Rendiconti di Matematica, Serie VII
Volume 25, Roma (2005), 69-84

# An algorithm for estimating the optimal regularization parameter by the L-curve

G. RODRIGUEZ – D. THEIS

Abstract: *In this paper we introduce a new algorithm to estimate the optimal regularization parameter in truncated singular value decomposition (TSVD) regularization methods for the numerical solution of severely ill-posed linear systems. The algorithm couples a geometrical approach to identify the corner of the L-curve associated to the problem with some heuristic rules. Numerical results are reported to highlight the performance of the algorithm with respect to other methods for the selection of the regularization parameter.*

## 1 – Introduction

A linear system of equations

$$A\mathbf{x} = \mathbf{b}$$

is considered *severely* ill-conditioned when the condition number

$$\kappa(A) := \|A\| \, \|A^{-1}\|,$$

in a given matrix norm, is of the same order of magnitude, or larger, than the reciprocal of the relative precision on the entries of the matrix $A$ and of the right hand side vector $\mathbf{b}$.

Key Words and Phrases: *ill-conditioned linear systems – Regularization – Truncated singular value decomposition (TSVD) – L-curve.*
A.M.S. Classification: 65F05 – 65F20 – 65F22

In this situation, no general purpose method is able to produce acceptable results, since ill-conditioning causes a huge amplification of errors in the solution of the linear system. Often, the only possibility for partially recovering the solution is the exploitation of *a priori* informations. Regularization methods consist of techniques that take advantage of informations on the regularity of the solution. Usually, these informations are formally expressed as the request for the solution $\mathbf{x}$ to belong to the kernel of a certain linear regularization operator $H$ (in most cases, the discrete approximation of a differential operator). This approach is particularly effective when the solution of the linear system may be thought of as the sampling of a function which exhibits some degree of regularity. An example of great applicative interest is given by the linear systems arising in the discretization of first kind Fredholm integral equations with discrete data

$$\int_\Omega k(u_i, v)\, f(v)\, dv = g(u_i), \quad i = 1, \dots, m.$$

Each regularization method depends on at least one parameter, whose tuning is crucial for the quality of the numerical solution, since it balances the request of approximately satisfying the linear system with the regularity constraint.

The three most widely used regularization techniques are the *Truncated (Generalized) Singular Value Decomposition* (TSVD/TGSVD) [8], *Tikhonov regularization* [20], [21], [6] and regularizing iterative methods [4], [7]. A complete survey of the various regularization strategies and the available methods for the estimation of the optimal regularization parameter can be found in [12].

In this paper we will concentrate on the first mentioned regularization method, and on a particular strategy for choosing its parameter, the *L-curve* method. In Section 2 the TSVD and TGSVD are recalled, while in Section 3 the L-curve method is described. In Section 4 we introduce a new algorithm for the localization of the corner of the L-curve, which has already been applied, in a preliminary version, in some previous researches [2], [19]. Finally, in Section 5 the performance of the new algorithm is assessed on a set of test linear systems, and in Section 6 plans for future work are discussed.

## 2 – The truncated (G)SVD

Let us consider, as a model problem, the overdetermined linear system

$$A\mathbf{x} = \mathbf{b} \tag{2.1}$$

where $A \in \mathbb{R}^{m\times n}$, $m \geq n$, is a full-rank matrix.

The *singular value decomposition* (SVD) of $A$ [1], [5] is given by

$$U^T A V = \begin{bmatrix} \Sigma \\ 0 \end{bmatrix}, \quad \Sigma = \mathrm{diag}(\sigma_1, \dots, \sigma_n), \tag{2.2}$$

where $\sigma_1 \geq \cdots \geq \sigma_n > 0$ are the singular values and the orthogonal matrices

$$U = [\mathbf{u}_1, \ldots, \mathbf{u}_m] \quad \text{and} \quad V = [\mathbf{v}_1, \ldots, \mathbf{v}_n]$$

contain the left and right singular vectors, respectively. Then, the least squares solution of (2.1) can be expressed in the form

$$\mathbf{x} = \sum_{i=1}^{n} \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{v}_i. \tag{2.3}$$

Severe ill-conditioning can be restated by saying that $A$ is numerically rank-deficient, i.e. there exists an integer $k \leq n$ such that, for a given tolerance $\epsilon$,

$$\sigma_{k+1}, \ldots, \sigma_n < \epsilon.$$

This integer is, in fact, the numerical $\epsilon$-rank of $A$, usually defined as

$$\operatorname{rank}_\epsilon(A) := \min_{\|E\|_2 \leq \epsilon} \operatorname{rank}(A + E).$$

When a singular value $\sigma_i$ is approximately zero, the corresponding singular vector $\mathbf{v}_i$ belongs to the numerical kernel of $A$ and we expect its coefficient in (2.3) to be negligible. If the system (2.1) is compatible this is certainly true, but the presence of noise on $\mathbf{b}$ may cause a huge growth in the norm of the solution $\mathbf{x}$.

