Numer Algor (2009) 51:61–76
DOI 10.1007/s11075-008-9243-2

ORIGINAL PAPER

Error estimates for the regularization of least squares problems

C. Brezinski · G. Rodriguez · S. Seatzu

Received: 14 August 2008 / Accepted: 26 September 2008 /
Published online: 18 October 2008

Abstract The a posteriori estimate of the errors in the numerical solution of ill-conditioned linear systems with contaminated data is a complicated problem. Several estimates of the norm of the error have been recently introduced and analyzed, under the assumption that the matrix is square and nonsingular. In this paper we study the same problem in the case of a rectangular and, in general, rank-deficient matrix. As a result, a class of error estimates previously introduced by the authors (Brezinski et al., Numer Algorithms, in press, 2008) are extended to the least squares solution of consistent and inconsistent linear systems. Their application to various direct and iterative regularization methods are also discussed, and the numerical effectiveness of these error estimates is pointed out by the results of an extensive experimentation.

Keywords Least squares problems · Regularization · Error estimation

This work was supported by MIUR under the PRIN grant no. 2006017542-003, and the University of Cagliari.

C. Brezinski
Laboratoire Paul Painlevé, UMR CNRS 8524, Université des Sciences et Technologies de Lille, 59655 Villeneuve d'Ascq Cedex, France
e-mail: Claude.Brezinski@univ-lille1.fr

G. Rodriguez (✉) · S. Seatzu
Dipartimento di Matematica e Informatica, Università di Cagliari, Viale Merello 92, 09123 Cagliari, Italy
e-mail: rodriguez@unica.it

S. Seatzu
e-mail: seatzu@unica.it

1 Introduction

In this paper, we study the effectiveness of estimates of the norm of the error for the approximate solution of the least squares problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \|A\mathbf{x}-\mathbf{b}\|, \tag{1.1}$$

where $A\in\mathbb{R}^{m\times n}$ is ill-conditioned and $\mathbf{b}\in\mathbb{R}^m$ is a data vector which we consider to be contaminated by random noise with zero mean, so that the use of a regularization method is indispensable to obtain feasible results. We impose no restrictions on m and n, thus allowing either $m\geq n$ or $m<n$. In (1.1), as in the sequel, the symbol $\|\cdot\|$ denotes both the Euclidean vector norm and the corresponding induced matrix norm.

Let $\mathbf{x}$ be an approximate solution of (1.1) and $\mathbf{r}=\mathbf{b}-A\mathbf{x}$ the associate residual vector. Denoting by $\mathbf{x}^\dagger$ the minimal-norm least squares (or *normal*) solution of (1.1), we will introduce an estimate for the norm of the error vector $\mathbf{e}=\mathbf{x}^\dagger-\mathbf{x}$.

An estimate of this kind for $\|\mathbf{e}\|$ was introduced by Auchmuty in [1] and further analyzed and extended in [6] and [3]. In [4], these error estimates were generalized and some of their properties were studied. In particular, it was shown that the following estimates are valid for any real number ν:

$$e_\nu^2 = c_0^{\nu-1}\left(c_1^2\right)^{3-\nu}c_2^{\nu-4} \simeq \|\mathbf{e}\|^2, \tag{1.2}$$

$$\widetilde{e}_\nu^2 = c_0^{\nu-1}\left(c_1^2\right)^{3-\nu}\widetilde{c}_2^{\nu-4} \simeq \|\mathbf{e}\|^2, \tag{1.3}$$

$$\widehat{e}_\nu^2 = c_0^{\nu+1}\left(c_1^2\right)^{2-\nu}c_2^{\nu-3} \simeq \mathbf{e}^T A\mathbf{e}, \tag{1.4}$$

with

$$\begin{aligned} c_0&=\|\mathbf{r}\|^2, & c_1&=\mathbf{r}^T A\mathbf{r},\\ c_2&=\|A\mathbf{r}\|^2, & \widetilde{c}_2&=\|A^T\mathbf{r}\|^2. \end{aligned} \tag{1.5}$$

In the same paper, these error estimates were applied to the choice of the regularization parameter in Tikhonov regularization, under the assumption of a square nonsingular linear system, and to the determination of a stopping criterion for the conjugate gradient method when the matrix A is symmetric positive definite.

The extension of the error estimates introduced in [4] to least squares problems will be discussed in Section 2. We will show how they can be constructed and illustrate some of their properties. In Section 3, we will discuss their application to various regularization methods like truncated SVD or GSVD, Tikhonov regularization and some iterative methods. The results of our numerical experiments will be reported in Section 4.

2 Error estimates for least squares problems

Let us consider the least squares problem (1.1), with $A \in \mathbb{R}^{m\times n}$ and $\operatorname{rank}(A) = r \leq \min(m, n)$. As it is well known, the normal solution of (1.1) is the unique vector $\mathbf{x}^\dagger$ that solves the minimization problem

$$\min_{\mathbf{x}\in\mathcal{S}} \|\mathbf{x}\|, \qquad \mathcal{S} = \left\{\mathbf{x} \in \mathbb{R}^n \mid A^T A\mathbf{x} = A^T\mathbf{b}\right\}.$$

Such a solution can be written in the form

$$\mathbf{x}^\dagger = A^\dagger\mathbf{b} = \left(A^T A\right)^\dagger A^T\mathbf{b},$$

where $A^\dagger$ indicates the *pseudoinverse* (*Moore–Penrose inverse*) of A [2].

For the sake of clarity, we resume some basic results. Let $A = U\Sigma V^T$ be the singular value decomposition (SVD) [8] of A, with

$$\begin{aligned} U &= [\mathbf{u}_1, \ldots \mathbf{u}_m], & \mathbf{u}_i &\in \mathbb{R}^m, \quad U^T U = I_m, \\ V &= [\mathbf{v}_1, \ldots \mathbf{v}_n], & \mathbf{v}_i &\in \mathbb{R}^n, \quad V^T V = I_n, \\ \Sigma &= \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{m\times n}, & \Sigma_r &= \operatorname{diag}(\sigma_1, \ldots, \sigma_r), \end{aligned}$$

and $\sigma_1 \geq \cdots \geq \sigma_r > 0$.

