

Erratum to: Inexact constraint preconditioners for linear systems arising in interior point methods

Luca Bergamaschi · Jacek Gondzio ·
Manolo Venturin · Giovanni Zilli

Published online: 13 November 2009
© Springer Science+Business Media, LLC 2009

Erratum to: Comput Optim Appl (2007) 36: 137–147
DOI 10.1007/s10589-006-9001-0

1 Introduction

It was brought to our attention by Professor Valeria Simoncini that our proof of Theorem 2.1 [2, p. 139] is incorrect. Indeed, on page 140 we write the inequality (2.4) which is correct because we work with \tilde{A} which is an $m \times n$ matrix and $m \leq n$ and $\text{rank}(A) = m$ hence $\frac{\|\tilde{A}^T y\|}{\|y\|} \geq \tilde{\sigma}_1$. However, on page 141 we write the inequality (2.6). We try to use a similar argument but now instead of \tilde{A}^T we have \tilde{A} and $x \in \mathbb{R}^n$. Hence the inequality $\frac{\|\tilde{A}x\|}{\|x\|} \geq \tilde{\sigma}_1$ (which we use to prove inequality (2.6)) is incorrect. In this short note we restate Theorem 2.1 of [2] and give a new proof.

The online version of the original article can be found under doi:[10.1007/s10589-006-9001-0](https://doi.org/10.1007/s10589-006-9001-0).

L. Bergamaschi (✉) · G. Zilli
Department of Mathematical Methods and Models for Scientific Applications, University of Padua,
Padua, Italy
e-mail: berga@dmsa.unipd.it

G. Zilli
e-mail: zilli@dmsa.unipd.it

J. Gondzio
School of Mathematics, University of Edinburgh, Edinburgh, Scotland, UK
e-mail: J.Gondzio@ed.ac.uk

M. Venturin
Department of Computer Science, University of Verona, Verona, Italy
e-mail: manolo.venturin@univr.it

Notation: We consider the preconditioning of the KKT system $Hw = b$ by the matrix \tilde{P} , where

$$H = \begin{bmatrix} Q & A^T \\ A & \end{bmatrix} \quad \text{and} \quad \tilde{P} = \begin{bmatrix} D & \tilde{A}^T \\ \tilde{A} & \end{bmatrix}, \tag{1}$$

where $Q \in \mathbb{R}^{n \times n}$ is the Hessian of Lagrangian, $A \in \mathbb{R}^{m \times n}$ is the (exact) Jacobian of constraints, $D \in \mathbb{R}^{n \times n}$ is a positive definite approximation of the Hessian and $\tilde{A} \in \mathbb{R}^{m \times n}$ is a full row rank approximation of the Jacobian ($\text{rank}(\tilde{A}) = m \leq n$). Following [2] we define $E = A - \tilde{A}$, $\text{rank}(E) = p$. Here $\tilde{\sigma}_1$ is the smallest singular value of $\tilde{A}D^{-1/2}$. Further we introduce two error terms:

$$e_Q = \|E_Q\| = \|D^{-1/2}QD^{-1/2} - I\| \quad \text{and} \quad e_A = \frac{\|ED^{-1/2}\|}{\tilde{\sigma}_1} \tag{2}$$

which measure the errors of Hessian and Jacobian approximations, respectively. The distance $|\epsilon|$ of the complex eigenvalues from one, with $\epsilon = \lambda - 1$, will be bounded in terms of these two quantities.

2 Correction

Theorem 2.1 in [2] should be replaced by the following one.

Theorem 2.1 *Assume A and \tilde{A} have maximum rank. If the eigenvector is of the form $(0, y)^T$ then the eigenvalues of $\tilde{P}^{-1}H$ are either one (with multiplicity at least $m - p$) or possibly complex and bounded by $|\epsilon| \leq e_A$. Corresponding to eigenvectors of the form $(x, y)^T$ with $x \neq 0$ the eigenvalues are*

- (1) *equal to one (with multiplicity at least $m - p$), or*
- (2) *real positive and bounded by*

$$\lambda_{\min}(D^{-1/2}QD^{-1/2}) \leq \lambda \leq \lambda_{\max}(D^{-1/2}QD^{-1/2}), \quad \text{or}$$

- (3) *complex, satisfying*

$$|\epsilon_R| \leq e_Q + e_A, \tag{3}$$

$$|\epsilon_I| \leq e_Q + e_A, \tag{4}$$

where $\epsilon = \epsilon_R + i\epsilon_I$.

Proof The eigenvalues of $\tilde{P}^{-1}H$ are the same as those of $\bar{P}^{-1}\bar{H}$ where $\bar{P} = D\tilde{P}D$ and $\bar{H} = DHD$ and

$$D = \begin{bmatrix} D^{-1/2} & 0 \\ 0 & I \end{bmatrix}.$$

They must satisfy

$$\begin{cases} Ku + B^T y = \lambda u + \lambda B^T y - \lambda F^T y, \\ Bu = \lambda Bu - \lambda Fu \end{cases} \tag{5}$$

where $K = D^{-1/2} Q D^{-1/2}$, $B = A D^{-1/2}$, $F = E D^{-1/2}$, $u = D^{1/2} x$. The eigenvalue problem can also be stated, setting $\tilde{B} = \tilde{A} D^{-1/2} = B - F$ and $\epsilon = \lambda - 1$,

$$\begin{cases} \epsilon u + \epsilon \tilde{B}^T y = (K - I)u + F^T y, \\ \epsilon \tilde{B} u = F u. \end{cases} \tag{6}$$

Let us observe that $K - I$ and F are the errors of approximation of the Hessian and Jacobian in (1), respectively.

