A discrete L-curve for the regularization of ill-posed inverse problems

G. Landi*

May 10, 2012

Abstract

In many applications, the discretization of continuous ill-posed inverse problems results in discrete ill-posed problems whose solution requires the use of regularization strategies. The L-curve criterium is a popular tool for choosing good regularized solutions, when the data noise norm is not a priori known. In this work, we propose replacing the original ill-posed inverse problem with a noise-independent equality constrained one and solving the corresponding first-order equations by the Newton method. The sequence of the computed iterates defines a new discrete L-curve. By numerical results, we show that good regularized solutions correspond with the corner of this L-curve.

Keywords: L-curve; ill-posed inverse problems; regularization; Newton method.

1 Introduction

In this work, we consider the linear system of equations

$$Ax = b \tag{1}$$

where $A \in R^{m\times n}$, $m \geq n$, is an ill-conditioned matrix and the right-hand side b is contaminated by noise and measurement errors. The least-squares solution of (1) may be corrupted with noise and regularization is necessary in order to reduce its sensitivity to noise [11, 12, 13, 43].

This paper is concerned with the problem of computing a regularized solution of (1) when the corresponding noise-free linear system is supposed to be consistent and the data error norm is not a priori known.

Among others, the Tikhonov method, the Conjugate Gradient method applied to the normal equations (CGNR), the Truncated Singular Value Decomposition (TSVD) method and the Truncated Generalized Singular Value Decomposition (TGSVD) method are widely used regularization techniques

*Department of Mathematics, University of Bologna, e-mail: landig@dm.unibo.it

[24, 22, 18, 13, 15]. In order to compute stable solutions, these methods need suitable values of a parameter, called regularization parameter.

Let $L \in \mathbb{R}^{p\times n}$, $p \leq n$, be a certain regularizing matrix containing prior information about the exact solution such as, for example, the identity matrix or the discrete approximation of the first or second order derivative operator. Tikhonov method [15, 43, 42] obtains a regularized solution x_k as

$$x_k = \arg\min_x \ \|Ax - b\|^2 + k^2\|Lx\|^2$$

for a proper choice of the regularization parameter $k > 0$. (Throughout this paper, $\|\cdot\|$ denotes the Euclidean norm.)

The CGNR method [37, 16, 18] determines a regularized solution x_k by minimizing the residual norm $\|b-Ax\|$ in the Krylov subspace $x_0+\mathcal{K}_k\big(A^TA, A^T(b-Ax_0)\big)$ for a suitable value of the iteration number k, representing the regularization parameter.

Let $A = U\Sigma V^T = \sum_{i=1}^n u_i\sigma_i v_i^T$ be the SVD decomposition [14] of A. The TSVD method [27, 24] obtains a regularized solution

$$x_k = \sum_{i=1}^{k} \frac{u_i^T b}{\sigma_i} v_i, \quad k < n$$

for a suitable value of the truncation index k, playing the role of the regularization parameter.

Consider the GSVD decomposition [14] of the matrix pair (A, L):

$$A = U \begin{pmatrix} \Sigma & 0 \\ 0 & I_{n-p} \end{pmatrix} Z^{-1}, \quad L = V(M, 0)Z^{-1}. \tag{2}$$

The TGSVD method [20, 21] determines a sequence of solutions

$$x_k = \sum_{i=p-k+1}^{p} \frac{u_i^T b}{\sigma_i} z_i + \sum_{i=p+1}^{n} u_i^T b z_i, \quad k < n.$$

As in the TSVD regularization method, the truncation index k is the regularizing parameter giving the TGSVD regularized solution.
In all the previous methods, a sequence of regularized solutions $\{x_k\}$ is obtained for different values of the regularization parameter $k > 0$. The solution quality strongly depends on the choice of the regularization parameter, controlling the amount of regularization. Observe that k is a discrete parameter in the CGNR, TSVD and TGSVD methods, while it is continuous in the Tikhonov method.

The L-curve criterion [26, 22] is a popular practical method for the regularization parameter choice which does not require any prior information about the error norm. This criterion is basically based on the plot, in a log-log scale, of the regularized solution norm $\|Lx_k\|$ ($L = I_n$, for the CGNR method) versus the residual norm $\|Ax_k - b\|$, for several values of the regularization parameter k. Very often, this curve is L-shaped and its corner corresponds to an optimal value of the regularization parameter. Intuitively, the L-curve describes, in a graphical way, the behavior of the least squares minimization problem as a function of k and its corner represents a compromise between the fitting to the data and the smoothing of the solution.

The main limitation in the use of the L-curve is that its computation is potentially an expensive task for large-size problems especially for Tikhonov and TSVD methods where the repeated solution of the corresponding regularization problems is required for many values of the regularization parameter. However, the L-curve criterion is used for the regularization of linear systems arising in many real-world applications where the error norm is usually not available; for this reason, the search for effective algorithms for its computation is actually an active area of research and in recent years some variants of the L-curve have been proposed. In [38], the so-called residual L-curve is proposed for TSVD regularization. This L-curve, which differs from that described in [26], is the graph of the residual error norm $\|b - \sum_{j=1}^{n}(u_j^T b)v_j\|$ versus the iteration number.In [5, 8], a new approach is described where an approximated value of the Tikhonov regularization parameter is determined from a L-ribbon containing the L-curve and computed by partial Lanczos bidiagonalizations. In [6, 8], a similar technique, based on partial Lanczos bidiagonalizations, is proposed for the computation of a curvature-ribbon containing the graph of the curvature of the L-curve. In [39], a new L-curve is proposed for Tikhonov regularization. This new L-curve is based on the graph of the approximated solution norm versus the regularization parameter and seems to perform better that the traditional L-curve for very smooth solution. In [44], the so-called f-slope method is presented where the regularization parameter of Tikhonv method is selected as the point on the graph of the points $(\ln(1/k), \|x_k\|)$ at which the slope is the flattest. Also this method requires the solution of many regularization problems for different values of the regularization parameter. In [9] an approximation to an analytic expression of the L-curve is presented and, for large-scale problems, a new method is proposed for the determination of Tikhonov regularization parameter. This method creates a new L-curve from the residual and solution norms obtained from CG iterations.

Indeed, the aforementioned recent literature indicates a great interest in the development of L-curve-based regularization algorithms.

The main contribution of this work consists in proposing a new discrete L-curve for computing regularized solutions of (1). In particular, starting from an equality constrained optimization problem, a sequence of approximations x_k, $k = 1, 2, \ldots$, is obtained such that the solution norm $\|Lx_k\|$ increases and the residual norm $\|Ax_k - b\|$ decreases with k. A new discrete L-curve is obtained by connecting the points

$$\Big(\log_{10}\|Ax_k - b\|, \log_{10}\|Lx_k\|\Big) \tag{3}$$

and a suitable regularized solution is computed as the solution x_k corresponding to the L-curve corner. The new method is also well-suited for linear systems whose coefficient matrix is very large or it is available only as an operator for matrix-vector products.

As far as the detection of the L-curve corner is concerned, several algorithms have been proposed in the literature. For continuous L-curves, like in Tikhonov regularization, the method proposed in [26] by Hansen and O'Leary determines the corner as the point of maximum curvature. For discrete L-curves, like in the CGNR and T(G)SVD methods, popular corner detecting schemes are the cubic spline-based method of Hansen and O'Leary [26], the adaptive pruning method of Hansen, Jensen and Rodriguez [25] and the triangle method of Castellanos, Gómez and Guerra [10]. The development of an efficient algorithm for locating the L-curve vertex is beyond the scope of this work. Even if in our numerical experiments the vertex can be detected quite easily by visual inspection, in order to locate it with an automatic procedure, both the triangle method of Castellanos et al. [10] and the pruning method of Hansen et al. [25] have been used. The results of these algorithms are compared in the numerical experiments.

In order to evaluate the effectiveness of the proposed L-curve, several comparisons with well-known L-curve-based strategies have been made and various numerical experiments have been performed.

The sequel is organized as follows. The new discrete L-curve is described in section 2 and the numerical results are presented in section 3; finally, conclusions are given in section 4.

2 The L-curve

In order to compute the sequences $\{\|Lx_k\|\}$ and $\{\|Ax_k - b\|\}$ defining the new L-curve, we consider the equality constrained minimization problem

$$\begin{array}{ll} \text{minimize} & \dfrac{1}{2}\|Lx\|^2 \\ \text{subject to} & \dfrac{1}{2}\|Ax - b\|^2 = \dfrac{\gamma^2}{2} + \varepsilon \end{array} \tag{4}$$

where γ is the residual of the minimal-norm least-squares solution x_{LS} to (1) and ε is a small positive parameter. The parameter ε is indeed introduced in order to overcome the non-regularity of the minimizer x_{LS} of the objective function of (4) subject to $\|Ax - b\|^2/2 = \gamma^2/2$, and it is chosen so that $\gamma^2/2 + \varepsilon$ is smaller than $\rho^2/2$, where ρ is the data noise norm. Even under the hypothesis of no knowledge about the noise norm, this is not a restrictive condition because, in practical applications, ε can be chosen close to the machine precision.