To obtain a better estimate of the least squares solution the *truncated SVD* (TSVD) solution is often used. It is given by

$$\mathbf{x}_k = \sum_{i=1}^{k} \frac{\mathbf{u}_i^T \mathbf{b}}{\sigma_i} \mathbf{v}_i \tag{2.4}$$

and it coincides with the minimum 2-norm solution to the least squares problem

$$\min \|A_k \mathbf{x} - \mathbf{b}\|_2, \tag{2.5}$$

where $A_k$ is the best rank $k$ approximation to $A$ in the 2-norm, obtainable by substituting $\sigma_j = 0$, $j = k+1, \ldots, n$, in (2.2).

It is then crucial to correctly tune the value of the regularization parameter $k$ in order to avoid the numerical explosion of the norm of the solution while preserving, at the same time, all of its significant components.

When there is the *a priori* information that the solution (approximately) belongs to the kernel of a certain regularization matrix $H$, that is the number $\|H\mathbf{x}\|_2$ is small, it is more effective to compute the solution of (2.5) which minimizes the semi-norm $\|H\mathbf{x}\|_2$, instead than the norm $\|\mathbf{x}\|_2$.

The solution to this *non-standard* regularization problem can be obtained by the transformation $\mathbf{y} = H\mathbf{x}$, but while the case of $H$ square nonsingular can be easily managed, if the regularization matrix is a non square $p \times n$ matrix ($p < n$) with rank $p$, the computation is a bit more cumbersome. A method for taking this problem to standard form has been described in [9] and starts by expressing the least-squares solution in the form

$$\mathbf{x} = H_A^{\dagger}\mathbf{y} + \mathbf{x}_0, \tag{2.6}$$

where $\mathbf{x}_0$ is in the null space of $H$ and the matrix $H_A^{\dagger}$ is the *A-weighted pseudo inverse* of $H$, defined in the following.

Let the *generalized singular value decomposition* (GSVD) of the matrix pair $(A, H)$ [1], [5] be the factorization

$$U^T A Z = \begin{bmatrix} D_A & 0 \\ 0 & I_{n-p} \\ 0 & 0 \end{bmatrix}$$

$$V^T H Z = [\, D_H \quad 0 \,]$$

with

$$D_A = \mathrm{diag}(d_1, \ldots, d_p), \quad 0 < d_1 \leq \cdots \leq d_p \leq 1,$$

$$D_H = \mathrm{diag}(h_1, \ldots, h_p), \quad 1 \geq h_1 \geq \cdots \geq h_p > 0,$$

and $d_i^2 + h_i^2 = 1$, $i = 1, \ldots, p$. The matrices $U$ and $V$ are orthogonal, $Z = [\mathbf{z}_1, \ldots, \mathbf{z}_n]$ is nonsingular and the ratios $\gamma_i = d_i/h_i$, $i = 1, \ldots, p$, are called the *generalized singular values* of $(A, H)$. Then, we define

$$H_A^{\dagger} := Z \begin{bmatrix} D_H^{-1} \\ 0 \end{bmatrix} V^T$$

and it is immediate to observe that it is a right inverse of $H$.

By substituting $\mathbf{y} = H\mathbf{x}$, with $\mathbf{x}$ given by (2.6), we come to the following standard form regularization problem: find the minimum norm solution to

$$\min \|\bar{A}_k \mathbf{y} - \bar{\mathbf{b}}\|_2,$$

where $\bar{A}_k$ is the best rank $k$ approximation to $\bar{A} = AH_A^{\dagger}$ and $\bar{\mathbf{b}} = \mathbf{b} - A\mathbf{x}_0$. The solution to this problem, known as the *truncated GSVD (TGSVD) solution*, is given by

$$\mathbf{x}_k = \sum_{i=p-k+1}^{p} \frac{\mathbf{u}_i^T \mathbf{b}}{d_i} \mathbf{z}_i + \sum_{i=p+1}^{n} (\mathbf{u}_i^T \mathbf{b}) \mathbf{z}_i. \tag{2.7}$$

Again, to get a meaningful solution it is essential to correctly estimate the value of the regularization parameter $k$.

Several criteria are available for this task, some requiring the knowledge of the standard deviation of the noise affecting the data and some not requiring it. We will discuss the *L-curve* method in the next section, here we briefly recall some of the other most widely used techniques.

