Analytic and numerical comparison of linear direct and iterative regularization methods

Torsten Hein *

November 2, 2004

Abstract

This paper deals with analytic and numerical study of linear regularization methods. Thereby Tikhonov regularization and Landweber iteration are compared. In particular the regularization with semi-norms for both methods are considered. A heuristic approach for improving the results for Landweber iteration are introduced. The analytic considerations are illustrated by a detailed numerical study.

Keywords: linear ill-posed problem, Moore-Penrose inverse, Tikhonov regularization, Landweber iteration, semi-norms, convergence rates, integral equation

MSC 2000 Subject Classification: 65D25, 65F20, 65J20, 65R20

1 Introduction

Let $A : \mathcal{X} \longrightarrow \mathcal{Y}$ denotes a linear operator between the infinite-dimensional Hilbert spaces $\mathcal{X}$ and $\mathcal{Y}$. We consider the linear equation

$$A\,x = y^\delta, \qquad x \in \mathcal{X}, \tag{1}$$

with given noisy data y^δ. We do not know the exact data y. Only an a-priori information

$$\|y^\delta - y\|_{\mathcal{Y}} \le \delta$$

of the noise level is known. A solution of (1) we denote with x^δ whereas x is a solution of (1) for exact given data y (provided it exists).

Determining x by given observation y is called the inverse problem. Unfortunately many application in practice are ill-posed. In particular the operator A has a non-closed range. Even if A injective and the inverse operator $A^{-1} : \mathcal{R}(A) \subset \mathcal{Y} \longrightarrow \mathcal{X}$ exists it remains unbounded. Hence a solution x^δ of (1) does not depend continuously on the given data y^δ.

*Technische Universität Chemnitz, Fakultät für Mathematik, 09107 Chemnitz, Deutschland. Email: `Torsten.Hein@mathematik.tu-chemnitz.de`

Consequently for $\delta \to 0$ and $y^\delta \to y$ we cannot guarantee the convergence $x^\delta \to x = A^{-1}y$. A visible ill-posedness effect are high oscillating solutions x^δ whereas the exact solution x is a smooth function.

To overcome these problems regularization strategies were developed. Thereby we divide into direct and iterative regularization methods. Most popular is Tikhonov regularization which is widely studied in literature. We refer e.g. to [1, chapter 5] or [4, section 4.1.2]. Thereby equation (1) is replaced by the equation

$$(A^*A + \alpha\, I)\, x = A^* y^\delta, \qquad x \in \mathcal{X} \tag{2}$$

or more generally

$$(A^*A + \alpha\, L^*L)\, x = A^* y^\delta, \qquad x \in \mathcal{X} \tag{3}$$

with a given penalty operator $L : \mathcal{X} \longrightarrow \mathcal{Z}$. The space $\mathcal{Z}$ is again a Hilbert space. We can imagine the identity I respectively the operator L as penalty operators. In particular choosing L as a differential operator is a common practice. That leads to the consequence that a solution of (3) is smoother than a solution of (2). The background is the following: In dependence of the noise level δ and the regularization parameter α we denote a solution of (2) with x_α^δ whereas a solution of (3) is denoted with $\tilde{x}_\alpha^\delta$. Choosing (2) we obtain solutions x_α^δ which are usually less oscillating than x^δ but the oscillations cannot be damped completely. The reason is that the identical operator I penalized high values of x and not the oscillation properties. By the application of (3) with L as a differential operator oscillations in the solution are penalized. Consequently this leads to a smoother and less oscillating solution $\tilde{x}_\alpha^\delta$ of (3). The right choice of L leads to better reconstruction chances of the sought solution x.

The most well-known representative of iterative regularization methods is Landweber iteration. For $\gamma \in (0, 2\,\|A\|^{-1})$ the k-th component of the iteration sequence is given by

$$x_0^\delta := 0,\ x_k^\delta := x_{k-1}^\delta - \gamma\, A^* \left(A\, x_{k-1}^\delta - y^\delta\right). \tag{4}$$

The regularization consists in the right choice of the index k where the iteration process is abandoned. The method (4) can be considered as alternative for the Tikhonov regularization method (2). On the other hand it seems interesting to generalize Landweber iteration (4) in that way that we can – simular to (3) – insert a penalty operator L which leads to better regularization properties. Whereas Landweber method of the form (4) is widely discussed in literature (see [1, section 6.1] or [4, section 4.1.3]) the generalization of the idea of introducing a penalty operator L to Landweber iteration is less common practice. In [1, Chapter 8] a general approach of introducing penalty operators L for a whole family of (direct and iterative) regularization methods is studied. The application of this idea to Landweber iteration is shortly illustrated in [1, Example 8.12]

The aim of our paper is an analytical and numerical comparison of Tikhonov regularization and Landweber iteration. We will demonstrate on a simple, well-studied example (see e.g. [2]) that by using a penalty operator L we obtain similar good regularization properties for Landweber iteration than in the case of Tikhonov regularization.

The paper is organized as follows: In section 2 we introduce the idea of [1, Chapter 8] again and show how to applicate it to Landweber iteration. We will give convergence and convergence rate results for both Tikhonov regularization and Landweber iteration. The necessary source conditions for obtaining convergence rates are derived for a specific sample operator A and penalty operator L. Section 3 is dedicated a detailed numerical comparison of Tikhonov regularization and Landweber iteration considering a simple integral operator as example. Afterwards another approach for Landweber iteration is suggested which removes or weakens the source conditions on the boundary to receive better regularization results. The idea is again illustrated by numerical results. In the Appendix we describe shortly the used numerical algorithms for applying the regularization strategies.

2 Regularization with Semi-norms

For practical reasons we reformulate equation (1) as least-square problem

$$\|A\,x - y^\delta\|_{\mathcal{Y}}^2 \to \min \quad \text{subject to } x \in \mathcal{X}. \tag{5}$$

The matter is based on two facts: On the one hand it is often more simple to guarantee the existence of a solution of (5) than a solution of (1). On the other hand many numerical algorithms for finding a solution of (1) in fact solve the problem (5).

We remind the concept of the Moore-Penrose inverse of the linear operator A, which is fundamental in the linear regularization theory (see e.g. [1, section 2.1]).

Definition 2.1 *The Moore-Penrose inverse $A^\dagger : \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp \subseteq \mathcal{Y} \longrightarrow \mathcal{X}$ is defined as*

$$A^\dagger y := x^*, \tag{6}$$

where $x^* \in \mathcal{X}$ is the unique element which satisfies

$$\|x^*\|_{\mathcal{X}} = \min\left\{\|x\|_{\mathcal{X}} \,:\, x \text{ solves (5) with } y^\delta \text{ replaced by } y\right\}. \tag{7}$$

Simply spoken, x^* is that solution of the least-square problem (5) which has the smallest norm (in X).

Since $\mathcal{R}(A) \neq \overline{\mathcal{R}(A)}$ the operator $A^\dagger$ is not continuous (see e.g. [1, Proposition 2.4]). In particular, a solution

$$x^\delta = A^\dagger y^\delta$$

of (1) does not depend continuously on the given data. Therefore we approximate $A^\dagger$ by a family $\{R_\alpha\}_{\alpha>0}$ of continuous operators which satisfy

$$R_\alpha y \to A^\dagger y \quad \text{for} \quad \alpha \to 0$$

for all $y \in \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$. This is what we call regularization. For further readings on general regularization theory we refer e.g. to [1], [4] and [7].

As already mentioned in the introduction we consider two different regularization strategies. Probably most popular is Tikhonov regularization where we approximate $A^\dagger$ by

$$R_\alpha := (A^*A + \alpha\, I)^{-1}A^*. \tag{8}$$

Alternatively there can be found another definition of this regularization method. We define x_α^δ as solution of the minimization problem

$$\|A\,x - y^\delta\|_{\mathcal{Y}}^2 + \alpha\,\|x\|_{\mathcal{X}}^2 \to \min \quad \text{subject to } x \in \mathcal{X}. \tag{9}$$

Using the necessary optimality condition for (9) we conclude that x_α^δ solves the normal equation (2) which is exactly (8). Since the functional in (9) is strictly convex for $\alpha > 0$ a solution x_α^δ of (9) respectively (2) is uniquely determined.

An example of iterative regularization methods is Landweber iteration, which is given in (4). Principally this iteration (4) describes a gradient method for the least-square problem (5). The difference we can find in the stop criterion. For gradient methods an arbitrary small tolerance $\varepsilon > 0$ is chosen and the iteration (4) is stopped when

$$\|A\, x_k^\delta - y^\delta\|_{\mathcal{Y}} \le \varepsilon < \|A\, x_{k-1}^\delta - y^\delta\|_{\mathcal{Y}}.$$

Landweber iteration consists of the idea that we choose the stop index k of the iteration not too large. In particular the tolerance ε is chosen large enough.

