BIT Numer Math (2014) 54:201–218
DOI 10.1007/s10543-013-0452-2

An algorithm for solving the indefinite least squares problem with equality constraints

Nicola Mastronardi · Paul Van Dooren

Received: 31 January 2013 / Accepted: 27 September 2013 / Published online: 11 October 2013

Abstract An algorithm for computing the solution of indefinite least squares problems and of indefinite least squares problems with equality constrained is presented. Such problems arise when solving total least squares problems and in H^{∞}-smoothing.

The proposed algorithm relies only on stable orthogonal transformations reducing recursively the associated augmented matrix to proper block anti-triangular form. Some numerical results are reported showing the properties of the algorithm.

Keywords Indefinite matrix · Indefinite least squares · Equality constraints

Mathematics Subject Classification (2010) 65F20 · 65G05 · 15A06

Communicated by Miloud Sadkane.

The work of the first author is partly supported by the GNCS INdAM project "Strategie risolutive per sistemi lineari di tipo KKT con uso di informazioni strutturali". The work of the second author is partly supported by the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office. The scientific responsibility rests with its authors.

N. Mastronardi (✉)
Istituto per le Applicazioni del Calcolo "M. Picone", Consiglio Nazionale delle Ricerche,
sede di Bari, via Amendola 122D, 70126 Bari, Italy
e-mail: n.mastronardi@ba.iac.cnr.it

P. Van Dooren
Department of Mathematical Engineering, Catholic University of Louvain,
Avenue Georges Lemaitre 4, 1348 Louvain-la-Neuve, Belgium

1 Introduction

An algorithm for computing the solution of indefinite least squares (ILS) problems and of indefinite least squares problems with equality constraints (ILSEC) is described in this paper.

Given a matrix $A \in \mathbb{R}^{(p+q)\times n}$, a vector $\mathbf{b} \in \mathbb{R}^{p+q}$, and

$$\Sigma_{pq} = \begin{bmatrix} I_p & \\ & -I_q \end{bmatrix}, \tag{1.1}$$

with I_k the identity matrix of order k, the ILS problem is formulated as follows:

$$\min_{\mathbf{x}} (\mathbf{b} - A\mathbf{x})^T \Sigma_{pq} (\mathbf{b} - A\mathbf{x}). \tag{1.2}$$

This problem is considered in [2, 3] where some numerical methods for computing the solution are proposed. The problem arises in solving total least squares problems [3, 4, 14] and in H^∞-smoothing [3, 7]. It is shown in [2] that the ILS problem has a unique solution if and only if $A^T \Sigma_{pq} A$ is positive definite. This means that $p \geq n$ and that $A(1:p, 1:n)$ has full column rank n and so has A. Denote the residual by $\mathbf{r} = \mathbf{b} - A\mathbf{x}$, the solution of (1.2) satisfies the augmented system

$$\begin{bmatrix} \Sigma_{pq} & A \\ A^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{s} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \end{bmatrix}, \tag{1.3}$$

with $\Sigma_{pq}\mathbf{s} = \mathbf{r}$. Furthermore, given a matrix $B \in \mathbb{R}^{s\times n}$ and a vector $\mathbf{d} \in \mathbb{R}^s$, the ILSEC problem can be formulated as follows :

$$\min_{\mathbf{x}} (\mathbf{b} - A\mathbf{x})^T \Sigma_{pq} (\mathbf{b} - A\mathbf{x}) \quad \text{subject to } B\mathbf{x} = \mathbf{d}. \tag{1.4}$$

As noted in [1], the ILSEC problem has a solution assuming that the following conditions

$$\operatorname{rank}(B) = s, \qquad \mathbf{x}^T A^T \Sigma_{pq} A\mathbf{x} > 0, \quad \mathbf{x} \in \ker(B) \tag{1.5}$$

hold. The first condition implies that the constraint equation admits a solution. The second one, imposing the positive definiteness of $A^T \Sigma_{pq} A$ on the nullspace of B, ensures that (1.4) has a unique solution [1]. Moreover, since $A^T \Sigma_{pq} A$ has rank at most p and $\dim(\operatorname{null}(B)) = n - s$, then

$$p \geq n - s. \tag{1.6}$$

The solution of the problem (1.4) satisfies the augmented system

$$\begin{bmatrix} \mathbf{0} & \mathbf{0} & B \\ \mathbf{0} & \Sigma_{pq} & A \\ B^T & A^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \lambda \\ \mathbf{s} \\ \mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{d} \\ \mathbf{b} \\ \mathbf{0} \end{bmatrix} \quad \Leftrightarrow \quad M\mathbf{y} = \mathbf{f}, \tag{1.7}$$

where $-\lambda$ is the vector of the Lagrange multipliers.

The problem of computing the solution of (1.7) is considered in [1, 9, 10, 13]. Recently, an algorithm to reduce an indefinite symmetric matrix to a matrix in proper block anti-triangular form via orthogonal transformations has been described [11]. In this paper we describe an algorithm for computing the solution of both ILS and ILSEC problems by solving the augmented systems (1.3) and (1.7) using this block anti-triangular form. The main idea is to first reduce recursively the coefficient matrix of the augmented system, already in block anti-triangular form to proper block anti-triangular form via orthogonal transformations. Then the obtained linear system is solved.

The algorithm described in the paper computes the solution for the general ILSEC problem. The corresponding algorithm for computing the solution of the ILS problem can be easily derived from the latter one.

The paper is organized as follows. After having introduced some definitions and notations in Sect. 2, the algorithm to reduce the augmented matrix to proper block anti-triangular form via orthogonal transformations is described in Sect. 3, followed by the numerical results and the conclusions.

2 Notations

- The *inertia* of a symmetric matrix $A \in \mathbb{R}^{n\times n}$ is denoted by $\text{Inertia}(A) = (n_-, n_0, n_+)$, where n_-, n_0 and n_+ are the number of negative, zero and positive eigenvalues of A, respectively, and $n_- + n_0 + n_+ = n$.
- The identity matrix and the anti-triangular unit matrix of order n are denoted by I_n and E_n, respectively, i.e.,

$$I_n = \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} \in \mathbb{R}^{n\times n}, \qquad E_n = \begin{bmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{bmatrix} \in \mathbb{R}^{n\times n}.$$

- We denote by `aqrl` the function computing the lower *anti-QR* factorization of $X \in \mathbb{R}^{m\times n}$, $m \geq n$, $\text{rank}(X) = n$, i.e., $[Q, L] = \texttt{aqrl}(X)$, with $Q \in \mathbb{R}^{m\times m}$ orthogonal and

$$X = QL, \qquad L = \begin{bmatrix} \mathbf{0} \\ \hat{L} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \lrcorner\!\!\diagup \end{bmatrix} \begin{matrix} \}m-n, \\ \}n, \end{matrix}$$

$\hat{L} \in \mathbb{R}^{n\times n}$ lower anti-triangular, i.e., all the entries above the main anti-diagonal are zero. The matrix $\hat{L}$ is nonsingular if X has full column rank n. In case $m < n$, the matrix $L = \hat{L}$ is lower anti-trapezoidal, i.e., $\hat{L}(i, j) = 0$, for $i + j < m + 1$,

$$\hat{L} = [\diagup\!\!\square].$$

- We denote by `aqru` the function computing the upper *anti-QR* factorization of $X \in \mathbb{R}^{m\times n}$, $m \geq n$, $\text{rank}(X) = n$, i.e., $[Q, R] = \texttt{aqru}(X)$, with $Q \in \mathbb{R}^{m\times m}$ orthogonal and

$$X = QR, \qquad R = \begin{bmatrix} \hat{R} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \diagup\!\!\!\!\!\triangle \\ \mathbf{0} \end{bmatrix} \begin{matrix} \}n, \\ \}m-n, \end{matrix}$$

$\hat{R} \in \mathbb{R}^{n \times n}$ upper anti-triangular, i.e., all the entries below the main anti-diagonal are zero. The matrix $\hat{R}$ is nonsingular if X has full column rank n. In case $m < n$, the matrix $R = \hat{R}$ is upper anti-trapezoidal, i.e., $\hat{R}(i, j) = 0$, for $i + j > n + 1$,

$$\hat{R} = [\,\square\!\!\diagup\,].$$

– We denote by $\mathbf{0}_{i,j}$ and $\mathbf{0}_k$ the $i \times j$ zero matrix and the square zero matrix of order k, respectively. Moreover, we omit the indexes if the size of the matrix can be deduced from the context.

