ON A QUADRATIC EIGENPROBLEM OCCURRING IN REGULARIZED TOTAL LEAST SQUARES

JÖRG LAMPE AND HEINRICH VOSS*

Key words. total least squares, regularization, quadratic eigenvalue problem, maxmin characterization

AMS subject classification. 65F15, 65F22

Abstract. In a recent paper Sima, Van Huffel and Golub [*Regularized total least squares based on quadratic eigenvalue problem solvers. BIT Numerical Mathematics 44, 793 - 812 (2004)*] suggested a computational approach for solving regularized total least squares problems via a sequence of quadratic eigenvalue problems. Taking advantage of a variational characterization of real eigenvalues of nonlinear eigenproblems we prove the existence of a right most eigenvalue. For large problems we improve the approach of Sima et al. considerably using thick and early updates in a nonlinear Arnoldi method.

1. Introduction. Many problems in data estimation are governed by overdetermined linear systems

$$Ax \approx b, \quad A \in \mathbb{R}^{m\times n},\ b \in \mathbb{R}^m, m \geq n \tag{1.1}$$

where both the matrix A and the right hand side b are contaminated by some noise. An appropriate approach to this problem is the total least squares (TLS) method which determines perturbations $\Delta A \in \mathbb{R}^{m\times n}$ to the coefficient matrix and $\Delta b \in \mathbb{R}^m$ to the vector b such that

$$\|[\Delta A, \Delta b]\|_F^2 = \text{min!} \quad \text{subject to } (A + \Delta A)x = b + \Delta b \tag{1.2}$$

where $\|\cdot\|_F$ denotes the Frobenius norm of a matrix (cf. [7, 17]).

In this paper we consider ill-conditioned problems which arise, for example, from the discretization of ill-posed problems such as integral equations of the first kind (cf. [4, 8, 11]). Then least squares or total least squares methods for solving (1.1) often yield physically meaningless solutions, and regularization is necessary to stabilize the solution.

Motivated by Tikhonov regularization a well established approach is to add a quadratic constraint to problem (1.2) yielding the regularized total least squares (RTLS) problem

$$\|[\Delta A, \Delta b]\|_F^2 = \text{min!} \quad \text{subject to } (A + \Delta A)x = b + \Delta b,\ \|Lx\| \leq \delta. \tag{1.3}$$

Here δ is a regularization parameter, and $L \in \mathbb{R}^{k\times n}$, $k \leq n$ defines a (semi-) norm on the solution through which the size of the solution is bounded or a certain degree of smoothness can be imposed on the solution. Stabilization by introducing a quadratic constraint was extensively studied in [3, 6, 9, 14, 15, 16]. Tikhonov regularization was considered in [2].

Based on the singular value decomposition of $[A, b]$ methods were developed for solving the TLS problem (1.2) [7, 17], and even a closed formula for its solution is known if $[A, b]$ has full column rank. However, this approach can not be generalized

*Institute of Numerical Simulation, Hamburg University of Technology, D-21071 Hamburg, Germany ({joerg.lampe,voss}@tu-harburg.de)

to the RTLS problem (1.3). Golub, Hansen and O'Leary [6] presented and analyzed the properties of regularization of TLS. Guo and Renaut [9, 14] took advantage of the fact, that the RTLS problem (1.3) is equivalent to the minimization of the Rayleigh quotient of the augmented matrix $[A,b]^T[A,b]$ subject to the regularization constraint, and suggested an algorithm based on inverse iteration. Inspired by the fact that quadratically constrained least squares problems can be solved by a quadratic eigenvalue problem [5], Sima, Van Huffel, and Golub [15, 16] developed an iterative method for solving (1.3), where in each step the right most eigenvalue and corresponding eigenvector of a quadratic eigenproblem has to be determined. Beck, Ben-Tal, and Teboulle [3] proved that the (non–convex) RTLS problem (1.3) has a global optimal solution, and they derived an efficient method for its solution by solving a sequence of convex minimization problems parametrized by a single parameter.

In this paper we investigate the method introduced in [16]. Taking advantage of a maxmin principle for non–overdamped nonlinear eigenvalue problems we derive a characterization of the right most eigenvalue of the appertaining quadratic eigenvalue problems. For large dimensions n we study iterative projection methods for the sequence of quadratic eigenproblems. Reusing informations from the previous quadratic problems and early updates (i.e. inexact solutions of the quadratic problems) we arrive at a substantial improvement of the method in [16].

The paper is organized as follows. In the next section we briefly summarize the mathematical formulation of the RTLS problem and the approach of Sima, Van Huffel, and Golub for solving it. Section 3 contains a short review of the maxmin characterization of extreme real eigenvalues of symmetric nonlinear eigenvalue problems, and in Section 4 we derive a characterization of the right most eigenvalue of the quadratic eigenvalue problems which occur in the method suggested in [16]. For small dimensions the quadratic eigenvalue problems can be solved by linearization. For problems of large dimension we discuss in Section 5 the use of two Krylov subspace solvers introduced by Li and Ye [13] and Bai and Su [1], and the nonlinear Arnoldi method [18]. A numerical example from the "Regularization Tools" [10] is presented in Section 6 demonstrating that the formulation of the RTLS problem using quadratic eigenvalue problems is a powerful approach.