The *Generalized Cross Validation* (GCV) [3], [22] is a statistical method which estimates the optimal value of the regularization parameter, under the assumption that the data vector $\mathbf{b}$ is affected by normally distributed noise, by minimizing the functional

$$V(k) = \frac{\frac{1}{m}\|(I - A(k))\mathbf{b}\|^2}{\left[\frac{1}{m}\operatorname{trace}(I - A(k))\right]^2}. \tag{2.8}$$

The *influence matrix* $A(k)$ is defined by the identity

$$A\mathbf{x}_k = A(k)\mathbf{b}.$$

The GCV has some computationally relevant properties and, moreover, is a *predictive mean-square error criteria* [22], in the sense that it estimates the minimizer of the residual function

$$T(k) = \frac{1}{m}\|A(\mathbf{x}_k - \mathbf{x})\|^2.$$

If the standard deviation $\sigma$ of the noise on the data is known, the following unbiased estimate [15] for the function $T(k)$ is also available

$$\hat{T}(k) = \frac{1}{m}\|(I - A(k))\mathbf{b}\|^2 - \frac{\sigma^2}{m}\operatorname{trace}(I - A(k))^2 + \frac{\sigma^2}{m}\operatorname{trace} A^2(k).$$

*Mallows' criterion* chooses the value of $k$ which minimizes $\hat{T}(k)$. Another technique which makes use of the value of $\sigma$ is *Morozov discrepancy principle* [17], which takes as optimal the value of $k$ that satisfies the equation

$$\frac{1}{m}\|(I - A(k))\mathbf{b}\|^2 = \sigma^2.$$

However, numerical experiments reported in the literature (see e.g. [3], [18]) showed that these two procedures do not give better results than GCV, even when $\sigma$ is exactly known.

## 3 – The L-curve method

The *L-curve* method [10], [14] consists of the analysis of the piecewise linear curve whose break-points are

$$(x_i, y_i) = (\log_{10} \|A\mathbf{x}_i - \mathbf{b}\|_2, \log_{10} \|H\mathbf{x}_i\|_2), \quad i = 1, \dots, p$$

($p$ is the row dimension of the regularization matrix $H$).

This curve, in most cases, exhibits a typical "L" shape, and the optimal value of the regularization parameter $k$ is considered to be the one corresponding to the *corner* of the "L" (see Figure 1).

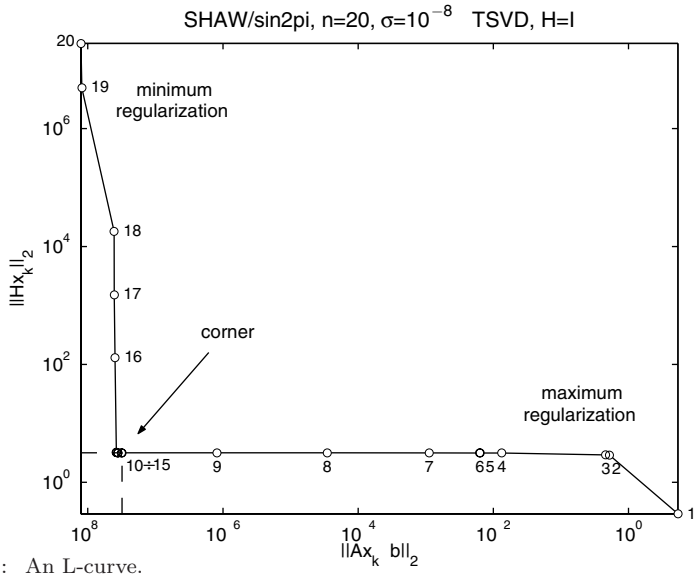


Fig. 1: An L-curve.

This choice is justified by the fact that while the regularized solution $\mathbf{x}_k$ of (2.1) coincides with the least-squares solution $\mathbf{x}$ when $k = p$, the ill-conditioning of $A$ causes a strong growth in the weighted semi-norm $\|H\mathbf{x}_k\|_2$ when $k$ exceeds a certain threshold (which is, in fact, the numerical $\epsilon$-rank of $A$ for a well-chosen $\epsilon$). The corner of the L-curve marks this transition, since it represents a compromise between the minimization of the norm of the residual and the semi-norm of the solution. This is particular evident in Figure 1: the horizontal branch of the "L" is dominated by the regularization error, while the vertical branch shows the sharp increase in the semi-norm caused by propagation errors.

We spend some words to explain the symbols used in this and in the following figures. The heading of the graph displays informations on the test problem and on the method used for its solution. In this case we created a test linear system with the SHAW matrix, taken from [11], and with sample solution *sin2pi*, given

by $x_i = \sin\frac{2\pi i}{n}$, $i = 1, \ldots, n$ (other test matrices and solutions will be introduced in Section 5). The dimension of the system is $n$ and the data vector $\mathbf{b}$ is affected by white noise with variance $n\sigma^2$. This system, as $H = I$, has been solved by TSVD. Each point on the graph stands for the particular regularized solution $\mathbf{x}_k$ whose index labels the point.

A numerical algorithm for the detection of the corner of the L-curve has been introduced in [14]. When the regularization method depends on a continuous parameter $\lambda$, like in Tikhonov regularization, then the L-curve is a smooth function, possibly twice differentiable, and this method selects the value which maximizes the curvature $\kappa(\lambda)$ of the L-curve. If, on the contrary, the regularization parameter is discrete, like in T(G)SVD or in iterative regularization methods, the algorithm selects the parameter closest to the point of maximum curvature of a cubic spline curve with knots resulting from a local smoothing of the L-curve points.