3 Reduction of the augmented matrix to anti-triangular form

Given a nonsingular indefinite symmetric matrix $A \in \mathbb{R}^{n \times n}$ with inertia $\mathrm{In}(A) = (n_-, 0, n_+)$, an orthogonal matrix Q can be computed such that $\hat{A} = Q^T A Q$ is in proper lower block anti-triangular form [11, 12],

$$\hat{A} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & Y^T \\ \mathbf{0} & X & Z^T \\ Y & Z & W \end{bmatrix} \begin{matrix} \} = \min\{n_-, n_+\} \\ \} = |n_- - n_+| \\ \} = \min\{n_-, n_+\} \end{matrix} \qquad (3.1)$$

with W symmetric, Y nonsingular lower anti-triangular, i.e., all the entries above the main anti-diagonal are zero, and $X = LL^T$ symmetric positive definite if $n_+ > n_-$, $X = -LL^T$ symmetric negative definite if $n_- > n_+$, L nonsingular lower triangular, $X = [\ ]$ if $n_- = n_+$.

In a similar way, an orthogonal matrix $\hat{Q}$ can be computed such that $\hat{A} = \hat{Q}^T A \hat{Q}$ is in proper upper block anti-triangular form [11, 12],

$$\hat{A} = \begin{bmatrix} \hat{W} & \hat{Z} & \hat{Y} \\ \hat{Z}^T & \hat{X} & \mathbf{0} \\ \hat{Y}^T & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{matrix} \} = \min\{n_-, n_+\} \\ \} = |n_- - n_+| \\ \} = \min\{n_-, n_+\} \end{matrix} \qquad (3.2)$$

with $\hat{W}$ symmetric, $\hat{Y}$ nonsingular upper anti-triangular, i.e., all the entries below the main anti-diagonal are zero, and $\hat{X} = \hat{L}\hat{L}^T$ symmetric positive definite if $n_+ > n_-$, $\hat{X} = -\hat{L}\hat{L}^T$ symmetric negative definite if $n_- > n_+$, L nonsingular lower triangular, $X = [\ ]$ if $n_- = n_+$.

The solution of a linear system with the coefficient matrix in proper lower (upper) block anti-triangular form can be computed with a cost depending quadratically on the size of the matrix [11, 12].

The main idea of the algorithm we want to propose is to recursively transform the nonsingular coefficient matrix of the linear system (1.7), already in block anti-triangular form, in proper block anti-triangular form by orthogonal transformations.

Let us partition $\mathbf{y}$ and $\mathbf{f}$ in (1.7) as follows,

$$\mathbf{y} = \begin{bmatrix} \lambda \\ \mathbf{s} \\ \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \begin{matrix} \}s, \\ \}p+q, \\ \}n-s, \\ \}s, \end{matrix} \quad \text{with } \mathbf{x} = \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix}, \qquad \mathbf{f} = \begin{bmatrix} \mathbf{d} \\ \mathbf{b} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \begin{matrix} \}s, \\ \}p+q, \\ \}n-s, \\ \}s. \end{matrix} \tag{3.3}$$

Let us consider the anti-QR factorization of B^T,

$$\left[\tilde{Q}^{(1)}, \tilde{L}_1\right] = \texttt{aqrl}\left(B^T\right), \quad \tilde{Q}^{(1)} = [Q_{\mathcal{N}}, Q_{\mathcal{R}}], \tag{3.4}$$

$\tilde{Q}_1 \in \mathbb{R}^{n\times n}$ orthogonal, with $Q_{\mathcal{N}} \in \mathbb{R}^{n\times(n-s)}$ spanning the nullspace of B, $Q_{\mathcal{R}} \in \mathbb{R}^{n\times s}$ spanning the range of B^T, and $\tilde{L}_1 = \left[\begin{smallmatrix} \mathbf{0} \\ Y_1^T \end{smallmatrix}\right] \in \mathbb{R}^{n\times s}$, $Y_1 \in \mathbb{R}^{s\times s}$ nonsingular lower anti-triangular. Let

$$Q^{(1)} = \begin{bmatrix} I_{p+q+s} & \\ & \tilde{Q}^{(1)} \end{bmatrix}.$$

Then (1.7) is transformed into the following equivalent linear system,

$$M^{(1)}\mathbf{y}^{(1)} = \mathbf{f}^{(1)}, \tag{3.5}$$

with

$$M^{(1)} = Q^{(1)^T} M Q^{(1)} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} & Y_1 \\ \mathbf{0} & \Sigma_{pq} & AQ_{\mathcal{N}} & AQ_{\mathcal{R}} \\ \mathbf{0} & Q_{\mathcal{N}}^T A^T & \mathbf{0} & \mathbf{0} \\ Y_1^T & Q_{\mathcal{R}}^T A^T & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{matrix} \}s, \\ \}p+q, \\ \}n-s, \\ \}s, \end{matrix}$$

$$\mathbf{y}^{(1)} = \begin{bmatrix} \lambda \\ \mathbf{s} \\ \tilde{\mathbf{x}}_1 \\ \tilde{\mathbf{x}}_2 \end{bmatrix} = Q^{(1)^T}\mathbf{y} = \begin{bmatrix} \lambda \\ \mathbf{s} \\ \tilde{Q}^{(1)^T} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \end{bmatrix}, \quad \tilde{\mathbf{x}}_1 \in \mathbb{R}^{n-s}, \ \tilde{\mathbf{x}}_2 \in \mathbb{R}^s, \tag{3.6}$$

and $\mathbf{f}^{(1)} = Q^{(1)^T}\mathbf{f} = \mathbf{f}$, since, by (3.3), the last n entries of $\mathbf{f}$ are zero.

Let $A_1 = AQ_{\mathcal{N}}$ and $A_2 = AQ_{\mathcal{R}}$. We can already compute $\tilde{\mathbf{x}}_2$ from the lower anti-triangular linear subsystem

$$Y_1\tilde{\mathbf{x}}_2 = \mathbf{d} \tag{3.7}$$

in (3.5), and, using $\tilde{\mathbf{b}} = \mathbf{b} - A_2\tilde{\mathbf{x}}_2$, update the right-hand-side,

$$\tilde{\mathbf{f}}^{(1)} = \begin{bmatrix} \mathbf{d} \\ \tilde{\mathbf{b}} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{d} \\ \begin{bmatrix} \mathbf{b} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} A_2 \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} \tilde{\mathbf{x}}_2 \end{bmatrix}.$$

Therefore, we now need to solve the linear system

$$\begin{bmatrix} \mathbf{0} & \Sigma_{pq} & A_1 \\ \mathbf{0} & A_1^T & \mathbf{0} \\ Y_1^T & A_2^T & \mathbf{0} \end{bmatrix} \begin{bmatrix} \lambda \\ \mathbf{s} \\ \tilde{\mathbf{x}}_1 \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{b}} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix}. \tag{3.8}$$

If we first solve the "smaller" linear system

$$M^{(2)}\mathbf{y}^{(2)} = \mathbf{f}^{(2)}, \tag{3.9}$$

with

$$M^{(2)} = \begin{bmatrix} \Sigma_{pq} & A_1 \\ A_1^T & \mathbf{0} \end{bmatrix}, \qquad \mathbf{y}^{(2)} = \begin{bmatrix} \mathbf{s} \\ \tilde{\mathbf{x}}_1 \end{bmatrix}, \qquad \mathbf{f}^{(2)} = \begin{bmatrix} \tilde{\mathbf{b}} \\ \mathbf{0} \end{bmatrix},$$

then λ can be computed from $Y_1^T \lambda = -A_2^T \mathbf{s}$.

Since

$$M^{(2)} = \begin{bmatrix} I & \\ A_1^T \Sigma_{pq} & I \end{bmatrix} \begin{bmatrix} \Sigma_{pq} & \\ & -A_1^T \Sigma_{pq} A_1 \end{bmatrix} \begin{bmatrix} I & \Sigma_{pq} A_1 \\ & I \end{bmatrix},$$

it follows from (1.5) that $\mathrm{Inertia}(M^{(2)}) = (q+n-s, 0, p)$ and $\mathrm{Inertia}(M) = \mathrm{Inertia}(M^{(1)}) = (q+n, 0, p+s)$.

We now consider the case $q \geq n-s$.