Observe that problem (4) is noise-independent since the right-hand side of the equality constraint does not contain any information about the noise affecting the data. For this reason, problem (4) is different from the following optimization problems usually introduced in the literature for the regularization of discrete ill-posed problems :

$$\begin{array}{ll} \text{minimize} & \dfrac{1}{2}\|Lx\|^2 \\ \text{subject to} & \dfrac{1}{2}\|Ax - b\|^2 = \dfrac{1}{2}R_{\text{exact}} \end{array} \tag{5}$$

and

$$\begin{array}{ll} \text{minimize} & \dfrac{1}{2}\|Ax-b\|^2 \\ \text{subject to} & \dfrac{1}{2}\|Lx\|^2 = \dfrac{1}{2}L_{\text{exact}} \end{array} \tag{6}$$

where $R_{\text{exact}} = \|Ax_{\text{exact}} - b\|^2$ and $L_{\text{exact}} = \|Lx_{\text{exact}}\|^2$ with x_{exact} the unknown exact solution of (1) (see, for example, [41, 7, 4, 30]). Problems (5) and (6) require some a priori information about x_{exact} and their solutions are regularized solutions of (1). On the contrary, the solution of (4) is not regularized since its residual norm is forced by the equality constraint to be smaller than the error norm and close to the residual norm of the least squares solution. Problem (4) has been introduced in [28, 29, 31] for the regularization of discrete ill-posed problems when the error norm is supposed to be known. In [28, 29, 31] a Lagrange method is applied to (4) and a regularized solution is obtained by stopping its iterations according to the discrepancy principle when an iterate has been determined whose residual norm is less than the error norm. In this work, we suppose the error norm to be not explicitly known and we derive a new L-curve in order to determine a good regularized solution of (1).

Let $h(x)$ be the constraint function

$$h(x) = \frac{1}{2}\|Ax-b\|^2 - \frac{\gamma^2}{2} - \varepsilon$$

and let

$$\mathcal{L}(x,\lambda) = \frac{1}{2}\|Lx\|^2 + \lambda h(x)$$

be the Lagrangian function where $\lambda \in \mathbb{R}$ is the Lagrange multiplier.

The solution x^* of (4) is unique and has a positive Lagrange multiplier λ^* [1, 3] such that the pair (x^*, λ^*) satisfies the first-order necessary conditions

$$\begin{cases} \nabla_x \mathcal{L}(x,\lambda) = 0 \\ h(x) = 0. \end{cases} \tag{7}$$

Conditions (7) are also sufficient for optimality [1, 3].

2.1 Computing the sequence $\{x_k\}$

The Newton method with a line search is applied to the first-order equations (7) in order to compute the sequence $\{x_k\}$. In particular, given an initial iterate (x_0, λ_0), an iteration of the Newton method to solve (7) is stated as

$$\begin{aligned} x_{k+1} &= x_k + \alpha_k \Delta x_k \\ \lambda_{k+1} &= \lambda_k + \alpha_k \Delta\lambda_k \end{aligned} \tag{8}$$

where $(\Delta x_k, \Delta\lambda_k)$ is the Newton step solving the linear system

$$\begin{pmatrix} \nabla^2_{xx}\mathcal{L}(x_k,\lambda_k) & \nabla_x h(x_k) \\ \nabla_x h(x_k)^T & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = -\begin{pmatrix} \nabla_x \mathcal{L}(x_k,\lambda_k) \\ h(x_k) \end{pmatrix} \tag{9}$$

and the step-length α_k is chosen in order to guarantee a sufficient decrease of the merit function

$$m(x,\lambda) = \frac{1}{2}\Big(\|\nabla_x \mathcal{L}(x,\lambda)\|^2 + w|h(x)|^2\Big), \quad w > 0. \tag{10}$$

The parameter w acts as a penalty parameter since for large w the constraint violations along the search direction are penalized. The step-length α_k is therefore chosen as the first number of the sequence $\{1, 2^{-1}, 2^{-2}, \ldots\}$ satisfying the Armijo condition

$$m(x_k+\alpha_k\Delta x_k, \lambda_k+\alpha_k\Delta\lambda_k) \leq m(x_k,\lambda_k)+\mu\alpha_k\left(\Delta x_k, \Delta\lambda_k\right)^T \nabla m(x_k,\lambda_k) \tag{11}$$

for some fixed $\mu \in (0,1)$.

Observe that $m(x,\lambda)$ is the sum of the squares of the equations of the system

$$\begin{cases} \nabla_x \mathcal{L}(x,\lambda) = 0 \\ \sqrt{w}h(x) = 0 \end{cases} \tag{12}$$

that is, $m(x,\lambda)$ is the most widely used merit function for (12) [35]. Moreover, observe that the solution of (9) is also the Newton step for (12). Therefore, the sequence $\{(x_k,\lambda_k)\}$ converges to the solution (x^*,λ^*) of (7) under standard assumption for the convergence of Newton method with line search [2, 32, 35].

The following result summarizes the above discussion.

Theorem 2.1. *Consider the Newton iteration* (8) *where* $(\Delta x_k, \Delta\lambda_k)$ *is the solution to* (9) *and* α_k *is chosen to satisfies* (11). *Assume that the points* $(\Delta x_k, \Delta\lambda_k)$ *exist and that lie in a compact set. Then any limit point of* $\{(x_k,\lambda_k)\}$ *satisfies conditions* (7).

(For a proof please refer to [2, 32, 35]). Under the conditions of the previous theorem, the sequence $\{(x_k,\lambda_k)\}$ converges to a solution of (4). A suitable initial iterate (x_0,λ_0) can be computed as follows. Let B be an approximation to the hessian of the objective function $L^T L$ containing its diagonal part. The initial iterate (x_0,λ_0) is chosen such that:

$$\begin{aligned} &x_0 = 0, \\ &\lambda_0 = \Big(\nabla_x h(x_0)^T B^{-1} \nabla_x h(x_0)\Big)^{-1}\Big[h(x_0) - \nabla_x h(x_0)^T B^{-1} B x_0\Big]. \end{aligned} \tag{13}$$

The formula for λ_0 is used in the Multiplier Update Methods (§14.4 of [32]) for updating the Lagrange multiplier. This value of λ_0 is obtained by approximating $\nabla_{xx}\mathcal{L}(x_0,\lambda_0)$ with B and by fully solving the Newton system (9) with respect to λ. Notice that computing $\Big(\nabla_x h(x_0)^T B^{-1} \nabla_x h(x_0)\Big)^{-1}$ merely requires the inversion of a scalar.

2.2 Computing the search direction $(\Delta x_k, \Delta\lambda_k)$

In this subsection, a brief discussion about efficient methods for solving (9) is given. Firstly, let us recall that the coefficient matrix of (9) is symmetric and always indefinite [35].

If the dimension of (9) is small, a direct method, such as the Gaussian Elimination method with partial pivoting, can be used.

If the dimension is large, iterative methods can be efficiently employed to inexactly solve (9). To take advantage of the symmetry of the system, the MINimal RESidual (MINRES) method of Paige and Saunders [36] can be used. In order to guarantee that the computed step is a descent direction for the merit function (10), the inexact Newton method is applied to (12) instead of (7). The inner iterative solver is stopped when a step $(\Delta x_k, \Delta\lambda_k)$ has been determined such that

$$\|r_k\| \leq \eta_k \left\| \begin{pmatrix} \nabla_x \mathcal{L}(x_k, \lambda_k) \\ \sqrt{w} h(x_k) \end{pmatrix} \right\| \tag{14}$$

where r_k is the residual

$$r_k = \begin{pmatrix} \nabla^2_{xx}\mathcal{L}(x_k, \lambda_k) & \nabla_x h(x_k) \\ \sqrt{w}\nabla_x h(x_k)^T & 0 \end{pmatrix} \begin{pmatrix} \Delta x_k \\ \Delta\lambda_k \end{pmatrix} + \begin{pmatrix} \nabla_x \mathcal{L}(x_k, \lambda_k) \\ \sqrt{w} h(x_k) \end{pmatrix}$$

and the sequence $0 < \eta_k \leq \eta < 1$, for all k.

