\rho) &= L(u,q,\rho) + \langle F(u,q,\rho),\delta\rangle_V + \frac{1}{2}\langle A_\rho\delta,\delta\rangle_V \\
&= L(u,p,\rho) + \rho\|Bu-g\|^2_{Q'} + \langle F(u,p,\rho),\delta\rangle_V + \rho\langle B^T I_Q^{-1}(Bu-g),\delta\rangle_V + \frac{1}{2}\langle A_\rho\delta,\delta\rangle_V \\
&\geq L(u,p,\rho) + \rho\|Bu-g\|^2_{Q'} + \langle F(u,p,\rho),\delta\rangle_V + \rho\langle B^T I_Q^{-1}(Bu-g),\delta\rangle_V \\
&\quad + \frac{\lambda}{2}\|\delta\|^2_V + \frac{\rho}{2}\|B\delta\|^2_{Q'}.
\end{aligned}
$$

Noticing that

$$
\frac{\rho}{2}\|Bv-g\|^2_{Q'} = \rho\langle B^T I_Q^{-1}(Bu-g),\delta\rangle_V + \frac{\rho}{2}\|Bu-g\|^2_{Q'} + \frac{\rho}{2}\|B\delta\|^2_{Q'},
$$

we get

$$
L(v,q,\rho) \geq L(u,p,\rho) + \langle F(u,p,\rho),\delta\rangle_V + \frac{\lambda}{2}\|\delta\|^2_V + \frac{\rho}{2}\|Bv-g\|^2_{Q'} + \frac{\rho}{2}\|Bu-g\|^2_{Q'}. \tag{14}
$$

Let us prove (i). From (9) and the Cauchy–Schwarz inequality, we get

$$
\langle F(u,p,\rho),\delta\rangle_V \geq -\nu\|Bu-g\|_{Q'}\|\delta\|_V.
$$

Then from (14), simple manipulations yield

$$
\begin{aligned}
L(v,q,\rho) &\geq L(u,p,\rho) + \left(\frac{\lambda}{2}\|\delta\|^2_V - \nu\|Bu-g\|_{Q'}\|\delta\|_V + \frac{\nu^2}{2\lambda}\|Bu-g\|^2_{Q'}\right) \\
&\quad - \frac{\nu^2}{2\lambda}\|Bu-g\|^2_{Q'} + \frac{\rho}{2}\|Bv-g\|^2_{Q'} + \frac{\rho}{2}\|Bu-g\|^2_{Q'} \\
&\geq L(u,p,\rho) + \frac{1}{2}\left(\rho - \frac{\nu^2}{\lambda}\right)\|Bu-g\|^2_{Q'} + \frac{\rho}{2}\|Bv-g\|^2_{Q'}.
\end{aligned}
$$

Let us prove (ii). From (11) and the Cauchy–Schwarz inequality, we get

$$
\langle F(u,p,\rho),\delta\rangle_V \geq -\eta\|\delta\|_V. \tag{15}
$$

Then from (14), simple manipulations yield

$$
\begin{aligned}
L(v,q,\rho) &\geq L(u,p,\rho) + \left(\frac{\lambda}{2}\|\delta\|^2_V - \eta\|\delta\|_V + \frac{\eta^2}{2\lambda}\right) - \frac{\eta^2}{2\lambda} + \frac{\rho}{2}\|Bv-g\|^2_{Q'} + \frac{\rho}{2}\|Bu-g\|^2_{Q'} \\
&\geq L(u,p,\rho) + \frac{\rho}{2}\|Bv-g\|^2_{Q'} + \frac{\rho}{2}\|Bu-g\|^2_{Q'} - \frac{\eta^2}{2\lambda}.
\end{aligned}
$$

Finally, let us prove (iii). Let $\widetilde{u}^*$ solves the following auxiliary problem

$$\min_{u\in V} \widetilde{L}(u,p,\rho), \tag{16}$$

let $Bw_0 = g$ and denote $\widetilde{\delta} := \widetilde{u}^* - u$. If (11) holds, then using (16), the quadratic Taylor expansion, (15) and Assumption 2.1 yield