Let us partition A_1, $\mathbf{s}$ and $\tilde{\mathbf{b}}$ as follows,

$$A_1 = \begin{bmatrix} A_{11} \\ A_{12} \end{bmatrix} \begin{matrix} \}p, \\ \}q, \end{matrix}, \qquad \mathbf{s} = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \mathbf{s}_3 \\ \mathbf{s}_4 \end{bmatrix} \begin{matrix} \}p-n+s, \\ \}n-s, \\ \}n-s, \\ \}q-n+s, \end{matrix} \qquad \tilde{\mathbf{b}} = \begin{bmatrix} \tilde{\mathbf{b}}_1 \\ \tilde{\mathbf{b}}_2 \\ \tilde{\mathbf{b}}_3 \\ \tilde{\mathbf{b}}_4 \end{bmatrix} \begin{matrix} \}p-n+s, \\ \}n-s, \\ \}n-s, \\ \}q-n+s. \end{matrix}$$

Compute the lower anti-QR factorization of A_{11}, $[Q_{11}, L_{11}] = \mathtt{aqrl}(A_{11})$ and the QR factorization of A_{12}, $[Q_{12}, R_{12}] = \mathtt{qr}(A_{12})$.

Let

$$\tilde{Q}^{(2)} = \begin{bmatrix} Q_{11} & \\ & Q_{12} \end{bmatrix} \begin{matrix} \}p \\ \}q \end{matrix} \quad \text{and} \quad Q^{(2)} = \begin{bmatrix} \tilde{Q}^{(2)} & \\ & I_{n-s} \end{bmatrix}.$$

Then (3.9) is transformed into the equivalent linear system

$$M^{(3)}\mathbf{y}^{(3)} = \mathbf{f}^{(3)}, \tag{3.10}$$

where

$$M^{(3)} = Q^{(2)^T} M^{(2)} Q^{(2)} = \left[\begin{array}{c|c|c|c|c} I_{p-n+s} & & & & \mathbf{0} \\ \hline & I_{n-s} & & & L_{11} \\ \hline & & -I_{n-s} & & R_{12} \\ \hline & & & -I_{q-n+s} & \mathbf{0} \\ \hline \mathbf{0} & L_{11}^T & R_{12}^T & \mathbf{0} & \mathbf{0} \end{array}\right],$$

$$\mathbf{y}^{(3)} = \begin{bmatrix} \mathbf{s}_1^{(1)} \\ \mathbf{s}_2^{(1)} \\ \mathbf{s}_3^{(1)} \\ \mathbf{s}_4^{(1)} \\ \tilde{\mathbf{x}}_1 \end{bmatrix} = Q^{(2)^T} \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \mathbf{s}_3 \\ \mathbf{s}_4 \\ \tilde{\mathbf{x}}_1 \end{bmatrix} = \begin{bmatrix} Q_{11}^T \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{bmatrix} \\ Q_{12}^T \begin{bmatrix} \mathbf{s}_3 \\ \mathbf{s}_4 \end{bmatrix} \\ \tilde{\mathbf{x}}_1 \end{bmatrix}, \tag{3.11}$$

$$\mathbf{f}^{(3)} = \begin{bmatrix} \mathbf{f}_1^{(3)} \\ \mathbf{f}_2^{(3)} \\ \mathbf{f}_3^{(3)} \\ \mathbf{f}_4^{(3)} \\ \mathbf{0} \end{bmatrix} = Q^{(2)^T} \begin{bmatrix} \tilde{\mathbf{b}}_1 \\ \tilde{\mathbf{b}}_2 \\ \tilde{\mathbf{b}}_3 \\ \tilde{\mathbf{b}}_4 \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} Q_{11}^T \begin{bmatrix} \tilde{\mathbf{b}}_1 \\ \tilde{\mathbf{b}}_2 \end{bmatrix} \\ Q_{12}^T \begin{bmatrix} \tilde{\mathbf{b}}_3 \\ \tilde{\mathbf{b}}_4 \end{bmatrix} \\ \mathbf{0} \end{bmatrix}.$$

Due to the structure of $M^{(3)}$, it follows from (3.10) that we can compute $\mathbf{s}_1^{(1)}$ and $\mathbf{s}_4^{(1)}$,

$$\mathbf{s}_1^{(1)} = \mathbf{f}_1^{(3)}, \qquad \mathbf{s}_4^{(1)} = -\mathbf{f}_4^{(3)}, \tag{3.12}$$

and "shrink" (3.10) to

$$\tilde{M}^{(3)} \tilde{\mathbf{y}}^{(3)} = \tilde{\mathbf{f}}^{(3)}, \tag{3.13}$$

with

$$\tilde{M}^{(3)} = \left[\begin{array}{c|c|c} I_{n-s} & & L_{11} \\ \hline & -I_{n-s} & R_{12} \\ \hline L_{11}^T & R_{12}^T & \mathbf{0} \end{array}\right], \qquad \tilde{\mathbf{y}}^{(3)} = \begin{bmatrix} \mathbf{s}_2^{(1)} \\ \mathbf{s}_3^{(1)} \\ \tilde{\mathbf{x}}_1 \end{bmatrix}, \qquad \tilde{\mathbf{f}}^{(3)} = \begin{bmatrix} \mathbf{f}_2^{(3)} \\ \mathbf{f}_3^{(3)} \\ \mathbf{0} \end{bmatrix},$$

with $\text{Inertia}(\tilde{M}^{(3)}) = (q+n-s, 0, p) - (q-n+s, 0, p-n+s) = (2(n-s), 0, n-s)$.

Let

$$\tilde{Q}^{(3)} = \begin{bmatrix} \frac{1}{\sqrt{2}} I_{n-s} & \frac{1}{\sqrt{2}} E_{n-s} \\ -\frac{1}{\sqrt{2}} E_{n-s} & \frac{1}{\sqrt{2}} I_{n-s} \end{bmatrix} \quad \text{and} \quad Q^{(3)} = \begin{bmatrix} \tilde{Q}^{(3)} & \\ & I_{n-s} \end{bmatrix}. \tag{3.14}$$

Then (3.13) is transformed into the following linear system,

$$M^{(4)}\mathbf{y}^{(4)} = \mathbf{f}^{(4)}, \tag{3.15}$$

with

$$M^{(4)} = Q^{(3)^T}\tilde{M}^{(3)}Q^{(3)} = \begin{bmatrix} & E_{n-s} & \tilde{Y}_2 \\ E_{n-s} & & R_2 \\ \tilde{Y}_2^T & R_2^T & \mathbf{0} \end{bmatrix},$$

$$\begin{aligned}\mathbf{y}^{(4)} &= \begin{bmatrix} \tilde{\mathbf{s}}_2^{(1)} \\ \tilde{\mathbf{s}}_3^{(1)} \\ \tilde{\mathbf{x}}_1 \end{bmatrix} = Q^{(3)^T}\begin{bmatrix} \mathbf{s}_2^{(1)} \\ \mathbf{s}_3^{(1)} \\ \tilde{\mathbf{x}}_1 \end{bmatrix} = \begin{bmatrix} \tilde{Q}^{(3)^T}\begin{bmatrix} \mathbf{s}_2^{(1)} \\ \mathbf{s}_3^{(1)} \end{bmatrix} \\ \tilde{\mathbf{x}}_1 \end{bmatrix} \\ &= \begin{bmatrix} 1/\sqrt{2}(\mathbf{s}_2^{(1)} - \mathtt{flipud}(\mathbf{s}_3^{(1)})) \\ 1/\sqrt{2}(\mathbf{s}_3^{(1)} + \mathtt{flipud}(\mathbf{s}_2^{(1)})) \\ \tilde{\mathbf{x}}_1 \end{bmatrix},\end{aligned} \tag{3.16}$$

$$\begin{aligned}\mathbf{f}^{(4)} &= \begin{bmatrix} \tilde{\mathbf{f}}_2^{(3)} \\ \tilde{\mathbf{f}}_3^{(3)} \\ \mathbf{0} \end{bmatrix} = Q^{(3)^T}\begin{bmatrix} \mathbf{f}_2^{(3)} \\ \mathbf{f}_3^{(3)} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \tilde{Q}^{(3)^T}\begin{bmatrix} \mathbf{f}_2^{(3)} \\ \mathbf{f}_3^{(3)} \end{bmatrix} \\ \mathbf{0} \end{bmatrix} \\ &= \begin{bmatrix} 1/\sqrt{2}(\mathbf{f}_2^{(3)} - \mathtt{flipud}(\mathbf{f}_3^{(3)})) \\ 1/\sqrt{2}(\mathbf{f}_3^{(3)} + \mathtt{flipud}(\mathbf{f}_2^{(3)})) \\ \mathbf{0} \end{bmatrix},\end{aligned}$$

and $\tilde{Y}_2 = 1/\sqrt{2}(L_{11} - E_{n-s}R_{12})$ and $R_2 = 1/\sqrt{2}(R_{12} + E_{n-s}L_{11})$, and the `matlab` function `flipud`($\mathbf{x}$) returning $\mathbf{x}$ with the entries flipped in the up-down direction. There are different approaches to reduce the coefficient matrix in (3.10) to anti-triangular form. Here we describe one based on the multiplication of a sequence of $n-s$ Householder matrices.