Let $\mathbf{y}$ be an arbitrary vector. Then

$$\begin{aligned} A\mathbf{y} &= \sum_{i=1}^{r} \sigma_i \left(\mathbf{v}_i^T\mathbf{y}\right)\mathbf{u}_i, \\ A^T\mathbf{y} &= \sum_{i=1}^{r} \sigma_i \left(\mathbf{u}_i^T\mathbf{y}\right)\mathbf{v}_i. \end{aligned} \tag{2.1}$$

Denoting by $\mathbf{x}$ any approximate solution of (1.1), we consider the error vector $\mathbf{e} = \mathbf{x}^\dagger - \mathbf{x}$, the corresponding residual $\mathbf{r} = \mathbf{b} - A\mathbf{x}$ and the *normal residual* $\boldsymbol{\rho} = A^T\mathbf{r}$, that is the residual of the normal equations

$$A^T A\mathbf{x} = A^T\mathbf{b}. \tag{2.2}$$

These vectors are connected by the relation

$$\mathbf{e} = \left(A^T A\right)^\dagger \boldsymbol{\rho} = \left(A^T A\right)^\dagger A^T\mathbf{r}. \tag{2.3}$$

Letting $\alpha_i = \mathbf{u}_i^T\mathbf{r}$, the relations (2.1) yield

$$\|\mathbf{e}\|^2 = \sum_{i=1}^{r} \frac{\alpha_i^2}{\sigma_i^2}. \tag{2.4}$$

In a similar way, we have

$$d_0 = \|\mathbf{r}\|^2 \quad = \|U^T\mathbf{r}\|^2 \quad = \sum_{i=1}^{m} \alpha_i^2,$$

$$d_1 = \|\boldsymbol{\rho}\|^2 \quad = \|A^T\mathbf{r}\|^2 \quad = \sum_{i=1}^{r} \sigma_i^2\alpha_i^2,$$

$$d_2 = \|A\boldsymbol{\rho}\|^2 = \|AA^T\mathbf{r}\|^2 = \sum_{i=1}^{r} \sigma_i^4\alpha_i^2. \tag{2.5}$$

As described in [4] in the case of a square nonsingular system, we approximate the sums in (2.4) and (2.5) by keeping only one term, that is

$$\|\mathbf{e}\|^2 \simeq \sigma^{-2}\alpha^2, \quad d_0 = \alpha^2, \quad d_1 = \sigma^2\alpha^2, \quad d_2 = \sigma^4\alpha^2.$$

The system of 3 equations in 2 unknowns (α and σ) expressing d_0, d_1 and d_2 has infinitely many solutions, which depend upon a real parameter. Solving it, we obtain the following estimate for the norm of the error

$$\|\mathbf{e}\|^2 \quad \simeq \quad \eta_\nu^2 = d_0^{\nu-1}d_1^{5-2\nu}d_2^{\nu-3}, \quad \nu \in \mathbb{R}. \tag{2.6}$$

Notice that, for $\nu = 3$, Auchmuty's estimate [1] is recovered

$$\eta_3 = \widetilde{e}_3 = \frac{\|\mathbf{r}\|^2}{\|A^T\mathbf{r}\|}.$$

Another interesting estimate, which appears to be effective in our experiments, is

$$\eta_2 = \frac{\|\mathbf{r}\| \cdot \|A^T\mathbf{r}\|}{\|AA^T\mathbf{r}\|}.$$

Setting $\rho = d_0d_2/d_1^2$, (2.6) becomes

$$\eta_\nu^2 = \rho^\nu\eta_0^2.$$

Since $\rho \geq 1$ by Schwartz inequality, we can conclude that η_ν^2 is a non decreasing function of ν and that there exists a value ν_e such that $\eta_{\nu_e}^2 = \|\mathbf{e}\|^2$. This value is given by the formula (obviously useless in practice)

$$\nu_e = 2\ln(\|\mathbf{e}\|/\eta_0)/\ln\rho.$$

Moreover, generalizing an inequality obtained by Auchmuty, we have

$$\eta_\nu^2 \leq \|\mathbf{e}\|^2, \quad \text{for } \nu \leq 3.$$

The representation (2.3) of the error allows us to apply the estimate $\widetilde{e}_\nu^2$ given by (1.3) directly to the normal equations (2.2), even when A is rank-deficient. In this way, we obtain

$$\|\mathbf{e}\|^2 \quad \simeq \quad \xi_\nu^2 = d_1^{\nu-1}d_2^{6-2\nu}d_3^{\nu-4},$$

where d_1 and d_2 are defined as in (2.5), and

$$d_3 = \|A^T A\boldsymbol{\rho}\|^2 = \sum_{i=1}^{r} \sigma_i^6 \alpha_i^2. \tag{2.7}$$

3 Application to regularization methods

Since the matrix A in (1.1) is assumed to be severely ill-conditioned, the error estimate (2.6) has to be computed very accurately to obtain satisfactory results. In fact, our numerical experiments showed that it is possible to greatly improve its performance, as a regularization parameter selection method, by suitably adapting its expression to the chosen regularization technique. For this reason, we will now develop particular representations of (2.6), specially suited to some classical regularization methods.

3.1 TSVD and TGSVD

Let $r = \text{rank}(A)$. The truncated SVD (TSVD) solution $\mathbf{x}_k$ of (1.1) is

$$\mathbf{x}_k = \sum_{i=1}^{k} \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{v}_i, \quad k = 1, \ldots, r.$$

The integer k acts as a regularization parameter which filters the components of the solution corresponding to the smallest singular values, while $\mathbf{x}_r$ equals the normal solution $\mathbf{x}^\dagger$.

