Irregular Solutions of an Ill-Posed Problem

Peter Linz
University of California at Davis

Richard L.C. Wang
Clearsight Systems Inc.

Abstract: *Tikhonov regularization is a popular and effective method for the approximate solution of ill-posed problems, including Fredholm equations of the first kind. The Tikhonov method works well when the solution of the equation is well-behaved, but fails for solutions with irregularities, such as jump discontinuities. In this paper we develop a method that overcomes the limitations of the standard Tikhonov regularization. We present a criterion by which approximate solutions can be evaluated and use it in a search method that is effective in locating points of irregular behavior. Once the points of irregularity have been found, the solution can be recovered with good accuracy.*

Keywords: *integral equations, ill-posed problems, regularization*

1. Introduction

The numerical solution of the Fredholm equation of the first kind

$$\int_0^1 k(t,s)x(s)ds = g(t), \quad 0 \le t \le 1 \tag{1}$$

can be attempted by replacing the integral by a quadrature with weights $w_1, w_2, \dots, w_n$ on the points $s_1, s_2, \dots s_n$ to give

$$\sum_{j=1}^{n} w_j k(t, s_j) x(s_j) \cong g(t). \tag{2}$$

To get a solvable system, we satisfy (2) at collocation points $t_1, t_2, \dots t_n$; this results in the $n \times n$ linear system

$$\mathbf{Ax} = \mathbf{g} \tag{3}$$

where $[\mathbf{A}]_{ij} = w_j k(t_i, s_j)$ and $[\mathbf{g}]_i = g(t_i)$. [1]

When the kernel in (1) is well behaved, the equation is ill-posed and discretization gives a very ill-conditioned system. Solving (3) directly usually yields highly oscillatory results that are of no use. The traditional remedy is to regularize (3) by adding some sort of stabilizing factor.

---

[1] $[\mathbf{A}]_{ij}$ denotes the element in row *i* and column *j* of the matrix **A**.

One method of stabilizing (3) is the standard Tikhonov regularization (TREG). Instead of trying to solve (3) as it stands, Tikhonov regularization minimizes

$$r(\mathbf{x}) = \| \mathbf{Ax} - \mathbf{g} \|^2 + \alpha \| \mathbf{Bx} \|^2 . \tag{4}$$

Here α is the *regularization parameter* and **B** is the *smoothness operator* designed to filter out undesirable components of the solution. When the 2-norm is used in (4), the minimizing solution $\mathbf{x}_\alpha$ can be found from the matrix equation

$$(\mathbf{A}^T \mathbf{A} + \alpha \mathbf{B}^T \mathbf{B}) \mathbf{x}_\alpha = \mathbf{A}^T \mathbf{g} . \tag{5}$$

With properly chosen regularization parameters and operators, TREG can give good results.

Regularization for ill-posed problems has been thoroughly studied in the last forty years and is now well understood. For a review of the main results, consult the book by Engl, Hanke and Neubauer [1].

2. Assessing Approximate Solutions

When ill-posed problems are solved in the presence of significant computational or experimental errors, the solutions obtained can be quite sensitive to the selection of the numerical method, the regularization parameters α, and the smoothing matrix **B**. Non-identical results can occur in all numerical computations, but for well-posed problems the discrepancies are usually insignificant. For ill-posed problems though, the situation is quite different and we may get several approximate solutions that deviate from each other not only quantitatively, but show rather significant qualitative differences.

Suppose we are given two approximate solutions $\mathbf{x}_1$ and $\mathbf{x}_2$, both of which look reasonable. They both are plausible in the sense that they satisfy our notion of being physically realistic (as measured by $\| \mathbf{Bx} \|$), and are acceptable, in that they satisfy the original equation within a tolerance depending on the computational and experimental errors. If $\mathbf{x}_1$ and $\mathbf{x}_2$ are not nearly the same, how do we choose between them?

In the absence of any problem-specific information, we can only use criteria that can be derived from the computed solution. Two obvious measures are:

(a) A smoothness measure $\rho(\mathbf{x}_\alpha)$ that measures the degree to which the solution satisfies our idea of physical plausibility. An obvious choice is

$$\rho(\mathbf{x}_\alpha) = \| \mathbf{Bx}_\alpha \| .$$

Normally, we want $\rho(\mathbf{x}_\alpha)$ to be as small as possible.