$$\begin{aligned} 0 \geq L(\widetilde{u}^*,p,\rho) - L(u,p,\rho) &= \langle F(u,p,\rho),\widetilde{\delta}\rangle_V + \frac{1}{2}\langle A_\rho\widetilde{\delta},\widetilde{\delta}\rangle_V \geq -\eta\|\widetilde{\delta}\|_V + \frac{\lambda}{2}\|\widetilde{\delta}\|_V^2 + \frac{\rho}{2}\|B\widetilde{\delta}\|_{Q'}^2 \\ &= \left(\frac{\eta^2}{2\lambda} - \eta\|\widetilde{\delta}\|_V + \frac{\lambda}{2}\|\widetilde{\delta}\|_V^2\right) - \frac{\eta^2}{2\lambda} + \frac{\rho}{2}\|B\widetilde{\delta}\|_{Q'}^2 \geq -\frac{\eta^2}{2\lambda}. \end{aligned}$$

As $L(\widetilde{u}^*,p,\rho) \leq L(w_0,p,\rho) = h(w_0)$, from the latter inequality we conclude that

$$L(u,p,\rho) \leq h(w_0) + (L(u,p,\rho) - L(\widetilde{u}^*,p,\rho)) \leq h(w_0) + \frac{\eta^2}{2\lambda}.$$

■

Corollary 2.1. *Let $\{u^{(k)}\}$, $\{p^{(k)}\}$ and $\{\rho^{(k)}\}$ be generated by Algorithm 1 with $\eta > 0$, $\beta > 1$, $\nu > 0$, $\rho^{(0)} > 0$ and $p^{(0)} \in Q$.*

(i) If $k > 0$ and $\rho^{(k-1)} \geq \frac{\nu^2}{\lambda}$, then

$$L(u^{(k)},p^{(k)},\rho^{(k)}) \geq L(u^{(k-1)},p^{(k-1)},\rho^{(k-1)}) + \frac{\rho^{(k)}}{2}\|Bu^{(k)} - g\|_{Q'}^2. \tag{17}$$

(ii) For any $k > 0$

$$L(u^{(k)},p^{(k)},\rho^{(k)}) \geq L(u^{(k-1)},p^{(k-1)},\rho^{(k-1)}) + \frac{\rho^{(k-1)}}{2}\|Bu^{(k-1)} - g\|_{Q'}^2 + \frac{\rho^{(k)}}{2}\|Bu^{(k)} - g\|_{Q'}^2 - \frac{\eta^2}{2\lambda}. \tag{18}$$

(iii) For any $k > 0$ and $w_0 \in V$ such that $Bw_0 = g$

$$L(u^{(k)},p^{(k)},\rho^{(k)}) \leq h(w_0) + \frac{\eta^2}{2\lambda}. \tag{19}$$

Proof Let us substitute $u := u^{(k-1)}$, $v := u^{(k)}$, $p := p^{(k-1)}$, $q := p^{(k)}$ and $\rho := \rho^{(k-1)}$ into Lemma 2.2. Notice also that

$$L(u^{(k)}, p^{(k)}, \rho^{(k)}) = L(u^{(k)}, p^{(k)}, \rho^{(k-1)}) + \frac{\rho^{(k)} - \rho^{(k-1)}}{2} \|Bu^{(k)} - g\|^2_{Q'}.$$

Then the proof of (17) and (18) follows directly from (10) and (12), respectively. The statement (19) is a straightforward consequence of (13).

For the proof, see also the proof of Lemma 4.1 in [10]. ■

Theorem 2.1. *Let $\{u^{(k)}\}$, $\{p^{(k)}\}$ and $\{\rho^{(k)}\}$ be generated by Algorithm 1 with $\eta > 0$, $\beta > 1$, $\nu > 0$, $\rho^{(0)} > 0$ and $p^{(0)} \in Q$. Let $s \geq 0$ be the smallest integer such that $\beta^s \rho^{(0)} \geq \frac{\nu^2}{\lambda}$.*

(i) The sequence $\{\rho^{(k)}\}$ is then bounded and

$$\rho^{(k)} \leq \beta^s \rho^{(0)}. \tag{20}$$

(ii) Let $u^* \in V$ solves (1), then

$$\sum_{k=1}^{\infty} \frac{\rho^{(k)}}{2} \|Bu^{(k)} - g\|^2_{Q'} \leq h(u^*) - L(u^{(0)}, p^{(0)}, \rho^{(0)}) + (1+s)\frac{\eta^2}{2\lambda}. \tag{21}$$

Proof The statement (i) follows directly from Corollary 2.1(i), from the definition of the update of $\rho^{(k)}$ in Algorithm 1 and from (20).