2. Regularized total least squares. We briefly summarize the mathematical formulation of the RTLS problem and the approach of Sima, Van Huffel, and Golub for solving it. It is well known (cf. [17], and [3] for a different derivation) that the RTLS problem (1.3) is equivalent to

$$\frac{\|Ax-b\|^2}{1+\|x\|^2} = \min! \quad \text{subject to } \|Lx\|^2 \le \delta^2. \tag{2.1}$$

We assume that the regularization parameter $\delta > 0$ is less than $\|Lx_{TLS}\|$, where x_{TLS} denotes the solution of the total least squares problem (1.2) (otherwise no regularization would be necessary). Then the constraint $\|Lx\| \le \delta$ is active, and we may replace (2.1) by

$$\frac{\|Ax-b\|^2}{1+\|x\|^2} = \min! \quad \text{subject to } \|Lx\|^2 = \delta^2, \tag{2.2}$$

the first order optimality conditions of which are

$$B(x)x + \lambda L^T Lx = d(x), \quad \|Lx\|^2 = \delta^2 \tag{2.3}$$

with

$$B(x) = \frac{1}{1+\|x\|^2}\Big(A^TA - f(x)I)\Big),\ f(x) = \frac{\|Ax-b\|^2}{1+\|x\|^2},\ d(x) = \frac{A^Tb}{1+\|x\|^2}. \tag{2.4}$$

Sima, Van Huffel and Golub [16] suggested Algorithm 1 called RTLSQEP for solving (2.3), and they proved its convergence if the initial vector is chosen suitably.

Algorithm 1 RTLSQEP

1: Let x be an initial vector. Compute $B = B(x)$ and $d = d(x)$
2: set $r = 1$
3: **while** $r >$ tol **do**
4: Solve

$$Bx + \lambda L^TLx = d, \quad \|Lx\| = \delta \tag{2.5}$$

for (x, λ) corresponding to the largest $\lambda \in \mathbb{R}$
5: compute $B = B(x)$ and $d = d(x)$
6: compute $r = \|Bx + \lambda L^TLx - d\|$
7: **end while**

The first order conditions (2.5) were suggested to be solved via a quadratic eigenvalue problem in every step which motivates the name RTLSQEP of the algorithm.

If L is quadratic and nonsingular, then with $z = Lx$ problem (2.5) is equivalent to

$$Wz + \lambda z := L^{-T}BL^{-1}z + \lambda z = L^{-T}d, \quad z^Tz = \delta^2. \tag{2.6}$$

Assuming that $W + \lambda I$ is positive definite, and denoting by $u = (W + \lambda I)^{-2}h$, one gets $h^Tu = z^Tz = \delta^2$, and $h = \delta^{-2}hh^Tu$ yields that $(W + \lambda I)^2u = h$ is equivalent to the quadratic eigenvalue problem

$$(W + \lambda I)^2u - \delta^{-2}hh^Tu = 0. \tag{2.7}$$

If λ is the right most eigenvalue of (2.7) and u is a corresponding eigenvector scaled such that $h^Tu = \delta^2$, then the solution of the original problem (2.5) is recovered from $z = (W + \lambda I)u$, and $x = L^{-1}z$.

For $k < n$, and $\operatorname{rank}(L) = k$ let $L^TL = USU^T$ be the spectral decomposition of L^TL. Then (2.5) is equivalent to

$$\frac{1}{1+\|x\|^2}\Big((AU)^T(AU) - f(x)I\Big)y + \lambda Sy = \frac{(AU)^Tb}{1+\|x\|^2}, \quad y^TSy = \delta^2 \tag{2.8}$$

with $y = U^Tx$. Partitioning the matrices and vectors in (2.8) in block form

$$(AU)^T(AU) = \begin{pmatrix} T_1 & T_2 \\ T_2^T & T_4 \end{pmatrix},\ S = \begin{pmatrix} S_1 & 0 \\ 0 & 0 \end{pmatrix},\ (AU)^Tb = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix},\ y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \tag{2.9}$$

where the leading blocks have dimension k, one gets

$$\frac{1}{1+\|x\|^2}\begin{pmatrix} T_1 - f(x)I_k & T_2 \\ T_2^T & T_4 - f(x)I_{n-k} \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \lambda\begin{pmatrix} S_1y_1 \\ 0 \end{pmatrix} = \frac{1}{1+\|x\|^2}\begin{pmatrix} c_1 \\ c_2 \end{pmatrix}. \tag{2.10}$$

Solving the second component for y_2

$$y_2 = (T_4 - f(x)I_{n-k})^{-1}(c_2 - T_2^T y_1)$$

and substituting in the first component one gets

$$\frac{1}{1+\|x\|^2}\Big(T_1 - f(x)I_k - T_2(T_4 - f(x)I_{n-k})^{-1}T_2^T\Big)y_1 + \lambda S_1 y_1 = \frac{1}{1+\|x\|^2}(c_1 - T_2(T_4 - f(x)I_{n-k})^{-1}c_2). \tag{2.11}$$