Finally, let $Q(x_k, \lambda_k)$ be an easily invertible approximation to $\nabla^2_{xx}\mathcal{L}(x_k, \lambda_k)$. Eventually, $Q(x_k, \lambda_k) = \nabla^2_{xx}\mathcal{L}(x_k, \lambda_k)$. In this case, (9) can be approximated with

$$\begin{pmatrix} Q(x_k, \lambda_k) & \nabla_x h(x_k) \\ \nabla_x h(x_k)^T & 0 \end{pmatrix} \begin{pmatrix} \Delta x \\ \Delta\lambda \end{pmatrix} = - \begin{pmatrix} \nabla_x \mathcal{L}(x_k, \lambda_k) \\ h(x_k) \end{pmatrix} \tag{15}$$

and an explicit solution of (15) can be computed as (§14.1 of [32]):

$$\begin{aligned} \Delta\lambda_k &= -\Big(D(x_k, \lambda_k)\Big)^{-1}\Big(\nabla_x h(x_k)^T \big(Q(x_k, \lambda_k)\big)^{-1}\nabla_x \mathcal{L}(x_k, \lambda_k) - h(x_k)\Big) \\ \Delta x_k &= -\Big(Q(x_k, \lambda_k)\Big)^{-1}\Big(\nabla_x h(x_k)\Delta\lambda_k + \nabla_x \mathcal{L}(x_k, \lambda_k)\Big) \end{aligned} \tag{16}$$

where

$$D(x_k, \lambda_k) = \nabla_x h(x_k)^T \big(Q(x_k, \lambda_k)\big)^{-1} \nabla_x h(x_k)$$

is a scalar.

2.3 Constructing the L-curve and locating its corner

Let $r_k = b - Ax_k$ be the residual at the k-th iterate. Given $x_0 = 0$, the iterative procedure (8) defines a solution norm sequence $\{\|Lx_k\|\}_{\mathbb{N}}$ and a residual norm sequence $\{\|r_k\|\}_{\mathbb{N}}$ respectively converging to $\|Lx^*\|$ and $\|r^*\|$ such that

$$\|Lx^*\| \geq \|Lx_0\| = 0, \tag{17}$$

$$\|r^*\| \leq \|r_0\|. \tag{18}$$

In fact, x_0 is the minimum of the unconstrained problem

$$\min_x \|Lx\|^2$$

but x_0 is not the solution of (4) since

$$\frac{1}{2}\|Ax_0 - b\|^2 = \frac{1}{2}\|b\|^2 > \frac{1}{2}\gamma^2$$

and

$$\frac{1}{2}\|Ax_0 - b\|^2 = \frac{1}{2}\|b\|^2 > \frac{1}{2}\gamma^2 + \varepsilon$$

if ε is sufficiently small. Thus, $x^* \neq x_0$ and $\|r_0\| = \|b\| > \sqrt{\gamma^2 + 2\varepsilon} = \|r^*\|$.

Even if the sequences $\{\|Lx_k\|\}_{\mathbb{N}}$ and $\{\|r_k\|\}_{\mathbb{N}}$ do not have a monotonic behavior, it is possible to show that a non empty, eventually finite, index subset $\mathcal{K}$ can be always extracted such that $\{\|Lx_k\|\}_{\mathcal{K}}$ and $\{\|r_k\|\}_{\mathcal{K}}$ are respectively increasing and decreasing with $k \in \mathcal{K}$.

Let us consider the sequences

$$\begin{aligned} a_k &:= \|Lx_0\| - \|Lx_k\| \\ b_k &:= \|r_0\| - \|r_k\| \end{aligned}$$

They respectively converge to $a^* = \|Lx_0\| - \|Lx^*\| < 0$ and $b^* = \|r_0\| - \|r^*\| > 0$. Then, two positive integers $\overline{k}_1 \in \mathbb{N}$ and $\overline{k}_2 \in \mathbb{N}$ exist such that

$$a_k < 0, \forall k \geq \overline{k}_1 \quad \text{and} \quad b_k > 0, \quad \forall k \geq \overline{k}_2. \tag{19}$$

By setting $\overline{k} = \max\{\overline{k}_1, \overline{k}_2\}$ we get

$$a_k < 0, \quad \text{and} \quad b_k > 0, \quad \forall k \geq \overline{k} \tag{20}$$

and

$$\|Lx_0\| < \|Lx_{\overline{k}}\|, \quad \text{and} \quad \|r_0\| > \|r_{\overline{k}}\|, \quad \forall k \geq \overline{k}. \tag{21}$$

Thus $\{0, \overline{k}\} \subseteq \mathcal{K}$. This shows that a non empty index subset $\mathcal{K}$ always exists and, in the worst case, it reduces to $\{0, \overline{k}\}$.

In order to create the index subset $\mathcal{K}$, the proposed algorithm checks the monotonicity of the solution and residual norm sequences and discards those points where the monotonicity condition is not satisfied.

The algorithm for the index subset $\mathcal{K}$ starts setting $0 \in \mathcal{K}$. Given $k \in \mathcal{K}$, let $\phi(k) \in \mathbb{N}$ be the least index with

$$\|Lx_{\phi(k)}\| > \|Lx_k\| \quad \text{and} \quad \|r_{\phi(k)}\| < \|r_k\|. \tag{22}$$

Then $\phi(k)$ is the term in $\mathcal{K}$ which follows k.

The proposed new discrete L-curve is the two-dimensional graph of the points

$$\Big(\log_{10}\|Ax_k - b\|, \log_{10}\|Lx_k\|\Big), \quad k \in \mathcal{K}. \tag{23}$$

Very often, this graph looks like the letter L and we propose to choose the iterate x_k corresponding to vertex point of the L-curve (23) as an approximated solution to (1).

Basically, the L-curve criterium is used here to single out one of the Newton iterates that may be a good regularized solution, even if the limit of the entire Newton sequence is not a useful one. A justification for why one of the Newton iterate can be used as a regularized solution can be given as follows.

The optimal regularized solution (in the sense of Tikhonov regularization) solves the constrained minimization problem

$$\begin{aligned} &\text{minimize} && \frac{1}{2}\|Lx\|^2 \\ &\text{subject to} && \frac{1}{2}\|Ax-b\|^2 = \frac{\rho^2}{2}. \end{aligned} \tag{24}$$

For iterative regularization, an optimal parameter choice strategy is the *discrepancy principle* stating that the iterations of the regularization method should be terminated as soon as an iterate x_k has been determined whose residual norm $\|Ax_k - b\|$ is less than ρ.

Given $x_0 = 0$, the sequence x_k converges to the solution of (4) and a finite index ℓ exists such that

$$\|Ax_\ell - b\| < \rho.$$

(This is true because $\|Ax_0 - b\|^2/2 = \|b\|^2/2 > \|\rho\|^2/2 > \gamma^2/2 + \varepsilon$, if ε is sufficiently small.) Such an iterate x_ℓ satisfies the discrepancy principle and, as shown in [31], it is a useful regularized solution. An extensive numerical experimentation shows that x_ℓ is close to the L-curve corner even if sometimes the index ℓ is not included in $\mathcal{K}$ by the proposed automatic procedure for constructing $\mathcal{K}$.

2.4 The overall algorithm

The overall algorithm for computing a regularized solution to (1) can be summarized as follows.

Algorithm 2.1.

Step 1: Computation of the sequence $\{\mathbf{x_k}\}$, $\mathbf{k} = \mathbf{0}, \mathbf{1}, \ldots$

Compute (x_0, λ_0) by (13). Choose $\mu \in (0, 1)$.
Repeat for $k = 0, 1, 2, \ldots$ until convergence

1. Compute $(\Delta x_k, \Delta\lambda_k)$ by solving (9);
2. Find the smallest number $\{1, 2^{-2}, \ldots, 2^{-i}, \ldots\}$ satisfying (11);
3. Set $x_{k+1} = x_k + \alpha_k \Delta x_k$ and $\lambda_{k+1} = \lambda_k + \alpha_k \Delta\lambda_k$;
   set $k = k + 1$;

Step 2: Computation of the index subset $\mathcal{K}$

Set $\underline{r} = \|r_0\|$ and $\overline{r} = \|Lx_0\|$; set $\mathcal{K} = \{0\}$.
For $j = 1, \ldots, k$ repeat

if $\|Lx_j\| > \overline{r}$ and $\|r_j\| < \underline{r}$

set $\mathcal{K} = \mathcal{K} \cup \{j\}$, $\overline{r} = \|Lx_j\|$ and $\underline{r} = \|r_j\|$

Step 3: Computation of the regularized solution $\mathbf{x_{reg}}$

Determine the index k^* corresponding to corner of the L-curve defined by the points $\Big(\log_{10}\|Ax_k - b\|, \log_{10}\|Lx_k\|\Big)$, $k \in \mathcal{K}$.
Set $x_{\text{reg}} = x_{k^*}$.