Let us prove (ii). Let $\mathcal{I}$ denote the set of indices k_i, $i = 1, 2, \ldots, s$ for which the update $\rho^{(k_i+1)} := \beta\rho^{(k_i)}$ realizes. For $k > 0$ either $k+1 \notin \mathcal{I}$ and by (17):

$$\frac{\rho^{(k)}}{2} \|Bu^{(k)} - g\|^2_{Q'} \leq L(u^{(k)}, p^{(k)}, \rho^{(k)}) - L(u^{(k-1)}, p^{(k-1)}, \rho^{(k-1)})$$

or $k+1 \in \mathcal{I}$ and by (18):

$$\begin{aligned}\frac{\rho^{(k)}}{2} \|Bu^{(k)} - g\|^2_{Q'} &\leq \frac{\rho^{(k-1)}}{2} \|Bu^{(k-1)} - g\|^2_{Q'} + \frac{\rho^{(k)}}{2} \|Bu^{(k)} - g\|^2_{Q'} \\ &\leq L(u^{(k)}, p^{(k)}, \rho^{(k)}) - L(u^{(k-1)}, p^{(k-1)}, \rho^{(k-1)}) + \frac{\eta^2}{2\lambda}.\end{aligned}$$

Summing all the terms up and using (19) concludes the proof

$$\sum_{k=1}^{n} \frac{\rho^{(k)}}{2} \|Bu^{(k)} - g\|_{Q'}^2 \leq L(u^{(n)}, p^{(n)}, \rho^{(n)}) - L(u^{(0)}, p^{(0)}, \rho^{(0)}) + s\frac{\eta^2}{2\lambda}$$
$$\leq h(u^*) - L(u^{(0)}, p^{(0)}, \rho^{(0)}) + (1+s)\frac{\eta^2}{2\lambda}.$$

■

Corollary 2.2. *Let* $\{u^{(k)}\}$, $\{p^{(k)}\}$ *and* $\{\rho^{(k)}\}$ *be generated by Algorithm 1 with* $\eta > 0$, $\beta > 1$, $\nu > 0$, $\rho^{(0)} > 0$ *and* $p^{(0)} := 0 \in Q$. *Let* $s \geq 0$ *be the smallest integer such that* $\beta^s \rho^{(0)} \geq \frac{\nu^2}{\lambda}$ *and let*

$$C := \frac{2}{\rho^{(0)}} \left[h(u^*) + L(u^{(0)}, p^{(0)}, \rho^{(0)}) + (1+s)\frac{\eta^2}{2\lambda} \right]. \tag{22}$$

Then for each $\varepsilon > 0$ *there is an index* $k \leq (C/\varepsilon^2) + 1$ *such that*

$$\nu^{-1} \|F(u^{(k)}, p^{(k)}, \rho^{(k)}\|_{V'} \leq \|Bu^{(k)} - g\|_{Q'} \leq \varepsilon. \tag{23}$$

Proof We will follow the proof of Theorem 5.1 in [10].

Using (21), note that for any index k

$$\frac{\rho^{(0)}k}{2} \min\{\|Bu^{(i)} - g\|_{Q'}^2 \,:\, i = 1, \ldots, k\} \leq \sum_{i=1}^{k} \frac{\rho^{(i)}}{2} \|Bu^{(i)} - g\|_{Q'}^2 \leq \sum_{i=1}^{\infty} \frac{\rho^{(i)}}{2} \|Bu^{(i)} - g\|_{Q'}^2$$
$$\leq h(u^*) - L(u^{(0)}, p^{(0)}, \rho^{(0)}) + (1+s)\frac{\eta^2}{2\lambda} = \frac{\rho^{(0)}}{2} C. \tag{24}$$

Taking for l the smallest integer that satisfies $l \geq C/\varepsilon^2$ and denoting by k the index that minimizes $\{\|Bu^{(i)} - g\|_{Q'}^2 \,:\, i = 1, \ldots, l\}$, it follows from (24) that

$$\|Bu^{(k)} - g\|_{Q'}^2 = \min\{\|Bu^{(i)} - g\|_{Q'}^2 \,:\, i = 1, \ldots, l\} \leq \frac{C}{l} \leq \varepsilon^2.$$