This method has some drawbacks, especially when applied to a discrete L-curve.

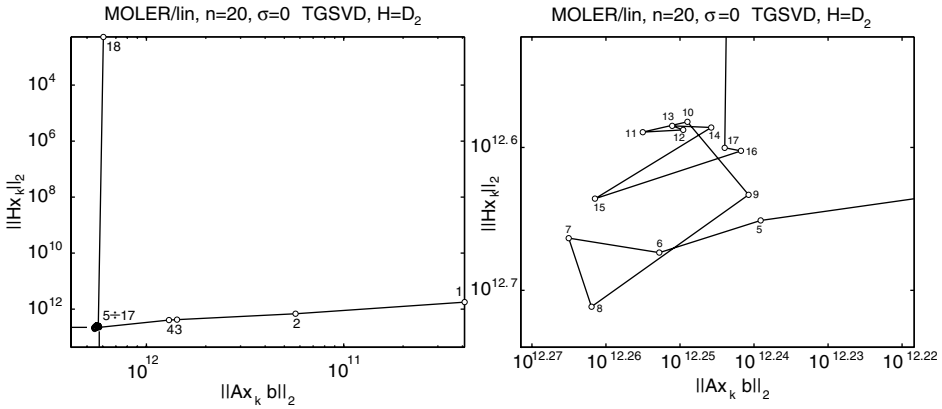


Fig. 2: A cluster of points.

In fact, in T(G)SVD regularization methods the points of the L-curve tend to cluster in a neighborhood of the corner. In this situation, errors due to floating point computations may produce false corners and loss of convexity, as illustrated in Figure 2 which shows an experimental L-curve together with a close-up of a neighborhood of its corner. The effect is that the spline which fits the L-curve often presents unexpected oscillations near the corner, which lead to an inaccurate estimate. This is particularly dangerous when the algorithm returns an over-estimation of the optimal value of $k$, which often causes a large increase in the norm of the computed solution.

Moreover, in many practical situations, some of which are depicted in Figure 3, the L-curve totally looses its "L" shape, making it difficult to choose a good value of the parameter without resorting to some heuristic rule.

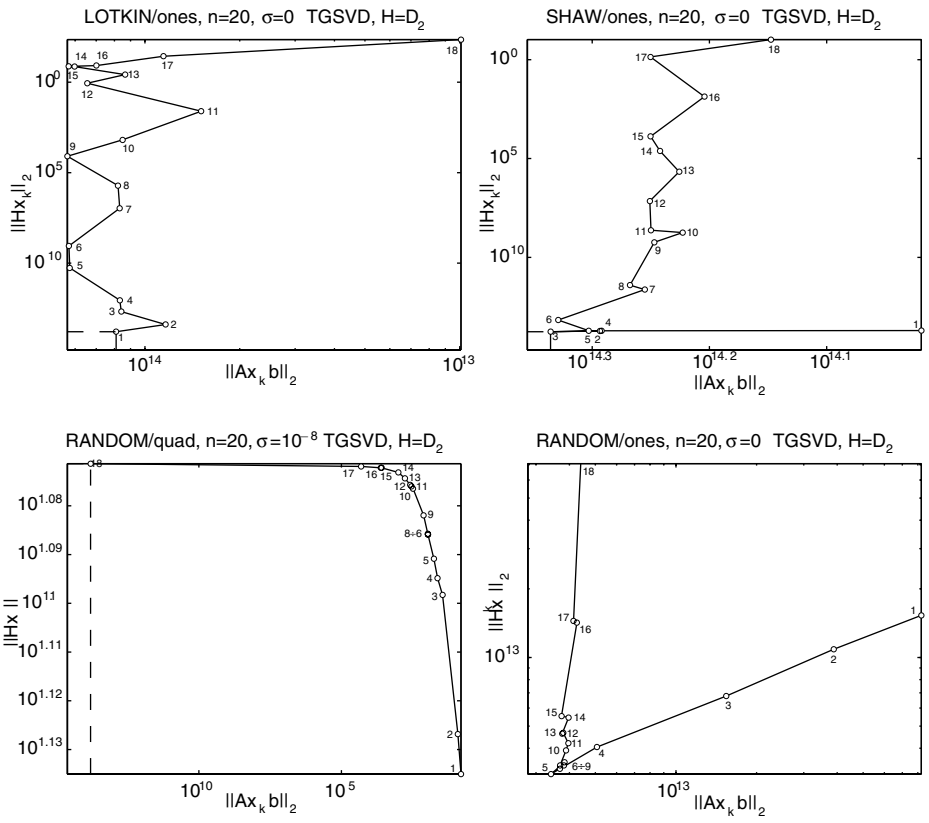


Fig. 3: Some bad L-curves.

We remark, anyway, that these L-curves are still rather informative about the problems we are trying to solve, and their interpretation will allow us, in the next section, to implement an effective strategy which exploits the informations they contain.