Let

$$M_0^{(4)} = M^{(4)} \quad \text{and} \quad C_0 = \begin{bmatrix} E_{n-s} \\ \tilde{Y}_2^T \end{bmatrix} = \begin{bmatrix} \diagup \\ \triangle \end{bmatrix}.$$

At the i-th step, $i = 1, 2, \ldots, n-s$, a Householder transformation $\tilde{H}_i \in \mathbb{R}^{2(n-s)\times 2(n-s)}$ is applied to $C_{i-1} = \tilde{H}_{i-1}\tilde{H}_{i-2}\cdots\tilde{H}_1 C_0$,

$$C_i = \tilde{H}_i C_{i-1},$$

such that the rows $(n-s)-i+1, (n-s)-i+2, \ldots, n-s, 2(n-s)-i+1$ are modified and the entries $(n-s)-i+1, (n-s)-i+2, \ldots, n-s$ of the i-th column

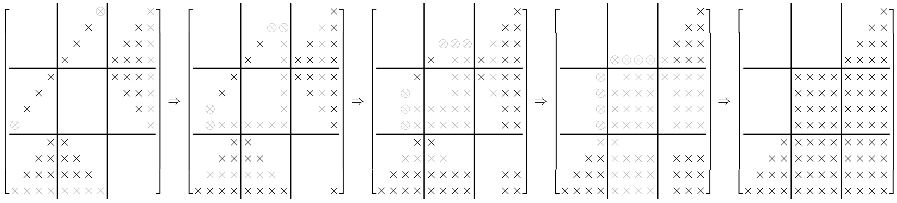

Fig. 1 Sequence of the matrices $M_i^{(4)}$, $i = 0, 1, \ldots, 4$. At the i-th step, $i = 1, \ldots, 4$, the entries to be annihilated are denoted by ⊗ and the entries to be modified by the orthogonal transformations are in *gray*

annihilated. Let

$$H_j = \begin{bmatrix} I_{n-s} & \\ & \tilde{H}_j \end{bmatrix}, \quad j = 1, \ldots, i,$$

and $M_{i-1}^{(4)} = H_{i-1} H_{i-2} \cdots H_1 M_0^{(4)} H_1^T \cdots H_{i-2}^T H_{i-1}^T$. Then

$$M_i^{(4)} = H_i M_{i-1}^{(4)} H_i^T,$$

has the rows (columns) $2(n-s)-i+1, 2(n-s)-i+2, 2(n-s), 3(n-s)-i+1$ modified and the entries $2(n-s)-i+1, 2(n-s)-i+2, 2(n-s)$ of the i-th column (row) annihilated. This process is graphically depicted in Fig. 1 for $n-s=4$.

Let $\tilde{Q}^{(4)} = \tilde{H}_1^T \tilde{H}_2^T \cdots \tilde{H}_{n-s}^T \in \mathbb{R}^{2(n-s)\times 2(n-s)}$ and $Q^{(4)} = H_1^T H_2^T \cdots H_{n-s}^T \in \mathbb{R}^{3(n-s)\times 3(n-s)}$.

Then the linear system (3.15) is transformed into the equivalent one

$$M^{(5)}\mathbf{y}^{(5)} = \mathbf{f}^{(5)}, \tag{3.17}$$

with $M^{(5)}$ having the following structure,

$$M^{(5)} = Q^{(4)^T} M^{(4)} Q^{(4)} = \begin{bmatrix} & & Y^{(5)} \\ & X^{(5)} & Z^{(5)} \\ Y^{(5)^T} & Z^{(5)^T} & W^{(5)} \end{bmatrix} \begin{matrix} \}n-s, \\ \}n-s, \\ \}n-s, \end{matrix}$$

$Y^{(5)} \in \mathbb{R}^{(n-s)\times(n-s)}$ nonsingular lower anti-triangular, $X^{(5)}, W^{(5)} \in \mathbb{R}^{(n-s)\times(n-s)}$ symmetric, and

$$\begin{aligned} \mathbf{y}^{(5)} &= \begin{bmatrix} \mathbf{y}_1^{(5)} \\ \mathbf{y}_2^{(5)} \\ \mathbf{y}_3^{(5)} \end{bmatrix} = Q^{(4)^T}\mathbf{y}^{(4)} = \begin{bmatrix} \tilde{\mathbf{s}}_2^{(1)} \\ \tilde{Q}^{(4)^T} \begin{bmatrix} \tilde{\mathbf{s}}_3^{(1)} \\ \tilde{\mathbf{x}}_1 \end{bmatrix} \end{bmatrix}, \\ \mathbf{f}^{(5)} &= \begin{bmatrix} \mathbf{f}_1^{(5)} \\ \mathbf{f}_2^{(5)} \\ \mathbf{f}_3^{(5)} \end{bmatrix} = Q^{(4)^T}\mathbf{f}^{(4)} = \begin{bmatrix} \tilde{\mathbf{f}}_2^{(3)} \\ \tilde{Q}^{(4)^T} \begin{bmatrix} \tilde{\mathbf{f}}_3^{(3)} \\ \mathbf{0} \end{bmatrix} \end{bmatrix}. \end{aligned} \tag{3.18}$$

Since $\mathrm{Inertia}(M^{(5)}) = \mathrm{Inertia}(\tilde{M}^{(3)}) = \mathrm{Inertia}(M^{(4)}) = (2(n-s), 0, n-s)$, then, by [6], the submatrix $X^{(5)}$ of $M^{(5)}$ is symmetric negative definite with Cholesky factorization $X^{(5)} = -L^{(5)} L^{(5)^T}$, $L^{(5)} \in \mathbb{R}^{(n-s)\times(n-s)}$ nonsingular lower triangular.

We can now compute the solution of the linear system (3.17) in the following steps:

– solve the lower anti-triangular linear system

$$Y^{(5)} \mathbf{y}_3^{(5)} = \mathbf{f}_1^{(5)};$$

– update of the right-hand-side:

$$\begin{bmatrix} \tilde{\mathbf{f}}_1^{(5)} \\ \tilde{\mathbf{f}}_2^{(5)} \\ \tilde{\mathbf{f}}_3^{(5)} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_1^{(5)} \\ \begin{bmatrix} \mathbf{f}_2^{(5)} \\ \mathbf{f}_3^{(5)} \end{bmatrix} - \begin{bmatrix} Z^{(5)} \\ W^{(5)} \end{bmatrix} \mathbf{y}_3^{(5)} \end{bmatrix};$$

– solve the linear system $X^{(5)} \mathbf{y}_2^{(5)} = \tilde{\mathbf{f}}_2^{(5)}$,

$$\begin{aligned} L^{(5)} \mathbf{t} &= -\tilde{\mathbf{f}}_2^{(5)} \\ L^{(5)^T} \mathbf{y}_2^{(5)} &= \mathbf{t}; \end{aligned}$$

– update of the right-hand-side:

$$\begin{bmatrix} \tilde{\mathbf{f}}_1^{(5)} \\ \tilde{\mathbf{f}}_2^{(5)} \\ \tilde{\mathbf{f}}_3^{(5)} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{f}}_1^{(5)} \\ \tilde{\mathbf{f}}_2^{(5)} \\ \tilde{\mathbf{f}}_3^{(5)} - Z^{(5)^T} \mathbf{y}_2^{(5)} \end{bmatrix};$$