Hence, problem (2.8) is equivalent to the quadratic eigenvalue problem (2.7), where

$$W = \frac{1}{1+\|x\|^2}S_1^{-1/2}\Big(T_1 - f(x)I_k - T_2(T_4 - f(x)I_{n-k})^{-1}T_2^T\Big)S_1^{-1/2} \tag{2.12}$$

$$h = \frac{1}{1+\|x\|^2}S_1^{-1/2}\Big(c_1 - T_2(T_4 - f(x)I_{n-k})^{-1}c_2\Big). \tag{2.13}$$

If (λ, u) is the eigenpair corresponding to the right most eigenvalue and u is normalized such that $u^T h = \delta^2$, and $z = (W + \lambda I)u$, then the solution of (2.5) is recovered by $x = Uy$ where

$$y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} S_1^{-1/2}z \\ (T_4 - f(x)I)^{-1}(c_2 - T_2^T S_1^{-1/2}z) \end{pmatrix}. \tag{2.14}$$

3. A maxmin principle for nonlinear eigenproblems. In this section we provide a maxmin result for symmetric nonlinear eigenvalue problems which generalizes the well known variational characterization of Poincaré for linear problems and which will be used in Section 4 to prove the existence of a rightmost eigenvalue of the quadratic eigenproblem (2.7).

We consider the nonlinear eigenvalue problem

$$T(\lambda)x = 0 \tag{3.1}$$

where $T(\lambda) \in \mathbb{R}^{n\times n}$ is a family of real symmetric matrices for every λ in an open real interval J which may be unbounded.

For a linear symmetric and positive definite problem $Ax = \lambda Bx$ all eigenvalues are real. If they are ordered by magnitude regarding their multiplicity $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$, then it is well known that they can be characterized by the maxmin principle of Poincaré

$$\lambda_j = \max_{V\in S_j}\ \min_{x\in V,\ x\neq 0}\ \frac{x^H Ax}{x^H Bx}, \quad j = 1, 2, \ldots, n. \tag{3.2}$$

Here S_j denotes the set of all j dimensional subspaces of $\mathbb{C}^n$.

Similar results also hold for certain nonlinear, symmetric eigenvalue problems (cf. [19]). We assume that for every fixed $x \in \mathbb{C}^n\backslash\{0\}$ the real function $f(\lambda; x) := x^H T(\lambda)x$ is continuously differentiable in J, and that the real equation

$$f(\lambda; x) = 0 \tag{3.3}$$

has at most one solution in J. Then equation (3.3) implicitly defines a functional p on some subset D of $\mathbb{R}^n \setminus \{0\}$. For a linear problem $T(\lambda) := \lambda B - A$ this is exactly the Rayleigh quotient, and we therefore call p the Rayleigh functional of problem (3.1).

We assume further that

$$x^H T'(p(x))x > 0 \quad \text{for every } x \neq 0 \tag{3.4}$$

which generalizes the definiteness requirement for B in the linear case.

The following Theorem is a special case of the general maxmin principle for nonlinear eigenproblems proved in [19] (cf. Theorems 2.1 and 2.9 in [19]):

THEOREM 3.1. *Assume that the general conditions given above are satisfied, and suppose that*

$$\sup_{x \in D} \; p(x) \in J. \tag{3.5}$$

Then the following assertions hold:

(i) Problem (3.1) *has m eigenvalues in J if and only if there is an m dimensional subspace $V \in S_m$ which is contained in $D \cup \{0\}$.*

(ii) The eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_m$, $m \leq n$, ordered by magnitude satisfy the maxmin characterization

$$\lambda_j = \max_{V \in S_j, V \cap D \neq \emptyset} \; \min_{x \in V \cap D} \; p(x), \quad j = 1, 2, \ldots, m. \tag{3.6}$$

(iii) $\tilde{\lambda}$ is the jth largest eigenvalue of (3.1) *in J if and only if 0 is the jth smallest eigenvalue of the linear eigenproblem $T(\tilde{\lambda})y = \mu y$.*

(iv) The maximum in (3.6) *is attained for the invariant subspace of $T(\lambda_j)$ corresponding to its jth smallest eigenvalues.*

4. The quadratic eigenproblem occurring in RTLSQEP. We consider the quadratic eigenvalue problem

$$T(\lambda)x = \Big((W + \lambda I)^2 - \delta^{-2} h h^T\Big)x \tag{4.1}$$

where $W \in \mathbb{R}^{k \times k}$ is a symmetric matrix and $h \in \mathbb{R}^k$.

In this case

$$f(\lambda, x) = x^H T(\lambda) x = \lambda^2 \|x\|_2^2 + 2\lambda x^H W x + \|Wx\|_2^2 - |x^H h|^2/\delta^2, \; x \neq 0 \tag{4.2}$$

is a parabola which attains its minimum at

$$\lambda = -\frac{x^H W x}{x^H x}. \tag{4.3}$$

Hence, we choose $J = (-\lambda_{\min}, \infty)$ where $\lambda_{\min}$ is the minimum eigenvalue of W.

Then $f(\lambda, x) = 0$ has at most one solution $p(x) \in J$ for every $x \neq 0$, and the Rayleigh functional p of (4.1) corresponding to J is defined. Obviously it holds that $x^H T'(p(x))x > 0$ for every $x \in D$, and the general conditions of the maxmin characterization are satisfied.