Both regularization strategies (9) and (4) are based on approximation of the Moore-Penrose inverse $A^\dagger$. On the other hand – choosing the smallest norm of an element as criterion for the best least-square solution of (5) often seems to be restrictive or arbitrarily. In particular Tikhonov regularization is defined in literature often in the form (3) which differs from (9) respectively (2) (see e.g. [7, Chapter 2]). Let $L : \mathcal{D}(L) \subseteq \mathcal{X} \longrightarrow \mathcal{Z}$ be a linear operator which maps into the Hilbert space $\mathcal{Z}$. Then we define a Tikhonov regularized solution $\tilde{x}_\alpha^\delta$ of (1) as solution of the the minimization problem

$$\|A\, x - y^\delta\|_{\mathcal{Y}}^2 + \alpha \|L\, x\|_{\mathcal{Z}}^2 \to \min \quad \text{subject to } x \in \mathcal{D}(L), \tag{10}$$

or equivalently as solution of the equation (3). For example, choosing L as a differential operator is a common practice. Thereby oscillations in the solution x are penalized so that a solution $\tilde{x}_\alpha^\delta$ of (10) is generally smoother and less oscillating than a solution of (9). But also other criterions for the choice of the the operator L are possible.

It can be shown that under certain conditions on the operator L a solution $\tilde{x}_\alpha^\delta$ of (10) is unique and for $\alpha \to 0$ we have convergence $\tilde{x}_\alpha^\delta \to \tilde{x}^\delta$ where $\tilde{x}^\delta$ satisfies

$$\|L\, \tilde{x}^\delta\|_{\mathcal{Z}} = \min \left\{ \|L\, x\|_{\mathcal{Z}} \,:\, x \text{ solves } (5) \right\}. \tag{11}$$

To verify this more accurately we follow the concept of [1, Chapter 8] and point out the following requirements to the operator L.

(A1) $\mathcal{D}(L)$ is dense in $\mathcal{X}$,

(A2) L is closed and surjective,

(A3) $\dim \mathcal{N}(L) < \infty$ and $\mathcal{N}(A) \cap \mathcal{N}(L) = \{0\}$.

Under the conditions stated above an appropriate proposition can be formulated (see e.g. [1, Theorem 8.2]).

Proposition 2.1 *Assume $y^\delta \in A\left(\mathcal{D}(L)\right) \oplus \left(A\left(\mathcal{D}(L)\right)\right)^\perp$ and that the conditions (A1)-(A3) hold. Then there exists an unique element $\tilde{x}^\delta \in \mathcal{D}(L)$ which satisfies (11).*

In [7] more general conditions to the operators A and L are formulated to obtain a simular result. The idea of the proof is the following: Under conditions (A1)-(A3) we can introduce a scalar product

$$\langle x_1, x_2 \rangle_{\mathcal{D}(L)} := \langle A\, x_1, A\, x_2 \rangle_{\mathcal{Y}} + \langle L\, x_1, L\, x_2 \rangle_{\mathcal{Z}}, \qquad x_1, x_2 \in \mathcal{D}(L),$$

and a corresponding norm $\|\cdot\|_{\mathcal{D}(L)}$ on the subspace $\mathcal{D}(L)$ of $\mathcal{X}$. Then $\mathcal{D}(L)$ is a Hilbert space itself (see [1, Lemma 8.1]) and we consider instead of the operator A the operator $A_L : \mathcal{D}(L) \longrightarrow \mathcal{Y}$ which is the restriction of A to the Hilbert space $\mathcal{D}(L)$.

Proposition 2.1 motivates the introduction of an operator $A_L^\dagger : \mathcal{D}(A_L^\dagger) \subseteq \mathcal{Y} \longrightarrow \mathcal{X}$ by

$$A_L^\dagger y := \tilde{x}, \qquad y \in \mathcal{D}(A_L^\dagger),$$

with $\mathcal{D}(A_L^\dagger) := A\,(\mathcal{D}(L)) \oplus (A\,(\mathcal{D}(L)))^\perp$ and $\tilde{x}$ denotes the solution of (11) where y^δ is replaced y. The operator $A_L^\dagger$ is the Moore-Penrose inverse of A_L and also called the *L-pseudo inverse* of A (see e.g. [7, p.28]). Since $A_L^\dagger$ is not continuous for operators A with non-closed range (see [1, Theorem 8.2]) we again use a regularization approach for estimating $\tilde{x}^\delta = A_L^\dagger y^\delta$ in a stable way. Analogous to the regularization of the Moore-Penrose inverse $A^\dagger$ we introduce a family $\{\tilde{R}_\alpha\}_{\alpha>0}$ of continuous operators so that

$$\tilde{R}_\alpha y \to A_L^\dagger y \quad \text{for} \quad \alpha \to 0$$

for all $y \in \mathcal{D}(A_L^\dagger)$. For Tikhonov regularization e.g. we choose

$$\tilde{R}_\alpha := (A^* A + \alpha\, L^* L)^{-1} A^*.$$

This idea can be applied also to iterative regularization methods, e.g. Landweber iteration. It can be shown (see e.g. [5, Lemma 4.1]) that we have to modify (4) in that way that we now solve the iteration sequence

$$x_0^\delta := 0,\ x_k^\delta := x_{k-1}^\delta - \gamma\,(A^*A + L^*L)^{-1} A^* (A\, x_{k-1}^\delta - y^\delta). \tag{12}$$

Since we know that

$$A_L^* := (A^*A + L^*L)^{-1} A^*$$

(see [5]), formula (12) is a simple consequence of applying Landweber iteration onto A_L. In contrast to (4) formula (12) requires a solution of a linear equation in each iteration step, which means much higher numerical costs. To reduce the costs an alternative version of (12) can be found in [1, Example 8.12.].

We now consider convergence rates. Using a regularization technique we set $x_\alpha^\delta := R_\alpha y^\delta$ and consider how fast is the convergence

$$x_\alpha^\delta \to A^\dagger y \quad \text{for} \quad \delta \to 0.$$

To formulate convergence rate results we need some source conditions to the exact data y. In particular conditions of the form

$$A^\dagger y \in \mathcal{R}((A^*A)^\nu), \qquad \nu > 0, \tag{13}$$

are often considered to obtain convergence rates satisfying

$$\|x_\alpha^\delta - A^\dagger y\|_{\mathcal{X}} = \mathcal{O}\left(\delta^{2\,\nu/(2\,\nu+1)}\right). \tag{14}$$

We give the following well-known results (see e.g. [1, Corollary 4.4]).

Proposition 2.2

(i) Assume (13) holds for a constant $0 < \nu \le 1$, $\alpha \sim \delta^{2/(2\,\nu+1)}$ and x_α^δ is the Tikhonov regularized solution of (1) then (14) is valid.

(ii) *Assume (13) holds for a constant* $\nu > 0$, $\alpha \sim \delta^{2/(2\nu+1)}$ *and* x_α^δ *is the Landweber regularized solution of (1) with* $\alpha = n^{-1}$ *then (14) is valid.*

Note that for the case $\nu = 1/2$ the source condition $x \in \mathcal{R}((A^*A)^{1/2})$ is equivalent to $x \in \mathcal{R}(A^*)$.

We consider two examples. For $\nu = 1/2$ and $\alpha \sim \delta$ we have a convergence rate $\mathcal{O}(\sqrt{\delta})$ for both Tikhonov regularization and Landweber iteration. For $\nu = 1$ and $\alpha \sim \delta^{2/3}$ we have a convergence rate $\mathcal{O}(\delta^{2/3})$ which is also the best possible rate for Tikhonov regularization.

Now we will give a result which generalizes these rates for regularizing the L-pseudo inverse $A_L^\dagger$ (see [1, Remark 8.9 + Proposition 8.10]). The following statement pointed out the source conditions to $A_L^\dagger y$ to obtain a convergence rate

$$\|\tilde{x}_\alpha^\delta - A_L^\dagger y\|_{\mathcal{X}} = \mathcal{O}\left(\delta^\nu\right)$$

for the two special cases $\nu = 1/2$ respectively $\nu = 2/3$.