The inequality

$$\nu^{-1} \|F(u^{(k)}, p^{(k)}, \rho^{(k)}\|_{V'} \leq \|Bu^{(k)} - g\|_{Q'}$$

is given by Algorithm 1. ■

3. Multigrid preconditioning

Now, let us consider a discretization of the mixed system (3). We introduce spd preconditioners $\widehat{A}^{-1}$ and $\widehat{I_Q}^{-1}$ to the matrices A and I_Q, respectively, and we consider their symmetric factorization $\widehat{A}^{-1} = \widehat{A_{\frac{1}{2}}}^{-1}\widehat{A_{\frac{1}{2}}}^{-T}$ and $\widehat{I_Q}^{-1} = \widehat{I_{Q\frac{1}{2}}}^{-1}\widehat{I_{Q\frac{1}{2}}}^{-T}$. Then, the preconditioned discretized mixed system reads as follows:

$$\begin{pmatrix} \widehat{A_{\frac{1}{2}}}^{-1} A \widehat{A_{\frac{1}{2}}}^{-T} & \widehat{A}_{\frac{1}{2}}^{-1} B^T \widehat{I_{Q\frac{1}{2}}}^{-T} \\ \widehat{I_{Q\frac{1}{2}}}^{-1} B \widehat{A_{\frac{1}{2}}}^{-T} & 0 \end{pmatrix} \begin{pmatrix} \widehat{u}^* \\ \widehat{p}^* \end{pmatrix} = \begin{pmatrix} \widehat{A_{\frac{1}{2}}}^{-1} f \\ \widehat{I_{Q\frac{1}{2}}}^{-1} g \end{pmatrix}, \quad \begin{array}{l} u^* := \widehat{A_{\frac{1}{2}}}^{-T} \widehat{u}^*, \\ p^* := \widehat{I_{Q\frac{1}{2}}}^{-T} \widehat{p}^*. \end{array} \tag{25}$$

Now we are working in the Euclidean spaces $V := \mathbb{R}^n$ and $Q := \mathbb{R}^m$. The preconditioned augmented Lagrangian is easier to express in the untransformed variables

$$\widetilde{L}(u, p, \rho) := \widehat{L}(\widehat{u}, \widehat{p}, \rho) = h(u) + (Bu - g)^T p + \frac{\rho}{2}(Bu - g)^T \widehat{I_Q}^{-1}(Bu - g),$$

as well as the related Fréchet derivative

$$\widehat{F}(\widehat{u}, \widehat{p}, \rho) := \nabla_{\widehat{u}} \widehat{L}(\widehat{u}, \widehat{p}, \rho) = \widehat{A_{\frac{1}{2}}}^{-1} \widetilde{F}(u, p, \rho),$$

where $\widetilde{F}(u, p, \rho) := Au - f + B^T p + \rho B^T \widehat{I_Q}^{-1}(Bu - g)$. The evaluations of the dual norms becomes as follows:

$$\|\widehat{F}(\widehat{u}, \widehat{p}, \rho)\|_{\mathbb{R}^n} = \|\widetilde{F}(u, p, \rho)\|_{\widehat{V}'} := \sqrt{\widetilde{F}(u, p, \rho)^T \widehat{A}^{-1} \widetilde{F}(u, p, \rho)},$$

$$\|\widehat{I_{Q\frac{1}{2}}}^{-1} B \widehat{A_{\frac{1}{2}}}^{-T} \widehat{u} - \widehat{I_{Q\frac{1}{2}}}^{-1} g\|_{\mathbb{R}^m} = \|Bu - g\|_{\widehat{Q}'} := \sqrt{(Bu - g)^T \widehat{I_Q}^{-1}(Bu - g)}.$$

The preconditioned version of Algorithm 1 is depicted in Algorithm 2.

It is well-known [4, 18] that the choice of multigrid preconditioning is optimal, i.e. it guarantees uniform bounds of the spectra $\sigma(\widehat{A}^{-1}A) = \sigma(\widehat{A_{\frac{1}{2}}}^{-1} A \widehat{A_{\frac{1}{2}}}^{-T})$ and $\sigma(\widehat{I_Q}^{-1} I_Q) = \sigma(\widehat{I_{Q\frac{1}{2}}}^{-1} I_Q \widehat{I_{Q\frac{1}{2}}}^{-T})$.