Using the associated residual $\mathbf{r}_k = \mathbf{b} - A\mathbf{x}_k$, we can express (1.5) and (2.5) in the following form

$$d_0 = c_0 = \|\mathbf{r}_k\|^2 = \sum_{i=k+1}^{r} \left(\mathbf{u}_i^T \mathbf{b}\right)^2 + \gamma^2, \tag{3.1}$$

$$d_1 = \widetilde{c}_2 = \|A^T \mathbf{r}_k\|^2 = \sum_{i=k+1}^{r} \sigma_i^2 \left(\mathbf{u}_i^T \mathbf{b}\right)^2, \tag{3.2}$$

$$d_2 = \|AA^T \mathbf{r}_k\|^2 = \sum_{i=k+1}^{r} \sigma_i^4 \left(\mathbf{u}_i^T \mathbf{b}\right)^2. \tag{3.3}$$

Letting $U = [U_1, U_2]$, with $U_1 = [\mathbf{u}_1, \ldots \mathbf{u}_r]$, we have

$$\gamma^2 = \|U_2^T \mathbf{b}\|^2 = \mathbf{b}^T \left(I - U_1 U_1^T\right) \mathbf{b},$$

with $\gamma = 0$ whenever $r = m$.

We note that formula (2.7) becomes

$$d_3 = \|A^T A\boldsymbol{\rho}\|^2 = \sum_{i=k+1}^{r} \sigma_i^6 \left(\mathbf{u}_i^T \mathbf{b}\right)^2,$$

while, if $r = m = n$,

$$c_1 = \mathbf{r}_k^T A \mathbf{r}_k = \sum_{i,j=k+1}^{r} \sigma_j (\mathbf{u}_i^T \mathbf{v}_j)(\mathbf{u}_i^T \mathbf{b})(\mathbf{u}_j^T \mathbf{b}).$$

Let us now introduce a regularization matrix $H \in \mathbb{R}^{p\times n}$. We call *minimal H-norm solution* of (1.1) the unique vector $\mathbf{x}_H^\dagger$ which solves the problem

$$\min_{\mathbf{x}\in\mathcal{S}} \|H\mathbf{x}\|, \qquad \mathcal{S} = \{\mathbf{x} \in \mathbb{R}^n \mid A^T A\mathbf{x} = A^T\mathbf{b}\}. \tag{3.4}$$

We require that

$$\mathcal{N}(A) \cap \mathcal{N}(H) = \{0\} \tag{3.5}$$

($\mathcal{N}(A)$ is the null space of A), which corresponds to assuming that the augmented matrix $C^T = \begin{bmatrix} A^T & H^T \end{bmatrix}$ has full rank.

Following [17], with the only assumption that $\operatorname{rank}(C) = n$, we define the *generalized singular value decomposition* (GSVD) of the matrix pair (A, H) as the factorization

$$\begin{aligned} A &= U\Sigma_A Z^{-1} \\ H &= V\Sigma_H Z^{-1}, \end{aligned} \tag{3.6}$$

where U and V are orthogonal matrices of order m and p, respectively, and $Z = [\mathbf{z}_1, \ldots, \mathbf{z}_n]$ is nonsingular of order n. The matrices Σ_A, $m \times n$, and Σ_H, $p \times n$, have the form

$$\Sigma_A = \begin{bmatrix} O_A & & \\ & C & \\ & & I_A \end{bmatrix}, \qquad \Sigma_H = \begin{bmatrix} I_H & & \\ & S & \\ & & O_H, \end{bmatrix}$$

where

$$\begin{aligned} C &= \operatorname{diag}(c_1, \ldots, c_q), \qquad 0 < c_1 \le c_2 \le \cdots \le c_q < 1, \\ S &= \operatorname{diag}(s_1, \ldots, s_q), \qquad 1 > s_1 \ge s_2 \ge \cdots \ge s_q > 0, \end{aligned}$$

with $c_i^2 + s_i^2 = 1$, for $i = 1, \ldots, q$. As usual, the ordering is chosen such that the *generalized singular values* $\sigma_i = c_i/s_i$ are non decreasing in $i = 1, \ldots, q$.

Moreover, I_A and I_H are identities of order $r - q$ and $n - r$, respectively, where $r = \operatorname{rank}(A)$, and O_A and O_H are zero matrices of order $(m - r)\times (n - r)$ and $(p - n + r - q) \times (r - q)$, with possibly no rows and/or no columns. We note that $q = \operatorname{rank}(A) + \operatorname{rank}(H) - n$.

Using the GSVD (3.6), we can express the minimal H-norm solution as

$$\mathbf{x}_H^\dagger = \sum_{i=1}^{q} \frac{\mathbf{u}_{\tilde{m}+i}^T \mathbf{b}}{c_i} \mathbf{z}_{\tilde{n}+i} + \sum_{i=q+1}^{r} (\mathbf{u}_{\tilde{m}+i}^T \mathbf{b}) \mathbf{z}_{\tilde{n}+i},$$

where $\widetilde{m}$ is $m-r$ when $m>r$ and 0 otherwise, and $\widetilde{n}$ is defined similarly. Then, the truncated GSVD (TGSVD) solution $\mathbf{x}_k$ of (3.4) is

$$\mathbf{x}_k = \sum_{i=q-k+1}^{q} \frac{\mathbf{u}_{\widetilde{m}+i}^T \mathbf{b}}{c_i} \mathbf{z}_{\widetilde{n}+i} + \sum_{i=q+1}^{r} \left(\mathbf{u}_{\widetilde{m}+i}^T \mathbf{b}\right) \mathbf{z}_{\widetilde{n}+i},$$

where $k=0,1,\ldots,q$ is the regularization parameter.