(b) A discrepancy measure $\varepsilon(\mathbf{x}_\alpha)$ that indicates how closely the approximate solution satisfies the original equation. A simple discrepancy measure is

$$\varepsilon(\mathbf{x}_\alpha) = \| \mathbf{A}\mathbf{x}_\alpha - \mathbf{g} \|,$$

although other definitions are possible. Again, we would like $\varepsilon(\mathbf{x}_\alpha)$ to be as small as possible.

These two measures give us a simple way of comparing alternative solutions.

Definition 1**.** Suppose we have two approximate solutions $\mathbf{x}_1$ and $\mathbf{x}_2$ to equation (1). We say that $\mathbf{x}_1$ is *preferable* to $\mathbf{x}_2$ if

$$\rho(\mathbf{x}_1) < \rho(\mathbf{x}_2) \tag{6}$$

and

$$\varepsilon(\mathbf{x}_1) < \varepsilon(\mathbf{x}_2)\,. \tag{7}$$

If (6) is true, but (7) is not (or vice versa), we say that the two solutions are not comparable.

3. Regularizing Problems With Irregular Solutions

Tikhonov regularization filters out oscillatory components and so delivers smooth solutions. Eliminating rapidly oscillating components is reasonable since typical physical phenomena rarely exhibit such behavior, for example, when $x(s)$ represents a density function in some remote sensing situations. Unfortunately, Tikhonov regularization also inhibits other irregularities, such as jump discontinuities or rapid slope changes. There are physical situations where such discontinuities are possible, and their recovery may be an important issue. In these situations the usual TREG method does not work well. In the past there have been a few papers (e.g. [2]. [3], [4]) that address this issue, but no comprehensive theory has emerged and the emphasis has been on dealing with situations where the nature and position of the irregularity is known,

If we know the point where the abrupt behavior occurs, we can compensate for it. Suppose that the irregularity is at $t = a$. We write (1) as

$$\int_0^a k(t,s)x(s)ds + \int_a^1 k(t,s)x(s)ds = g(t)\,.$$

To avoid integrating through a point of discontinuity, we use separate quadratures for each part. To allow for rapid changes at $t = a$, we also use Tikhonov regularization separately in each subinterval. This will give smooth solutions in each subinterval, but allows for rapid changes at the point of discontinuity. We will call the resulting method regularization with discontinuities (REGDC). In this report we summarize our computational experiences with REGDC.

For the results in this paper we used a composite Simpson rule with specifiable panel size $p = 2 \times \text{mesh size}$. We used second-order Tikhonov regularization in which $\mathbf{B}$ is an $(n-2)\times n$ matrix of second differences, that is $[\mathbf{B}]_{i,i-1} = [\mathbf{B}]_{i,i+1} = 1$, $[\mathbf{B}]_{i,i} = -2$, $[\mathbf{B}]_{ij} = 0$, otherwise. For the test examples, some solution $x(s)$ was chosen, and the right side $g(t)$ was computed numerically to a high accuracy.

<u>Example 1</u>**.** This example involves a solution with a step function discontinuity

$$\begin{aligned} x(s) &= 0, \quad 0 \le s < 0.5 \\ &= 1, \quad 0.5 \le s \le 1. \end{aligned}$$

The kernel

$$k(t,s) = \frac{1}{1 + 10(t-s)^2}$$

is slightly peaked and the problem is ill-posed, but not extremely so.

TREG was used with panel size of 0.05 and $\alpha = 10^{-6}$. The results are shown in Figure 1. As is apparent, TREG gives some hint of irregular behavior around $t = 0.5$, but does not locate the point of the jump discontinuity nor does it uncover the qualitative behavior of the true solution.

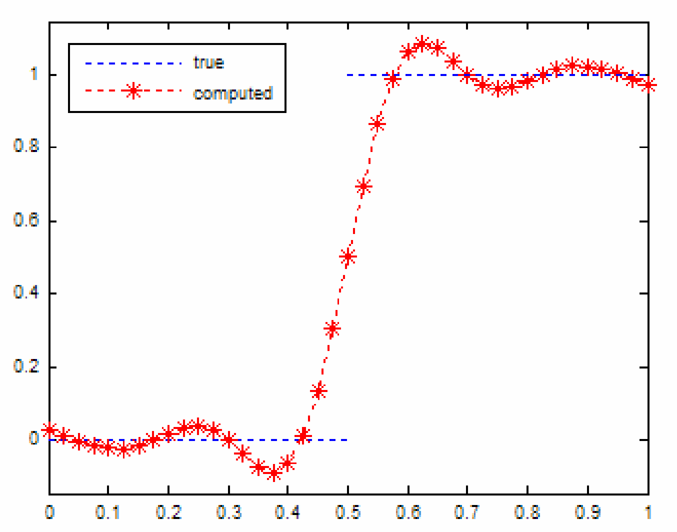

Figure 1. The TREG solution for Example 1 with $\alpha = 10^{-6}$

4. Locating Points of Irregular Behavior

In practice we do not usually know the points at which irregular behavior occurs. TREG may suggest a region, but to locate the trouble points, we need to switch to REGDC. We use a very simple approach. We search a promising region, using various trial values for *a*, comparing the results using the criteria established in Definition 1.