Algorithm 2 Preconditioned semi–monotonic augmented Lagrangians

Given $\eta > 0$, $\beta > 1$, $\nu > 0$, $\rho^{(0)} > 0$, $p^{(0)} \in Q$, precision $\varepsilon > 0$, feasibility precision $\varepsilon_{\text{feas}} > 0$

for $k := 0, 1, 2, \ldots$ **do**

Find $u^{(k)} : \|\widetilde{F}(u^{(k)}, p^{(k)}, \rho^{(k)})\|_{\widehat{V}'} \leq \min\left\{\nu\|Bu^{(k)} - g\|_{\widehat{Q}'}, \eta\right\}$

if $\|\widetilde{F}(u^{(k)}, p^{(k)}, \rho^{(k)})\|_{\widehat{V}'} \leq \varepsilon$ and $\|Bu^{(k)} - g\|_{\widehat{Q}'} \leq \varepsilon_{\text{feas}}$ **then**

break

end if

$p^{(k+1)} := p^{(k)} + \rho^{(k)}\widehat{I_Q}^{-1}(Bu^{(k)} - g)$

if $k > 0$ and $\widetilde{L}(u^{(k)}, p^{(k)}, \rho^{(k)}) < \widetilde{L}(u^{(k-1)}, p^{(k-1)}, \rho^{(k-1)}) + \frac{\rho^{(k)}}{2}\|Bu^{(k)} - g\|^2_{\widehat{Q}'}$ **then**

$\rho^{(k+1)} := \beta\rho^{(k)}$

else

$\rho^{(k+1)} := \rho^{(k)}$

end if

end for

$u^{(k)}$, $p^{(k)}$ is the solution.

Assumption 3.1. Let $\widehat{A}$ and $\widehat{I_Q}$ be effective spd preconditioners to A and I_Q, respectively, i.e. there exists $\widehat{\gamma} \in [0, 1)$ independent of the discretization such that

$$\forall v \in V : (1 - \widehat{\gamma})\langle \widehat{A}v, v\rangle_V \leq \langle Av, v\rangle_V \leq \|v\|^2_{\widehat{V}} := \langle \widehat{A}v, v\rangle_V$$

and there exists $\widehat{\delta} \in [0, 1)$ independent of the discretization such that

$$\forall q \in Q : (1 - \widehat{\delta})\langle \widehat{I_Q}q, q\rangle_Q \leq \|q\|^2_Q \leq \|q\|^2_{\widehat{Q}} := \langle \widehat{I_Q}q, q\rangle_Q.$$

Under Assumption 3.1 we have

$$\sigma(\widehat{A_{\frac{1}{2}}}^{-1} A \widehat{A_{\frac{1}{2}}}^{-T}) \subset [1 - \widehat{\gamma}, 1] \text{ and } \sigma(\widehat{I_{Q\frac{1}{2}}}^{-1} I_Q \widehat{I_{Q\frac{1}{2}}}^{-T}) \subset [1 - \widehat{\delta}, 1].$$

Thus, we have proven that the number of outer iterations of Algorithm 2 is uniformly bounded

$$k \leq \frac{2}{\rho^{(0)}} \left[h(u^*) + L(u^{(0)}, p^{(0)}, \rho^{(0)}) + (1+s)\frac{\eta^2}{2(1-\widehat{\gamma})} \right] / \min{}^2\{\nu\varepsilon, \varepsilon_{\text{feas}}\} + 1, \tag{26}$$

which guarantees optimality of the outer loop. At the same time, the penalty parameter is bounded by

$$\rho^{(k)} \leq \rho_{\max} := \max\left\{ \frac{\beta\nu^2}{1-\widehat{\gamma}}, \rho^{(0)} \right\}.$$

We conclude that all the operators are spectrally equivalent, i.e. for all indices k and $\forall v \in V$:

$$\langle Av, v\rangle_V \leq \langle (A + \rho^{(k)} B^T \widehat{I_Q}^{-1} B)v, v\rangle_V \leq (1 + \rho_{\max}\|B\|^2_{V\to Q'})\langle Av, v\rangle_V,$$

where $\|B\|_{V\to Q'} := \sup\limits_{v,q\neq 0} \frac{|\langle Bv,q\rangle_Q|}{\|v\|_V\|q\|_Q}$. Therefore, the spd operator

$$\widehat{A_\rho} := (1 + \rho_{\max}\|B\|^2_{V\to Q'})\,\widehat{A}$$

is an optimal preconditioner for the inner loop, which effectively works throughout all the outer iterations, i.e. there exists $\widehat{\gamma_\rho} \in [0,1)$ independent of the discretization so that for all indices k and $\forall v \in V$:

$$(1-\widehat{\gamma_\rho})\langle \widehat{A_\rho} v, v\rangle_V \leq \langle (A + \rho^{(k)} B^T \widehat{I_Q}^{-1} B)v, v\rangle_V \leq \langle \widehat{A_\rho} v, v\rangle_V \tag{27}$$

with

$$\widehat{\gamma_\rho} := 1 - \frac{1-\widehat{\gamma}}{1+\rho_{\max}\|B\|^2_{V\to Q'}}.$$