Let $x_{\min}$ be an eigenvector of W corresponding to $\lambda_{\min}$. Then

$$f(-\lambda_{\min}, x_{min}) = x_{\min}^H (W - \lambda_{\min} I)^2 x_{\min} - |x_{\min}^H h|^2/\delta^2 = -|x_{\min}^H h|^2/\delta^2 \leq 0$$

Hence, if $x_{min}^H h \neq 0$ then $x_{\min} \in D$.

If $x_{min}^H h = 0$, and the minimum eigenvalue $\mu_{\min}$ of $T(-\lambda_{\min})$ is negative, then for the corresponding eigenvector y_{min} it holds that

$$f(-\lambda_{\min}, y_{\min}) = y_{\min}^H T(-\lambda_{\min}) y_{\min} = \mu_{min} \|y_{\min}\|_2^2 < 0,$$

and $y_{\min} \in D$.

Finally, if $x_{min}^H h = 0$, and $T(-\lambda_{\min})$ is positive semi-definite, then

$$f(-\lambda_{\min}, x) = x^H T(-\lambda_{\min}) x \geq 0 \quad \text{for every } x \neq 0,$$

and $D = \emptyset$. In this case $-\lambda_{\min}$ is the largest real eigenvalue of (4.1).

Assume that $D \neq \emptyset$. For $x^H h = 0$ it holds that

$$f(\lambda, x) = \|(W + \lambda I)x\|_2^2 > 0 \quad \text{for every } \lambda \in J,$$

i.e. $x \notin D$. Hence, D does not contain a two-dimensional subspace of $\mathbb{C}^n$, and therefore J contains at most one eigenvalue of (4.1).

If $\lambda \in \mathbb{C}$ is a non-real eigenvalue of (4.1) and x a corresponding eigenvector, then

$$x^H T(\lambda) x = \lambda^2 \|x\|_2^2 + 2\lambda x^H W x + \|Wx\|_2^2 - |x^H h|^2/\delta^2 = 0.$$

Hence, the real part of λ satisfies

$$\text{real}(\lambda) = -\frac{x^H W x}{x^H x} \leq -\lambda_{\min}.$$

Thus we have proved the following characterization of the rightmost eigenvalue of problem (4.1):

THEOREM 4.1. *Let $\lambda_{\min}$ be the minimal eigenvalue of W, and $x_{\min}$ be a corresponding eigenvector. Let $J = (-\lambda_{\min}, \infty)$, and denote by p the Rayleigh functional of $T(\cdot)$ and by D its domain of definition.*

(i) If $x_{\min}^T h = 0$ and $T(-\lambda_{min})$ is positive semi-definite, then $\hat{\lambda} := -\lambda_{\min}$ is the maximal real eigenvalue of (4.1) and $x_{\min}$ is a corresponding eigenvector.

(ii) Otherwise, the maximal real eigenvalue is the unique eigenvalue $\hat{\lambda}$ of (4.1) in J, and it holds that

$$\hat{\lambda} = \max_{x \in D} \; p(x). \tag{4.4}$$

(iii) $\hat{\lambda}$ is the right most eigenvalue of (4.1), i.e.

$$\textit{real}(\lambda) \leq -\lambda_{\min} \leq \hat{\lambda} \quad \textit{for every eigenvalue } \lambda \textit{ of } (4.1).$$

In [16] it is remarked that the special structure of the quadratic eigenvalue problem (4.1) enforces the right-most eigenvalue to be real and positive. This is not true in general. If W is positive definite with eigenvalue $\lambda_j > 0$, then $-\lambda_j$ are the only eigenvalues of the quadratic eigenproblem $(W + \lambda I)^2 x = 0$, and if the term $\delta^{-2} h h^T$ is small enough, then problem (4.1) will have no positive eigenvalue. However, Theorem 4.1 demonstrates that there always exists a real right-most eigenvalue.

5. Computational considerations. An obvious approach for solving the quadratic eigenvalue problems

$$T_j(\lambda)u := \Big((W_j + \lambda I)^2 - \delta^{-2} h_j h_j^T\Big)u = 0 \tag{5.1}$$

at the j-th iteration step of RTLSQEP is linearization, i.e. solving the liner eigenproblem

$$\begin{pmatrix} -2W_j & -W_j^2 + \delta^{-2} h_j h_j^T \\ I & 0 \end{pmatrix} \begin{pmatrix} v \\ u \end{pmatrix} = \lambda \begin{pmatrix} v \\ u \end{pmatrix} \tag{5.2}$$

and choosing the maximal real eigenvalue, and the corresponding u–part of the eigenvector, which is an eigenvector of (5.1). This approach is reasonable if the dimension n of problem (2.2) is small.

For larger dimensions it is not efficient to determine the entire spectrum of (5.2), and to choose the eigenpair that is needed afterwards. In this case one could apply the implicitly restarted Arnoldi method implemented in ARPACK [12] (and included in MATLAB as function eigs) to determine the right most eigenvalue and corresponding eigenvector of (5.2). However, it is a drawback of linearization that symmetry properties that the quadratic problem has and which could be used in its numerical solution are destroyed.

Quadratic eigenvalue problems of large problem size occurring in the RTLSQEP algorithm can be solved efficiently by Krylov subspace based projection methods [1, 13] or the nonlinear Arnoldi method [18].