Proposition 2.3

(i) *We obtain for* $\alpha \sim \delta$ *a convergence rate* $\mathcal{O}(\sqrt{\delta})$ *for both Tikhonov regularization and Landweber iteration if*

$$A_L^\dagger y \in \mathcal{D}(L^*L) \quad \textit{and} \quad L^*L\, A_L^\dagger y \in \mathcal{R}(A^*). \tag{15}$$

(ii) *We obtain for* $\alpha \sim \delta^{2/3}$ *a convergence rate* $\mathcal{O}(\delta^{2/3})$ *for both Tikhonov regularization and Landweber iteration if*

$$A_L^\dagger y \in \mathcal{D}(L^*L) \quad \textit{and} \quad L^*L\, A_L^\dagger y \in \mathcal{R}\left(A^*A|_{\mathcal{D}(L)}\right). \tag{16}$$

How we choose the regularization parameter α? A simple and widely used technique is the discrepancy principle of Morozov. This method is says that we choose $\alpha = \alpha(\delta, y^\delta)$ so that

$$\alpha(\delta, y^\delta) := \min\{\alpha \,:\, \|A\,x_\alpha^\delta - y^\delta\|_{\mathcal{Y}} \le \tau\delta\} \tag{17}$$

for a constant $\tau > 0$. For Tikhonov regularization and Landweber iteration we have to choose $\tau > 1$ to obtain convergence and convergence rate results. We summarize them in the following proposition.

Proposition 2.4 *Assume* $y \in \mathcal{R}(A)$ *and* $\alpha(\delta, y^\delta)$ *is chosen by (17). Then the following statements hold:*

(i) *If (13) holds for a constant* $0 < \nu \le 1/2$ *and* x_α^δ *is the Tikhonov regularized solution of (1) then (14) is valid.*

(ii) *Assume (13) holds for a constant* $\nu > 0$ *and* x_α^δ *is the Landweber regularized solution of (1) with* $\alpha = n^{-1}$ *then (14) is valid.*

In particular for Tikhonov regularization the convergence rate $\mathcal{O}(\sqrt{\delta})$ which is obtained for $\nu = 1/2$ is the best possible one we can expect by using the discrepancy principle. Analogous for regularizing $A_L^\dagger$ we obtain a convergence rate $\mathcal{O}(\sqrt{\delta})$ for Tikhonov regularization and Landweber iteration provided the source condition (15) holds. Under the condition (16) a convergence rate $\mathcal{O}(\delta^{2/3})$ is only attainable by using the discrepancy principle (17) for Landweber iteration.

We conclude this section by considering a simple example. Let $A : L^2(0,1) \longrightarrow L^2(0,1)$ denotes the linear integral operator which is given by

$$(A\,x)(t) := \int_0^\tau x(\tau)\,d\tau, \qquad t \in [0,1]. \tag{18}$$

Then $A^* : L^2(0,1) \longrightarrow L^2(0,1)$ is given by

$$(A^* y)(t) := \int_\tau^1 y(\tau)\,d\tau, \qquad t \in [0,1].$$

Furthermore we define $L : \mathcal{D}(L) \subset L^2(0,1) \longrightarrow L^2(0,1)$ as

$$L\,x := x''. \tag{19}$$

Thereby $\mathcal{D}(L) := \{x \in L^2(0,1) \,:\, x'' \in L^2(0,1)\}$. Then L obviously satisfies the conditions (A1)-(A3). Considering the adjoint $L^* : \mathcal{D}(L^*) \subset L^2(0,1) \longrightarrow L^2(0,1)$ of L we obtain $\mathcal{D}(L^*) = \{x \in L^2(0,1) \,:\, x'' \in L^2(0,1),\, x(0) = x(1) = x'(0) = x'(1) = 0\}$ and

$$L^* x = x''.$$

Assume $y \in \mathcal{R}(A)$, then equation (1) has a unique solution $x \in L^2(0,1)$. What do the source conditions of the propositions 2.2 and 2.3 mean for the solution x of (1) in this specific case? We can easily give the following results.

(i) $x \in \mathcal{R}(A^*) \;\Leftrightarrow\; x' \in L^2(0,1)$ and $x(1) = 0$.

(ii) $x \in \mathcal{R}(A^*A) \;\Leftrightarrow\; x'' \in L^2(0,1)$, $x(0) = 0$ and $x'(1) = 0$.

(iii) $x \in \mathcal{D}(L^*L) \;\Leftrightarrow\; x^{(4)} \in L^2(0,1)$ and $x''(0) = x''(1) = x'''(0) = x'''(1) = 0$.

(iv) $L^*L\,x \in \mathcal{R}(A^*) \;\Leftrightarrow\; x^{(5)} \in L^2(0,1)$, $x \in \mathcal{D}(L^*L)$ and $x^{(4)}(1) = 0$.

(v) $L^*L\,x \in \mathcal{R}(A^*A|_{\mathcal{D}(L)}) \;\Leftrightarrow\; x^{(8)} \in L^2(0,1)$, $x \in \mathcal{D}(L^*L)$ and $x^{(5)}(1) = x^{(4)}(0) = 0$.

3 Numerical examples

Before we discuss the numerical realization of the regularization strategies we describe the used discretization of equation (1).

Let $n \in \mathbb{N}$ be a given number. We set

$$\Delta t := \frac{1}{n},\; t_j := \frac{j}{n},\; 0 \le j \le n, \;\text{ and }\; \tilde{t}_j := \frac{j}{n} - \frac{1}{2\,n},\; 1 \le j \le n.$$

Then we can give a discrete version $\underline{A} : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ by using the midpoint rule. We set

$$\underline{A} := \frac{1}{n} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 1 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & 0 \\ 1 & 1 & \cdots & 1 \end{pmatrix}.$$

Let $\underline{y}^\delta := (y_1^\delta, \ldots, y_n^\delta)^T \in \mathbb{R}^n$ be the given (noisy) data of measurements

$$y_i^\delta := y^\delta(t_i), \quad 1 \le i \le n.$$

Let $\underline{x} := (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ be a vector with

$$x_i := x(\tilde{t}_i), \quad 1 \le i \le n.$$

Then we try to find a solution $\underline{x}^\delta$ of the linear equation

$$\underline{A}\,\underline{x} = \underline{y}^\delta \tag{20}$$

which is a discrete version of (1). A solution of (1) is now approximated by a piecewise constant function x^δ with

$$x^\delta(t) = x_i^\delta, \qquad t \in (t_{i-1}, t_i], \qquad t \in (0,1].$$

The operator L of the second derivative we discretize as

$$\underline{L} := n^2 \begin{pmatrix} 1 & -2 & 1 & 0 & \cdots & 0 \\ 0 & 1 & -2 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \ddots & 0 & 1 & -2 & 1 \end{pmatrix} \in \mathbb{R}^{n-2,n}.$$

We will show a numerical result for the (exact) sample function

$$y(t) := \frac{1}{3}\,t^3 - \frac{1}{12}\,t^2 + \frac{7}{20}\,t, \qquad t \in [0,1].$$

Then the exact solution x_{ex} of (1) is given by

$$x_{ex} := t^2 - \frac{1}{4}\,t^2 + \frac{7}{20} = (t-0.5)^2 + 0.1, \qquad t \in [0,1], \tag{21}$$

which we now try to reconstruct by given noisy data. For $n = 50$ and a noise level $\delta = 0.01$ (in the Euclidean norm) we solve (20). Note that for further calculations we always use the same noisy data. The result can be found in Figure 1. As we see the solution $\underline{x}^\delta$ has an oscillating structure which is a result of the ill-posedness of equation (1). Now we apply regularization methods. At first we take Tikhonov regularization and Landweber iteration in the standard form (2) respectively (4). The algorithm is based on the singular value decomposition (SVD) of the matrix $\underline{A}$ (see [3] or Appendix A for a brief overview). In particular for Landweber iteration we can calculate the k-th element of the iteration sequence (4) explicitly. The regularization parameter α for Tikhonov regularization and k for Landweber iteration is chosen optimally. That means, the results shown in Figure 2 and Figure 5 are the best possible approximations x_α^δ respectively x_k^δ of x_{ex} we could obtain by using these regularization strategies. Two effects we see in both pictures: On the one hand we have an improvement of the reconstruction of the sought function x_{ex}. But otherwise the founded solutions x_α^δ and x_k^δ are still oscillating quite strongly. Hence an improvement of the regularization results are of high interest.