– solve the lower anti-triangular linear system

$$Y^{(5)^T} \mathbf{y}_1^{(5)} = \tilde{\mathbf{f}}_3^{(5)}.$$

From (3.18) we can compute

$$\mathbf{y}^{(4)} = \begin{bmatrix} \tilde{\mathbf{s}}_2^{(1)} \\ \tilde{\mathbf{s}}_3^{(1)} \\ \tilde{\mathbf{x}}_1 \end{bmatrix} = Q^{(4)} \mathbf{y}^{(5)} = \begin{bmatrix} \mathbf{y}_1^{(5)} \\ \tilde{Q}^{(4)} \begin{bmatrix} \mathbf{y}_2^{(5)} \\ \mathbf{y}_3^{(5)} \end{bmatrix} \end{bmatrix}.$$

If one is only interested in the computation of the solution $\mathbf{x}$ of (1.4), since $\tilde{\mathbf{x}}_2$ is already computed in (3.7), it can be obtained from (3.6) as

$$\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \tilde{Q}^{(1)} \begin{bmatrix} \tilde{\mathbf{x}}_1 \\ \tilde{\mathbf{x}}_2 \end{bmatrix}.$$

If one is also interested in the computation of the whole solution of the augmented system (1.7), from (3.16),

$$\begin{bmatrix} \mathbf{s}_2^{(1)} \\ \mathbf{s}_3^{(1)} \end{bmatrix} = \tilde{Q}^{(3)} \begin{bmatrix} \tilde{\mathbf{s}}_2^{(1)} \\ \tilde{\mathbf{s}}_3^{(1)} \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} \tilde{\mathbf{s}}_2^{(1)} + \texttt{flipud}(\tilde{\mathbf{s}}_3^{(1)}) \\ \tilde{\mathbf{s}}_3^{(1)} - \texttt{flipud}(\tilde{\mathbf{s}}_2^{(1)}) \end{bmatrix}.$$

Furthermore, from (3.11) and (3.12),

$$\mathbf{s} = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \mathbf{s}_3 \\ \mathbf{s}_4 \end{bmatrix} = \tilde{Q}^{(2)} \begin{bmatrix} \mathbf{s}_1^{(1)} \\ \mathbf{s}_2^{(1)} \\ \mathbf{s}_3^{(1)} \\ \mathbf{s}_4^{(1)} \end{bmatrix} = \begin{bmatrix} Q_{11} \begin{bmatrix} \mathbf{s}_1^{(1)} \\ \mathbf{s}_2^{(1)} \end{bmatrix} \\ Q_{12} \begin{bmatrix} \mathbf{s}_3^{(1)} \\ \mathbf{s}_4^{(1)} \end{bmatrix} \end{bmatrix}.$$

Finally, also λ can be computed from (3.8) by solving the lower anti-triangular linear system

$$Y_1^T \lambda = -A_2^T \mathbf{s}.$$

3.1 Computational complexity and implementation details

In Table 1 the number of floating point operations required to compute the solution $\mathbf{x}$ of (1.4) by the proposed algorithm is reported. Observe that, as usual, only the terms of higher order in n, s, p, q, of the floating point operations are displayed, neglecting the other terms.

About the storage requirement for the proposed algorithm, we need additional memory for storing $X^{(5)}$, $W^{(5)}$, and $\tilde{Q}^{(4)}$. Of course, instead of storing $\tilde{Q}^{(4)}$, we store the coefficients $\beta_i \in \mathbb{R}$ and the vectors $\mathbf{v}_i \in \mathbb{R}^{i+1}$ of the Householder matrices $\tilde{H}_i = I_{n-s} - \beta_i \mathbf{v}_i \mathbf{v}_i^T$, $i = 1, \ldots, n-s$. Since $X^{(5)}, W^{(5)} \in \mathbb{R}^{(n-s)\times(n-s)}$ and $\mathbf{v}_i \in \mathbb{R}^{i+1}$, $i = 1, \ldots, n-s$, and $p+q > p, n > n-s$, the additional memory required is negligible with respect to that required to store A and B.

The case $q < n-s$ is very similar but requires ranges for the Householder transformations that are adapted to the trapezoidal shape of R_{12}. Since the formulas are similar, we moved the description of this case to the appendix. Notice that the complexity and storage requirement are then also reduced.

Table 1 Number of floating point operations required to solve the (1.4) problem by the proposed algorithm

	# flops
Y_1	$2s^2(n-s/3)$
$[A_1, A_2]$	$2s(p+q)(n-s)$
L_{11}	$2(n-s)^2(p-(n-s)/3)$
R_{12}	$2(n-s)^2(q-(n-s)/3)$
$M^{(5)}$	$4(n-s)^3$
$L^{(5)}$	$(n-s)^3$

4 Numerical examples

In this section we report the results of some numerical experiments performed in `Matlab©R2010b` with machine precision $\varepsilon \approx 2.2 \times 10^{-16}$. For each example, the matrix A, given κ_A, is constructed as $A = Q_{\Sigma_{pq}} DU$, where $Q_{\Sigma_{pq}} \in \mathbb{R}^{(p+q)\times(p+q)}$ is a Σ_{pq}-orthogonal matrix, i.e., such that $Q_{\Sigma_{pq}}^T \Sigma_{pq} Q_{\Sigma_{pq}} = \Sigma_{pq}$, generated by the method described in [8], $D \in \mathbb{R}^{(p+q)\times n}$ is a diagonal matrix with decreasing diagonal values geometrically distributed between 1 and $1/\kappa_A$, and $U \in \mathbb{R}^{n\times n}$ is a random orthogonal matrix generated by the function `gallery('qmult',`n`)`. It turns out that the condition number[1] of A, $\kappa(A) \approx \kappa_A$. Furthermore, A is normalized so that $\|A\|_2 = 1$. The matrix $B \in \mathbb{R}^{s\times n}$, given its condition number κ_B, is constructed by using the matlab command $B =$ gallery('`randsvd`', $[s, n], \kappa_B$), so that $\|B\|_2 = 1$ and its singular values are geometrically distributed between 1 and $1/\kappa_B$.

The solution $\mathbf{x}$ of problem (1.7), depending on a parameter $c_1 \in \mathbb{R}$, is chosen as $\mathbf{x} = \tilde{Q}^{(1)}\mathbf{v}_1$, with $\mathbf{v}_1 = c_1 \times$ `randn`$(n, 1)$, where $= \tilde{Q}^{(1)} \in \mathbb{R}^{n\times n}$ is the Q factor of the anti-QR factorization of B^T. Furthermore, partitioning $\tilde{Q}^{(1)}$ as in (3.4), the vector $\mathbf{s}$, depending on a parameter $c_2 \in \mathbb{R}$, is chosen as $\mathbf{s} = c_2 V_2 \times$ `randn`$(s, 1)$, $V_2 \in \mathbb{R}^{(p+q)\times s}$ with orthogonal columns spanning the nullspace of $AQ_{\mathcal{N}}$. Hence, $\mathbf{d} = Y_1\mathbf{v}_1(n - s + 1 : n)$, $\mathbf{b} = \Sigma_{pq}\mathbf{s} + A\tilde{Q}^{(1)}\mathbf{v}_1$ and $\lambda = -Y_1^T \backslash (A_2^T \mathbf{s})$. Then the solution of the augmented system (1.7) is given by $\mathbf{y} = [\lambda^T, \mathbf{s}^T, \mathbf{x}^T]^T$.

We observe that, given the matrices Σ_{pq}, $A \in \mathbb{R}^{(p+q)\times n}$, $B \in \mathbb{R}^{s\times n}$, with $\|A\|_2, \|B\|_2 \geq 1$, the corresponding linear system (1.7) can be always scaled by the matrix $D_1 = \text{diag}\{\frac{\beta}{\alpha} I_s, I_{p+q}, \alpha I_n\}$, with $\alpha = 1/\|A\|_2$, $\beta = 1/\|B\|_2$, to have an equivalent linear system

$$M_S \mathbf{y}_S = \mathbf{f}_S,$$

where

$$M_S = D_1 M D_1 = \begin{bmatrix} 0 & 0 & B_S \\ 0 & \Sigma & A_S \\ B_S^T & A_S^T & 0 \end{bmatrix},$$

with $A_S = \alpha A$, $B_S = \beta B$, so that $\|A_S\|_2 = \|B_S\|_2 = 1$, $\mathbf{y}_S = D_1^{-1}\mathbf{y}$, and $\mathbf{f}_S = D_1\mathbf{f}$. In [13] it is shown that the matrix M_S is often better conditioned than M.

For all the experiments, we choose $n = 50$, $s = 20$, $p = 60$, $q = 40$.