In our numerical experiments, both the triangle method of Castellanos et al. [10] and the pruning method of Hansen et al. [25] have been used for locating the L-curve vertex.

The iterative procedure at step 1 of algorithm 2.1 is terminated according to the following stopping criteria:

i) when the step norm $\|(\Delta x_k, \Delta\lambda_k)\|$ has became larger than a given positive tolerance τ_1; i.e. when

$$\big\|(\Delta x_k, \Delta\lambda_k)\big\| \geq \tau_1 \tag{25}$$

ii) when the constraint value $h(x_k)$ is sufficiently close to zero; i.e. when

$$\big|h(x_k)\big| \leq \tau_2 \tag{26}$$

where τ_2 is a given positive tolerance;

iii) when a maximum number $K_{\max}$ of allowed iterations has been performed.

The stopping criteria (25) and (26) are motivated by the consideration that we are interested in a good approximation to the exact solution to (1) rather than (4). Therefore, an iterate x_k is considered a sufficiently good approximation to the solution of (4) when it satisfies the constraint to the accuracy τ_2. Moreover, as the iterates x_k approach the solution of (4), the matrix of (9) may become increasingly ill-conditioned. This ill-conditioning may cause the direction norm $\|(\Delta x_k, \Delta\lambda_k)\|$ to be extremely large. This consideration justifies the stopping criterium (25).

3 Numerical results

In this section, we describe the results of several numerical experiments in order to illustrate the performance of the proposed method. The numerical experiments have been executed on a Pentium IV PC using Matlab 7.0 (Release 14).

For all the tests, a maximum number of $K_{\max} = 200$ iterations has been set and the values $\tau_1 = 10^8$ and $\tau_2 = 10^{-8}$ have been fixed for the step norm and constraint tolerances, respectively.

For the first set of test problems, which are taken from the Regularization Tools package by P. C. Hansen [23, 24], we compare the new L-curve with the L-curve for Tikhonov, T(G)SVD and CGNR methods. For these last methods, the codes from [23] have been used for computing the L-curve, its corner and the corresponding regularized solution. The identity matrix and an approximation to the first and second order derivative operators have been considered as regularizing operators L. They are respectively referred as I, ∇ and Δ in the following.

The considered problems, whose size is $n = 128$, are listed in tables 1 and 2. For these experiments, the selected value of the penalty parameter is $w = 10^{10}$. Since the dimension of the problems is small, the Gaussian Elimination method is used for solving (9). Both the pruning method and the triangle method are used for detecting the corner of the new L-curve; these methods are respectively referred as "alg. 2.1 (p)" and "alg. 2.1 (t)" in the tables.

The relative error of the various regularized solutions is reported in tables 1 and 2. Particularly, table 1 shows the results obtained when the level δ of the white noise e degrading the noise-free right-hand side $\widetilde{b}$ is $\delta := \|e\|/\|\widetilde{b}\| = 0.005$, while, in table 2, the methods are compared for different noise levels ($\delta = 0.005, 0.01$) and different realization numbers of the noise. The numerical results in tables 1 and 2 show that algorithm 2.1 with the triangle method has the best overall performance. The overall behavior of the algorithm 2.1 with the pruning method is however better than that of the other classic methods (Tikhonov, T(G)SVD, CGNR).

Figure 1 depicts the results obtained for the *phillips* test problem when $L = \Delta$ and $\delta = 0.005$. In this case, the algorithm performs 32 outer iterations and both the triangle and the pruning methods locate the vertex at $k = 13$. Figure 1(a) shows the computed L-curve and the identified vertex (indicated by the triangle). Figure 1(b) plots, in a semi-log scale, the relative error of the computed approximations illustrating that the iteration (8) exhibits a semiconvergence behavior, i.e. the iterates initially approximate the exact solution, then they move away from it. The vertex of the L-curve corresponds to a solution with small relative error.

Figure 2(a) shows the L-curve for the *Baart* test problem ($L = \Delta$, $\delta = 0.005$). In the graph, the upward-pointing triangles indicate the points on the computed L-curve, i.e. the points $(\log_{10} \|r_k\|, \log_{10} \|Lx_k\|)$, $k \in \mathcal{K}$, while the downward-pointing triangles indicate the discarded points where the monotonicity condition is not satisfied. The circle (figure 2(b)) denotes the detected vertex by both the triangle and pruning methods. Figures 2(c) and 2(d) show the solution and residual norm sequences, respectively, while figure 2(e) depicts the relative error behavior.

Figures 3 and 4 show two L-curves which are problematic for Thikonov and TSVD methods. Also in these cases, the proposed L-curve is L-shaped and the corner is correctly detected.

A well-known limitation of the L-curve criterion for Tikhonov method [17] is that, when the exact solution of the problem is smooth, the computed solution may not converge to the exact solution as the noise level in the data goes to zero. In order to illustrate the behavior of the new L-curve in this circumstance, the *Heat* test problem from [23] has been considered for decreasing values of δ ($n = 128$, $L = I$). Particularly, the values $\delta = 5 \cdot 10^{-2}, 5 \cdot 10^{-3}, 5 \cdot 10^{-4}, 5 \cdot 10^{-5}$ have been set. Figure 5 shows the relative error of the approximations versus the noise level. The circles, triangles and squares denote the solutions computed by the proposed method with the triangle algorithm, by the TGSVD method and by Tikhonov method, respectively. Figure 5 indicates that the proposed method has good performance also for small values of δ.

In order to evaluate the efficiency and the effectiveness of the new L-curve for large-size problems, two image deblurring test problems have been considered.

The first one is the *Blur* test problem from [23]. The dimension of the problem is $n = 128 \times 128$ and the noise level is $\delta = 0.01$. The parameter of the merit function is $w = 1$. The MINRES method has been used for inexactly solving the Newton system (9) by stopping its iterations when condition (14) is satisfied with $\eta = 0.5$ or after a maximum of 300 allowed iterations. The performance of algorithm 2.1 has been compared with that of the CGNR method. For locating the corner of both the new and CGNR L-curves, the pruning and triangle algorithms have been used. Table 3 shows the numerical results provided by the two methods. Algorithm 2.1 has performed a total of 166 MINRES iterations, each requiring a matrix-vector product. No backtracks have been necessary (i.e. $\alpha_k = 1$ for all k). Each CGNR iteration requires two matrix-vector products and a maximum of 300 CGNR iterations have been performed. The results in table 3 indicate that algorithm 2.1 gives a regularized solution with a smaller relative error requiring less matrix-vector products than the CGNR method. Figure 6 shows the computed regularized solutions and figure 7 depicts the L-curves and the relative error graphs.

The last image deblurring test problem concerns the restoration of an image degraded by a space invariant point spread function (PSF) and gaussian noise.

The blurring matrix A, $n = 256 \times 256$, is a large ill-conditioned block Toeplitz matrix with Toeplitz blocks (BTTB) arising from the discretization of a first-kind integral equation. The test problem is a computer simulation of a field experiment showing a satellite as taken from a ground-based telescope. The set of data was developed at the US Air Force Phillips Laboratory, Laser and Imaging Directorate, Kirtland US Air Force Base, New Mexico [34, 40], and has been widely used in the literature for testing image restoration methods. The data set [33] contains the exact image (fig. 8(a)), the blurred and noisy one (fig. 8(b)) and the PSF. Due to the BTTB structure of A, matrix-vector products involving A can be done efficiently using the Fast Fourier Transform (FFT) [19].

Let $L = I$ and let C be a block circulant matrix with circulant blocks (BCCB) such that $C^T C$ is a BCCB approximation to $A^T A$. Then, the matrix

$$Q(\lambda) = I + \lambda C^T C \tag{27}$$

is again a BCCB matrix and is an approximation of the true hessian $\nabla^2_{xx}\mathcal{L}(x, \lambda)$. In particular, the matrix $Q(\lambda)$ is symmetric and positive definite and is easily inverted by using FFTs [19]. Therefore, by approximating the true hessian $\nabla^2_{xx}\mathcal{L}(x, \lambda)$ with $Q(\lambda)$, a quasi-Newton direction $(\Delta x_k, \Delta\lambda_k)$ is obtained using the formulas (16).

The value of the penalty parameter of the merit function is $w = 10^{12}$. Figure 8(c) illustrates the deblurred image corresponding to the corner of the L-curve. (We mention that, in figure 8, images intensities are displayed in “negative gray-scale”.) The L-curve is depicted in figure 8(d). The computation of the reconstructed image requires a total of 312 FFTs.

4 Conclusions

In this work, a new discrete L-curve has been presented for the computation of regularized solutions of discrete ill-posed inverse problems in the absence of prior information about the noise level on the data. The results of numerous numerical experiments are promising and indicate that the proposed method is able to compute good regularized solutions even for large-size problems.