Therefore now Tikhonov regularization and Landweber iteration of the form (3) respectively (12) are applied. For solving these problems numerically we use the generalized

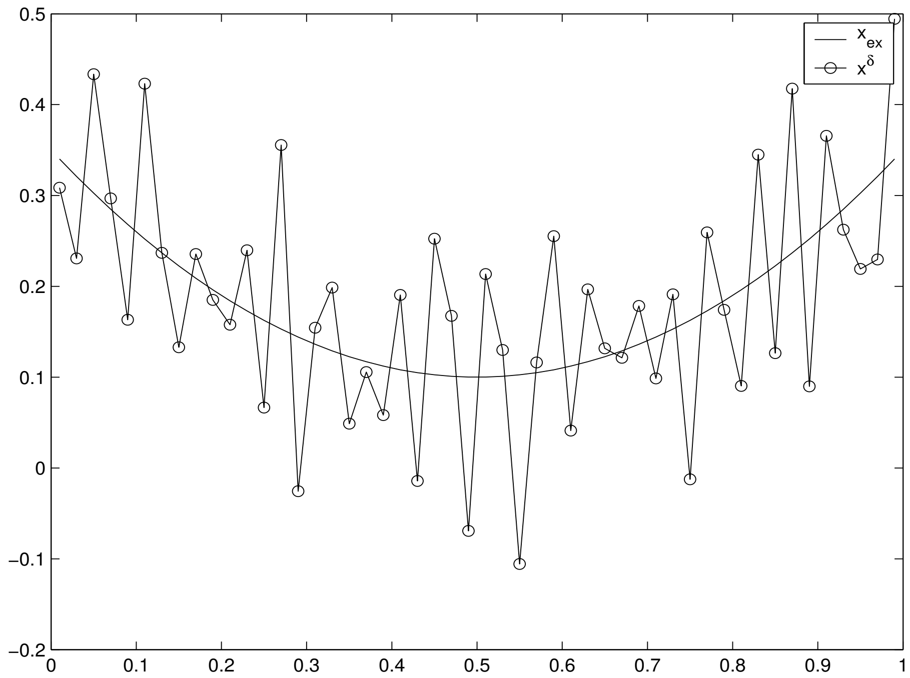

Figure 1: unregularized solution, $n = 50$, $\delta = 0.01$

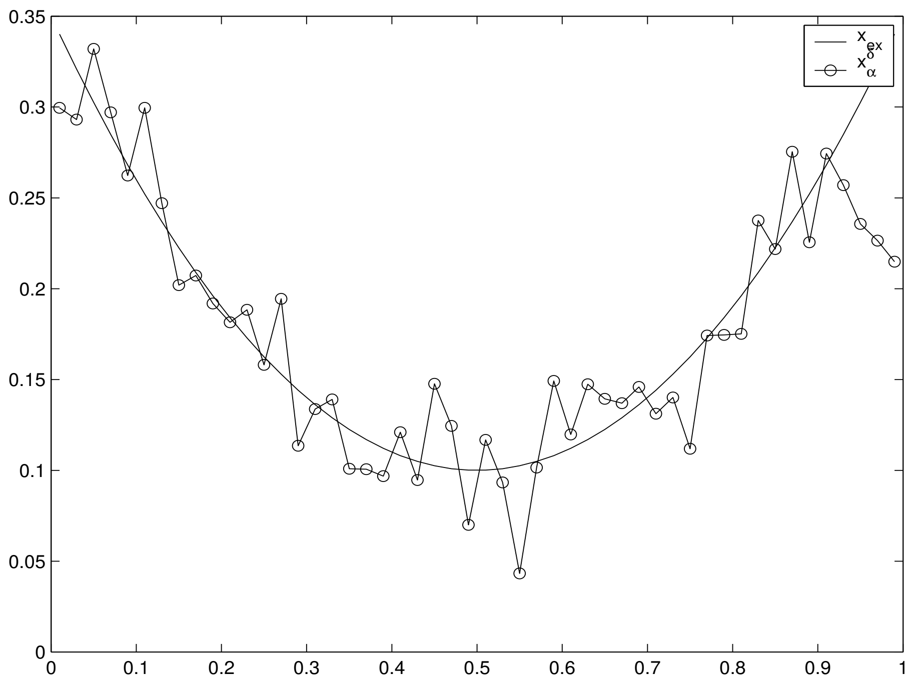

Figure 2: Tikhonov regularization, $n = 50$, $\delta = 0.01$, $\alpha_{op} = 5.5 \cdot 10^{-4}$

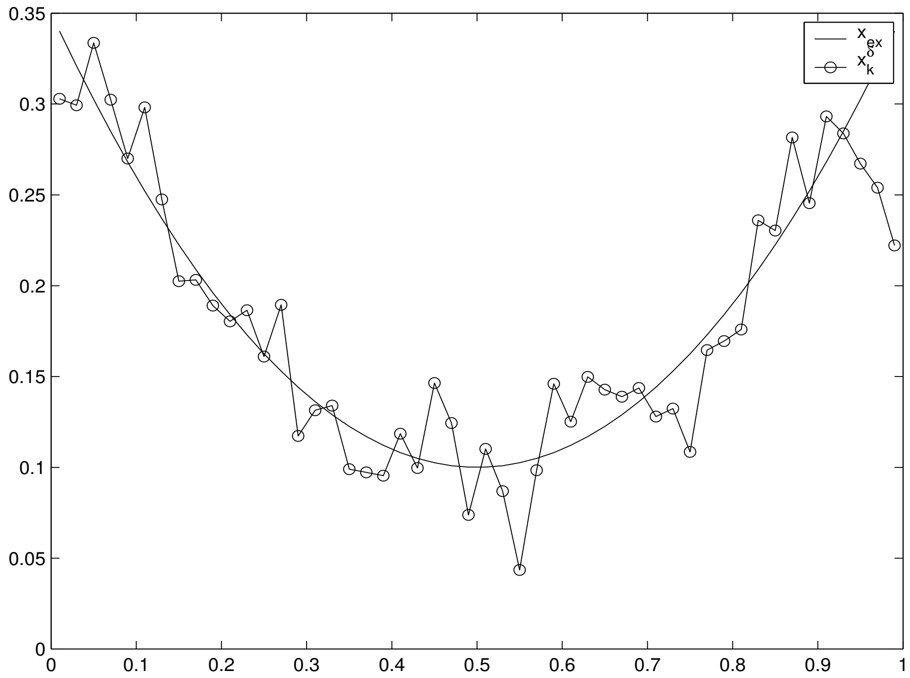

Figure 3: classical Landweber iteration, $n = 50$, $\delta = 0.01$, $k_{op} = 1381$

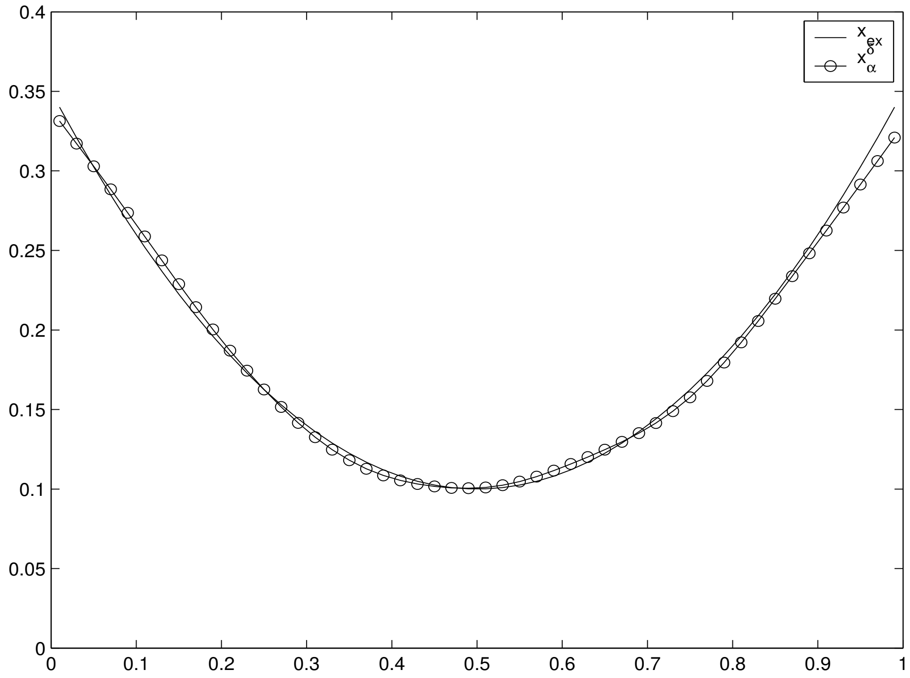

Figure 4: Tikhonov regularization with higher regularity, $n = 50$, $\delta = 0.01$, $\alpha_{op} = 4.6 \cdot 10^{-8}$