The two upper graphs in Figure 3, for example, exhibit a huge increase in the semi-norm of the solution and small changes in the residuals, mostly due to floating-point arithmetics. This situation is typical of linear systems whose solution exactly belongs to the kernel of the regularization matrix $H$ and for which just a few generalized singular values are sufficient to compute a good approximation of the solution. In real applications, to get good results we are interested in using a regularization matrix $H$ whose kernel contains the biggest possible component of the solution $\mathbf{x}$, so it is important that the parameter estimation routine could treat effectively this situation.

The lower left graph in Figure 3, instead, displays a monotonically decreasing residual associated to a negligible growth in the semi-norm, typical of a

well-conditioned (of mildly ill-conditioned) linear system, for which it is possible to use all the singular values in the computation of the solution. The last graph shows both the effects in the same test problem. Obviously, one would not apply regularization to a well conditioned matrix, but since in some applications the matrix $A$ is severely ill-conditioned only for a certain range of dimensions, or in correspondence of particular values of some constants, it would be desirable if the regularization method would automatically detect a well-conditioned matrix, tuning correspondingly the regularization parameter.

## 4 – The *corner* algorithm

The algorithm we propose couples a simple geometrical approach to locate the corner of a discrete L-curve, preceded by a suitable numerical pre-processing of cluster of points, to some empirical rules aimed to recognize two classes of L-curves "without a unique corner" for which it is possible to predict a good value of the regularization parameter. These rules have been devised through analyzing, and interpreting, a large number of experimental L-curves.

The algorithm takes in input the residuals and the semi-norms associated to each regularized solution, namely

$$\|\mathbf{b} - A\mathbf{x}_i\|, \ \|H\mathbf{x}_i\|, \quad i = 1, \dots, p, \tag{4.1}$$

and can be decomposed into three phases.

1. Initially, we try to understand if the solution is approximately in the kernel of the regularization operator $H$. It is important to check this condition first, because if affirmative the L-curve would not be "L" shaped, and the search for a corner would be useless (see upper graphs in Figure 3). The task is performed through detecting extremely small semi-norms, with a test of the type
$$\frac{\min \|H\mathbf{x}_i\|}{\max \|H\mathbf{x}_i\|} < \tau_1.$$
The subroutine applies this test in conjunction with
$$\min \frac{\|H\mathbf{x}_i\|}{\|\mathbf{x}_i\|} < \tau_2$$
whenever the 2-norms of the regularized solutions are made available.
The tolerances $\tau_1$ and $\tau_2$ are two of the four constants the algorithm depends on. They have been fixed to $10^{-12}$ and $10^{-4}$ respectively, working in double precision, and the results do not seem much sensitive on changes in these parameters.

If the detection of small semi-norms is successful, the algorithm stops returning the index of the smaller one as an estimate of the optimal regularization parameter.

2. If the previous test is not verified, we compute the points of the L-curve

$$P_i = (\log_{10} \|\mathbf{b} - A\mathbf{x}_i\|, \log_{10} \|H\mathbf{x}_i\|), \quad i = 1, \dots, p,$$

and the vectors

$$\mathbf{v}_i = P_{i+1} - P_i, \quad i = 1, \dots, p-1.$$

To eliminate clusters (Figure 2) we delete all the "short" vectors, i.e. those verifying the condition

$$\|\mathbf{v}_i\| < \tau_3,$$

leaving $q$ acceptable vectors ($q \leq p$). The constant $\tau_3$, which we fix at $\|P_p - P_1\|/(2p)$, is rather important. Taking smaller values can give very good results, but may also lead to dangerous over-estimates. We are currently analyzing the possibility of choosing the value of $\tau_3$ adaptively.
After normalizing the $q$ remaining vectors (we are only interested in their orientation) the situation is similar to the one depicted in Figure 4.

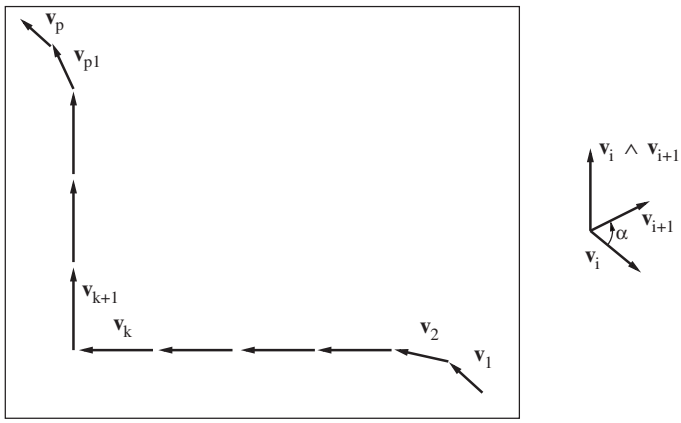


Fig. 4: L-curve and wedge products.

