Numerical stability of inexact saddle point solvers

Pavel Jiránek*, **Miroslav Rozložník****

Faculty of Mechatronics, Technical University of Liberec,
Hálkova 6, 461 17 Liberec, Czech Republic*

Institute of Computer Science, Czech Academy of Sciences,
Pod Vodárenskou věží 2, 182 07 Prague, Czech Republic**

pavel.jiranek@tul.cz*, miro@cs.cas.cz**

1. Introduction

We consider a saddle point problem with the symmetric 2×2 block form

$$\begin{pmatrix} A & B \\ B^T & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ 0 \end{pmatrix}, \qquad (1)$$

where A is an $n\times n$ symmetric positive definite matrix and B is an $n\times m$ matrix of a full column rank. Such systems arise in many applications including the finite element approximation of second order elliptic partial differential equations, weighted least squares and quadratic optimization with linear constraints; see [1] and the references therein.

We present our results on the numerical stability of two representatives of segregated methods for solving saddle point systems – the Schur complement reduction method and the null-space projection method, see [2, 3]. A segregated method transforms (1) into a reduced system for x or y which is solved by some iterative method. Such a scheme requires matrix-vector multiplications which need solutions of inner systems with A or B. We assume that these systems are solved inexactly with the uniformly bounded backward error. The remaining component of the approximate solution vector is found by a back-substitution into (1) which can be done in several ways. Here we denote them as the scheme A (generic update), B (direct substitution) and C (corrected direct substitution).

The errors due to the inexact solution of inner systems and the roundoff (unit roundoff is denoted as u) are propagated throughout the iteration process. This affects the limiting accuracy of the computed solution which we measure in terms of the residuals in outer iteration, the residuals in the saddle point system and the forward errors.

The theoretical results are illustrated on a simple numerical example with $n = 100$, $m = 20$, $A = \mathrm{tridiag}(1, 4, 1)$, and B and f are chosen randomly. We have $\kappa(A) = \|A\| \cdot \|A^{-1}\| = 7.1695 \cdot 0.4603$, $\kappa(B) = \|B\| \cdot \|B^{\dagger}\| = 5.9990 \cdot 0.4998$.

2. Schur complement reduction method

The Schur complement reduction method is based on the equivalent formulation of (1)

$$\begin{pmatrix} A & B \\ 0 & S \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} f \\ B^T A^{-1} f \end{pmatrix}.$$

The symmetric positive definite system with the Schur complement matrix $S \equiv B^T A^{-1} B$ is solved iteratively. Given an approximation y_k ($k = 0, 1, 2, \ldots$) to the solution vector component y, the corresponding approximation x_k to x is found by solving $Ax_k = f - By_k$. We consider three different back-substitution formulas summarized in the following algorithm:

choose y_0, solve $Ax_0 = f - By_0$
for $k = 0, 1, 2, \ldots$ do :
compute α_k and $p_k^{(y)}$
$y_{k+1} = y_k + \alpha_k p_k^{(y)}$
solve $Ap_k^{(x)} = -Bp_k^{(y)}$
back-substitution:
A: $x_{k+1} = x_k + \alpha_k p_k^{(x)}$,
B: solve $Ax_{k+1} = f - By_{k+1}$,
C: solve $Au_k = f - Ax_k - By_{k+1}$, $x_{k+1} = x_k + u_k$.
$r_{k+1}^{(y)} = r_k^{(y)} - \alpha_k B^T p_k^{(x)}$

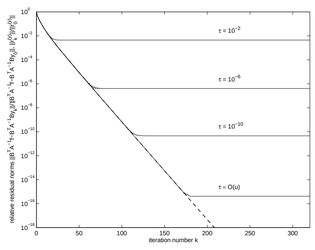
(inner iteration: solve $Ap_k^{(x)} = -Bp_k^{(y)}$ back-substitution; outer iteration: the whole loop)

Since it can be inefficient to solve the systems with A accurately, we assume the computed solution $\hat{v}$ of some system $Av = b$ is an exact solution of $(A + \Delta A)\hat{v} = b + \Delta b$ with $\|\Delta A\| \le \tau\|A\|$, $\|\Delta b\| \le \tau\|b\|$ and $\tau\kappa(A) \ll 1$.