We now analyze a number of cases depending on u and y .

1. $u = 0$.

Every vector of the form $\begin{pmatrix} 0 \\ y \end{pmatrix}$ where y is an eigenvector of the (non-square) generalized eigenproblem

$$B^T y = \lambda \tilde{B}^T y$$

is the eigenvector of (5) corresponding to λ . Since $\text{rank}(E) = p$ (hence $\text{rank}(F) = p$), among those vectors y there are $m - p$ satisfying $F^T y = 0$. The first equation of (5) reads

$$B^T y = \lambda B^T y$$

so that eigenvector $\begin{pmatrix} 0 \\ y \end{pmatrix}$ is associated to the unit eigenvalue.

We can bound the remaining such eigenvalues in terms of $\|F\|$ using the first equation of (6) as

$$|\epsilon| = \frac{\|F^T y\|}{\|\tilde{B}^T y\|} \leq \|F\| \frac{\|y\|}{\|\tilde{B}^T y\|} \leq \frac{\|F\|}{\tilde{\sigma}_1} = e_A. \tag{7}$$

2. $u \neq 0$.

There are at least $m - p$ linearly independent vectors u satisfying $Fu = 0$ and $Bu \neq 0$. For such vectors the second equation of (5) reads

$$Bu = \lambda Bu$$

giving again unit eigenvalues.

In the most general case where $Bu \neq 0$ and $Fu \neq 0$, we shall consider two possibilities: (a) real eigenvalues and (b) complex eigenvalues.

(a) *Real eigenvalues.* Let us multiply the first equation of (5) by u^T and the second one by y^T . We obtain the following system

$$\begin{cases} u^T K u + u^T B^T y = \lambda u^T u + \lambda u^T \tilde{B}^T y, \\ y^T B u = \lambda y^T \tilde{B} u. \end{cases} \tag{8}$$

If $\lambda \in \mathbb{R}$, subtracting the transpose of the second equation from the first one, we obtain

$$\lambda_{\min}(K) \leq \lambda = \frac{u^T K u}{u^T u} \leq \lambda_{\max}(K). \tag{9}$$

(b) *Complex eigenvalues.* To bound the complex eigenvalues we write in short (6) as

$$\epsilon N \begin{pmatrix} u \\ y \end{pmatrix} = M \begin{pmatrix} u \\ y \end{pmatrix}$$

and observe that matrix N^{-1} can be decomposed using the Cholesky factorization LL^T of the symmetric positive definite matrix $\tilde{B}\tilde{B}^T$.

$$\begin{aligned} N^{-1} &= \begin{pmatrix} I & \tilde{B}^T \\ \tilde{B} & 0 \end{pmatrix}^{-1} = \begin{pmatrix} I & -\tilde{B}^T L^{-T} \\ 0 & L^{-T} \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & -I \end{pmatrix} \begin{pmatrix} I & 0 \\ -L^{-1}\tilde{B} & L^{-1} \end{pmatrix} \\ &= UJU^T \end{aligned} \tag{10}$$

so that the eigenvalue problem is equivalent to

$$U^T M U w = \epsilon J w, \quad \text{with} \quad \begin{pmatrix} u \\ v \end{pmatrix} = U w. \tag{11}$$

It is useful to set $R = L^{-1}\tilde{B}$ since it is easily found that $\|R\| = 1$. Since

$$\begin{aligned} U^T M U &= \begin{pmatrix} I & 0 \\ -R & L^{-1} \end{pmatrix} \begin{pmatrix} E_Q & F^T \\ F & 0 \end{pmatrix} \begin{pmatrix} I & -R^T \\ 0 & L^{-T} \end{pmatrix} \\ &= \begin{pmatrix} E_Q & F^T \\ -RE_Q + L^{-1}F & -RF^T \end{pmatrix} \begin{pmatrix} I & -R^T \\ 0 & L^{-T} \end{pmatrix} \\ &= \begin{pmatrix} E_Q & -E_Q R^T + F^T L^{-T} \\ -RE_Q + L^{-1}F & RE_Q R^T - L^{-1}F R^T - RF^T L^{-T} \end{pmatrix} \end{aligned}$$

we rewrite equation (11)

$$\begin{pmatrix} E_Q & -E_Q R^T + F^T L^{-T} \\ -RE_Q + L^{-1}F & RE_Q R^T - L^{-1}F R^T - RF^T L^{-T} \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \epsilon \begin{pmatrix} w_1 \\ -w_2 \end{pmatrix}. \tag{12}$$