Li and Ye [13] presented a Krylov subspace projection method for a monic quadratic eigenproblem $(\lambda^2 I - \lambda P_1 - P_0)x = 0$ which does not use a linearization but works with the matrices P_1 and P_0 directly. The method has particularly favorable properties if some linear combination of P_1 and P_0 is a matrix of small rank q. Then with an Arnoldi-type process a matrix $Q \in \mathbb{R}^{n \times \ell+q+1}$ with orthonormal columns and two matrices $H_1 \in \mathbb{R}^{\ell+q+1 \times \ell}$ and $H_0 \in \mathbb{R}^{\ell+q+1 \times \ell}$ with lower bandwidth $q+1$ are determined such that

$$P_1 Q(:, 1:\ell) = Q(:, 1:\ell+q+1)H_1 \quad \text{and} \quad P_0 Q(:, 1:\ell) = Q(:, 1:\ell+q+1)H_0.$$

Then approximations to eigenpairs of the quadratic eigenproblem are obtained from its orthogonal projection onto $\operatorname{span}\{Q(:, 1:\ell)\}$ which reads

$$(\lambda^2 I_k - \lambda H_1(1:\ell, :) - H_0(1:\ell, :))z = 0.$$

For the quadratic eigenproblem (5.1) the algorithm of Li and Ye is applied with $P_1 = -W_j$ and $P_0 = -h_j h_j^T$ from which one gets the quadratic eigenproblem

$$(\theta^2 I - 2\theta H_1(1:\ell, :) - H_1(1:\ell+2, :)^T H_1(1:\ell+2, :) + \delta^{-2} H_0(1:\ell, :))z = 0, \tag{5.3}$$

which is an approximation to the orthogonal projection of (5.1) onto $\operatorname{span}\{Q(:, 1:\ell)\}$.

The Second Order Arnoldi Reduction (SOAR for short) introduced by Bai and Su [1] is based on the observation that the Krylov space of the linearization

$$\begin{pmatrix} P_1 & P_0 \\ I & O \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} = \lambda \begin{pmatrix} y \\ z \end{pmatrix}$$

of $(\lambda^2 I - \lambda P_1 - P_0)z = 0$ with initial vector $\begin{pmatrix} r_0 \\ 0 \end{pmatrix}$ has the form

$$\mathcal{K}_\ell = \left\{ \begin{pmatrix} r_0 \\ 0 \end{pmatrix}, \begin{pmatrix} r_1 \\ r_0 \end{pmatrix}, \begin{pmatrix} r_2 \\ r_1 \end{pmatrix}, \ldots, \begin{pmatrix} r_{\ell-1} \\ r_{\ell-2} \end{pmatrix} \right\}$$

where

$$\begin{aligned} r_1 &= P_1 r_0 \\ r_j &= P_1 r_{j-1} + P_0 r_{j-2}, \text{ for } j \geq 2. \end{aligned} \tag{5.4}$$

The entire information on $\mathcal{K}_\ell$ is therefore contained in the Second Order Krylov Space

$$\mathcal{G}_\ell(P_1, P_0) = \operatorname{span}\{r_0, r_1, \ldots, r_{\ell-1}\}.$$

Bai and Su [1] presented an Arnoldi type algorithm based on the two term recurrence (5.4) for computing an orthonormal basis $Q_\ell \in \mathbb{R}^{n\times\ell}$ of $\mathcal{G}_\ell(P_1, P_0)$, and the orthogonal projection of (5.1) onto $\mathcal{G}_\ell(P_1, P_0)$ then yields approximate solutions of the quadratic eigenproblem.

A further method for solving (5.1) which does not destroy its structure is the nonlinear Arnoldi method [18] in Algorithm 2. which applies to much more general nonlinear eigenvalue problems.

Algorithm 2 Nonlinear Arnoldi

start with initial basis V, $V^TV = I$
determine preconditioner $M \approx T(\sigma)^{-1}$, σ close to wanted eigenvalue
find largest real eigenvalue μ of $V^TT(\mu)Vy = 0$ and corresponding eigenvector y
set $u = Vy$, $r = T(\mu)u$
while $\|r\|/\|u\| > \epsilon$ **do**
  $v = Mr$
  $v = v - VV^Tv$
  reorthogonalize if necessary
  $\tilde{v} = v/\|v\|$, $V = [V, \tilde{v}]$
  find smallest eigenvalue μ of $V^TT(\mu)Vy = 0$ and corresponding eigenvector y
  set $u = Vy$, $r = T(\mu)u$
end while

In the RTLSQEP algorithm a sequence of quadratic eigenvalue problems has to be solved, and the convergence of the matrices and vectors

$$W_j = \frac{1}{1+\|x_j\|^2}\Big(C - f(x_j)\tilde{S} - D(T_4 - f(x_j)I_{n-k})^{-1}D^T\Big) \tag{5.5}$$

$$h_j = \frac{1}{1+\|x_j\|^2}\Big(g_1 - D(T_4 - f(x_j)I_{n-k})^{-1}c_2\Big). \tag{5.6}$$

with $C := S_1^{-1/2}T_1S_1^{-1/2}$, $\tilde{S} = S_1^{-1}$, $D = S_1^{-1/2}T_2$ and $g_1 = S_1^{-1/2}c_1$ suggest to reuse information from the previous steps when solving problem (5.1) in step j.