Each set of experiments consists of 16 runs, in each of these κ_A and κ_B are taken from the set $\{1e + 1, 1e + 2, 1e + 4, 1e + 8\}$. The solution computed by the proposed method, denoted by $\mathbf{x}_{NP}$, is compared to the one computed by `matlab` using the command "\", denoted by $\mathbf{x}_B$, and the solutions yielded by the methods `GQR-Cholesky` and `GHQR` described in [1], denoted by $\mathbf{x}_{GC}$ and $\mathbf{x}_{GH}$, respectively.

In each table we report the results for matrices A and B with different condition number. In particular, the condition number of the matrices A, B and M and are

[1] The condition number of a rectangular matrix $A \in \mathbb{R}^{m\times n}$, $m \geq n$, $\text{rank}(A) = n$, is defined as $\sigma_{\max}(A)/\sigma_{\min}(A)$ [5].

displayed in columns 2, 3 and 4, respectively. Moreover, the relative errors of the solution computed by "\" of `matlab`, by the `GQR-Cholesky` method, by the `GHQR` method and by the proposed method are reported in columns 5, 6, 7, and 8, respectively. In Table 2 are reported the results for $c_1 = c_2 = 1$, in Table 3 the results for $c_1 = 1$, $c_2 = 1e4$, and in Table 4 the results for $c_1 = 1e4$, $c_2 = 1$.

For each of the compared methods the forward errors are compatible with the condition of backward stability. Indeed, if these methods would be backward stable then the relative forward error would be bounded by the condition number $\kappa(M)$ times the machine accuracy ε (i.e., the relative backward error). One can check that this bound indeed holds for all the methods and for all the examples. Nevertheless, the `GQR-Cholesky` method, the `GHQR` method and the proposed one often outperform the `matlab` "\", but with varying success. This is due to the fact that these methods exploit the structure of the problem. Moreover, we can observe that in Table 4, when the condition number of A is large, `GQR-Cholesky` and `GHQR` behave better than the proposed algorithm. This is mainly due to the fact that the proposed algorithm modifies the initial zero blocks in the original structure of the matrix in (1.3). In such cases `GQR-Cholesky` and `GHQR` should be preferred. A new algorithm proposed in [13], also guarantees the backward error to respect the structure of the augmented system.

Table 2 Relative errors of the computed solutions with the $c_1 = 1 = c_2 = 1$

#	$\kappa(A)$	$\kappa(B)$	$\kappa(M)$	$\frac{\|\mathbf{x}-\mathbf{x}_B\|_2}{\|\mathbf{x}\|_2}$	$\frac{\|\mathbf{x}-\mathbf{x}_{GC}\|_2}{\|\mathbf{x}\|_2}$	$\frac{\|\mathbf{x}-\mathbf{x}_{GH}\|_2}{\|\mathbf{x}\|_2}$	$\frac{\|\mathbf{x}-\mathbf{x}_{NP}\|_2}{\|\mathbf{x}\|_2}$
1	2.88e+01	1.00e+01	7.45e+02	2.48e-14	1.13e-14	1.09e-14	1.79e-14
2	1.33e+02	1.00e+01	6.96e+03	1.84e-13	3.88e-14	3.49e-14	5.84e-14
3	1.27e+04	1.00e+01	2.32e+06	6.84e-12	1.73e-11	1.19e-11	1.36e-11
4	2.21e+08	1.00e+01	9.07e+11	3.91e-06	1.95e-06	1.27e-06	1.03e-05
5	4.24e+01	1.00e+02	1.46e+03	1.57e-14	2.38e-14	2.44e-14	4.15e-14
6	1.92e+02	1.00e+02	1.27e+04	7.74e-14	1.83e-13	1.19e-13	1.36e-13
7	2.99e+04	1.00e+02	9.68e+06	7.67e-11	5.14e-11	3.39e-11	1.51e-10
8	1.37e+08	1.00e+02	1.87e+11	8.62e-07	6.95e-07	5.34e-07	3.82e-06
9	2.07e+01	1.00e+04	2.76e+06	1.07e-11	5.07e-13	5.07e-13	8.32e-14
10	2.44e+02	1.00e+04	9.05e+04	1.09e-11	1.36e-11	1.36e-11	2.35e-13
11	1.58e+04	1.00e+04	9.06e+06	2.15e-11	6.78e-11	8.39e-11	9.97e-11
12	1.64e+08	1.00e+04	1.09e+11	4.63e-07	2.75e-07	2.86e-07	3.48e-07
13	1.44e+01	1.00e+08	6.81e+14	2.34e-04	2.34e-09	2.34e-09	5.12e-10
14	3.91e+02	1.00e+08	3.54e+12	3.51e-06	8.44e-08	8.44e-08	4.93e-10
15	2.67e+04	1.00e+08	2.48e+10	1.96e-06	5.84e-07	5.84e-07	7.75e-10
16	2.63e+08	1.00e+08	3.41e+12	9.40e-05	6.25e-04	6.22e-04	3.36e-05

Table 3 Relative errors of the computed solutions with the $c_1 = 1, c_2 = 1e4$

#	$\kappa(A)$	$\kappa(B)$	$\kappa(M)$	$\frac{\|\mathbf{x}-\mathbf{x}_B\|_2}{\|\mathbf{x}\|_2}$	$\frac{\|\mathbf{x}-\mathbf{x}_{GC}\|_2}{\|\mathbf{x}\|_2}$	$\frac{\|\mathbf{x}-\mathbf{x}_{GH}\|_2}{\|\mathbf{x}\|_2}$	$\frac{\|\mathbf{x}-\mathbf{x}_{NP}\|_2}{\|\mathbf{x}\|_2}$
1	2.69e+01	1.00e+01	5.82e+02	6.04e-11	4.07e-11	6.26e-11	1.18e-10
2	2.25e+02	1.00e+01	2.29e+04	9.00e-10	1.36e-09	8.20e-10	2.12e-09
3	3.30e+04	1.00e+01	1.39e+07	6.98e-07	3.76e-07	3.86e-07	8.34e-07
4	1.77e+08	1.00e+01	2.18e+12	4.35e-02	1.01e-01	1.01e-01	2.62e-02
5	1.99e+01	1.00e+02	4.30e+02	3.76e-10	1.57e-10	1.58e-10	8.65e-11
6	1.52e+02	1.00e+02	7.24e+03	2.21e-09	2.78e-09	2.93e-09	7.61e-10
7	3.19e+04	1.00e+02	1.56e+07	3.33e-07	4.34e-07	5.32e-07	1.83e-06
8	1.81e+08	1.00e+02	5.55e+11	8.03e-03	7.97e-03	7.37e-03	9.75e-03
9	7.54e+01	1.00e+04	1.94e+05	4.17e-08	3.68e-08	3.68e-08	8.76e-10
1	6.84e+02	1.00e+04	1.12e+05	3.28e-07	4.17e-07	4.21e-07	1.39e-08
11	1.72e+04	1.00e+04	2.22e+06	5.68e-07	4.05e-07	3.88e-07	1.66e-07
12	4.51e+08	1.00e+04	4.53e+12	1.36e-01	4.87e-02	1.06e-01	4.11e-01
13	4.77e+01	1.00e+08	3.65e+13	9.21e-02	1.12e-05	1.12e-05	2.60e-10
14	2.08e+02	1.00e+08	2.78e+13	9.74e-01	2.50e-04	2.50e-04	1.73e-09
15	2.83e+04	1.00e+08	8.03e+09	4.42e-03	4.23e-03	4.23e-03	7.07e-07
16	1.00e+08	1.00e+08	3.99e+11	4.40e-01	5.51e-02	5.27e-02	1.47e-02

Table 4 Relative errors of the computed solutions with the $c_1 = 1e4$, $c_2 = 1$