References

[1] M. Avriel. *Nonlinear programming: analysis and methods.* Dover Publications, Inc., New York, 2003.

[2] D. Bertsekas. *Nonlinear Programming.* Athena Scientific, (2nd Edition), 1999.

[3] D. Bertsekas. *Convex Analysis and Optimization.* Athena Scientific, 2003.

[4] P. Blomgren and T. F. Chan. Modular solvers for image restoration problems using the discrepancy principle. *Numer. Linear Algebra Appl.*, 9:347–358, 2002.

Problem	L	Method	Rel. Error	Problem	L	Method	Rel. Error
phillips	I	Alg. 2.1 (p)	$8.4647{\cdot}10^{-2}$	phillips	Δ	Alg. 2.1 (p)	$1.9483{\cdot}10^{-2}$
phillips	I	Alg. 2.1 (t)	$3.8481{\cdot}10^{-2}$	phillips	Δ	Alg. 2.1 (t)	$1.9483{\cdot}10^{-2}$
phillips	I	Tikh	$9.7262{\cdot}10^{-2}$	phillips	Δ	Tikh	$2.1824{\cdot}10^{-2}$
phillips	I	TSVD	$7.8504{\cdot}10^{-2}$	phillips	Δ	TSVD	$2.6728{\cdot}10^{-2}$
phillips	I	CGNR	$5.5095{\cdot}10^{-2}$				
laplace	I	Alg. 2.1 (p)	$6.6013{\cdot}10^{-2}$	baart	Δ	Alg. 2.1 (p)	$5.3844{\cdot}10^{-2}$
laplace	I	Alg. 2.1 (t)	$5.9597{\cdot}10^{-2}$	baart	Δ	Alg. 2.1 (t)	$5.3844{\cdot}10^{-2}$
laplace	I	Tikh	$7.7328{\cdot}10^{-2}$	baart	Δ	Tikh	$4.3711{\cdot}10^{-2}$
laplace	I	TSVD	$8.5457{\cdot}10^{-2}$	baart	Δ	TSVD	$1.1496{\cdot}10^{-1}$
laplace	I	CGNR	$8.7399{\cdot}10^{-2}$	baart		CGNR	$1.6682{\cdot}10^{-1}$
shaw	I	Alg. 2.1 (p)	$6.9759{\cdot}10^{-2}$	shaw	Δ	Alg. 2.1 (p)	$1.7318{\cdot}10^{-1}$
shaw	I	Alg. 2.1 (t)	$5.5822{\cdot}10^{-2}$	shaw	Δ	Alg. 2.1 (t)	$1.7318{\cdot}10^{-1}$
shaw	I	Tikh	$7.2107{\cdot}10^{-2}$	shaw	Δ	Tikh	$6.5400{\cdot}10^{-1}$
shaw	I	TSVD	$8.9495{\cdot}10^{-2}$	shaw	Δ	TSVD	$6.7228{\cdot}10^{-1}$
shaw	I	CGNR	$8.8296{\cdot}10^{-2}$				
foxgood	I	Alg. 2.1 (p)	$2.2334{\cdot}10^{-2}$	foxgood	Δ	Alg. 2.1 (p)	$6.5893{\cdot}10^{-2}$
foxgood	I	Alg. 2.1 (t)	$3.7051{\cdot}10^{-2}$	foxgood	Δ	Alg. 2.1 (t)	$6.5893{\cdot}10^{-2}$
foxgood	I	Tikh	$4.3181{\cdot}10^{-2}$	foxgood	Δ	Tikh	$6.8360{\cdot}10^{+2}$
foxgood	I	TSVD	$5.9874{\cdot}10^{-2}$	foxgood	Δ	TSVD	$8.5244{\cdot}10^{+2}$
foxgood	I	CGNR	$6.0319{\cdot}10^{-2}$				
gravity	I	Alg. 2.1 (p)	$6.0659{\cdot}10^{-2}$	gravity	Δ	Alg. 2.1 (p)	$2.2772{\cdot}10^{-2}$
gravity	I	Alg. 2.1 (t)	$3.6171{\cdot}10^{-2}$	gravity	Δ	Alg. 2.1 (t)	$2.2772{\cdot}10^{-2}$
gravity	I	Tikh	$4.6718{\cdot}10^{-2}$	gravity	Δ	Tikh	$1.2743{\cdot}10^{-2}$
gravity	I	TSVD	$5.1280{\cdot}10^{-2}$	gravity	Δ	TSVD	$8.4308{\cdot}10^{-3}$
gravity	I	CGNR	$4.8276{\cdot}10^{-2}$				
deriv2	∇	Alg. 2.1 (p)	$4.5232{\cdot}10^{-2}$	deriv2	Δ	Alg. 2.1 (p)	$2.8450{\cdot}10^{-2}$
deriv2	∇	Alg. 2.1 (t)	$4.5232{\cdot}10^{-2}$	deriv2	Δ	Alg. 2.1 (t)	$2.8450{\cdot}10^{-2}$
deriv2	∇	Tikh	$4.1228{\cdot}10^{-2}$	deriv2	Δ	Tikh	$4.2472{\cdot}10^{-1}$
deriv2	∇	TSVD	$4.6417{\cdot}10^{-2}$	deriv2	Δ	TSVD	$1.0303{\cdot}10^{-1}$
deriv2	I	CGNR	$2.3945{\cdot}10^{-1}$				
heat	I	Alg. 2.1 (p)	$1.2275{\cdot}10^{-1}$	heat	Δ	Alg. 2.1 (p)	9.6856e-002
heat	I	Alg. 2.1 (t)	$1.4663{\cdot}10^{-1}$	heat	Δ	Alg. 2.1 (t)	9.6856e-002
heat	I	Tikh	$1.7140{\cdot}10^{-1}$	heat	Δ	Tikh	2.6483e+002
heat	I	TSVD	$3.7584{\cdot}10^{-1}$	heat	Δ	TSVD	1.1823e-001
heat	I	CGNR	$2.2695{\cdot}10^{-1}$				

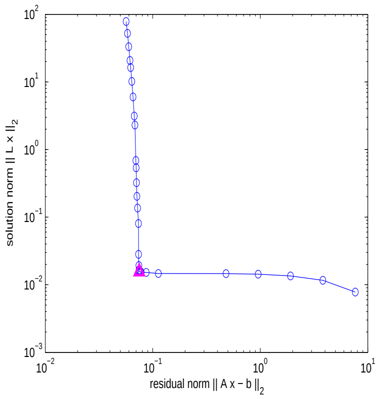
Table 1: Comparison between the proposed L-curve and the L-curve for Tikhonov, T(G)SVD and CGNR methods for $\delta = 0.005$.

