

Georg-August-Universität Göttingen



Optimal Regularization with Two Interdependent Regularization Parameters

Frank Bauer, Olha Ivanyshyn

Preprint Nr. 2006-19

Preprint-Serie des
Instituts für Numerische und Angewandte Mathematik
Lotzestr. 16-18
D - 37083 Göttingen

Optimal Regularization with Two Interdependent Regularization Parameters

Frank Bauer¹, Olha Ivanyshyn¹

¹ University of Göttingen, Institute for Numerical and Applied Mathematics,
Lotzestr. 16-18, 37083 Göttingen, Germany

E-mail: bauer@mathematik.uni-kl.de, ivanyshy@math.uni-goettingen.de

Abstract. We will consider the situation of ill-posed problems with regularization methods which demand the choice of interdependent regularization parameters. We show that one can adapt the Lepskij-type balancing principle to this situation.

AMS classification scheme numbers: 65J20

1. Introduction

In the practice of inverse problems rather often one needs to reconstruct two or more quantities simultaneously originating from two completely different spaces. I.e. assume that we have an ill-posed operator equation

$$\tilde{A}(u_1, u_2) = y \quad (1)$$

where $\tilde{A} : \mathcal{X}_1 \times \mathcal{X}_2 \rightarrow \mathcal{Y}$ with separable Hilbert spaces \mathcal{X}_1 , \mathcal{X}_2 and \mathcal{Y} .

Due to measurement errors we just have access to the perturbed data y^δ . Therefore we need to apply regularization techniques. Examples are blind deconvolution [1] or determining cracks out of a background medium [2] respectively inclusions [3].

When we penalize each of the quantities u_1 , u_2 in (1) in a Tikhonov like fashion [4] we get that the solution of (2) depends nonlinearly on both regularization parameters α and β ; however this dependence is unknown. The choice of the regularization parameters is crucial.

$$(u_{1,\alpha}^\delta, u_{2,\beta}^\delta) = \underset{(u_1, u_2) \in \mathcal{X}_1 \times \mathcal{X}_2}{\operatorname{argmin}} \|\tilde{A}(u_1, u_2) - y^\delta\|_{\mathcal{Y}}^2 + \alpha \|u_1\|_{\mathcal{X}_1}^2 + \beta \|u_2\|_{\mathcal{X}_2}^2 \quad (2)$$

Parameter choice methods as proposed in [5, 6] are not applicable in this situation where either the regularization parameters could be chosen independently or an explicit dependence structure is known. In [7] it is proposed to use either L-curve or generalized cross-validation like methods to tackle this problem, however no convergence proofs for these algorithms are presented.

The Lepskij-type balancing principle [8] has already been proven to work in a number of situations with one regularization parameter [9, 10]. We will introduce in section 2 how this principle can be adapted to this two parameter setting and prove optimal convergence rates (Corollary 2.20). As a starting point we will assume to have a valid regularization procedure, i.e. we demand Assumptions 2.4 and 2.5 to be satisfied. In [11] one can find a proof for the one parameter case.

The balancing principle requires the knowledge of the measurement noise behavior with respect to the regularization operators. A proposal how to compute this quantity in practice is presented in section 3.

Two numerical examples, one proof of concept and a crack detection problem, will be shown in the end.

2. Balancing Principle

In the following we will assume to have a (possibly non-linear) not continuously invertible operator $A : \mathcal{X} \rightarrow \mathcal{Y}$ where \mathcal{X} is a metric space with metric $d(\cdot, \cdot)$. We want to solve the equation

$$Ax = y. \quad (3)$$

Furthermore we assume to have a regularization method with two regularization parameters which can be indexed by $(n_1, n_2) \in \mathbb{N} \times \mathbb{N}$ where \mathbb{N} denotes the positive integers. This is not a severe restriction because we can just do the calculations for a finite number of regularization parameters. From now on for the sake of simplicity we will call this pair of indexes (n_1, n_2) regularization parameters.

Definition 2.1 *The regularized solution of (3) with respect to the noisy set of input data y^δ and the pair of regularization parameters (n_1, n_2) is denoted by $x_{(n_1, n_2)}^\delta$.*

The regularized solution with respect to the noise free set of input data $y = y^0$ and the pair of regularization parameters (n_1, n_2) is denoted by $x_{(n_1, n_2)}^0$. The exact solution is denoted by x .