For the two Krylov subspace methods the only degree of freedom is the choice of the initial vector, and we therefore start in step j with the solution u_{j-1} of the preceding step. The nonlinear Arnoldi method allows thick starts, i.e. solving $T_j(\lambda)u = 0$ in

step j of RTLSQEP we start Algorithm 2 with the orthonormal basis V that was used in the preceding step when determining the solution $u_{j-1} = Vz$ of $V^T T_{j-1}(\lambda)Vz = 0$.

The projected problem

$$V^T T_j(\lambda)Vz = ((W_j + \lambda I)V)^T((W_j + \lambda I)V)z - \delta^{-2}(h_j^T V)^T(h_j^T V)z = 0 \qquad (5.7)$$

can be determined efficiently, if the matrices CV, SV, $D^T V$ and $g_1^T V$ are known. These are obtained on-the-fly appending one column and component to the current matrix and vector, respectively, in every iteration step of the nonlinear Arnoldi method. Notice, that the explicit form of the matrices C, D and g_1 are not needed to execute these multiplications. Moreover we can take advantage of further updates multiplying $((W_j + \lambda I)V)^T((W_j + \lambda I)V)$.

Since it is very inexpensive to obtain updates of $W_j V$ and $h_j^T V$ from the preceding matrices $W_{j-1}V$ and $h_{j-1}^T V$ (cf. (5.5) and (5.6)) we decided to terminate the inner iteration long before convergence. Our numerical experiments demonstrated that computing time could be reduced further terminating the inner iteration if the residual of the quadratic eigenvalue was less than 10^{-1}.

6. Numerical experiments. In order to evaluate the performance of the RTLSQEP method for large dimensions where the quadratic eigenproblems are solved by iterative projection methods we use test examples from Hansen's *Regularization Tools* [10]. Two functions, *baart* and *shaw*, which are both discretizations of Fredholm integral equations of the first kind, are used to generate matrices $A_{\text{true}} \in \mathbb{R}^{n\times n}$, right hand sides $b_{\text{true}} \in \mathbb{R}^n$ and solutions $x_{\text{true}} \in \mathbb{R}^n$ such that

$$A_{\text{true}} x_{\text{true}} = b_{\text{true}}.$$

In all cases the matrices A_{true} and $[A_{\text{true}}, b_{\text{true}}]$ are ill-conditioned.

We added white noise of level σ to the data, and obtained systems $Ax \approx b$ to be solved where $A = A_{\text{true}} + \sigma E$ and $b = b_{\text{true}} + \sigma e$, and the elements of E and e are independent random variables with zero mean and variance 1. The matrix $L \in \mathbb{R}^{n-1\times n}$ approximates the first order derivative, and δ is chosen to be $\delta = \|Lx_{\text{true}}\|$.

TABLE 6.1
Average computing time: Example baart

σ	n	Li & Ye	SOAR	NL Arn.	early update
10^{-2}	1000	0.91	0.81	1.12	0.31
	2000	3.67	3.35	3.26	1.41
	4000	19.52	18.80	8.69	4.91
10^{-3}	1000	0.75	0.66	0.54	0.24
	2000	2.89	2.62	2.27	1.22
	4000	13.03	12.31	6.24	3.41

The numerical test were run on a PentiumR4 computer with 3.4 GHz and 3GB RAM under MATLAB R2006a. Tables 6.1 and 6.2 contain the CPU times in seconds averaged over 100 random simulations for dimensions $n = 1000$, $n = 2000$, and $n = 4000$ with noise levels $\sigma = 10^{-2}$ and $\sigma = 10^{-3}$ for *baart* and *shaw*, respectively. The outer iteration was terminated if the residual norm of the first order condition was less than 10^{-10}. In the Li & Ye algoritm we always performed 20 inner iterations (because this turned out to be efficient), whereas for the SOAR algorithm we terminated the inner iteration if the the right most eigenvalue of the projected problem had converged,

TABLE 6.2
Average computing time: Example shaw

σ	n	Li & Ye	SOAR	NL Arn.	early up.
10^{-2}	1000	0.99	1.05	0.45	0.39
	2000	3.77	3.93	1.62	1.36
	4000	16.87	18.41	5.53	4.98
10^{-3}	1000	0.54	0.83	0.48	0.47
	2000	2.23	3.24	1.63	1.63
	4000	10.41	15.70	5.48	6.00

and in the nonlinear Arnoldi method we based the termination on the size of the residual norm because the residuals have to be evaluated anyhow to expand the search space. The preconditioner in the nonlinear Arnoldi method was chosen to be the diagonal matrix $M = \mathrm{diag}(T(\sigma))^{-1}$ with an initial guess σ, and it turned out that the method even converged without preconditioning, i.e. $M = I$.

The second order Krylov subspace solvers [1, 13] show a similar behavior while the nonlinear Arnoldi method due to the thick initialization of the inner iterations is significantly faster, and the early termination of the inner iteration accelerates the nonlinear Arnoldi method further (substantially for *baart* and slightly for *shaw*).