Problem	**L**	δ	**Method**	**Relative error**			
phillips	I	5^{-3}	Alg. 2.1 (p)	$8.4647{\cdot}10^{-2}$	$1.0571{\cdot}10^{-1}$	$1.0743{\cdot}10^{-1}$	$2.0191{\cdot}10^{-1}$
phillips	I	5^{-3}	Alg. 2.1 (t)	$3.8481{\cdot}10^{-2}$	$2.2504{\cdot}10^{-2}$	$2.9531{\cdot}10^{-2}$	$2.9143{\cdot}10^{-2}$
phillips	I	5^{-3}	Tikh	$9.7262{\cdot}10^{-2}$	$1.1255{\cdot}10^{-1}$	$1.2127{\cdot}10^{-1}$	$1.2073{\cdot}10^{-1}$
phillips	I	5^{-3}	TSVD	$7.8504{\cdot}10^{-2}$	$1.5590{\cdot}10^{-1}$	$1.4029{\cdot}10^{-1}$	$1.8950{\cdot}10^{-1}$
phillips	I	5^{-3}	CGNR	$5.5095{\cdot}10^{-2}$	$1.6461{\cdot}10^{-1}$	$1.6877{\cdot}10^{-1}$	$1.6670{\cdot}10^{-1}$
phillips	I	10^{-2}	Alg. 2.1 (p)	$7.9808{\cdot}10^{-2}$	$3.2283{\cdot}10^{-2}$	$1.5043{\cdot}10^{-1}$	$2.1111{\cdot}10^{-1}$
phillips	I	10^{-2}	Alg. 2.1 (t)	$2.7552{\cdot}10^{-2}$	$2.3077{\cdot}10^{-2}$	$3.9067{\cdot}10^{-2}$	$4.0705{\cdot}10^{-2}$
phillips	I	10^{-2}	Tikh	$6.7984{\cdot}10^{-2}$	$8.6439{\cdot}10^{-2}$	$1.1126{\cdot}10^{-1}$	$1.1099{\cdot}10^{-1}$
phillips	I	10^{-2}	TSVD	$4.2033{\cdot}10^{-2}$	$4.7851{\cdot}10^{-2}$	$9.2222{\cdot}10^{-2}$	$8.4988{\cdot}10^{-2}$
phillips	I	10^{-2}	CGNR	$7.6210{\cdot}10^{-2}$	$7.0767{\cdot}10^{-2}$	$1.0689{\cdot}10^{-1}$	$9.8137{\cdot}10^{-2}$
baart	Δ	5^{-3}	Alg. 2.1 (p)	$5.3844{\cdot}10^{-2}$	$4.4948{\cdot}10^{-2}$	$2.6890{\cdot}10^{-2}$	$5.3586{\cdot}10^{-2}$
baart	Δ	5^{-3}	Alg. 2.1 (t)	$5.3844{\cdot}10^{-2}$	$4.4948{\cdot}10^{-2}$	$2.6890{\cdot}10^{-2}$	$5.3586{\cdot}10^{-2}$
baart	Δ	5^{-3}	Tikh	$4.3711{\cdot}10^{-2}$	$3.1495{\cdot}10^{-2}$	$3.0801{\cdot}10^{-2}$	$3.8117{\cdot}10^{-2}$
baart	Δ	5^{-3}	TSVD	$1.1496{\cdot}10^{-1}$	$6.9363{\cdot}10^{-2}$	$5.7942{\cdot}10^{3}$	$1.4572{\cdot}10^{7}$
baart	I	5^{-3}	CGNR	$1.6682{\cdot}10^{-1}$	$1.6573{\cdot}10^{-1}$	$1.6620{\cdot}10^{-1}$	$1.6524{\cdot}10^{-1}$
baart	Δ	10^{-2}	Alg. 2.1 (p)	$3.6082{\cdot}10^{-2}$	$2.6851{\cdot}10^{-2}$	$2.7208{\cdot}10^{-2}$	$5.9060{\cdot}10^{-2}$
baart	Δ	10^{-2}	Alg. 2.1 (t)	$3.6082{\cdot}10^{-2}$	$2.6851{\cdot}10^{-2}$	$2.7208{\cdot}10^{-2}$	$5.9060{\cdot}10^{-2}$
baart	Δ	10^{-2}	Tikh	$6.0663{\cdot}10^{-2}$	$3.7392{\cdot}10^{-2}$	$3.1423{\cdot}10^{-2}$	$5.5460{\cdot}10^{-2}$
baart	Δ	10^{-2}	TSVD	$1.3912{\cdot}10^{1}$	$1.4109{\cdot}10^{-1}$	$1.1588{\cdot}10^{4}$	$2.9145{\cdot}10^{7}$
baart	I	10^{-2}	CGNR	$1.6902{\cdot}10^{-1}$	$1.6575{\cdot}10^{-1}$	$1.6671{\cdot}10^{-1}$	$1.6550{\cdot}10^{-1}$

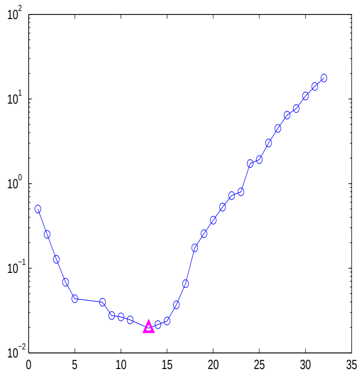
Table 2: Comparison between the proposed L-curve and the L-curve for Tikhonov, T(G)SVD and CGNR methods for $\delta = 0.005$ and $\delta = 0.01$ and four different realization numbers of the noise.

Problem	Method	Rel. Error
blur	alg. 2.1 (p)	0.1733
blur	alg. 2.1 (t)	0.1875
blur	CGNR (p)	0.5208
blur	GCNR (t)	0.1880
satellite	alg. 2.1 (p,t)	0.3650

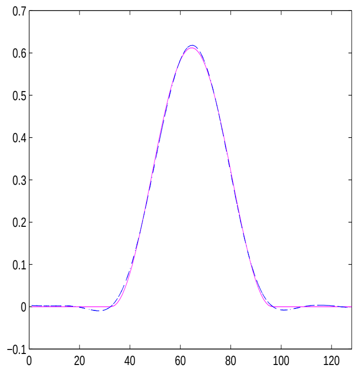
Table 3: Numerical results of the image deblurring test problems.



(a) Computed L-curve. The located vertex is indicated by the triangle.

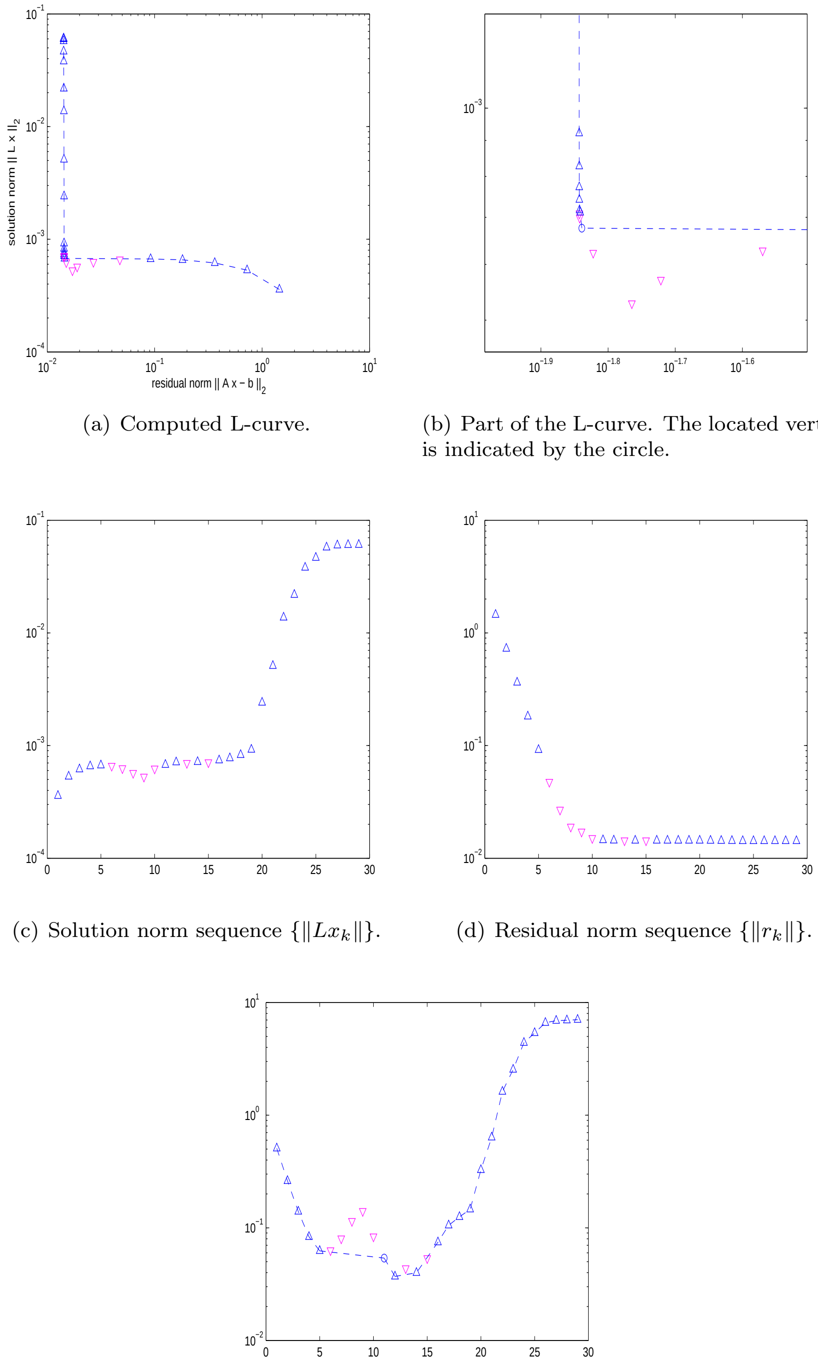


(b) Relative error graph. The relative error of the computed solution is indicated by the triangle.



(c) Computed regularized solution (dashed line) and exact solution (continuous line).

Figure 1: *Phillips* test problem ($L = \Delta$ and $\delta = 0.005$).



(a) Computed L-curve.

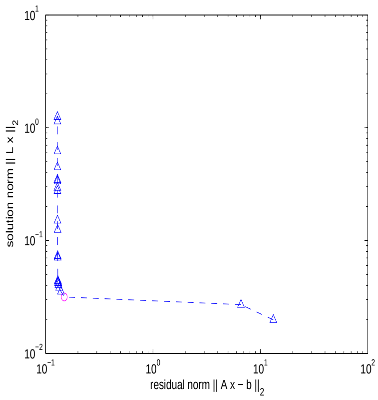
(b) Part of the L-curve. The located vertex is indicated by the circle.