If we travel along the L-curve visiting the vectors $\mathbf{v}_i$ in ascending order, the corner is characterized by an angle $\alpha \simeq -\frac{\pi}{2}$ between $\mathbf{v}_k$ and $\mathbf{v}_{k+1}$. Then, the search for the corner can be carried out by finding the minimum $z$-coordinate of the wedge products between two succeeding vectors (See Figure 4)

$$w_i = (\mathbf{v}_i \wedge \mathbf{v}_{i+1})_z = \|\mathbf{v}_i\| \cdot \|\mathbf{v}_{i+1}\| \cdot \sin\alpha, \quad i = 1, \dots, q.$$

The following elementary property of wedge products

$$(\mathbf{v}_i \wedge \mathbf{v}_{i+1})_z = \det([\,\mathbf{v}_i \quad \mathbf{v}_{i+1}\,])$$

allows to compute the numbers $w_i$ easily and with a small computational effort.

3. The minimum of the wedge products is accepted as a corner only if it verifies the condition

$$\min_{i=1,\dots,p} w_i < \tau_4, \quad (\tau_4 = -0.5). \tag{4.2}$$

The value of $\tau_4$, like the first two constants of the algorithm, does not seem to be very critical for the performance of the method.

If condition (4.2) is not verified, the L-curve is considered to be *without a corner* and we check for the presence of a well conditioned (or mildly ill-conditioned) matrix by detecting a small change in the extremal semi-norms

$$|\log_{10}\|H\mathbf{x}_p\| - \log_{10}\|H\mathbf{x}_1\|\,| < 10.$$

In this case we return $k = p$ as the optimal parameter.

The failure of the last test is an error condition, which we still keep in the subroutine with the hope to trap unforeseen situations, i.e. L-curves not belonging to the three classes we have considered, and to further improve the algorithm.

The outline of the algorithm is reported in Figure 5. A Matlab [16] function is available upon request (send an email to `rodriguez@unica.it`).

We end up with a computational remark. It is known that the residuals and the semi-norms (4.1) can be expressed in terms of the singular system of the matrix $A$. In fact, from (2.4) and (2.7) it follows

$$\|\mathbf{x}_i\|^2 = \sum_{\ell=1}^{i} \left(\frac{\mathbf{u}_\ell^T\mathbf{b}}{\sigma_\ell}\right)^2 \quad \text{and} \quad \|H\mathbf{x}_i\|^2 = \sum_{\ell=p-i+1}^{p} \left(\frac{\mathbf{u}_\ell^T\mathbf{b}}{\gamma_\ell}\right)^2.$$

It is also possible, taking into account that

$$\mathbf{b} - A\mathbf{x}_k = A(\mathbf{x} - \mathbf{x}_k) = A(\mathbf{x}_p - \mathbf{x}_k),$$

to obtain a similar formula for the residuals. Anyway, even if this formula is computationally less expensive, we noticed that the residuals computed in this way are sometimes *too well computed* and give worse results, for what concerns the estimation of the regularization parameter, with respect to the residuals computed by implementing their definition. We feel that the reason for this is that the residuals should be affected by propagation errors in the same amount as the solution is, since we want to extract from them informations about the quality of results.

1. input $\|\mathbf{b} - A\mathbf{x}_i\|$, $\|H\mathbf{x}_i\|$, $\|\mathbf{x}_i\|$, $i = 1, \ldots, p$
2. $\tau_1 = 10^{-12}$, $\tau_2 = 10^{-4}$
3. if $\dfrac{\min\limits_{i=1,\ldots,p} \|H\mathbf{x}_i\|}{\max\limits_{i=1,\ldots,p} \|H\mathbf{x}_i\|} < \tau_1$ and $\min\limits_{i=1,\ldots,p} \dfrac{\|H\mathbf{x}_i\|}{\|\mathbf{x}_i\|} < \tau_2$
   1. $k = \arg \min\limits_{i=1,\ldots,p} \|H\mathbf{x}_i\|$
4. else
   1. for $i = 1, \ldots, p$
      1. $P_i = (\log_{10} \|\mathbf{b} - A\mathbf{x}_i\|, \log_{10} \|H\mathbf{x}_i\|)$
   2. $\tau_3 = \dfrac{\|P_p - P_1\|}{2p}$, $\tau_4 = -0.5$
   3. $q = 0$
   4. for $i = 1, \ldots, p$
      1. $\mathbf{v} = P_{i+1} - P_i$
      2. if $\|\mathbf{v}\| > \tau_3$
         1. $q = q + 1$
         2. $\mathbf{v}_q = \dfrac{\mathbf{v}}{\|\mathbf{v}\|}$
   5. for $i = 1, \ldots, q-1$
      1. $w_i = \det\left(\begin{bmatrix}\mathbf{v}_i & \mathbf{v}_{i+1}\end{bmatrix}\right)$
   6. if $\min\limits_{i=1,\ldots,q} w_i < \tau_4$
      1. $k = \arg \min\limits_{i=1,\ldots,q} w_i$
   7. else
      1. if $|\log_{10} \|H\mathbf{x}_p\| - \log_{10} \|H\mathbf{x}_1\|| < 10$
         1. $k = p$
      2. else
         1. error 'corner not found'
5. output $k$

Fig. 5: The *corner* algorithm.