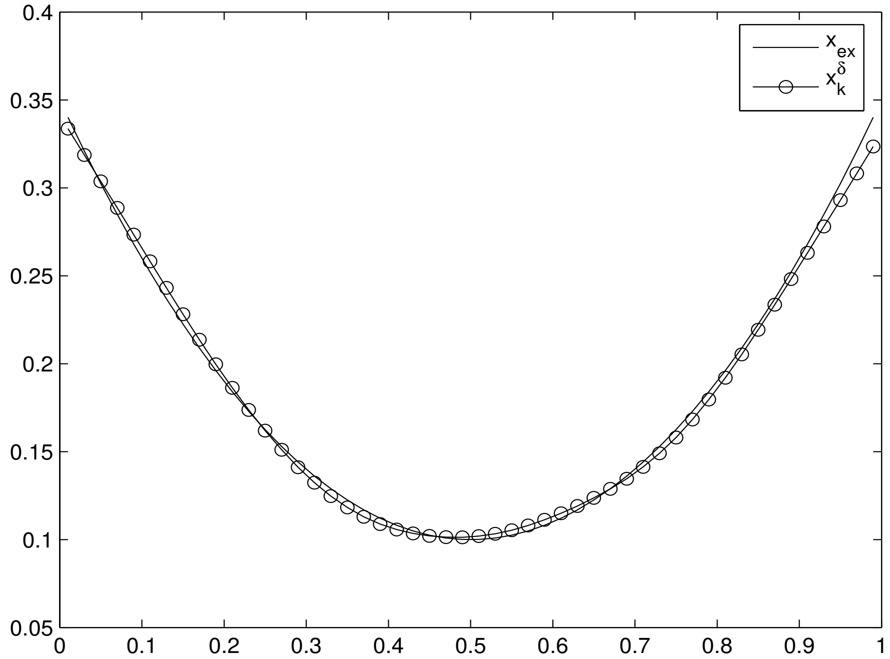

Figure 5: Landweber iteration with higher regularity, $n = 50$, $\delta = 0.01$, $k_{op} = 10\,525\,316$

singular value decomposition (GSVD) of the matrix pair $(\underline{A}, \underline{L})$ (see again [3] Appendix A. As by the SVD based algorithm above we can calculate each iteration step of the Landweber iteration (12) separately. We again choose the regularization parameter optimally and present the results in Figure 4 respectively Figure 5. The results show that we can improve the chances of the identifiability of the sought solution x_{ex} significantly by using the norm of derivatives as penalty term. Both Tikhonov regularization and Landweber iteration provide similar good results.

Also the numerical costs are in both cases comparably. The last fact could be obtained by applying the GSVD-based solution algorithm. Using the iteration sequence (12) we have to calculate $k_{op} = 10\,525\,316$ iteration steps to achieve the optimal regularization result. Then the costs of the Landweber iteration are clearly higher even we have to solve a linear equation for each α in Tikhonov regularization. This is a disadvantage when the (discrete) linear operator $\underline{A}$ is not given in matrix form (as for example in linear problem for differential equations). The same problem occurs when the ideas of introducing different penalty terms for nonlinear problems. The slow convergence of Landweber iteration can lead to high numerical costs. We finish the section with two error tables. Table 1 shows the regularization parameter and the regularization errors for the optimal parameter choice for all four regularization strategies. In a second Table 2 we show the obtained regularization parameter and errors when we apply the discrepancy principle (17). We see that both regularization strategies provide comparable result. Introducing a smoothing operator L gives much better reconstruction chances for both Tikhonov regularization and Landweber iteration. In all four cases the discrepancy principle leads to an over-regularization which means that the regularization parameter α is chosen somewhat too large. But this is a known property of the discrepancy principle.

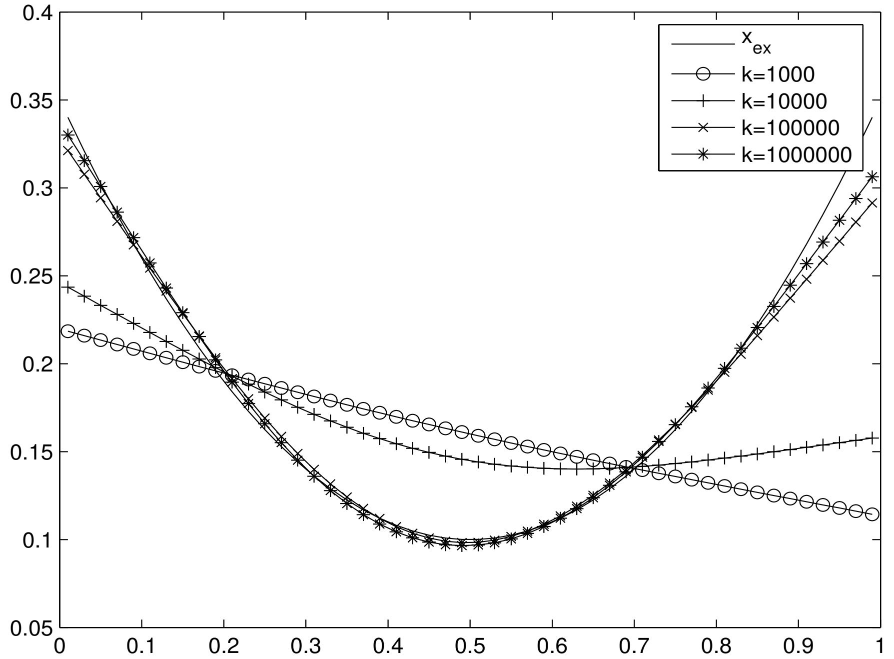

Figure 6: Landweber iteration with higher regularity, different k

We consider the last row of Table 1 and 2. Comparing the number of iterations we remark that the optimal iteration number k_{op} is more than 60 times larger then the iteration number obtained by the discrepancy principle. On the other hand the reconstruction error has been only halfed. The reason is the slow convergence of Landweber iteration with smoothing penalty operator for larger k. To illustrate the fact we show in Figure 6 the solutions x_k^δ for different k. Between $k = 10\,000$ and $k = 100\,000$ there are only small differences to be seen. In particular a slight improvement of the regularized solution x_k^δ we can obtain only by an extreme increasing of the number of iteration steps.

	(1)	(2)	(3)	(4)	(5)
reg.param.	–	$\alpha = 5.5 \cdot 10^{-4}$	$\alpha = 4.6 \cdot 10^{-8}$	$k = 1381$	$k = 10\,526\,316$
$\\|\underline{x}_\alpha^\delta - \underline{x}_{ex}\\|_2$	0.7287	0.2395	0.0378	0.2172	0.0338

(1) … without regularization
(2) … Tikhonov regularization (2)
(3) … Tikhonov regularization (10)
(4) … Landweber iteration (4)
(5) … Landweber iteration (12)

Table 1: reconstruction error for optimal parameter choice

	(1)	(2)	(3)	(4)	(5)
reg.param.	–	$\alpha = 6.9443 \cdot 10^{-4}$	$\alpha = 9.2891 \cdot 10^{-7}$	$k = 634$	$k = 169\,855$
$\Vert\underline{x}_\alpha^\delta - \underline{x}_{ex}\Vert_2$	0.7287	0.2429	0.0692	0.2495	0.0691

(1) ... without regularization
(2) ... Tikhonov regularization (2)
(3) ... Tikhonov regularization (10)
(4) ... Landweber iteration (4)
(5) ... Landweber iteration (12)

Table 2: reconstruction error choosing α/k with the discrepancy principle

4 Regularization in H^1

Assume – as in our example – the case $\mathcal{X} = L^2(0,1)$. A special choice of the penalty operator L are differential operators. In particular we try to improve the regularization chances by providing a higher regularity of the sought solution x. This approach suggest the alternative idea to replace the solution space $\mathcal{X} = L^2(0,1)$ by $H^s(0,1)$, $s > 0$. Consequently we introduce $\hat{A} : H^s(0,1) \longrightarrow \mathcal{Y}$ by

$$\hat{A} := A \circ I_{(s)},$$

where $I_{(s)} : H^s(0,1) \longrightarrow L^2(0,1)$ denotes the (compact) embedding operator from $H^s(0,1)$ into $L^2(0,1)$. Instead of (1) we now consider the equation

$$\hat{A}\,x = y^\delta, \qquad x \in H^s(0,1). \tag{22}$$

As $A^\dagger$ the Moore-Penrose inverse $\hat{A}^\dagger$ of $\hat{A}$ remains unbounded too. Therefore – for solving equation (22) – we have to apply regularization strategies again. Using the fact that for $x \in H^s(0,1)$ and $y \in \mathcal{Y}$

$$\hat{A}\,x = A\,I_{(s)}\,x = A\,x \quad \text{respectively} \quad \hat{A}^* y = \left(A \circ I_{(s)}\right)^* y = I_{(s)}^* A^* y$$

is valid we can Tikhonov regularization write as

$$x_\alpha^\delta := (I_{(s)}^* A^* A + \alpha\, I)^{-1} I_{(s)}^* A^* y^\delta. \tag{23}$$

Applying Landweber iteration the k-th element of the iteration process is given by

$$x_0 := 0, \quad x_k^\delta := x_{k-1}^\delta - \gamma\, I_{(s)}^* A^* \left(A\, x_{k-1}^\delta - y^\delta\right). \tag{24}$$

In both regularization methods (23) respectively (24) the additional adjoint operator $I_{(s)}^* : \mathcal{D}(I_{(s)}^*) \subseteq L^2(0,1) \longrightarrow H^s(0,1)$ has to be estimated. We give a simple example.