Example 2. For this case we used the kernel in Example 1, with

$$\begin{aligned} x(s) &= \sin(\pi s), \quad 0 \le s < 0.5 \\ &= e^{-s}, \quad\quad 0.5 \le s \le 1 \end{aligned}$$

TREG results suggest irregular behavior in $(0.4,0.6)$. A run was done with panel size of 0.05, $\alpha = 10^{-4}$, and the region $(0.3,0.65)$ was searched with incremental steps of 0.05. For each break point assumption the values of ε and ρ were computed. These are plotted in Figure 2. The assumption $a = 0.5$ gives a solution that is preferable to the other assumptions, leading us to accept it. The actual solution with $a = 0.5$ is in Figure 3. It recovers the point of discontinuity, and the approximation solution differs only slightly from the true answer.

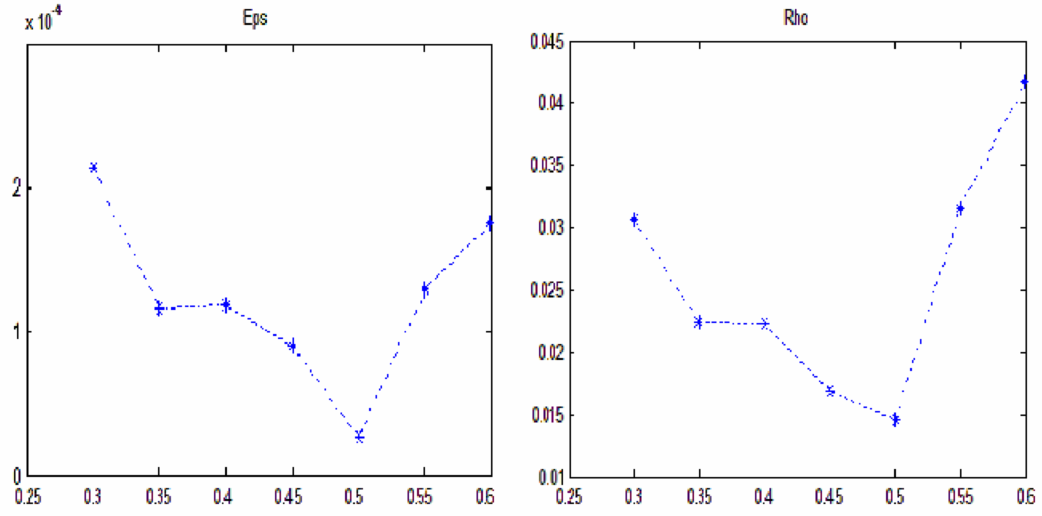

Figure 2. Computed ε and ρ for assumed break points

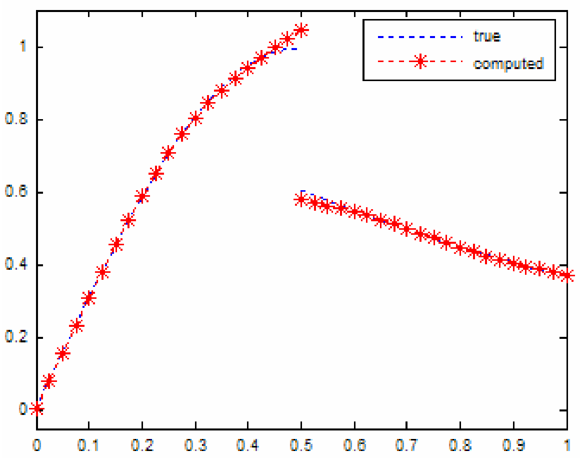

Figure 3. The REGDC solution for Example 2 with $\alpha = 10^{-4}$ at break point $a = 0.5$

For $a = 0.5$, REGDC gives the values $\varepsilon = 2.5\times10^{-5}$ and $\rho = 1.5\times10^{-2}$. The results for TREG are $\varepsilon = 2.1\times10^{-4}$ and $\rho = 5.1\times10^{-1}$. We conclude that the REGDC solution is preferable to the TREG solution.