4. Application to the Stokes Problem

To introduce the Stokes problem, we will follow the presentation in [23]. Let $\Omega \subset \mathbb{R}^2$ be a bounded polygonal domain and let $f \in \left[L^2(\Omega)\right]^2$. The Stokes problem with homogeneous

Dirichlet boundary condition is, for $i \in \{1, 2\}$, given by

$$-\triangle u_i + \frac{\partial}{\partial x_i} p = f_i \qquad \text{in } \Omega$$

$$\operatorname{div} u = 0 \qquad \text{in } \Omega$$

$$u_i = 0 \quad \text{on } \partial\Omega$$

The weak formulation leads to a saddle–point problem which fits to our notation as follows:

$$V := \left[H_0^1(\Omega)\right]^2, \quad Q := L^2(\Omega), \quad \langle Au, v\rangle_V := \int_\Omega \sum_{i=1}^{2} \nabla u_i \cdot \nabla v_i \, dx,$$

$$\langle f, v\rangle_V := \int_\Omega f \cdot v \, dx, \quad \langle Bu, q\rangle_Q := \int_\Omega \operatorname{div} u \, q \, dx, \quad g := 0.$$

Thus, A is the tensor–product Laplacian and I_Q is the L^2–inner product.

We discretize the domain Ω using a nested sequence of triangulations $(\mathcal{T}_l)$ so that $\mathcal{T}_{l+1}$ is obtained by connecting the midpoints of edges of the triangles in $\mathcal{T}_l$. We employ Crouzeix–Raviart elements, which are determined by the following nonconforming finite element spaces:

$$V_l := \Big\{ v \in \left[L^2(\Omega)\right]^2 : v|_T \text{ is linear for all } T \in \mathcal{T}_l,$$
$$v \text{ is continuous at the midpoints of interelement boundaries}$$
$$\text{and } v = 0 \text{ along } \partial\Omega \Big\},$$

$$Q_l := \left\{ q \in L^2(\Omega) : q|_T \text{ is constant for all } T \in \mathcal{T}_l \right\}.$$

Since I_{Q_l} is a diagonal matrix, the preconditioner is just the inverse matrix $\widehat{I_{Q_l}}^{-1} := I_{Q_l}^{-1}$. The construction of a multigrid preconditioner for $A_l + B_l^T I_{Q_l}^{-1} B_l$ follows from [6, 23]. In our case, the inter–grid transfer operator $I_{l-1}^l : V_{l-1} \times Q_{l-1} \to V_l \times Q_l$ is given by

$$I_{l-1}^l(v, z) := (J_{l-1}^l v, z)$$

with

$$J_{l-1}^l v(m_e) := \begin{cases} v(m_e) & \text{if } m_e \in \operatorname{int}(T) \text{ for some } T \in \mathcal{T}_{l-1} \\ \frac{1}{2}\left[v|_{T_1} + v|_{T_2}\right] & \text{if } e \subset T_1 \cap T_2 \text{ for some } T_1, T_2 \in \mathcal{T}_{l-1} \end{cases}$$

at midpoints m_e of internal edges e in $\mathcal{T}_l$.

Concerning the smoothers, we use a point additive and a block Gauss–Seidel smoother. The point one is constructed in the additive way out of the inverted diagonal entries of $A_l + B_l^T {I_Q}_l^{-1} B_l$, where as a relaxation parameter we use an upper estimate to the discrete spectrum derived from Gershgorin's theorem. The block smoother is built in the multiplicative way out of the entries of $A_l + B_l^T {I_Q}_l^{-1} B_l$ defined on patches around each inner triangular element. The latter is very much motivated by [1].