Figures 6.1 and 6.2 show the typical convergence behavior of the RTLSQEP method where the quadratic eigenvalue problems in the inner iteration are solved by the Krylov subspace methods presented by Li and Ye and by Bai and Su, and Figures 6.3 and 6.4 show the convergence history for the same problem where the quadratic eigenproblems are solved by the nonlinear Arnoldi method and its variant with early termination of the inner iteration, respectively. An asterisk marks the residual norm of a quadratic eigenvalue problem in an inner iteration, and a circle denotes the residual norm of the first order condition in an outer iteration. Notice, however, that the method presented by Li and Ye requires only one matrix vector product in every inner iteration (due to the fact that only the approximation (5.3) to the projection of (5.1) is considered), whereas SOAR and the nonlinear Arnoldi method need roughly two matrix vector products.

7. Conclusions. Regularized total least squares problems can be solved efficiently by the RTLSQEP method introduced by Sima, Van Huffel and Golub [16] via a sequence of quadratic eigenvalue problems. We characterized the right most eigenvalue and corresponding eigenvector of the associated eigenproblems. For problems of high dimension we compared the Krylov subspace methods for quadratic eigenproblems presented by Li and Ye and the one by Bai and Su, and the nonlinear Arnoldi method. Since the latter one can be initialized in every outer iteration taking advantage of previous iterations in an optimal way it turns out to be faster than the two former methods. Using early termination of the inner iteration the nonlinear Arnoldi method can be even accelerated.

REFERENCES

[1] Z. Bai and Y. Su. SOAR: A second order Arnoldi method for the solution of the quadratic eigenvalue problem. *SIAM J. Matr. Anal. Appl.*, 26:540 – 659, 2005.
[2] A. Beck and A. Ben-Tal. On the solution of the tikhonov regularization of the total least squares problem. *SIAM J. Optim.*, 17:98 – 118, 2006.
[3] A. Beck, A. Ben-Tal, and M. Teboulle. Finding a global optimal solution for a quadratically

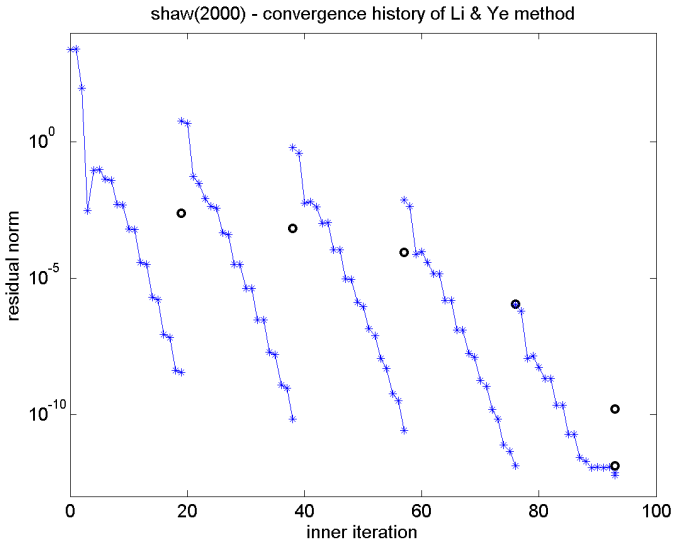

FIG. 6.1. *Convergence history solving quadratic eigenproblems by the Li & Ye algorithm*

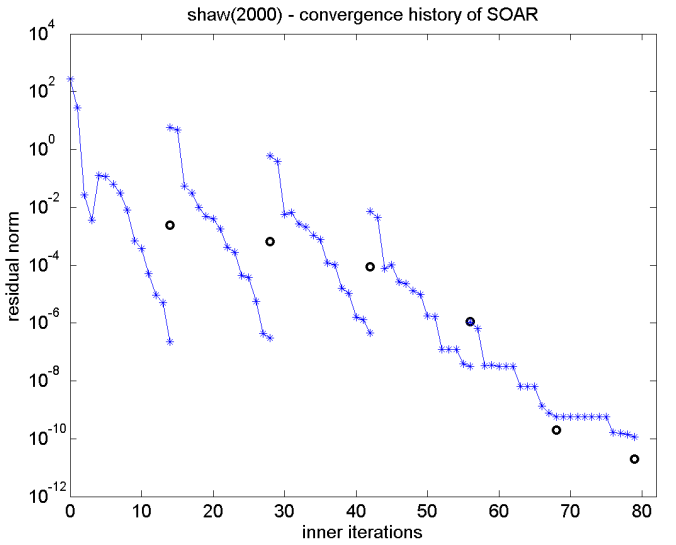

FIG. 6.2. *Convergence history solving quadratic eigenproblems by SOAR*

constrained fractional quadratic problem with applications to the regularized total least squares problem. *SIAM J. Matr. Anal. Appl.*, 28:425 – 445, 2006.

[4] H.W. Engl, M. Hanke, and A. Neubauer. *Regularization of Inverse Problems*. Kluwer, Dodrecht, The Netherlands, 1996.

[5] W. Gander, G.H. Golub, and U. von Matt. A constrained eigenvalue problem. *Lin.Alg.Appl.*, 114–115:815 – 839, 1989.