(c) Solution norm sequence $\{\|Lx_k\|\}$.

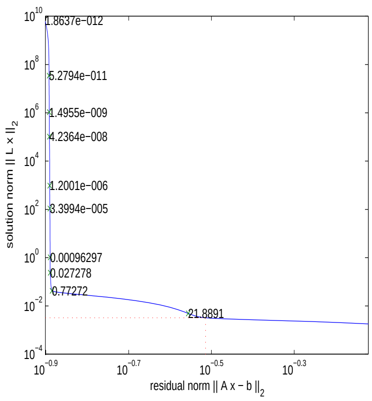
(d) Residual norm sequence $\{\|r_k\|\}$.

(e) Relative error. The relative error of the computed solution is indicated by the circle.

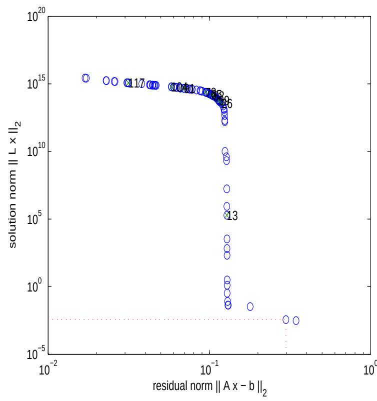
Figure 2: *Baart* test problem ($L = \Delta$ and $\delta = 0.005$). The upward-pointing triangles and the downward-pointing triangles correspond respectively to the indices $k \in \mathcal{K}$ and $k \notin \mathcal{K}$.



(a) Proposed L-curve. The located vertex is indicated by the circle.

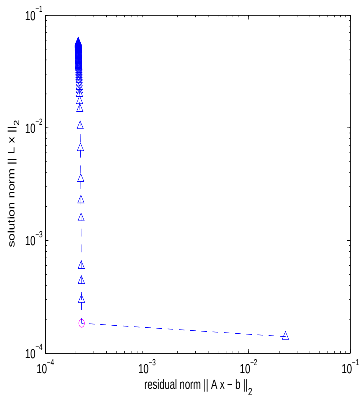


(b) L-curve for Tikhonov method.

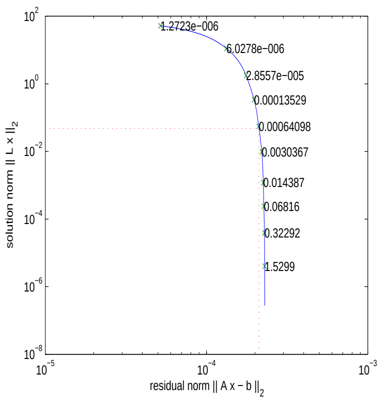


(c) L-curve for TGSVD method.

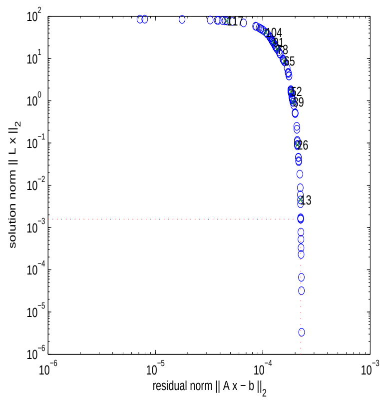
Figure 3: *Shaw* test problem ($L = \Delta$ and $\delta = 0.005$).



(a) Proposed L-curve. The located vertex is indicated by the circle.



(b) L-curve for Tikhonov method.



(c) L-curve for TGSVD method.

Figure 4: *Deriv2* test problem ($L = \Delta$ and $\delta = 0.005$).

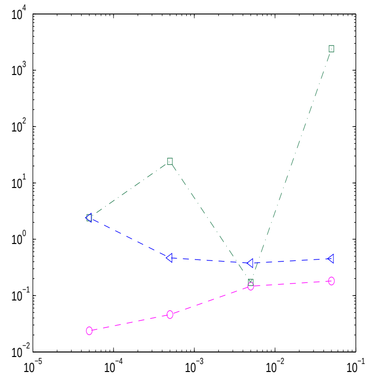


Figure 5: *Heat* test problem ($L = I$). Relative errors for varying δ. The circles, triangles and squares denote the solutions computed by the proposed method with the triangle algorithm, by the TGSVD method and by Tikhonov method, respectively.

[5] D. Calvetti, G. Golub, and L. Reichel. Estimation of the L-curve via Lanczos bidiagonalization. *BIT*, 39:603–619, 1999.

[6] D. Calvetti, P. C. Hansen, and L. Reichel. L-curve curvature bounds via Lanczos bidiagonalization. *ETNA*, 14:20–35, 2002.

[7] D. Calvetti and L. Reichel. Tikhonov regularization with a solution constraint. *SIAM J. Sci. Comput.*, 26(1):224–239, 2004.

[8] D. Calvetti, L. Reichel, and A. Shuibi. L-curve and curvature bounds for Tikhonov regularization. *ETNA*, 14:20–35, 2002.

[9] P. J. Mc Carthy. Direct analytic model of the L-curve for Tikhonov regularization parameter selection. *Inv. Prob.*, 19:643–663, 2003.

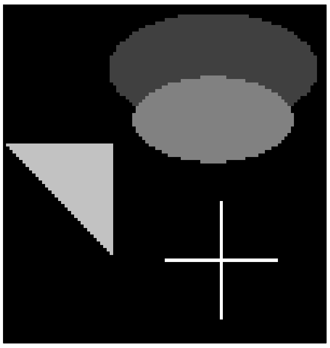
[10] J. L. Castellanos, S. Gómez, and V. Guerra. The triangle method for finding the corner of the L-curve. *Appl. Numer. Math.*, 43(4):359–373, 2002.

[11] L. Eldén. Algorithms for the regularization of ill-conditioned least squares problems. *BIT*, 17:134–145, 1977.

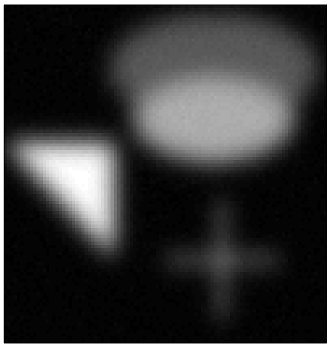
[12] H. W. Engl. Regularization methods for the stable solution of inverse problems. *Surv. Math. Ind.*, 3:71–143, 1993.

[13] H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 1996.

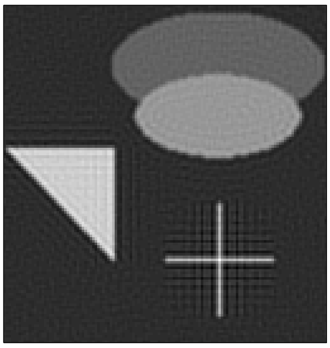
[14] G. H. Golub and C. F. Van Loan. *Matrix Computations*. the Johns Hopkins University Press, Baltimore, third edition edition, 1996.



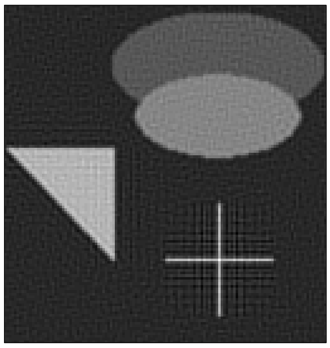
(a) Exact image.



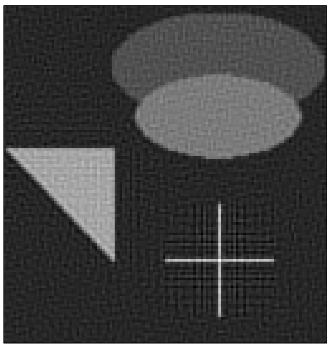
(b) Noisy blurred image.



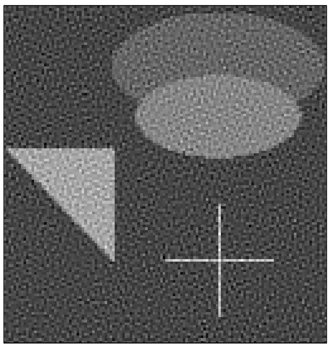
(c) Reconstruction obtained by algorithm 2.1 with the triangle method.



(d) Reconstruction obtained by algorithm 2.1 with the pruning method.