We tested the algorithm for an academical problem defined on $\Omega := (-1,1) \times (-1,1)$ and $f(x_1, x_2) := \mathrm{sign}(x_1)\,\mathrm{sign}(x_2)\,(1,1)$. Each computation was started with the initial values $u^{(0)} := 0$, $p^{(0)} := 0$ and $\rho^{(0)} := 1$. The other algorithmic parameters were chosen as follows: $\eta := 1$, $\beta := 10$ and $\nu := 1$ and the terminate relative precisions were $\varepsilon/\varepsilon^{(0)} = \varepsilon_{\mathrm{feas}}/\varepsilon_{\mathrm{feas}}^{(0)} = 10^{-3}$. For the multigrid we chose 3 pre– and post–smoothing steps. The penalty parameter ρ did not exceed 1000, which is related to (20). From columns 5 and 7 in Table I we can see the optimal behaviour for both the point additive and block Gauss–Seidel multigrid smoothers. In Fig. 1 there are the resulting velocity and pressure fields depicted.

Comments and conclusions

We have shown that the recently proposed semimonotonic augmented Lagrangian (SMALE) algorithm can be combined with the multigrid preconditioning to develop an optimal algorithm for the solution of the Stokes problem. The number of outer iterations is controlled by the initial penalty parameter which is automatically adjusted if it does not comply with the convergence theory. If the constraints are well conditioned, then it is possible to achieve fast convergence with a large penalty parameter due the gap in the spectrum of the Hessian of the augmented

level l	dimV_l	dimQ_l	point additive smoother		block multiplicative smoother	
			outer/PCG **iterations**	**total PCG iterations**	**outer**/PCG **iterations**	**total PCG iterations**
1	56	32	**6**/1,0,1,2,4,8	**16**	**6**/1,0,1,2,4,8	**16**
2	208	128	**6**/1,0,1,4,16,30	**52**	**6**/1,0,1,2,5,13	**22**
3	800	512	**5**/1,1,4,20,41	**67**	**6**/1,0,1,2,5,14	**23**
4	3136	2048	**5**/1,1,3,16,47	**68**	**6**/1,0,1,2,6,14	**24**
5	12416	8192	**5**/1,1,3,17,50	**72**	**6**/1,0,1,2,6,15	**25**
6	49418	32768	**5**/1,1,3,19,54	**77**	**6**/1,0,1,2,6,16	**26**

Table I. Numerical experiments

Lagrangian [8], however, the qualitative optimality results presented here are independent of the conditioning of the constraints. The algorithm may be adapted for the solution of strictly convex quadratic programming problems with bound and equality constraints [12]. The modified algorithm has already been applied to the development of optimal solvers for problems arising from discretization of variational inequalities [14].

ACKNOWLEDGEMENTS

This research has been supported by the Czech Ministry of Education under the grant AVČR 1ET400300415 and by the Czech Grant Agency under the grant GAČR 201/05/P008.

REFERENCES

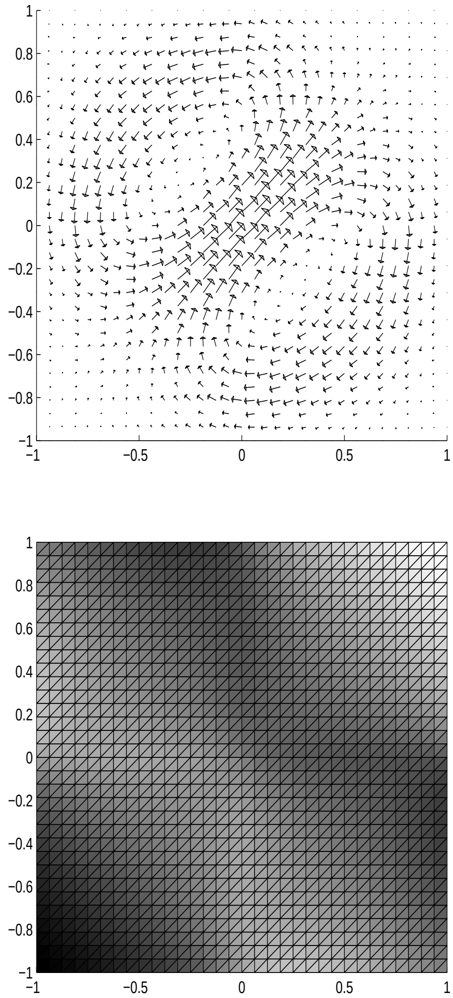

Figure 1. Resulting velocity and pressure fields

1. Arnold, D. N., Falk, R. S., Winther, R.: Multigrid in H(div) and H(curl). Numer. Math. **85**, 197–217 (2000).
2. Arrow, K., Hurwicz, L., Uzawa, H.: Studies in Linear and Nonlinear Programming, Stanford University Press, CA (1958).
3. Benzi, M., Golub, G.H., Liesen, J.: Numerical solution of saddle point problems, Acta Numer. **14**, 1–137

(2005).