#	$\kappa(A)$	$\kappa(B)$	$\kappa(M)$	$\frac{\|\mathbf{x}-\mathbf{x}_B\|_2}{\|\mathbf{x}\|_2}$	$\frac{\|\mathbf{x}-\mathbf{x}_{GC}\|_2}{\|\mathbf{x}\|_2}$	$\frac{\|\mathbf{x}-\mathbf{x}_{GH}\|_2}{\|\mathbf{x}\|_2}$	$\frac{\|\mathbf{x}-\mathbf{x}_{NP}\|_2}{\|\mathbf{x}\|_2}$
1	8.01e+01	1.00e+01	5.49e+03	5.70e-14	1.18e-14	4.14e-15	8.81e-14
2	3.73e+02	1.00e+01	3.71e+04	5.34e-14	1.52e-14	5.30e-15	2.18e-13
3	1.96e+04	1.00e+01	6.89e+06	5.56e-11	1.57e-13	3.76e-14	1.78e-11
4	1.07e+08	1.00e+01	2.08e+11	2.14e-06	5.95e-11	6.60e-11	2.12e-07
5	2.53e+01	1.00e+02	5.40e+02	9.74e-15	3.95e-15	4.14e-15	4.73e-15
6	3.01e+02	1.00e+02	2.60e+04	9.91e-14	1.11e-14	1.39e-14	1.89e-13
7	1.85e+04	1.00e+02	7.87e+06	9.90e-12	7.06e-14	5.52e-14	3.64e-11
8	2.73e+08	1.00e+02	8.04e+11	2.48e-06	9.97e-11	7.22e-11	1.07e-05
9	3.65e+01	1.00e+04	9.76e+05	2.65e-13	1.10e-13	1.09e-13	4.95e-14
10	3.50e+02	1.00e+04	1.19e+05	3.96e-13	3.90e-13	3.88e-13	3.29e-13
11	1.65e+04	1.00e+04	2.66e+06	9.56e-12	2.48e-13	2.57e-13	6.26e-12
12	2.85e+08	1.00e+04	4.67e+11	2.59e-07	2.42e-10	2.46e-10	3.03e-07
13	3.54e+01	1.00e+08	9.03e+13	4.20e-09	2.03e-10	2.03e-10	6.57e-10
14	1.00e+03	1.00e+08	8.46e+11	1.36e-09	3.06e-09	3.06e-09	3.48e-10
15	1.91e+04	1.00e+08	7.24e+10	1.77e-09	1.61e-09	1.61e-09	3.71e-10
16	1.25e+08	1.00e+08	9.02e+10	2.14e-07	5.38e-09	5.39e-09	3.52e-07

5 Conclusions

An algorithm for computing the solution of indefinite least squares problems and indefinite least squares problems with equality constraints is presented. The algorithm performs a similarity transformation on the associated augmented matrices to block anti-triangular form, relying only on Givens, and Householder transformations.

Some numerical examples are provided showing that the presented algorithm is numerically stable.

Appendix

We now describe how the linear system (3.9) can be solved when $q < n - s$.

Let s partition A_1, $\mathbf{s}$ and $\tilde{\mathbf{b}}$ as follows,

$$A_1 = \begin{bmatrix} A_{11} \\ A_{12} \end{bmatrix} \begin{matrix} \}p, \\ \}q, \end{matrix} \qquad \mathbf{s} = \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \mathbf{s}_3 \end{bmatrix} \begin{matrix} \}n-s, \\ \}p-n+s, \\ \}q, \end{matrix} \qquad \tilde{\mathbf{b}} = \begin{bmatrix} \tilde{\mathbf{b}}_1 \\ \tilde{\mathbf{b}}_2 \\ \tilde{\mathbf{b}}_3 \end{bmatrix} \begin{matrix} \}n-s, \\ \}p-n+s, \\ \}q. \end{matrix}$$

Compute the upper anti-QR factorization of A_{11}, $[Q_{11}, R_{11}] = \mathtt{aqru}(A_{11})$ and A_{12}, $[Q_{12}, R_{12}] = \mathtt{aqru}(A_{12})$.

Let

$$\tilde{Q}^{(2)} = \begin{bmatrix} Q_{11} & \\ & Q_{12} \end{bmatrix} \begin{matrix} \}p \\ \}q \end{matrix} \quad \text{and} \quad Q^{(2)} = \begin{bmatrix} \tilde{Q}^{(2)} & \\ & I_{n-s} \end{bmatrix}.$$

Then (3.9) is transformed into the equivalent linear system

$$M^{(3)}\mathbf{y}^{(3)} = \mathbf{f}^{(3)}, \tag{6.1}$$

where

$$M^{(3)} = Q^{(2)^T} M^{(2)} Q^{(2)} = \left[\begin{array}{c|c|c|c} I_{n-s} & & & R_{11} \\ & I_{p-n+s} & & \mathbf{0} \\ \hline & & -I_q & R_{12} \\ \hline R_{11}^T & \mathbf{0} & R_{12}^T & \mathbf{0} \end{array}\right],$$

$$\mathbf{y}^{(3)} = \begin{bmatrix} \mathbf{s}_1^{(1)} \\ \mathbf{s}_2^{(1)} \\ \mathbf{s}_3^{(1)} \\ \tilde{\mathbf{x}}_1 \end{bmatrix} = Q^{(2)^T} \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \\ \mathbf{s}_3 \\ \tilde{\mathbf{x}}_1 \end{bmatrix} = \begin{bmatrix} Q_{11}^T \begin{bmatrix} \mathbf{s}_1 \\ \mathbf{s}_2 \end{bmatrix} \\ Q_{12}^T \mathbf{s}_3 \\ \tilde{\mathbf{x}}_1 \end{bmatrix}, \tag{6.2}$$

$$\mathbf{f}^{(3)} = \begin{bmatrix} \mathbf{f}_1^{(3)} \\ \mathbf{f}_2^{(3)} \\ \mathbf{f}_3^{(3)} \\ \mathbf{0} \end{bmatrix} = Q^{(2)^T} \begin{bmatrix} \tilde{\mathbf{b}}_1 \\ \tilde{\mathbf{b}}_2 \\ \tilde{\mathbf{b}}_3 \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} Q_{11}^T \begin{bmatrix} \tilde{\mathbf{b}}_1 \\ \tilde{\mathbf{b}}_2 \end{bmatrix} \\ Q_{12}^T \tilde{\mathbf{b}}_3 \\ \mathbf{0} \end{bmatrix}.$$

De to the structure of $M^{(3)}$, from (6.1) we can compute $\mathbf{s}_2^{(1)}$,

$$\mathbf{s}_2^{(1)} = \mathbf{f}_2^{(3)}, \tag{6.3}$$

and “shrink” (6.1) to

$$\tilde{M}^{(3)} \tilde{\mathbf{y}}^{(3)} = \tilde{\mathbf{f}}^{(3)}, \tag{6.4}$$

with

$$\tilde{M}^{(3)} = \left[\begin{array}{cc|c} I_{n-s} & & R_{11} \\ \hline & -I_q & R_{12} \\ \hline R_{11}^T & R_{12}^T & \mathbf{0} \end{array}\right], \qquad \tilde{\mathbf{y}}^{(3)} = \begin{bmatrix} \mathbf{s}_1^{(1)} \\ \mathbf{s}_3^{(1)} \\ \tilde{\mathbf{x}}_1 \end{bmatrix}, \qquad \tilde{\mathbf{f}}^{(3)} = \begin{bmatrix} \mathbf{f}_1^{(3)} \\ \mathbf{f}_3^{(3)} \\ \mathbf{0} \end{bmatrix},$$

with Inertia$(\tilde{M}^{(3)}) = (q+n-s, 0, p) - (0, 0, p-n+s) = (q+n-s, 0, n-s)$.

The matrix $\tilde{M}^{(3)}$ can be reduced to upper block antitriangular form by a sequence of $n-s$ Householder transformations.

Let $\tilde{M}_0^{(3)} = \tilde{M}^{(3)}$.

At step i, $i = 1, \dots, q-1$, the matrix $\tilde{M}_{i-1}^{(3)} = H_{i-1} \cdots H_1 \tilde{M}_0^{(3)} H_1^T \cdots H_{i-1}^T$, is multiplied to the left by a Householder matrix $H_i \in \mathbb{R}^{(2(n-s)+q) \times (2(n-s)+q)}$ and to the right by the transpose of H_i, such that

$$\tilde{M}_i^{(3)} = H_i \tilde{M}_{i-1}^{(3)} H_i^T$$

has the rows (columns) $i, n-s+1, n-s+2, \dots, n-s+i$ modified and the entries $n-s+1, n-s+2, \dots, n-s+i$ in column (row) $2(n-s)+q-i+1$ annihilated.

Furthermore, at step i, $i = q, \dots, n-s$, the matrix $\tilde{M}_{i-1}^{(3)}$ is multiplied to the left by a Householder matrix $H_i \in \mathbb{R}^{(2(n-s)+q) \times (2(n-s)+q)}$ and to the right by the transpose of H_i, such that

$$\tilde{M}_i^{(3)} = H_i \tilde{M}_{i-1}^{(3)} H_{i-1}^T$$

has the rows (columns) $i, n-s+1, n-s+2, \dots, n-s+q$ modified and the entries $n-s+1, n-s+2, \dots, n-s+q$ in column (row) $2(n-s)+q-i+1$ annihilated.