Example 4.1 *For the case $s = 1$ the adjoint $I_{(1)}^* : \mathcal{D}(I_{(1)}^*) \subseteq L^2(0,1) \longrightarrow H^1(0,1)$ of the embedding operator from $H^1(0,1)$ into $L^2(0,1)$ is given by*

$$I_{(1)}^* y := x,$$

where x solves the ordinary differential equation

$$x - x'' = y, \quad x'(0) = x'(1) = 0,$$

$$\mathcal{R}(I_{(1)}^*) := \left\{x \in L^2(0,1) \,:\, x'' \in L^2(0,1),\ x'(0) = x'(1) = 0\right\}.$$

Besides oscillation properties of the solution x_α^δ there is another effect which destroys the quality of the regularized solution. We consider classical Landweber iteration (4). Obviously we have

$$x_k^\delta \in \mathcal{R}(A^*).$$

for all $k \in \mathbb{N}$. In our example (18) we can conclude $(x_k^\delta)' \in L^2(0,1)$ and

$$x_k^\delta(1) = 0. \tag{25}$$

In particular for obtaining pointwise convergence $x_k^\delta(t) \to x(t)$, $t \in [0,1]$ for $k \to \infty$ we have to suppose $x(1) = 0$. This effect can be seen numerically in Figure 3 on the right boundary of the interval $[0,1]$.

Applying the modified Landweber iteration (24) with $s = 1$ we get some additional boundary conditions. In particular we get the additional assumption

$$x \in \mathcal{R}\left(I_{(1)}^* A^*\right)$$

and hence $x''' \in L^2(0,1)$ as well as the boundary conditions

$$x_k^\delta(1) = (x_k^\delta)''(1) \;\text{ and }\; (x_k^\delta)'(0) = (x_k^\delta)'(1) = 0. \tag{26}$$

Therefore we conclude the following result. Replacing (4) by (24) with $s = 1$ we obtain two effects:

1. We increase the regularity of x_k^δ, in particular we obtain a regularized solution x_k^δ which is less oscillating.
2. For any regularized solution x_k^δ we have replaced the source condition (25) by the condition (26).

This effect we can illustrate in Figure 7. Note that the sample function x_{ex} from (21) satisfies neither the condition (25) nor (26). Using the example from the previous section we apply the iteration sequence (24) for $s = 1$. The number k of iterations is again chosen optimally. Comparing the result with Figure 3 we see two main differences. Choosing $H^1(0,1)$ as solutions space $\mathcal{X}$ leads to smoother a solutions x_k^δ. On the other hand there occurs a 'new' boundary effect for $t = 0$. In particular we have $(x_k^\delta)'(0) = 0$, which destroys the quality of the (pointwise) reconstruction of x_{ex}. Summarizing we obtain higher regularity by destroying the convergence at the point $t = 0$. To motivate the approach suggested below we give another example. We again examine the standard Landweber iteration (4) for the sample operator A which is given by (18). In order to eliminate the boundary effect (25) we consider the approximation error $\|x_k^\delta - x_{ex}\|_{\mathcal{X}}$ not on the whole interval $(0,1)$ but on the subinterval $(c, 1-c)$ for a constant $0 < c < 1/2$. Choosing $c = 0.1$ and k in that way that the error $\|x_k^\delta - x_{ex}\|_{L^2(0.1,0.9)}$ is minimized we get a result which is shown in Figure 8. After $k = 238$ iterations we obtain a solution which approximates the solution x_{ex} near the boundary worse than the optimally solution $x_{k_{op}}^\delta$ which is shown in Figure 3. On the other hand inside the interval $(0,1)$ the solution in Figure 8 approximates x_{ex} much better. That means – improving the reconstruction result at the boundary by a higher number of iterations leads to higher oscillations inside the interval $(0,1)$. This provides a worse reconstruction result there.

To remove these boundary effects we want to suggest a heuristic approach for solving (22). The idea is simple. We consider equation (22) not on the interval $(0,1)$ but on

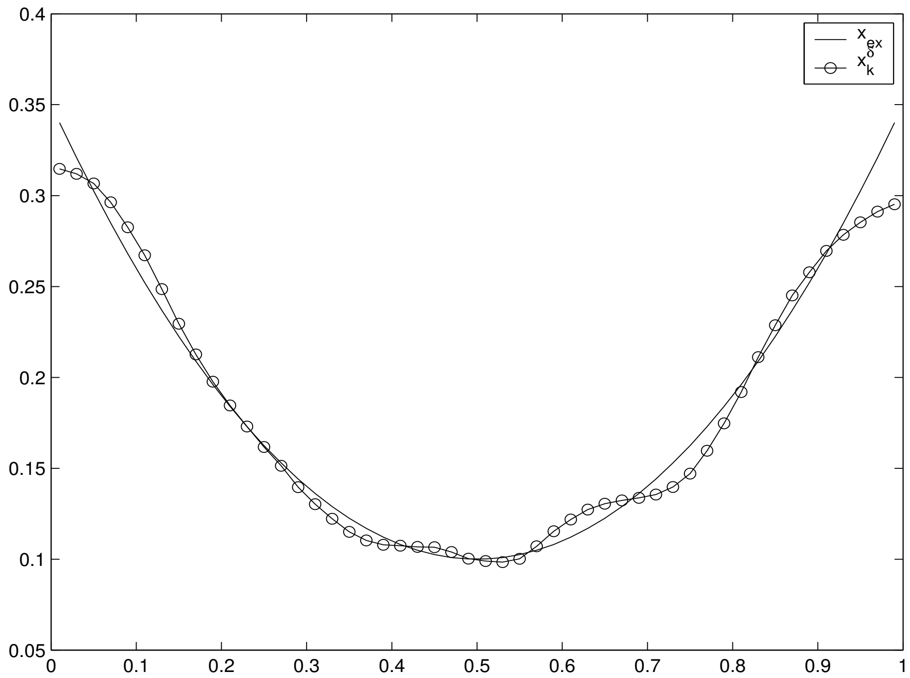

Figure 7: Landweber iteration in $H^1(0,1)$, $n = 50$, $\delta = 0.01$, $k_{op} = 186\,392$

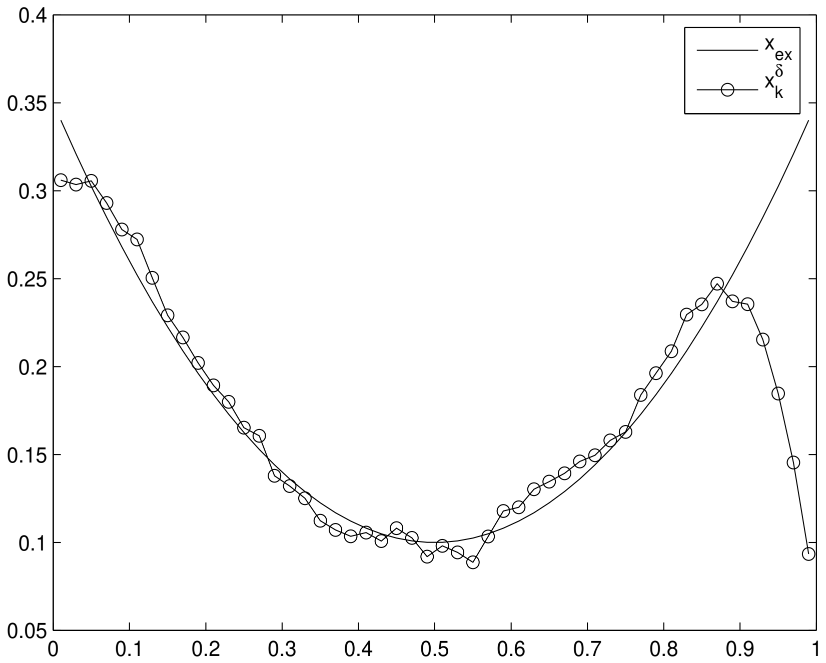

Figure 8: Landweber iteration on the interval $(0.1, 0.9)$, $n = 50$, $\delta = 0.01$, $k_{op} = 238$

a larger interval $(a,b) \supset (0,1)$. Thereby we have to continue the data y^δ in a suitable way. Solving the equation (22) on the interval (a,b) leads to the consequence that the boundary conditions (26) for $t=0$ and $t=1$ are removed to boundary condition for $t=a$ and $t=b$. We illustrate the approach on the concrete example.