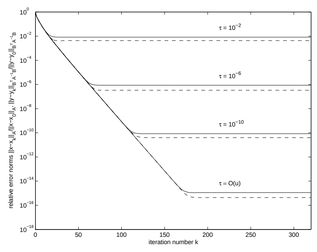
The gap between the residual in the Schur complement system and the updated residual $r_k^{(y)}$ can be bounded as

$$\|-B^T A^{-1} f + Sy_k - r_k^{(y)}\| \le \frac{O(\tau)\kappa(A)}{1-\tau\kappa(A)} \|A^{-1}\| \|B\| (\|f\| + \|B\| Y_k),$$

where $Y_k \equiv \max_{i\le k} \|y_i\|$. Hence the accuracy in the outer iteration does not depend on the back-substitution scheme and the ultimate level of the residual in the Schur complement system is proportional to τ.



Residual in the Schur complements system



Forward errors

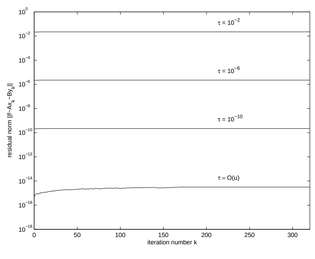
For the residuals in the system (1) we have the estimates

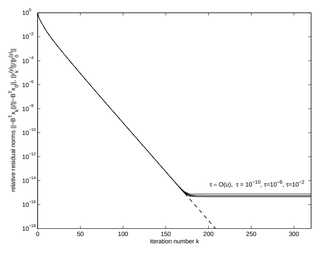
$$\|f - Ax_k - By_k\| \le \frac{O(\alpha_1)\kappa(A)}{1-\tau\kappa(A)} (\|f\| + \|B\| Y_k),$$
$$\| - B^T x_k - r_k^{(y)}\| \le \frac{O(\alpha_2)\kappa(A)}{1-\tau\kappa(A)} \|A^{-1}\| \|B\| (\|f\| + \|B\| Y_k),$$

where

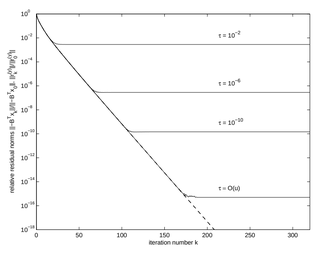
$\alpha_1 = \tau$, $\alpha_2 = u$ for the scheme A,
$\alpha_1 = \tau$, $\alpha_2 = \tau$ for the scheme B,
$\alpha_1 = u$, $\alpha_2 = \tau$ for the scheme C.

The accuracy of the residuals $f - Ax_k - By_k$ and $-B^T x_k$ in the saddle point system thus depends on the back-substitution scheme.

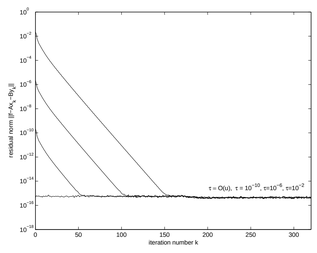




Scheme A



Scheme B



Scheme C

For the forward errors we have

$$\|x - x_k\| \le \gamma_1 \|f - Ax_k - By_k\| + \gamma_2 \| - B^T x_k\|,$$
$$\|y - y_k\| \le \gamma_2 \|f - Ax_k - By_k\| + \gamma_3 \| - B^T x_k\|,$$

where $\gamma_1 \equiv \sigma_{min}^{-1}(A)$, $\gamma_2 \equiv \kappa^{1/2}(A)\sigma_{min}^{-1}(B)$, $\gamma_3 \equiv \sigma_{min}^{-1}(S)$ and hence their are both proportional to τ independently on the back-substitution scheme.

3. Null-space projection method

The null-space projection method uses the second equation of (1) indicating that $x \in N(B^T)$. The solution vector component x can be thus found by solving the projected system $PAPx = Pf$, where P is the orthogonal projector onto $N(B^T)$. This symmetric positive semidefinite system is then solved iteratively producing iterates x_k ($k = 0, 1, 2, \ldots$). The corresponding iterate y_k is found as a least squares solution of $By_k \approx f - Ax_k$. Such a back-substitution can be performed in several ways and it is summarized in the following algorithm:

choose x_0, solve $By_0 \approx f - Ax_0$
compute α_k and $p_k^{(x)} \in N(B^T)$
$x_{k+1} = x_k + \alpha_k p_k^{(x)}$
solve $Bp_k^{(y)} \approx r_k^{(x)} - \alpha_k A p_k^{(x)}$
back-substitution:
A: $y_{k+1} = y_k + p_k^{(y)}$,
B: solve $By_{k+1} \approx f - Ax_{k+1}$,
C: solve $Bv_k \approx f - Ax_{k+1} - By_k$, $y_{k+1} = y_k + v_k$.
$r_{k+1}^{(x)} = r_k^{(x)} - \alpha_k A p_k^{(x)} - Bp_k^{(y)}$