Let $Q^{(3)} = H_1^T H_2^T \cdots H_{n-s}^T \in \mathbb{R}^{(2(n-s)+q) \times (2(n-s)+q)}$.

Then the linear system (6.4) is transformed into the equivalent one

$$M^{(4)} \mathbf{y}^{(4)} = \mathbf{f}^{(4)}, \tag{6.5}$$

with $M^{(4)}$ having the following structure,

$$M^{(4)} = Q^{(3)^T} \tilde{M}^{(3)} Q^{(3)} = \begin{bmatrix} W^{(4)} & Z^{(4)} & Y^{(4)} \\ Z^{(4)^T} & X^{(4)} & \\ Y^{(4)^T} & & \end{bmatrix} \begin{matrix} \}n-s, \\ \}q, \\ \}n-s, \end{matrix}$$

$Y^{(4)} \in \mathbb{R}^{(n-s)\times(n-s)}$ nonsingular upper anti-triangular, $X^{(4)}, W^{(4)} \in \mathbb{R}^{(n-s)\times(n-s)}$ symmetric, and

$$\mathbf{y}^{(4)} = \begin{bmatrix} \mathbf{y}_1^{(4)} \\ \mathbf{y}_2^{(4)} \\ \mathbf{y}_3^{(4)} \end{bmatrix} = Q^{(3)^T} \tilde{\mathbf{y}}^{(3)} = Q^{(3)^T} \begin{bmatrix} \mathbf{s}_1^{(1)} \\ \mathbf{s}_3^{(1)} \\ \tilde{\mathbf{x}}_1 \end{bmatrix}, \tag{6.6}$$

$$\mathbf{f}^{(4)} = \begin{bmatrix} \mathbf{f}_1^{(4)} \\ \mathbf{f}_2^{(4)} \\ \mathbf{f}_3^{(4)} \end{bmatrix} = Q^{(3)^T} \tilde{\mathbf{f}}^{(3)} = Q^{(3)^T} \begin{bmatrix} \mathbf{f}_1^{(3)} \\ \mathbf{f}_3^{(3)} \\ \mathbf{0} \end{bmatrix}.$$

Observe that $\mathbf{y}_3^{(4)} = \tilde{\mathbf{x}}_1$ and $\mathbf{f}_3^{(4)} = \mathbf{0}$, because of the structure of $Q^{(3)}$. Since $\text{Inertia}(M^{(4)}) = \text{Inertia}(\tilde{M}^{(3)}) = (q+n-s, 0, n-s)$, then, by [6], the submatrix $X^{(4)}$ of $M^{(4)}$ is symmetric negative definite with Cholesky factorization $X^{(4)} = -L^{(4)} L^{(4)^T}$, $L^{(4)} \in \mathbb{R}^{q\times q}$ nonsingular lower triangular.

We can now solve the linear system (3.17) in the following steps.

- Observe that $\mathbf{y}_1^{(4)} = \mathbf{0}$, since $\mathbf{f}_3^{(4)} = \mathbf{0}$. Therefore the $\mathbf{y}_1^{(4)} = \mathbf{0}$ is the solution of the upper anti-triangular linear system

$$Y^{(4)^T} \mathbf{y}_1^{(4)} = \mathbf{f}_3^{(4)};$$

- update the right-hand-side

$$\begin{bmatrix} \tilde{\mathbf{f}}_1^{(4)} \\ \tilde{\mathbf{f}}_2^{(4)} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_1^{(4)} \\ \mathbf{f}_2^{(4)} \end{bmatrix} - \begin{bmatrix} W^{(4)} \\ Y^{(4)^T} \end{bmatrix} \mathbf{y}_1^{(4)};$$

- solve the linear system $X^{(4)} \mathbf{y}_2^{(4)} = \tilde{\mathbf{f}}_2^{(4)}$,

$$L^{(4)} \mathbf{t} = -\tilde{\mathbf{f}}_2^{(4)}$$
$$L^{(4)^T} \mathbf{y}_2^{(4)} = \mathbf{t};$$

- solve the upper anti-triangular linear system

$$Y^{(4)} \mathbf{y}_3^{(4)} = Y^{(4)} \tilde{\mathbf{x}}_1 = \tilde{\mathbf{f}}_1^{(4)} - Z^{(4)} \mathbf{y}_2^{(4)}.$$

Once $\tilde{\mathbf{x}}_1$ is computed, the solution $\mathbf{x}$ of the problem (1.4) can be obtained as

$$\mathbf{x} = \tilde{Q}^{(1)} \begin{bmatrix} \tilde{\mathbf{x}}_1 \\ \tilde{\mathbf{x}}_2 \end{bmatrix}.$$

If one is also interested in the computation of the solution of the augmented system (1.7), from (6.6) $\tilde{\mathbf{y}}^{(3)} = \begin{bmatrix} \mathbf{s}_1^{(1)} \\ \mathbf{s}_3^{(1)} \\ \tilde{\mathbf{x}}_1 \end{bmatrix} = Q^{(3)}\mathbf{y}^{(4)}$ can be computed and, therefore, $\mathbf{y}^{(3)}$, since, by (6.3), $\mathbf{s}_2^{(1)}$ is already computed. Finally, from (6.2), $\mathbf{y}^{(2)} = Q^{(2)}\mathbf{y}^{(3)}$ can be computed.

About the computational complexity of this step, the computation of R_{11} and R_{12} requires $2(n-s)^2(p-(n-s)/3)$ and $2q^2(q+n-s)$ floating point operations, respectively. Moreover, the computation of $Y^{(4)}$ requires $4q(n-s-q)^2 + 2q^2(n-s) - 2/3q^3$ floating point operations.

References

1. Bojanczyk, A., Higham, N.J., Patel, H.: The equality constrained indefinite least squares problem: theory and algorithms. BIT Numer. Math. **41**(3), 505–517 (2003)
2. Bojanczyk, A., Higham, N.J., Patel, H.: Solving the indefinite least squares problem by hyperbolic QR factorization. SIAM J. Matrix Anal. Appl. **24**(4), 914–931 (2003)
3. Chandrasekaran, S., Gu, M., Sayed, A.H.: A stable and efficient algorithm for the indefinite linear least squares problem. SIAM J. Matrix Anal. Appl. **20**, 354–362 (1998)
4. Golub, G.H., Van Loan, C.F.: An analysis of the total least squares problem. SIAM J. Numer. Anal. **17**, 883–893 (1980)
5. Golub, G.H., Van Loan, C.F.: Matrix Computations, 3rd edn. Johns Hopkins University Press, Baltimore (1996)
6. Gould, N.I.M.: On practical conditions for the existence and uniqueness of solutions to the general equality quadratic programming problem. Math. Program. **32**, 90–99 (1985)
7. Hassibi, B., Sayed, A.H., Kailath, T.: Recursive linear estimation in Krein spaces–part I: theory. IEEE Trans. Autom. Control **41**, 18–33 (1996)
8. Higham, N.J.: J-Orthogonal matrices: properties and generation. SIAM Rev. **45**(3), 504–519 (2003)
9. Li, Q., Wang, M.: Algebraic properties and perturbation results for the indefinite least squares problem with equality constraints. Int. J. Comput. Math. **87**(2), 425–434 (2010)
10. Li, Q., Pan, B.Z., Wang, Q.: The hyperbolic elimination method for solving the equality constrained indefinite least squares problem. Int. J. Comput. Math. **87**(13), 2953–2966 (2010)
11. Mastronardi, N., Van Dooren, P.: The anti–triangular factorization of symmetric matrices. SIAM J. Matrix Anal. Appl. **34**(1), 173–196 (2013)
12. Mastronardi, N., Van Dooren, P.: Recursive approximation of the dominant eigenspace of an indefinite matrix. J. Comput. Appl. Math. **236**(16), 4090–4104 (2012)
13. Mastronardi, N., Van Dooren, P.: A structurally backward stable algorithm for solving the indefinite least squares problem with equality constraints. IMA J. Numer. Anal. (submitted)
14. Van Huffel, S., Vandewalle, J.: The Total Least Squares Problem: Computational Aspects and Analysis. SIAM, Philadelphia (1991)