Letting $\boldsymbol{\beta}_k = \left(\mathbf{u}_{m-r+1}^T \mathbf{b}, \ldots, \mathbf{u}_{m-r+q-k}^T \mathbf{b}, 0, \ldots, 0\right)^T \in \mathbb{R}^q$, for $k=0,1,\ldots,q$, we have

$$d_0 = c_0 = \|\boldsymbol{\alpha}\|^2 + \|\boldsymbol{\beta}_k\|^2, \tag{3.7}$$

$$d_1 = \widetilde{c}_2 = \|M_2 C \boldsymbol{\beta}_k\|^2, \tag{3.8}$$

$$d_2 = \|C M_2^T M_2 C \boldsymbol{\beta}_k\|^2 + \|M_3^T M_2 C \boldsymbol{\beta}_k\|^2, \tag{3.9}$$

$$c_1 = \left[\boldsymbol{\alpha}^T \ \boldsymbol{\beta}_k^T\right] U_1^T M_2 C \boldsymbol{\beta}_k, \quad (\text{if } m=n), \tag{3.10}$$

where (using the Matlab subindexing notation)

$$\boldsymbol{\alpha} = \left(\mathbf{u}_1^T \mathbf{b}, \ldots, \mathbf{u}_{m-r}^T \mathbf{b}\right),$$

$$U_1 = U(1:m, 1:m-r+q),$$

$$M_2^T = Z^{-1}(n-r+1:n-r+q, 1:n),$$

$$M_3^T = Z^{-1}(n-r+q+1:n, 1:n).$$

To choose a value for the regularization parameter in the truncated SVD or GSVD, we used formulae (3.1–3.3) and (3.7–3.9) to evaluate (2.6) for $k=0,1,\ldots,q$, assuming as optimal the value of k that minimizes $\eta_\nu(k)$.

3.2 Tikhonov regularization

In Tikhonov regularization, (1.1) is replaced by the following penalized least squares problem

$$\min_{\mathbf{x}\in\mathbb{R}^n} \left\{ \|A\mathbf{x}-\mathbf{b}\|^2 + \lambda^2 \|H\mathbf{x}\|^2 \right\}, \tag{3.11}$$

where λ is the regularization parameter and $H \in \mathbb{R}^{p\times n}$ a regularization matrix.

In [4], assuming A square nonsingular and $p \le n$, we proposed the following family of estimates for the Euclidean norm of the error

$$\widetilde{e}_\nu = \lambda^{-2} \cdot \|\mathbf{r}_\lambda\|^{\nu-1} \cdot \left(\mathbf{r}_\lambda^T E \mathbf{x}_\lambda\right)^{3-\nu} \cdot \|E\mathbf{x}_\lambda\|^{\nu-4}, \tag{3.12}$$

where, for a fixed λ, $\mathbf{x}_\lambda$ is the minimizer of (3.11), $\mathbf{r}_\lambda = \mathbf{b} - A\mathbf{x}_\lambda$, $E = H^T H$ and $\nu \in \mathbb{R}$. These estimates were obtained by replacing $A^T \mathbf{r}_\lambda$ by $\lambda^2 E \mathbf{x}_\lambda$ in (1.3) and (1.5). In the same paper, the estimates (3.12) were used to determine a suitable value for λ, and an extensive numerical experimentation highlighted the effectiveness of this method.

Here, we use the same approach as adopted in [4] to derive an error estimate in the general case of $A \in \mathbb{R}^{m\times n}$, without any restrictions on m, n

and p. Assuming that condition (3.5) holds, the solution of (3.11) is uniquely determined by the consistent linear system

$$\left(A^T A + \lambda^2 H^T H\right)\mathbf{x}_\lambda = A^T \mathbf{b}, \tag{3.13}$$

which implies the relation

$$A^T \mathbf{r}_\lambda = \lambda^2 E \mathbf{x}_\lambda.$$

Hence, substituting $A^T \mathbf{r}_\lambda$ by $\lambda^2 E \mathbf{x}_\lambda$ in (2.5), (2.6) becomes

$$\eta_\nu(\lambda) = \lambda^{4-2\nu} \cdot \|\mathbf{r}_\lambda\|^{\nu-1} \cdot \|E\mathbf{x}_\lambda\|^{5-2\nu} \cdot \|AE\mathbf{x}_\lambda\|^{\nu-3}, \quad \nu \in \mathbb{R}. \tag{3.14}$$

As already observed in [4] for the case of regular linear systems treated by Tikhonov's regularization method, this representation of the error estimate η_ν is numerically much more effective than (2.6).

To select a suitable value of λ, we start with a grid of values μ_i, $i = 1, \ldots, 5$, equally spaced on a chosen interval $[\log(\lambda_{\min}), \log(\lambda_{\max})]$. Then, for $\lambda_i = 10^{\mu_i}$, we compute the regularized solution $\mathbf{x}_{\lambda_i}$ and evaluate the norm of the error by formula (3.14). After selecting the parameter λ which minimizes $\eta_\nu(\lambda)$, we iteratively add points to the grid around the minimum, by bisection, until the smallest grid step exceeds a prescribed accuracy τ.

For each value of λ, when the dimension of the linear system is small enough, the regularized solution $\mathbf{x}_\lambda$ can be effectively computed either by the GSVD of (A, H) or, when $H = I$, by the SVD of A; see [11]. For large scale problems, the conjugate gradient method applied to the system (3.13) is the most natural choice.

3.3 Iterative regularization

It is well known that some iterative methods have a regularizing effect. This means that, when applied to an ill-conditioned linear system with a noisy right hand side, the error at the k-th iteration (with respect to the solution corresponding to exact data) first decreases and then diverges. So the index k plays the role of a regularization parameter and it must be carefully tuned.

A typical iterative approach for computing the normal solution of (1.1) is to apply the conjugate gradient method to the normal equations (2.2). The CGLS method [14] represents a clever implementation of this idea, and it is easy to organize this algorithm in order to obtain d_0, d_1 and d_2 in (2.6) inexpensively, so that the computation of $\eta_\nu(k)$, that is the value of η_ν for a given value of the iteration index k, does not increase the computational complexity of the method.