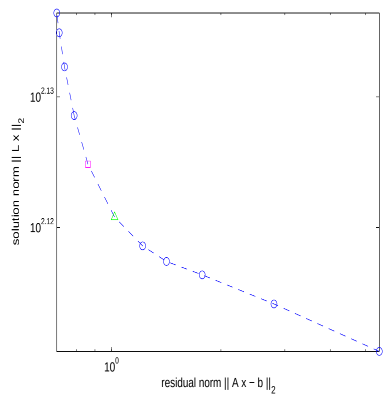


(e) Reconstruction obtained by the CGNR method with the triangle method.

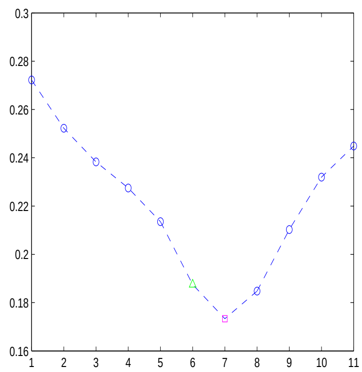


(f) Reconstruction obtained by the CGNR method with the pruning method.

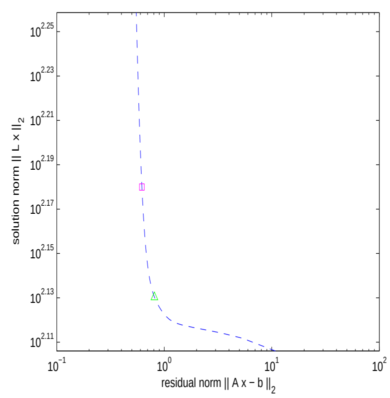
Figure 6: *Blur* test problem ($n = 128 \times 128$, $L = I$, $\delta = 0.01$).



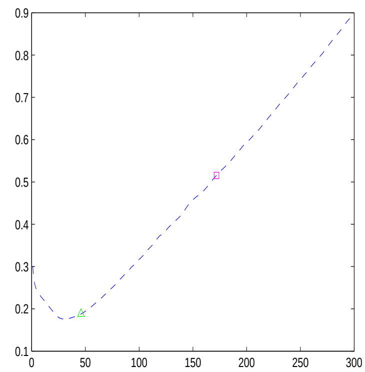
(a) L-curve.



(b) Relative error.

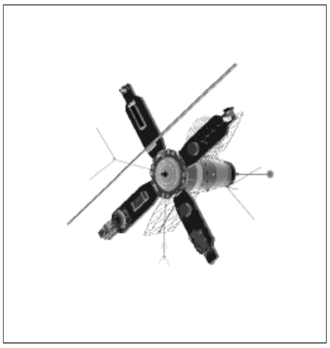


(c) L-curve.

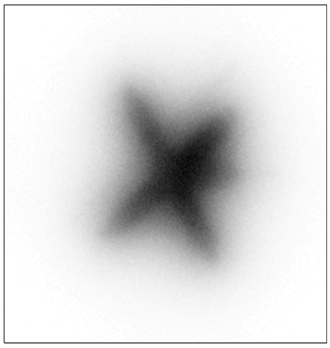


(d) Relative error.

Figure 7: *Blur* test problem ($n = 128 \times 128$, $L = I$, $\delta = 0.01$). Top line: proposed L-curve (left) and corresponding relative error graph (right); bottom line: CGNR L-curve (left) and corresponding relative error graph (right). The triangles and the squares indicate the iterates detected by the triangle and pruning algorithms, respectively.

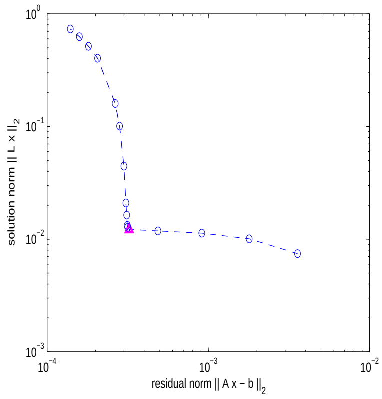


(a) Exact image.



(b) Noisy blurred image.

(c) Reconstruction obtained by algorithm 2.1.



(d) Proposed L-curve. The triangle and pruning algorithms detect the same iterate, which is indicated by the triangle.

Figure 8: *Satellite* test problem ($n = 256 \times 256$, $L = I$).

[15] C. W. Groetsch. *The theory of Tikhonov regularization for Fredholm equations of the first kind*. Pitman, Boston, 1983.

[16] M. Hanke. *Conjugate Gradient Type Methods for Ill-Posed Problems*. Pitman Research Notes in Mathematics, Longman House, Harlow, Essex, 1995.

[17] M. Hanke. Limitations of the L-curve method in ill-posed problems. *BIT*, 36(2):287–301, 1996.

[18] M. Hanke and P. C. Hansen. Regularization methods for large-scale problems. *Surv. Math. Ind.*, 3:253–315, 1993.

[19] M. Hanke and G. Nagy. Restoration of atmospherically blurred images by symmetric indefinite conjugate gradient techniques. *Inv. Prob.*, 12:157–173, 1996.

[20] P. C. Hansen. Regularization, GSVD and truncated GSVD. *BIT*, 29:491–504, 1989.

[21] P. C. Hansen. The discrete Picard condition for discrete ill-posed problems. *BIT*, 30:658–672, 1990.

[22] P. C. Hansen. Analysis of discrete ill-posed problems by means of the L-curve. *SIAM Rev.*, 34(4):561–580, 1992.

[23] P. C. Hansen. Regularization tools: A Matlab package for analysis and solution of discrete ill-posed problems. *Numer. Alg.*, 6:1–35, 1994.

[24] P. C. Hansen. *Rank-Deficient and Discrete Ill-Posed Problems*. SIAM, 1998.

[25] P. C. Hansen, T. K. Jensen, and G. Rodriguez. An adaptive pruning algorithm for the discrete L-curve criterion. *J. Comput. Appl. Math.*, 198(2):483–492, 2007.

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

[27] R. J. Hanson. A numerical method for solving Fredholm integral equations of the first kind using singular values. *SIAM J. Numer. Anal.*, 8:616–622, 1971.

[28] G. Landi. A fast truncated Lagrange method for large-scale image restoration problems. *Appl. Math. Comput.*, 186(2):1075–1082, 2007.

[29] G. Landi. A truncated Lagrange method for Total Variation-based image restoration. *J. Math. Imag. Vis.*, 28(2):113–123, 2007.

[30] G. Landi. The Lagrange method for the regularization of discrete ill-posed problems. *Comput. Optim. Appl.*, 39(3):347–368, 2008.

[31] G. Landi and E. Loli Piccolomini. An iterative Lagrange method for the regularization of discrete ill-posed inverse problems. *Computers and Mathematics with Applications*, 60(6):1723–1738, 2010.

[32] D. G. Luenberger. *Linear and nonlinear programming*. Kluwer Academic Publishers, (2nd Ed.), 2003.

[33] J. G. Nagy, K. Palmer, and L. Perrone. Iterative methods for image restoration: a matlab object oriented approach. *Numer. Alg.*, 36:73–93, 2003.

[34] J. G. Nagy, R. J. Plemmons, and T. C. Torgersen. Iterative image restoration using approximate inverse preconditioners. *IEEE Trans. Image Proc.*, 5(7):1151–1162, 1996.

[35] J. Nocedal and S. J. Wright. *Numerical optimization*. Springer-Verlag New York, Inc., 1999.

[36] C. C. Paige and M. A. Saunders. Solution of sparse indefinite systems of linear equations. *SIAM J. Numer. Anal.*, 12:617–629, 1975.

[37] C. C. Paige and M. A. Saunders. LSQR: Sparse linear equations and least squares problems. *ACM TOMS*, 4:195–209, 1982.

[38] L. Reichel and H. Sadok. A new L-curve for ill-posed problems. *J. Comput. Appl. Math.*, 219(2):493–508, 2008.

[39] M. Rezghi and S. M. Hosseini. A new variant of L-curve for Tikhonov regularization. *J. Comput. Appl. Math.*, 231(2):914–924, 2009.

[40] M. C. Roggemann and B. Welsh. *Imaging through torbolence*. CRC Press, Boca Raton, Florida, 1996.

[41] M. Rojas and D. C. Sorensen. A trust-region approach to the regularization of large-scale discrete forms of ill-posed problems. *SIAM J. Sci. Comput.*, 26(3):1843–1861, 2002.

[42] A. N. Tikhonov. Solution of incorrectly formulated problems and the regularization method. *Soviet Math. Dokl.*, 4:1035–1038, 1963.

[43] A. N. Tikhonov and V. Y. Arsenin. *Solutions of Ill-Posed Problems*. John Wiley, New York, 1977.

[44] L. Wu. A parameter choice method for Tikhonov regularization. *ETNA*, 16:107–128, 2003.