## 5 – Numerical experimentation

To investigate the performance of our algorithm we applied it to the estimatation of the optimal regularization parameter in a set of test problem, which we solved by TSVD or TGSVD when $H = I$ or $H \neq I$, respectively.

We considered eleven square test matrices, taken from the package *Regularization Tools* [11], [13] (HEAT(1), SHAW, SPIKES, BAART, ILAPLACE) and from Matlab [16] (HILBERT, PASCAL, LOTKIN, MOLER, PROLATE, RANDOM). For $n = 20$, most of these matrices are severely ill-conditioned (in the sense that their condition number exceeds the reciprocal of the *machine epsilon* $\epsilon_M \simeq 2.2 \cdot 10^{-16}$),

two of them are mildly ill-conditioned (MOLER and PROLATE) and one (the RANDOM matrix) is well conditioned. For $n = 80$ they are all severely ill-conditioned, except the RANDOM matrix.

For each test matrix, we constructed different linear systems by computing the right hand side $\mathbf{b}$ corresponding to the sample solutions listed in Table 1 (the *rtools* solution, which is the sample solution adopted in [11], is used only with the matrices coming from the *Regularization Tools*).

**Table 1.** Sample solutions of linear systems.

| | |
|---|---|
| *rtools* | defined as in [11] |
| *ones* | $x_i = 1$ |
| *lin* | $x_i = \dfrac{i}{n}$ |
| *quad* | $x_i = \left(i - \left\lfloor \dfrac{n}{2} \right\rfloor\right)^2 / \left\lceil \dfrac{n}{2} \right\rceil^2$ |
| *sin2pi* | $x_i = \sin \dfrac{2\pi(i-1)}{n}$ |
| *sinpi* | $x_i = \sin \dfrac{\pi(i-1)}{n}$ |
| *lin+sinpi* | $x_i = \dfrac{i}{n} + \sin \dfrac{\pi(i-1)}{n}$ |

The linear systems so obtained were solved both in the presence and in the absence of noise on the data. In practice, the data vector $\mathbf{b}$ was substituted by the vector $\tilde{\mathbf{b}}$, with components

$$\tilde{b}_i = b_i + \sigma\epsilon_i, \quad i = 1, \ldots, n,$$

with $\epsilon_i$ normally distributed with mean value 0 and variance 1. We considered $\sigma = 0, 10^{-8}, 10^{-4}$.

By this procedure we generated 213 test problems, which we solved for $n = 20$ and $n = 80$ either by TSVD ($H = I$) and by TGSVD for each of the regularization matrices $H = D_1, D_2, D_3$, being $D_k$ the Toeplitz matrix of dimension $(n-k) \times n$ whose first row is the discrete approximation of the $k$-th derivative.

For each test problem, the optimal regularization parameter was compared with the estimates furnished by our algorithm (labelled as *corner* in Tables 2 and 3), by the routine *l_curve* from [13], which is an implementation of the *maximum curvature* algorithm described in [14], and by the routine *GCV*, coming from the same package, which returns the minimizer of the functional (2.8).

**Table 2.** Numerical tests, $n = 20$.

| $H$ | corner | l_curve | GCV |
|---|---|---|---|
| $I$ | 102(8/2) | 38(76/37) | 94(42/35) |
| $D_1$ | 89(21/4) | 64(46/31) | 76(74/54) |
| $D_2$ | 92(30/20) | 50(50/39) | 39(136/114) |
| $D_3$ | 101(44/35) | 41(58/40) | 28(119/100) |

Table 2 lists the results obtained for $n = 20$ by applying the three mentioned methods with each of the four adopted regularization matrices. The first number in every entry of the table equals the number of tests in which the optimal parameter is exactly identified (the "full successes"), while the two numbers in parentheses indicate in how many tests an incorrect estimate of the regularization parameter produces an error in the solution which exceeds the optimal one by a factor $10^2$ and $10^4$, respectively (the "failures"). We remark that the algorithms should be considered successful at least when the error on the computed solution is smaller than $10^2$ times the optimal error. Table 3 shows the same results for $n = 80$.

**Table 3.** Numerical tests, $n = 80$.

| $H$ | corner | l_curve | GCV |
|---|---|---|---|
| $I$ | 90(39/33) | 50(66/40) | 72(62/51) |
| $D_1$ | 71(27/24) | 30(75/65) | 56(84/39) |
| $D_2$ | 73(25/17) | 20(125/98) | 56(74/62) |
| $D_3$ | 84(34/17) | 13(128/111) | 44(108/77) |

The numerical results show that the estimates furnished by our algorithm are significantly more trustworthy than the ones produced by the two other algorithms considered. Moreover, the number of the cases of total failure is rather small. We feel that the reason for this is also that our algorithm tends to under-estimate, rather than over-estimate, the optimal regularization parameter. These results confirm, in particular, the great efficacy of the L-curve as a tool for the analysis and the numerical resolution of ill-conditioned linear systems.