An alternative approach, although theoretically equivalent, is the LSQR method [18], on which we focus. The central part of this method consists of computing iteratively the factorization

$$AV_k = U_{k+1} B_k,$$

where U_{k+1} and V_k are matrices of dimension $m \times (k+1)$ and $n \times k$, respectively, with orthonormal columns, and B_k is a lower bidiagonal matrix of

dimension $(k+1)\times k$. This factorization allows us to express the norm of the k-th residual in the form

$$\|\mathbf{r}_k\| := \|\mathbf{b} - A\mathbf{x}_k\| = \|\beta_1\mathbf{e}_1 - B_k\mathbf{y}_k\|, \tag{3.15}$$

where $\beta_1 = \|\mathbf{b}\|$ and $\mathbf{y}_k \in \mathbb{R}^k$. The minimization of (3.15) is performed by suitably updating the QR factorization of B_k at each iteration.

The LSQR method can be adapted, inexpensively, in order to obtain the error estimate $\eta_\nu(k)$. In fact, using the notation introduced in [18],

$$\begin{aligned} d_0 &= \overline{\phi}_{k+1}, \\ d_1 &= \overline{\phi}_{k+1}\alpha_{k+1}|c_k|, \\ d_2 &= \overline{\phi}_{k+1}\alpha_{k+1}|c_k|\big(\alpha_{k+1}^2 + \beta_{k+2}^2\big), \end{aligned} \tag{3.16}$$

where the constants $\overline{\phi}_{k+1}$ are generated by the recursion

$$\overline{\phi}_{k+1} = s_k\overline{\phi}_k, \quad \text{with } \overline{\phi}_1 = \beta_1,$$

α_k and β_k are the entries on the main and the lower diagonal of B_k, respectively, and (c_k, s_k) are the parameters which define the plane rotation constructed at each iteration of the method. We note that the first two formulae in (3.16), coupled to an estimate of the Frobenius norm of A, were used in [18] as a stopping criterion for the method.

To obtain a regularized solution, we let the LSQR method perform a fixed number ℓ of iterations and then select the solution $\mathbf{x}_k$ corresponding to the minimum of $\eta_\nu(k)$, $k = 0, \ldots, \ell$. It is worthwhile to notice that, even if the LSQR method often gives good results without reorthogonalization, our numerical experiments show that a reliable evaluation of $\eta_\nu(k)$ by formulae (3.16) requires the reorthogonalization of the columns of V_k and U_{k+1} at each step.

4 Numerical simulation

In this section, we will show that the error estimates introduced above can be used to estimate the parameters typical of some regularization techniques. We first describe how we planned the numerical tests, in order to compare our parameter selection schemes with two of the most widely used methods, namely the L-curve [11, 13] and the generalized cross validation (GCV) [5, 7].

All the computations were performed with Matlab 7.5 [15] on an AMD64 computer running Debian Linux. We developed a small Matlab toolbox, called ErresTools, whose functions are listed in Table 1, and which is available upon request. The Regularization Tools ver. 4.0 [10], by P.C. Hansen, was used for computations involving standard regularization methods and parameter estimation techniques.

We selected ten test matrices, either mildly or severely ill-conditioned, among those available in the Matlab `gallery` function and in the Regulariza-

Table 1 Functions in the toolbox

`enu`	Compute estimates (1.2) and (1.3)
`ehatnu`	Compute estimate (1.4)
`enutik`	Compute estimate (3.12)
`etanu`	Compute estimate (2.6)
`etanutik`	Compute estimate (3.14)
`errestsv`	Estimate the optimal parameter for TSVD/TGSVD
`errestik`	Estimate the optimal parameter for Tikhonov regularization

tion Tools. These matrices are listed in Table 2, together with their condition numbers for dimension 20.

For each matrix, we computed a right hand side $\|\mathbf{b}\|$ corresponding to each of the seven model solutions listed in Table 3, adding a vector of Gaussian random errors with zero mean and scaled to have norm $\varepsilon\|\mathbf{b}\|$. Each test was repeated for $\varepsilon = 10^{-6}, 10^{-4}, 10^{-2}, 10^{-1}$ and for five realizations of the noise. Moreover three regularization matrices were used: the identity matrix and the discretization matrices corresponding to the discretization of the first and second derivatives. This leads to a total of 4,200 different test linear systems.

We used various techniques to choose the regularization parameter and compared the corresponding solutions with the optimal solution, i.e., the one which leads to the smallest error. For the methods which depend upon a discrete regularization parameter, the optimal solution is simply the one which minimizes the 2-norm error with respect to all the admissible parameters. For Tikhonov regularization, denoting by $\mathbf{x}$ the exact solution, we consider optimal the minimizer of

$$\psi(\rho) = \|\mathbf{x} - \mathbf{x}_\lambda\|, \quad \text{with } \lambda = 10^\rho,$$

obtained by the `fminsearch` function of Matlab, using as a starting point the logarithm of the parameter which produces the lowest error, among the methods to be compared.

If $\mathbf{x}_{\text{est}}$ represents the solution corresponding to an estimated regularization parameter and $\mathbf{x}_{\text{opt}}$ the optimal one, we adopted the ratio

$$Q(\mathbf{x}_{\text{est}}) = \frac{\|\mathbf{x} - \mathbf{x}_{\text{est}}\|}{\|\mathbf{x} - \mathbf{x}_{\text{opt}}\|} \tag{4.1}$$

as a quality index. It is always greater than or equal to 1, and it equals 1 only if the estimated parameter leads to the optimal error.