Note that the diagonal blocks are symmetric. Now multiply the first block of equations of (12) by w_1^H and the second by w_2^H thus obtaining

$$\begin{aligned} w_1^H E_Q w_1 + w_1^H (-E_Q R^T + F^T L^{-T}) w_2 &= \epsilon \|w_1\|^2, \\ w_2^H (-E_Q R^T + F^T L^{-T})^T w_1 + w_2^H RE_Q R^T w_2 - 2\Re(w_2^H L^{-1}F R^T w_2) &= -\epsilon \|w_2\|^2. \end{aligned} \tag{13}$$

Subtracting the two equations yields two equations for the real and the imaginary part respectively (assuming $\|w\| = 1$).

$$\begin{aligned} w_1^H E_Q w_1 - w_2^H R E_Q R^T w_2 + 2\Re(w_2^H L^{-1} F R^T w_2) &= \epsilon_R, \\ 2\Im(w_1^H (-E_Q R^T + F^T L^{-T}) w_2) &= \epsilon_I. \end{aligned} \tag{14}$$

If, instead, we add together the two equations in (13) we get for the imaginary part

$$0 = \epsilon_I (\|w_1\|^2 - \|w_2\|^2)$$

which gives for every complex eigenvalue $\|w_1\|^2 = \|w_2\|^2 = \frac{1}{2}$. Hence (14) provides a bound for the real and imaginary part of ϵ in terms of $e_Q = \|E_Q\|$ and e_A .

$$\begin{aligned} |\epsilon_R| &\leq |w_1^H E_Q w_1| + |w_2^H R E_Q R^T w_2| + 2|w_2^H L^{-1} F R^T w_2| \\ &\leq \|E_Q\| (\|w_1\|^2 + \|w_2\|^2) + 2\|w_2\|^2 \|L^{-1} F R^T\| = e_Q + e_A, \\ |\epsilon_I| &\leq 2\|w_1\| \|w_2\| \|(-E_Q R^T + F^T L^{-T})\| \leq e_Q + e_A. \end{aligned} \tag{15}$$

□

Our new proof has followed the ideas from the paper of Benzi and Simoncini [1]. The major difference is that in [1] the preconditioner uses the *exact* Jacobian A , while our analysis applies to the case when the *approximate* Jacobian \tilde{A} is used.

In the special case of linear programming or separable quadratic programming, the Hessian of Lagrangian Q is a diagonal matrix, hence the preconditioner uses exact Hessian $D = Q$. The analysis simplifies in this case.

Corollary 2.2 *Assume that \tilde{A} has maximum rank. The eigenvalues of $\tilde{P}^{-1}H$ are either one or bounded by*

$$|\lambda - 1| \leq e_A.$$

Proof The eigenvalues of $\tilde{P}^{-1}H$ can be characterized in the same way as in Theorem 2.1 for the case $u = 0$ (they are either unit or bounded by $|\epsilon| < e_A$). If $u \neq 0$, the real eigenvalues must satisfy (9) with $K = I$, from which $\lambda = 1$; while for the complex ones, using $E_Q = 0$ and again $\|w_1\| = \|w_2\|$, the first equation of (13) simplifies to

$$|\epsilon| = \frac{|w_1^H F^T L^{-T} w_2|}{\|w_1\|^2} \leq \frac{\|w_1\| e_A \|w_2\|}{\|w_1\|^2} = e_A. \tag{15}$$

□

We summarize the classification of eigenvalues of preconditioned matrix in Table 1.

Acknowledgements We are grateful to Professor Valeria Simoncini for bringing to our attention that the proof of Theorem 2.1 [2, p. 139] was incorrect.

Table 1 Types of eigenvalues in $\tilde{P}^{-1}H$

Case	Eigenvector	Nonseparable case		Separable case	
		Eigenvalue	Bound	Eigenvalue	Bound
1.	$u = 0, F^T y = 0$	real	$\lambda = 1$	real	$\lambda = 1$
	$u = 0, F^T y \neq 0$	real/complex	$ \lambda - 1 \leq e_A$	real/complex	$ \lambda - 1 \leq e_A$
2.	$u \neq 0, Bu \neq 0, Fu = 0$	real	$\lambda = 1$	real	$\lambda = 1$
	$u \neq 0, Bu \neq 0, Fu \neq 0$	real	$\lambda \in [\lambda_{\min}(K), \lambda_{\max}(K)]$	real	$\lambda = 1$
	$u \neq 0, Bu \neq 0, Fu \neq 0$	complex	$\begin{cases} \Re(\lambda) - 1 < e_Q + e_A \\ \Im(\lambda) < e_Q + e_A \end{cases}$	complex	$ \lambda - 1 \leq e_A$

References

1. Benzi, M., Simoncini, V.: On the eigenvalues of a class of saddle point matrices. *Numer. Math.* **103**, 173–196 (2006)
2. Bergamaschi, L., Gondzio, J., Venturin, M., Zilli, G.: Inexact constraint preconditioners for linear system arising in interior point methods. *Comput. Optim. Appl.* **36**, 137–147 (2007)