We apply Landweber iteration and consider the residuum

$$r_k^\delta := A\,x_k^\delta - y^\delta \in L^2(0,1)$$

of the k-th iteration step. We generate a continuation $\tilde r_k^\delta \in L^2(a,b)$ for example in the way that

$$\tilde r_k^\delta := \begin{cases} r_k^\delta, & \text{on } (0,1), \\ 0, & \text{else}. \end{cases} \tag{27}$$

This approach can be motivated by a simple consideration. Let $\tilde A : \mathcal{D}(\tilde A) \subseteq L^2(a,b) \longrightarrow L^2(a,b)$ be any continuation of A on the interval (a,b). Then $\tilde A^* \tilde r_k^\delta$ is the gradient of the functional

$$x \in L^2(a,b) \mapsto J(x) := \frac{1}{2}\left\|\tilde A\,x - y^\delta\right\|^2_{L^2(0,1)} \tag{28}$$

at the point $x = x_k^\delta$. For arbitrary $h \in L^2(a,b)$ we derive

$$\begin{aligned} \nabla J(x_k^\delta)\,h &:= \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\left(\frac{1}{2}\left\|\tilde A\,(x_k^\delta + \varepsilon\,h) - y^\delta\right\|^2_{L^2(0,1)} - \frac{1}{2}\left\|\tilde A\,x_k^\delta - y^\delta\right\|^2_{L^2(0,1)}\right) \\ &= \int_0^1 \left(\tilde A\,x_k^\delta - y^\delta\right)\tilde A\,h\,dt \\ &= \int_0^1 r_k^\delta\,\tilde A\,h\,dt = \int_a^b \tilde r_k^\delta\,\tilde A\,h\,dt = \int_a^b \tilde A^* \tilde r_k^\delta\,h\,dt \end{aligned}$$

and hence

$$\nabla J(x_k^\delta) = \tilde A^* \tilde r_k^\delta$$

by the Riesz representation theorem. Using this relationship we write Landweber iteration as

$$x_0 := 0,\ x_k^\delta := x_{k-1}^\delta - \gamma\,\tilde I^*_{(s)} \tilde A^* \tilde r_{k-1}\Big|_{(0,1)}$$

where $\tilde r_k^\delta$ is defined via (27) and $\tilde I^*_{(s)}$ is the adjoint embedding operator from $H^s(a,b)$ into $L^2(a,b)$.

Example 4.2 *For our example (18) we set*

$$D(\tilde A) := \left\{x \in L^2(a,b)\ :\ x(t) = 0 \quad a.e.\ \ on\ \ (a,0)\right\}$$

and define

$$(\tilde A\,x)(t) := \int_a^t x(t)\,dt, \qquad t \in [a,b],$$

which seems to be natural. Then we obtain

$$(\tilde A^* r_k^\delta)(t) := \begin{cases} 0, & t > 1, \\ (A^* r_k^\delta)(t), & t \in [0,1], \\ (A^* r_k^\delta)(0), & t < 0. \end{cases}$$

	optimal k		discrepancy principle	
$\mathcal{X}$	k	$\Vert \underline{x}_k^\delta - \underline{x} \Vert_2$	k	$\Vert \underline{x}_k^\delta - \underline{x} \Vert_2$
$L^2(0,1)$	1381	0.2172	634	0.2495
$H^1(0,1)$	186 392	0.0791	35 440	0.1108
$H^1(-0.5,1.5)$	46 124	0.0343	14 353	0.0660
$H^1(-1,3)$	35 979	0.0300	12 880	0.0551

Table 3: reconstruction errors for modified Landweber iteration

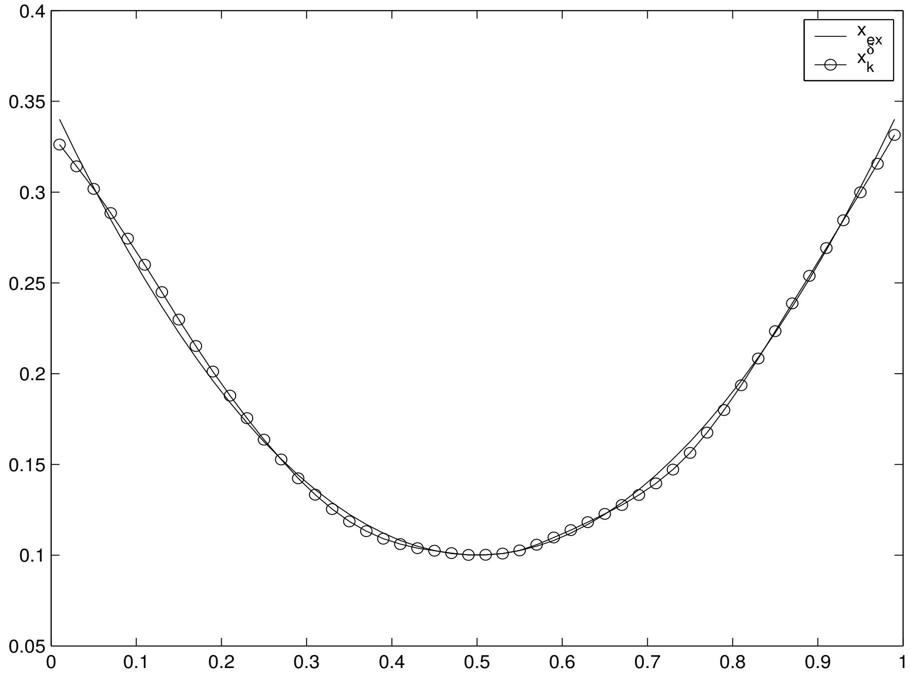

Figure 9: Landweber iteration, $n = 50$, $\delta = 0.01$, $(a, b) = (-1, 3)$, $k_{op} = 35\,979$

Hence the second boundary condition of (26) for $t = 0$ respectively $t = 1$ are replaced by a boundary conditions for $t = a$ respectively $t = b$. The less restrictive boundary condition $x_k^\delta(1) = (x_k^\delta)''(1)$ remains valid.

We give the following result in Figure 9. We choose the interval $(a, b) = (-1, 3)$ which is three times larger than the original one. As we can see the boundary effects which occurred in Figure 5 are vanished. The reconstruction on the boundary are also quite well. Another effect is noticeably. Considering the number of necessary iterations k_{op} we see that for $(a, b) = (-1, 3)$ we need less than $1/5$ of iterations to receive a better regularization result. To specify that fact we present Table 3. For different interval (a, b) we give the number k of iterations steps obtained by choosing $k = k_{op}$ optimally and by the discrepancy principle as well as the corresponding reconstruction errors. We see that by increasing the the interval (a, b) we obtain better reconstruction results by a lower number of iterations. On the other hand the numerical costs increase in each iteration step. The reason is calculation of the adjoint $\tilde{I}^*_{(1)}$ by constant step width of the discretization.

Remark 4.1 *The approach remains somewhat heuristic. The reason therefore is the fact* *that $\mathcal{D}(\tilde{A})$ contains only functions which vanish on the interval $(a,0)$. Since we consider the embedding $H^1(a,b) \to L^2(a,b)$ we have continuous functions x which leads to $x(0)=0$ for all $x \in \mathcal{D}(\tilde{A})$. On the other hand the numerical results show that the approach is suited quite well to improve the regularization results.*

A SVD and GSVD

A powerful tool to implement regularization strategies numerically is the SVD of the matrix $\underline{A}$ in the case of approximating $A^\dagger$. For regularization with semi-norms the GSVD of the matrix pair $(\underline{A}, \underline{L})$ can be applied. For the sake of completeness we repeat the definitions of both SVD and GSVD. We start with the SVD.