Table 2 Test matrices and condition numbers for $n = 20$

Baart	$1.9 \cdot 10^{17}$	Pascal	$1.3 \cdot 10^{20}$
Hilbert	$1.2 \cdot 10^{18}$	Phillips	$4.0 \cdot 10^{03}$
Ilaplace(3)	$2.7 \cdot 10^{30}$	Prolate	$5.6 \cdot 10^{13}$
Lotkin	$9.0 \cdot 10^{18}$	Shaw	$1.1 \cdot 10^{16}$
Moler	$1.7 \cdot 10^{13}$	Wing	$3.2 \cdot 10^{18}$

Table 3 Solution vectors $\mathbf{x} = (x_1, \ldots, x_n)^T$

given	Default solution for problems from [10], `shaw` solution for the others
ones	$x_i = 1$
linear	$x_i = \frac{i}{n}$
parabola	$x_i = \left(\left(i - \lfloor \frac{n}{2} \rfloor\right) / \lceil \frac{n}{2} \rceil\right)^2$
sin2pi	$x_i = \sin \frac{2\pi(i-1)}{n}$
linear+sin2pi/4	$x_i = \frac{i}{n} + \frac{1}{4} \sin \frac{2\pi(i-1)}{n}$
step	$x_i = 0, i \leq \lfloor \frac{n}{2} \rfloor$ $x_i = 1, i > \lfloor \frac{n}{2} \rfloor$

The two methods compared to our estimates are the L-curve [11, 13] and the generalized cross validation (GCV) [5, 7]. Like our estimates, they do not require any information on the noise. To apply both methods, we used the implementations contained in the Regularization Tools [10], forcing the `l_curve` routine to always use the *corner* algorithm [12] for discrete L-curves, which proved to be more robust. Regarding the GCV, we will also investigate, in the sequel, the performance of another algorithm for its computation [9], particularly suited to large scale problems.

Table 4 shows the results obtained by selecting the truncation parameter k in TSVD/TGSVD using the error estimate η_2, the L-curve and the GCV, for 8,400 test linear systems obtained by generating our standard problem set for the dimensions 40×20 and 100×50. In the table, we report the number of *failures* and *severe failures*, that is the number of ratios (4.1) larger than 10 and 100, respectively.

Table 4 Results for TSVD/TGSVD

	η_2		L-curve		GCV	
	> 10	> 100	> 10	> 100	> 10	> 100
Matrix						
Baart	122	16	280	232	300	210
Hilbert	27	0	219	204	86	42
Ilaplace(3)	8	0	149	133	155	104
Lotkin	13	0	277	266	80	28
Moler	19	0	251	144	66	15
Pascal	469	212	331	178	67	41
Phillips	32	0	210	83	54	5
Prolate	23	4	167	48	209	128
Shaw	47	0	206	147	215	146
Wing	8	0	442	419	260	142
Total	768	232	2,532	1,854	1,492	861
(8,400)	9%	3%	30%	22%	18%	10%
H						
I	136	0	396	173	374	243
D_1	241	96	793	613	458	268
D_2	391	136	1,343	1,068	660	350
Total	768	232	2,532	1,854	1,492	861

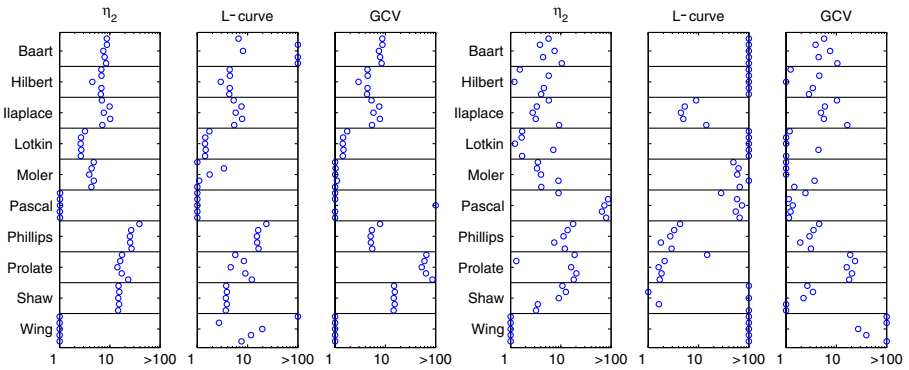

Fig. 1 Results for TSVD/TGSVD, $\varepsilon = 10^{-2}$, $H = I$(left), D_2(right)

The data contained in Table 4 are not sufficient to assert the accuracy of the solution vectors computed, as the severity of a *failure* depends on the value of the optimal error. Anyway, the superiority of η_2 over the other methods considered, on this set of examples, is evident. The lower table, which represents the data with respect to the three regularization matrices, shows that this is true especially when H is not the identity matrix, that is when the TGSVD is applied.

To better represent this behavior, we display, in Fig. 1, the values of the ratio Q for each test matrix and for one realization of the noise. The values of Q larger than 100 are levelled to this value. The tests are performed for linear systems of dimension 100×50 with the *given* solution (Table 3) and $\varepsilon = 10^{-2}$. The results produced by our error estimates are more or less equivalent to those determined by the GCV. For $H = I$, the error estimate η_2 and the L-curve give comparable results, while, when H is the discretization of the second derivative, the results produced by η_2 are more reliable.

The 8,400 test linear systems used for Table 4 are all consistent. The same set of test problems was constructed by generating inconsistent linear systems with a residual $\tau = \|\mathbf{b} - A\mathbf{x}\|$. In Table 5, the number of failures in the three cases $\tau = 0, 1, 10$ is displayed. Even if the results worsen as the residual grows, the performance of our error estimate is comparable to that of GCV and far better than those furnished by the L-curve.