Note that in Figure 2 the conjecture $a = 0.5$ simultaneously minimizes ε and ρ and so leads to a solution preferable to those obtained with other assumptions on the location of *a*. This kind of situation appears frequently, often exhibiting a distinct v-shaped graphs, with the minima at the same value of *a*. We will say a solution has the *v-characteristic* if the $\varepsilon - \rho$ curves have a v-shaped form with both vertices at the same value of the breakpoint *a*. The v-characteristic is particularly noticeable for mildly ill-posed problems.

<u>Example 3</u>**.** The kernel

$$k(t,s) = 1 - |t - s|$$

is continuous but not differentiable, so this case can be said to be mildly ill-posed. For

$$\begin{aligned} x(s) &= \sin(\pi s), \quad 0 \le s < 0.5 \\ &= e^{-s}, \qquad 0.5 \le s \le 1 \end{aligned}$$

a panel size of 0.05 and $\alpha = 10^{-4}$, we computed the approximate solution in the region $0.46 \le a \le 0.53$ in steps of 0.01. The resulting $\varepsilon - \rho$ curves, shown in Figure 4, exhibit a strong v-characteristic with vertices at $a = 0.5$. The computed solution for this is identical with the true solution to plotting accuracy.

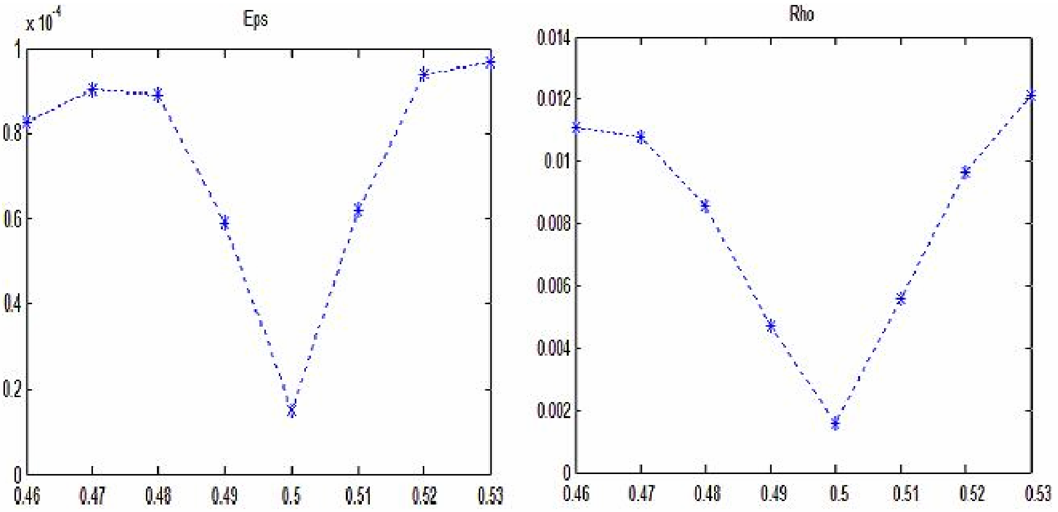

Figure 4. Computed ε and ρ for assumed break points

5. Summary

The examples given here are typical of a wide range of similar results we have obtained. For problems that are not extremely ill-posed the v-curve method normally locates the singularities accurately and gives approximations that agree with the exact solutions to plotting accuracy. For very ill-posed problems the method occasionally fails in the sense that the minima of the ε and ρ curves occur at slightly different points. In that case there is some ambiguity in the location of the discontinuity. But in all tested cases, the results obtained by REGDC were better than the results obtained by TREG.

We have also used the method for locating derivative continuities and it appears to be equally effective there. On the whole, the method described in this paper is a reliable way for locating points of irregular behavior in the solution of ill-posed problems.

References

[1] Engl, Heinz W., Martin Hanke, and Andreas Neubauer: *Regularization of Inverse Problems*; Kluwer Academic Publishers, 1996

[2] Linz, P.: A New Numerical Method for Ill-Posed Problems: *Inverse Problems* 10, (1994), L1-L6

[3] Simcik, Luke: Resolving Non-smooth Solutions to Discretized Ill-Posed Problems, Ph.D. Dissertation, University of California, Davis, 1995

[4] Wahba, Grace: Three Topics in Ill-Posed Problems, in H.W. Engl and C.W. Groetsch; *Inverse and Ill-Posed Problems*, Academic Press, 1987