## 6 – Future work

In this section we expose the lines of research which we consider important in order to further improve the performance of our algorithm and to extend its range of applicability.

First of all, we are developing an adaptive procedure to choose the values of the constants on which the algorithm depends, in particular the one ($\tau_3$) whose value seems to be the most sensible for the performance of the method.

We also plan to carry on a wider numerical experimentation, with the hope to identify some particular test problems leading to L-curves which our method actually does not recognize, i.e. cases which fall into the final error condition of the algorithm.

Finally we wish to extend the method in order to apply it to iterative regularization methods and to Tikhonov regularization. The difficulty, in the first case, is that the discrete regularization parameter, namely the iteration index, does not have an upper bound, unlike in SVD methods, and we think that a part of the algorithm should be repeated at each iteration to be able to track the overcoming of the corner of the L-curve.

In the second case, where the parameter is a positive real number, we plan to start with a coarse discretization of the L-curve and to add adaptively more points in a neighborhood of the corner until its position is identified up to a prescribed accuracy.

## REFERENCES

[1] Å. Björck: *Numerical Methods for Least Squares Problems*, SIAM, Philadelphia, 1996.

[2] C. Brezinski – M. Redivo-Zaglia – G. Rodriguez – S. Seatzu: *Multi-parameter regularization techniques for ill-conditioned linear systems*, Numer. Math., **94** (2003), 203-228.

[3] P. Craven – G. Wahba: *Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of Generalized Cross Validation*, Numer. Math., **31** (1979), 377-403.

[4] H.W. Engl – M. Hanke – A. Neubauer: *Regularization of Inverse Problems*, Kluwer, Dordrecht, the Netherlands, 1996.

[5] G.H. Golub – C.F. Van Loan: *Matrix Computations*, The John Hopkins University Press, Baltimore, third edition, 1996.

[6] Ch.W. Groetsch: *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Research Notes in Mathematics, Vol. 10, Pitman, London, 1984.

[7] M. Hanke: *Conjugate Gradient Type Methods for Ill-Posed Problems*, Pitman Research Notes in Mathematics, Vol. 327, Longman, Harlow, U.K., 1995.

[8] P.C. Hansen: *The truncated SVD as a method for regularization*, BIT, **27** (1987), 543-553.

[9] P.C. Hansen: *Regularization, GSVD and truncated GSVD*, BIT, **29** (1989), 491-504.

[10] P.C. Hansen: *Analysis of discrete ill-posed problems by means of the L-curve*, SIAM Review, **34** (1992), 561-580.

[11] P.C. Hansen: *Regularization Tools: A Matlab package for analysis and solution of discrete ill–posed problems*, Numerical Algorithms, **6** (1994), 1-35.

[12] P.C. Hansen: *Rank-deficient and discrete ill-posed problems*, SIAM Monographs on Mathematical Modeling and Computation. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1998.

[13] P.C. Hansen: *Regularization Tools version 3.0 for Matlab 5.2*, Numerical Algorithms, **20** (1999), 195-196.

[14] P.C. Hansen – D.P. O'Leary: *The use of the L-curve in the regularization of discrete ill-posed problems*, SIAM J. Sci. Comput., **14** (1993), 1487-1503.

[15] C. Mallows: *Some comments on $C_p$*, Technometrics, **15** (1973), 661-675.

[16] The MathWorks Inc.: South Natick, MA, MATLAB Version 6.5, 2002.

[17] V.A. Morozov: *The choice of parameter when solving functional equations by regularization*, Dokl. Akad. Nauk SSSR, **175** (1962), 1225-1228.

[18] G. Rodriguez – S. Seatzu: *Numerical solution of the finite moment problem in a reproducing kernel Hilbert space*, J. Comp. Appl. Math., **33** (3) (1990), 233-244.

[19] G. Rodriguez – S. Seatzu – D. Theis: *A new technique for ill-conditioned linear systems*, Numerical Algorithms, **33** (2003), 433-442.

[20] A.N. Tikhonov: *Regularization of incorrectly posed problems*, Soviet Math. Dokl., **4** (1963), 1624-1627.

[21] A.N. Tikhonov: *Solution of incorrectly formulated problems and the regularization method*, Soviet Math. Dokl., **4** (1963), 1036-1038.

[22] G. Wahba: *Spline Models for Observational Data*, CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 59, SIAM, Philadelphia, PA, 1990.

*Lavoro pervenuto alla redazione il 15 febbraio 2003*
*ed accettato per la pubblicazione il 6 aprile 2004.*
*Bozze licenziate il 14 febbraio 2005*

Indirizzo degli Autori:

G. Rodriguez, D. Theis – Dipartimento di Matematica e Informatica – Università di Cagliari – viale Merello 92 – 09123 Cagliari, Italy
E-mail: rodriguez@unica.it theis@bugs.unica.it

This work was partially supported by MIUR, grant number 2002014121.