Table 5 TSVD/TGSVD for inconsistent linear systems

	η_2		L-curve		GCV	
	> 10	> 100	> 10	> 100	> 10	> 100
$\tau = 0$	768	232	2,532	1,854	1,492	861
	9%	3%	30%	22%	18%	10%
$\tau = 1$	2,332	744	4,195	3,021	1,911	594
	28%	9%	50%	36%	23%	7%
$\tau = 10$	2,501	911	4,804	3,631	2,267	805
	30%	11%	57%	43%	27%	10%

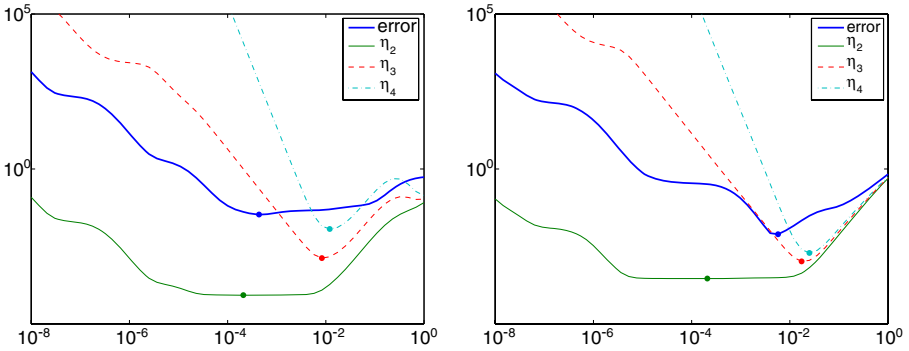

Fig. 2 Results for the *Baart* (*left*) and *Shaw* (*right*) test problems

An extensive experimentation of the effectiveness of formulae (1.2–1.4) for the selection of the parameter in Tikhonov regularization has already been executed in [4]. The performance of the estimates (2.6) is similar, so, in this paper, we prefer to analyze some particular examples.

The parameter ν in (2.6) can take every real value, but we found out that, for solving discrete ill-posed problems by TSVD/TGSVD or Tikhonov regularization, it is better to choose $\nu \geq 2$. This is probably due to the need of giving a sufficient weight to the residual $\|\mathbf{r}\|$ in the expression of η_ν. In Fig. 2, we plot the graphs of η_ν, $\nu = 2, 3, 4$, and of the error for the *Baart* and the *Shaw* test problems of dimension 40×20, with $\varepsilon = 10^{-4}$. The solution is computed by using the SVD of the matrix A in Tikhonov regularization, with $H = I$.

In Fig. 3 (left), we compare η_2 with the GCV algorithm described in [9], the only one, as far as we know, suited to large scale problems. We would like to acknowledge the support of Gene Golub, during earlier stages of this work, in sending us the software developed for the paper [9]. We solve the *Phillips* test problem of dimension 200×100, with $\varepsilon = 10^{-2}$ and $H = D_1$. The

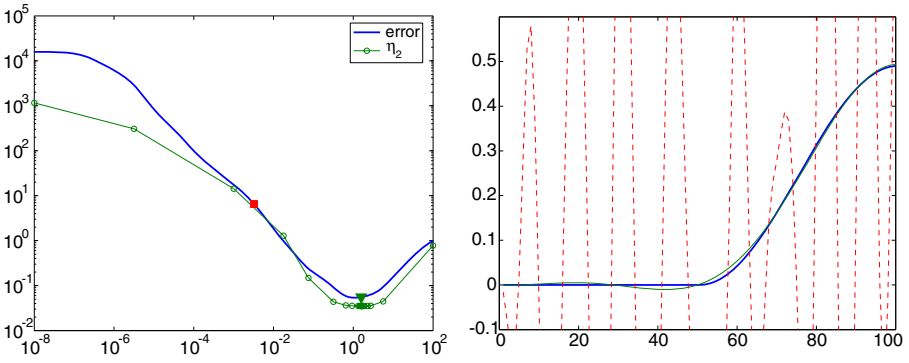

Fig. 3 Error and solution for the *Phillips* test problem

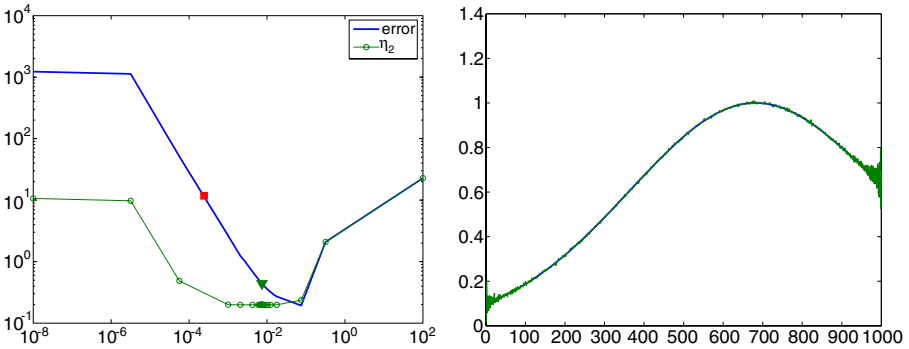

Fig. 4 Error and solution for the *Prolate* test problem

small triangle on the error curve marks the error produced by the minimizer of η_2, while the square is the error resulting from the `gcv_lanczos` routine from [9]. The circle markers on the η_2 curve represent the values effectively computed by the algorithm described in Section 3.2. In the right graph of Fig. 3, we plot the true solution (thick line) and those determined by η_2 (thin line) and `gcv_lanczos` (dashed line).

In Fig. 4, we repeat the experiment for the *Prolate* test problem [19] of dimension $2,000 \times 1,000$, with $\varepsilon = 10^{-2}$ and $H = I$. For the sake of clarity, in the right graph, only the true solution and the one which minimizes η_2 are plotted. In the two preceding Figures, the solution was computed by solving the Tikhonov normal equations by the conjugate gradient method.