[6] G.H. Golub, P.C. Hansen, and D.P. O'Leary. Tikhonov regularization and total least squares. *SIAM J. Matr. Anal. Appl.*, 21:185 – 194, 1999.

[7] G.H. Golub and C.F. Van Loan. *Matrix Computations*. The John Hopkins University Press,

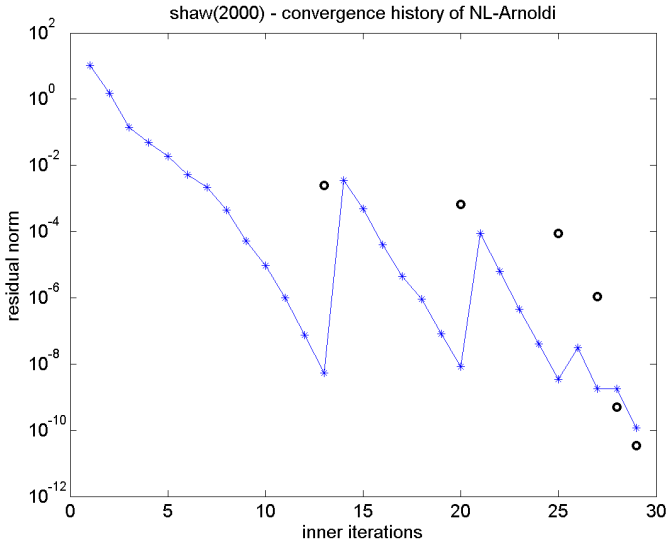

FIG. 6.3. *Convergence history solving quadratic eigenproblems by the Nonlinear Arnoldi method*

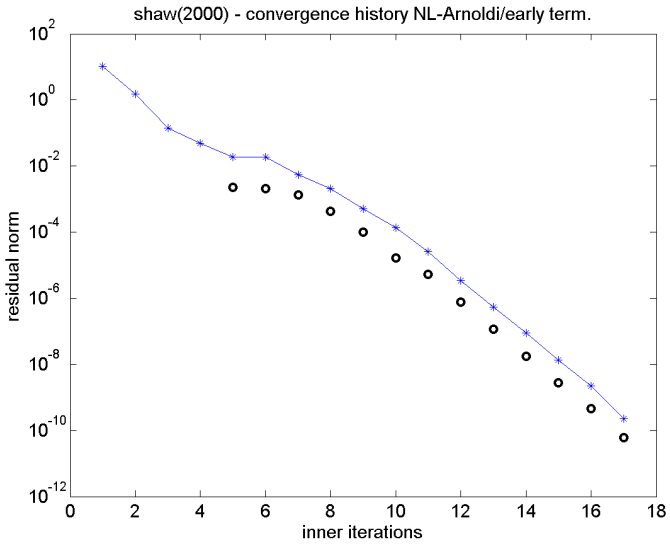

FIG. 6.4. *Convergence history solving quadratic eigenproblems by the Nonlinear Arnoldi method with early termination*

Baltimore and London, 3rd edition, 1996.

[8] C.W. Groetsch. *Inverse Problems in the Mathematical Sciences*. Vieweg, Wiesbaden, Germany, 1993.

[9] H. Guo and R.A. Renaut. A regularized total least squares algorithm. In S. Van Huffel and P. Lemmerling, editors, *Total Least Squares and Errors-in-Variable Modelling*, pages 57 – 66, Dodrecht, The Netherlands, 2002. Kluwer Academic Publisher.

[10] P.C. Hansen. Regularization tools, a Matlab package for analysis of discrete regularization problems. *Numer. Alg.*, 6:1 – 35, 1994.

[11] P.C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion*. SIAM, Philadelphia, 1998.

[12] R.B. Lehoucq, D.C. Sorensen, and C. Yang. *ARPACK Users' Guide. Solution of Large-Scale Eigenvalue Problems with Implicitly Restarted Arnoldi Methods*. SIAM, Philadelphia, 1998.

[13] R.-C. Li and Q. Ye. A Krylov subspace method for quadratic matrix polynomials with application to constrained least squares problems. *SIAM J. Matrix Anal. Appl.*, 25:405 – 428, 2003.

[14] R.A. Renaut and H. Guo. Efficient algorithms for solution of regularized total least squares. *SIAM J. Matr. Anal. Appl.*, 26:457 – 476, 2005.

[15] D.M. Sima. *Regularization Techniques in Model Fitting and Parameter Estimation*. PhD thesis, Katolieke Universiteit Leuven, Leuven, Belgium, 2006.

[16] D.M. Sima, S. Van Huffel, and G.H. Golub. Regularized total least squares based on quadratic eigenvalue problem solvers. *BIT Numerical Mathematics*, 44:793 – 812, 2004.

[17] S. Van Huffel and J. Vandevalle. *The Total Least Squares Problems: Computational Aspects and Analysis*, volume 9 of *Frontiers in Applied Mathematics*. SIAM, Philadelphia, 1991.

[18] H. Voss. An Arnoldi method for nonlinear eigenvalue problems. *BIT Numerical Mathematics*, 44:387 – 401, 2004.

[19] H. Voss and B. Werner. A minimax principle for nonlinear eigenvalue problems with applications to nonoverdamped systems. *Math. Meth. Appl. Sci.*, 4:415–424, 1982.