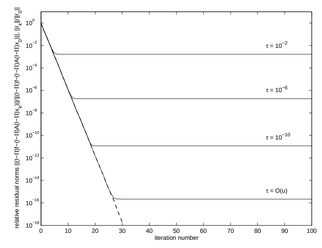
(inner iteration: solve $Bp_k^{(y)} \approx r_k^{(x)} - \alpha_k A p_k^{(x)}$ back-substitution; outer iteration: the whole loop)

The algorithm relies on the effective solution of least squares problems with B. We assume that the computed solution $\hat{w}$ of some problem $Bw \approx c$ is an exact solution of $(B + \Delta B)\hat{w} = c + \Delta c$, where $\|\Delta B\| \le \tau\|B\|$, $\|\Delta c\| \le \tau\|c\|$, $\tau\kappa(B) \ll 1$.

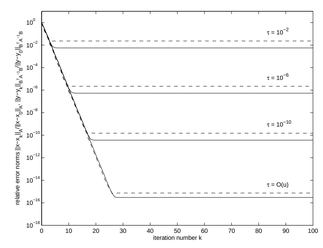
The gap between the residual in the projected system and its approximation $r_k^{(x)}$ can be estimated as

$$\|Pf - PAPx_k - r_k^{(x)}\| \le \frac{O(\tau)\kappa(B)}{1-\tau\kappa(B)} (\|f\| + \|A\| X_k),$$

where $X_k \equiv \max_{i\le k} \|x_i\|$ and thus the accuracy in the projected system is ultimately proportional to τ and it does not depend on the actual scheme.



Residual in the projected system



Forward errors

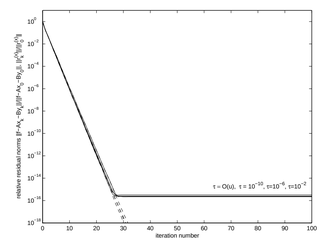
For the residuals in (1) we have the bounds

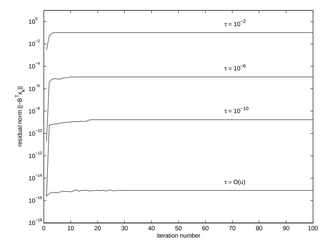
$$\|f - Ax_k - By_k - r_k^{(x)}\| \le \frac{O(\alpha_3)\kappa(B)}{1-\tau\kappa(B)} (\|f\| + \|A\| X_k),$$
$$\| - B^T x_k\| \le \frac{O(\tau)\kappa(B)}{1-\tau\kappa(B)} \|B\| X_k$$

with

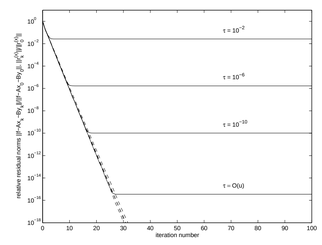
$\alpha_3 = u$ for the scheme A,
$\alpha_3 = \tau$ for the scheme B,
$\alpha_3 = u$ for the scheme C.

Hence the ultimate level of $f - Ax_k - By_k$ depends on the actual back-substitution scheme which is not the case here for the residual $-B^T x_k$ affected only by the departure of direction vectors $p_k^{(x)}$ from $N(B^T)$.

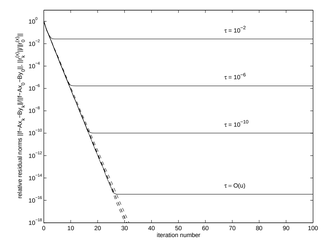




Scheme A



Scheme B



Scheme C

The forward errors behave similarly as in the Schur complement reduction method – they do not depend on the actual scheme and are ultimately proportional to τ.

Acknowledgments

This work was supported by the MSMT CR under the project 1M0554 “Advanced Remedial Technologies”, by the project 1ET400300415 within the National Program of Research “Information Society” and by the Institutional Research Plan AV0Z10300504 “Computer Science for the Information Society: Models, Algorithms, Applications”.

References

[1] M. Benzi, G. H. Golub, and J. Liesen. Numerical solution of saddle point problems. *Acta Numerica*, pages 1–137, 2005.

[2] P. Jiránek and M. Rozložník. Limiting accuracy of segregated solution methods for nonsymmetric saddle point problems. 2007. To appear in J. Comp. Appl. Math.

[3] P. Jiránek and M. Rozložník. Maximum attainable accuracy of inexact saddle point solvers. 2007. To appear in SIAM J. Matrix Anal. Appl.