We also applied our estimate to stop the iteration of the LSQR method, as explained in Section 3.3. It turns out that it is often successful when a small number of iteration is required to get a good approximation of the solution. In Fig. 5 (left), we report the error and the estimates η_1 and η_2 with respect to the iteration index k for the *Baart* problem of dimension 200×100, $\varepsilon = 10^{-4}$. The

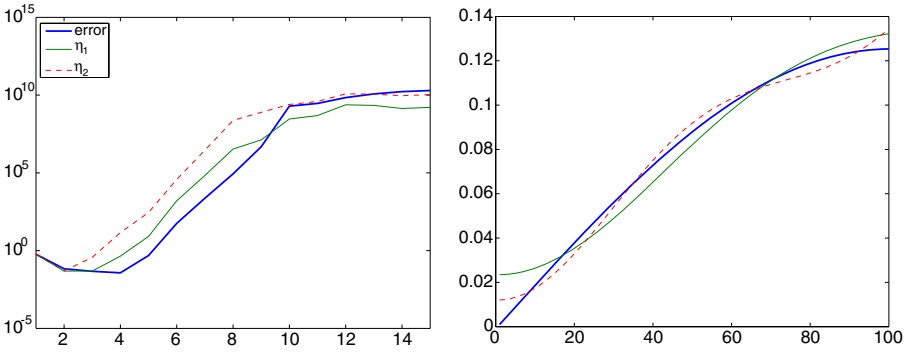

Fig. 5 Error and solution for the *Baart* test problem

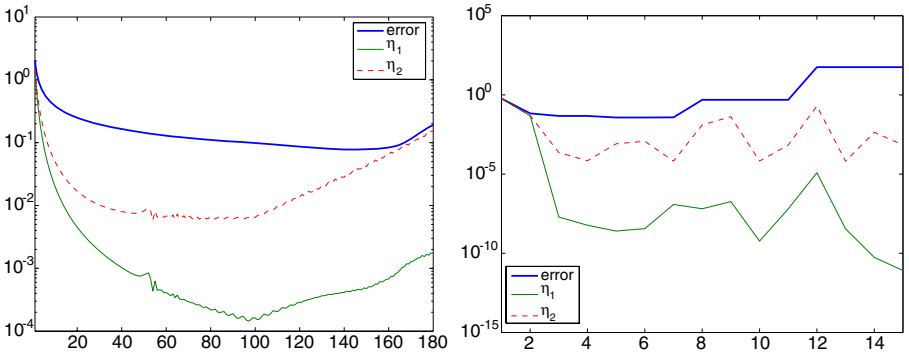

Fig. 6 Results for the *Gaussian* matrix and the *Baart* test problem

right graph shows the true solution (thick line), the optimal one (thin line) and the one corresponding to the minimizer of η_1 (dashed line).

On the contrary, Fig. 6 exhibits some potential problems in applying our estimates to iterative methods. The left graph shows the error and its estimates for a linear system involving the *Gaussian* matrix [16] with parameter 0.01, dimension $4,000 \times 2,000$, solution $\mathbf{x} = (1, \ldots, 1)^T$ and $\varepsilon = 10^{-4}$. When the convergence is slow, as in this case, our error estimates often oscillate irregularly, occasionally making it difficult to locate the minimum.

Another important issue is reorthogonalization: when it is not implemented in the LSQR method, formulae (3.16) are totally inaccurate, making our estimates useless. This is evident in Fig. 6 (right), where we plot the error and its estimates obtained by solving the same problem as in Fig. 5 by LSQR without reorthogonalization.

References

1. Auchmuty, G.: A posteriori error estimates for linear equations. Numer. Math. **61**, 1–6 (1992)
2. Björck, Å.: Numerical Methods for Least Squares Problems. SIAM, Philadelphia (1996)
3. Brezinski, C.: Error estimates for the solution of linear systems. SIAM J. Sci. Comput. **21**, 764–781 (1999)
4. Brezinski, C., Rodriguez, G., Seatzu, S.: Error estimates for linear systems with applications to regularization. Numer. Algorithms **49** (2008, in press)
5. Craven, P., Wahba, G.: Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross validation. Numer. Math. **31**, 377–403 (1979)
6. Galantai, A.: A study of Auchmuty's error estimate. Comput. Math. Appl. **42**, 1093–1102 (2001)
7. Golub, G.H., Heath, M., Wahba, G.: Generalized cross-validation as a method for choosing a good ridge parameter. Technometrics **21**(2), 215–223 (1979)
8. Golub, G.H., Van Loan, C.F.: Matrix Computations. The John Hopkins University Press, Baltimore (1989)
9. Golub, G.H., von Matt, U.: Generalized cross-validation for large-scale problems. J. Comput. Graph. Stat. **6**(1), 1–34 (1997)

10. Hansen, P.C.: Regularization tools version 4.0 for matlab 7.3. Numer. Algorithms **46**, 189–194 (2007)
11. Hansen, P.C.: Rank-Deficient and Discrete Ill-Posed Problems: Numerical Aspects of Linear Inversion. SIAM, Philadelphia (1998)
12. Hansen, P.C., Jensen, T.K., Rodriguez, G.: An adaptive pruning algorithm for the discrete L-curve criterion. J. Comput. Appl. Math. **198**, 483–492 (2006)
13. Hansen, P.C., O'Leary, D.P.: The use of the L-curve in the regularization of discrete ill-posed problems. SIAM J. Sci. Statist. Comput. **14**, 1487–1503 (1993)
14. Hestenes, M.R., Stiefel, E.: Methods of conjugate gradients for solving linear systems. J. Res. Natl. Bur. Stand. **49**, 409–436 (1952)
15. Matlab ver. 7.5, The MathWorks. Inc., Natick, MA (2007)
16. van der Mee, C.V.M., Seatzu, S.: A method for generating infinite positive self-adjoint test matrices and Riesz bases. SIAM J. Matrix Anal. Appl. **26**, 1132–1149 (2005)
17. Paige, C.C., Saunders, M.A.: Towards a generalized singular value decomposition. SIAM J. Numer. Anal. **18**(3), 398–405 (1981)
18. Paige, C.C., Saunders, M.A.: LSQR: an algorithm for sparse linear equations and sparse least squares. ACM Trans. Math. Softw. **8**, 43–71 (1982)
19. Varah, J.M.: The Prolate matrix. Linear Algebra Appl. **187**, 269–278 (1993)