Proposition A.1 *Let $\underline{A} \in \mathbb{R}^{m,n}$ be a rectangular or square matrix with $m \geq n$. Then there exists matrices $\underline{U} = (\underline{u}_1|\dots|\underline{u}_n) \in \mathbb{R}^{m,n}$, $\underline{V} = (\underline{v}_1|\dots|\underline{v}_n) \in \mathbb{R}^{n,n}$ with orthonormal columns and $\underline{\Sigma} = \operatorname{diag}(\sigma_1,\dots,\sigma_n) \in \mathbb{R}^{n,n}$ with $\sigma_1 \geq \dots \geq \sigma_n \geq 0$ so that the decomposition*

$$\underline{A} = \underline{U}\,\underline{\Sigma}\,\underline{V}^T \tag{29}$$

of $\underline{A}$ holds.

The GSVD is defined as follows.

Proposition A.2 *Let $\underline{A} \in \mathbb{R}^{m,n}$ and $\underline{L} \in \mathbb{R}^{p,n}$ with $m \geq n \geq p$, where $\underline{L}$ has full rank. Then there exist matrices $\underline{U} = (\underline{u}_1|\dots|\underline{u}_n) \in \mathbb{R}^{m,n}$, $\underline{V} = (\underline{v}_1|\dots|\underline{v}_p) \in \mathbb{R}^{p,p}$ with orthonormal columns, $\underline{\Sigma} = \operatorname{diag}(\sigma_1,\dots,\sigma_p) \in \mathbb{R}^{p,p}$ respectively $\underline{M} = \operatorname{diag}(\mu_1,\dots,\mu_p) \in \mathbb{R}^{p,p}$ with $1 \geq \sigma_1 \geq \dots \geq \sigma_p \geq 0$ respectively $0 \leq \mu_1 \leq \dots \leq \mu_p \leq 1$ and*

$$\sigma_i^2 + \mu_i^2 = 1, \quad 1 \leq i \leq p,$$

and $\underline{X} = (\underline{x}_1|\dots|\underline{x}_n) \in \mathbb{R}^{n,n}$ so that

$$\underline{A} = \underline{U} \begin{pmatrix} \underline{\Sigma} & 0 \\ 0 & \underline{I}_{n-p} \end{pmatrix} \underline{X}^{-1} \quad \textit{respectively} \quad \underline{L} = \underline{V}\,(\underline{M}\,|\,0)\,\underline{X}^{-1}. \tag{30}$$

The quotients

$$\gamma_i := \frac{\sigma_i}{\mu_i}, \qquad 1 \leq i \leq p,$$

are called the general singular values of the matrix pair $(\underline{A}, \underline{L})$. The columns of $\underline{X}$ are $\underline{A}^T\underline{A}$-orthogonal, in particular

$$\underline{X}^T \underline{A}^T \underline{A}\,\underline{X} = \begin{pmatrix} \underline{\Sigma}^2 & 0 \\ 0 & \underline{I}_{n-p} \end{pmatrix} \quad \textit{and} \quad \underline{X}^T \underline{L}^T \underline{L}\,\underline{X} = \begin{pmatrix} \underline{M}^2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Knowing the SVD then we can discrete Moore-Penrose inverse $\underline{A}^\dagger$ write explicitly as

$$\underline{x} = \underline{A}^\dagger \underline{y} = \sum_{j=1}^{n} \frac{\underline{u}_j^T \underline{y}}{\sigma_j}\,\underline{v}_j. \tag{31}$$

For given GSVD of the pair $(\underline{A}, \underline{L})$ the L-pseudo inverse $\underline{A}_L^\dagger$ of $\underline{A}$ is given by

$$\underline{x} = \underline{A}_L^\dagger \underline{y} = \sum_{j=p+1}^{n} \left(\underline{u}_j^T \underline{y}\right) \underline{x}_j + \sum_{j=1}^{p} \frac{\underline{u}_j^T \underline{y}}{\sigma_j} \underline{x}_j. \tag{32}$$

Many regularization methods can also be described with so-called filter functions $f(\alpha, \sigma)$ (see [3] or [4, section 4.1.1]). For example Tikhonov regularization is defined by the filter function

$$f(\alpha, \sigma) := \frac{\sigma^2}{\sigma^2 + \alpha}$$

and we can write

$$\underline{x}_\alpha^\delta := \sum_{j=1}^{n} f(\alpha, \sigma_j) \frac{\underline{u}_j^T \underline{y}}{\sigma_j} \underline{v}_j = \sum_{j=1}^{n} \frac{\sigma_j}{\sigma_j^2 + \alpha} \left(\underline{u}_j^T \underline{y}\right) \underline{v}_j \tag{33}$$

respectively

$$\tilde{\underline{x}}_\alpha^\delta := \sum_{j=p+1}^{n} \left(\underline{u}_j^T \underline{y}\right) \underline{x}_j + \sum_{j=1}^{p} f(\alpha, \gamma_j) \frac{\underline{u}_j^T \underline{y}}{\sigma_j} \underline{x}_j = \sum_{j=p+1}^{n} \left(\underline{u}_j^T \underline{y}\right) \underline{x}_j + \sum_{j=1}^{p} \frac{\sigma_j}{(\gamma_j^2 + \alpha)\,\mu_j^2} \left(\underline{u}_j^T \underline{y}\right) \underline{x}_j.$$

Although Landweber iteration is defined classical as iterative regularization method $\underline{x}_k^\delta$ respectively $\tilde{\underline{x}}_k^\delta$ can be calculated directly by applying SVD respectively GSVD and the filter function

$$f(k, \sigma) := 1 - \left(1 - \gamma\, \sigma^2\right)^k.$$

Determine the SVD respectively the GSVD is useful when we have to estimate the regularization parameter α. We give the algorithm of applying the discrepancy principle of Morozov for the discrete version of Tikhonov regularization (2).

Algorithm: discrepancy principle for Tikhonov regularization

1. calculate SVD of $\underline{A}$
2. set $\alpha := \alpha_0$
3. calculate $\underline{x}_\alpha^\delta$ by formula (33). Is $\|\underline{A}\, \underline{x}_\alpha^\delta - \underline{y}^\delta\|_{\leq} \delta$ then STOP
4. $\alpha = \alpha\, c$ $(c < 1)$, goto 3.

That means we have to calculate the SVD of $\underline{A}$ once and can solve the discrete version of (2) by simple matrix$\times$vector multiplications for each α. We avoid costs of the solution of a linear equations system for each α.

By a slight modification we can apply the algorithm above also for Landweber iteration (4).

Algorithm: discrepancy principle for Landweber iteration

1. calculate SVD of $\underline{A}$
2. set $\alpha := \alpha_0$
3. set $k := \mathrm{round}(\alpha^{-1})$

4. calculate $\underline{x}_k^\delta$ by

$$\underline{x}_k^\delta := \sum_{j=1}^{n} \frac{1-(1-\gamma\,\sigma_j^2)^k}{\sigma_j}\,\left(\underline{u}_j^T\,\underline{y}\right)\underline{v}_j$$

   Is $\|\underline{A}\,\underline{x}_k^\delta - \underline{y}^\delta\|_{\leq}\delta$ then STOP

5. $\alpha = \alpha\,c$ $(c < 1)$, goto 3.

In particular we do not calculate $\underline{x}_k^\delta$ for each k. This helps to overcome the slow convergence of the Landweber iteration process. As we see the numerical costs are similar for both regularization methods by using the SVD. The application of the GSVD occurs analogous.

References

[1] Engl, H.W.; Hanke, M.; Neubauer, A.: *Regularization of Inverse problems.* Kluwer Academic Publishers (1996).

[2] Hanke, M.; Scherzer, O.: *Inverse problems light: Numerical differentiation.* Am.Math.Mon. 108, No.6, 2001, pp.512-521.

[3] Hansen, P.C.: *Rank-Deficient and Discrete Ill-posed Problems.* Philadelphia SIAM 1998.

[4] Hofmann, B.: *Mathematik Inverse Probleme.* B.G. Teubner Stuttgart · Leipzig, 1999.

[5] Locker, J.; Prenter, P.: *Regularization with differential operators I: General theory.* J.Math.Anal.Appl. 74 (1980), pp.505-529.

[6] Neubauer, A.: *When do Sobolev spaces form a Hilbert scale?* Proc.Am.Math.Soc. 103, 1988, pp.557-562.

[7] Morozov, V.A.: *Regularization methods for ill-posed problems.* CRC Press (1993).