4. Bramble, J. H.: Multigrid Methods. Pitman Research Notes in Mathematics **294**, John Wiley & Sons (1993).
5. Bramble, J.H., Pasciak, J.E., Vassilev, A.T.: Analysis of the inexact Uzawa algorithm for saddle point problems, SIAM J. Numer. Anal. **34**, 1072–1092 (1997).
6. Brenner, S. C.: A nonconforming mixed multigrid method for the pure displacement problem in linear elasticity. SIAM J. Numer. Anal. **30**(1), 116–135 (1993).
7. Dostál, Z.: Domain decomposition and PDE with periodic boundary conditions, In: Keyes, D.E., Chan, T.F., Meurant, G., Scroggs, J.S., Voigt, R.G. (eds.), Domain Decomposition Methods for Partial Differential Equations, 258–261, SIAM, Philadelphia (1992).
8. Dostál, Z.: On preconditioning and penalized matrices, Num. Lin. Alg. Appl. **6** (1999), 109–114.
9. Dostál, Z., Friedlander, A., Santos: Augmented Lagrangians with adaptive precision control for quadratic programming with simple bounds and equality constraints, SIAM J. Optim. **13**, 1120–1140 (2003).
10. Dostál, Z.: Semi-monotonic inexact augmented Lagrangians for quadratic programming with equality constraints. Optimization Methods and Software **20**(6), 715–727 (2005).
11. Dostál, Z.: An optimal algorithm for a class of equality constrained quadratic programming problems with bounded spectrum, to appear in Computational Optimization and Applications.
12. Dostál, Z.: Inexact semi-monotonic augmented Lagrangians with optimal feasibility convergence for quadratic programming with simple bounds and equality constraints, SIAM J. Numerical Analysis **43**(1), (2005) 96–115.
13. Dostál, Z., Horák, D.: Scalability and FETI based algorithm for large discretized variational inequalities, Mathematics and Computers in Simulation **61**, 347–353 (2003).
14. Dostál, Z., Horák, D.: Scalable FETI for numerical solution of variational inequalities, to appear in SIAM J. Numerical Analysis.
15. Dreyer, T., Maar, B., Schulz, V.: Multigrid optimization in applications, J. Comp. Appl. Math. **120**, 67–84 (2000).
16. Fortin, M., Glowinski, R.: Augmented Lagrangian Methods: Application to the Numerical Solution of Boundary Value Problems, North Holland, Amsterdam (1983).
17. Glowinski, R., Le Tallec, P.: Augmented Lagrangians and Operator Splitting Methods, SIAM, Philadelphia (1989).
18. Hackbusch, W.: Multi-grid Methods and Applications, Springer, Berlin (1985).

19. Kruis, J., Matouš, K., Dostál, Z., Solving laminated plates by domain decomposition, Advances in Engineering Software **33**, 445–452 (2002).
20. Le Tallec, P., Sassi, T., Domain decompotisision for non–matching grids: augmented Lagrangians approach, Math. Comp. **64**, 1367–1396 (1995).
21. Hu, Q., Zou, J.: Two new variants of nonlinear inexact Uzawa algorithms for saddle-point problems. Numer. Math. **93**, 333–359 (2002).
22. Hu, Q., Zou, J.: Nonlinear inexact Uzawa algorithms for linear and nonlinear saddle–point problems, SIAM J. Optim. **16**, 798–825 (2006).
23. Schöberl, J., Zulehner, W.: On Schwarz-type smoothers for saddle-point problems. Numer. Math. **95**, 377–399 (2003).
24. Vanka, S.P.: Block–implicit multigrid solution of Navier–Stokes equations in primitive variables, J. Comput. Phys. **65**, 138–158 (1986).
25. Zulehner, W.: A class of smoothers for saddle point problems, Computing **65**, 227–246 (2000).
26. Zulehner, W.: Analysis of iterative methods for saddle point problems: A unified approach, Math. Comp. **71**, 479–505 (2001).