In regularization methods with one parameter one has, e.g., that for small regularization parameters (that means regularization parameters with small indexes) the solution is too smooth and measurement noise does not play a role whereas for large parameters noise dominates the solution. In order to introduce a similar concept we will need a half order on the regularization parameters (n_1, n_2) .

Definition 2.2 (Half Order) *The pair (n_1, n_2) of regularization parameters is smaller than or equal to (m_1, m_2) , denoted by $(n_1, n_2) \leq (m_1, m_2)$ iff both $n_1 \leq m_1$ and $n_2 \leq m_2$.*

If neither $(n_1, n_2) \leq (m_1, m_2)$ nor $(n_1, n_2) \geq (m_1, m_2)$ the pairs are not comparable and we denote this by $(n_1, n_2) \not\leq (m_1, m_2)$.

The set of comparable regularization parameters with respect to (n_1, n_2) is denoted by $\mathcal{C}_{(n_1, n_2)}$.

Remark 2.3 *For all pairs of regularization parameters (n_1, n_2) , (m_1, m_2) and (k_1, k_2) the following statements hold:*

- $(1, 1) \leq (n_1, n_2)$
- $(n_1, n_2) \in \mathcal{C}_{(m_1, m_2)}$ iff $(m_1, m_2) \in \mathcal{C}_{(n_1, n_2)}$
- if $(n_1, n_2) \leq (m_1, m_2)$ and $(m_1, m_2) \leq (k_1, k_2)$ then $(n_1, n_2) \leq (k_1, k_2)$

First of all we will ask for big regularization parameters to control the bias, i.e. the error we would make in a noise-free situation:

Assumption 2.4 (Smoothness) *Assume that there exists a monotonically decreasing (with respect to the above half-order) continuous function $\psi_x : \mathbb{N} \times \mathbb{N} \rightarrow [0, \infty]$ with*

$$d(x, x_{(n_1, n_2)}^0) \leq \psi_x(n_1, n_2). \quad (4)$$

Furthermore there exists a global constant $c_1 > 1$ such that

$$\psi_x(n_1, n_2) \leq c_1 \psi_x(n_1 + 1, n_2)$$

and

$$\psi_x(n_1, n_2) \leq c_1 \psi_x(n_1, n_2 + 1).$$

Now we will ask for small regularization parameters to control the error:

Assumption 2.5 (Error Behavior) *Assume that there exists a monotonically decreasing (with respect to the above half-order) continuous function $\rho : \mathbb{N} \times \mathbb{N} \rightarrow [0, \infty]$ with*

$$d(x_{(n_1, n_2)}^0, x_{(n_1, n_2)}^\delta) \leq \rho(n_1, n_2)^{-1}. \quad (5)$$

Furthermore there exists a global constant $c_2 > 1$ such that

$$\rho(n_1, n_2) \leq c_2 \rho(n_1 + 1, n_2)$$

and

$$\rho(n_1, n_2) \leq c_2 \rho(n_1, n_2 + 1).$$

Using the triangle inequality we get

Corollary 2.6 *For all (n_1, n_2) it holds*

$$d(x, x_{(n_1, n_2)}^\delta) \leq \psi_x(n_1, n_2) + \rho(n_1, n_2)^{-1}$$

We will restrict ourselves to a finite subset $\mathcal{R} = \{1, \dots, N_1\} \times \{1, \dots, N_2\}$ of all possible regularization parameters with $N_1, N_2 \in \mathbb{N}$. Denote accordingly $\mathcal{C}_{(n_1, n_2)}^* = \mathcal{C}_{(n_1, n_2)} \cap \mathcal{R}$. In order to guarantee that we can do proper regularization in this set we will assume

Assumption 2.7 (Consistency)

- $\psi_x(1, 1)\rho(1, 1) > 1$
- For all $n_1 \leq N_1$ and $n_2 \leq N_2$ it holds that $\psi_x(n_1, N_2)\rho(n_1, N_2) < 1$ and $\psi_x(N_1, n_2)\rho(N_1, n_2) < 1$

This assumption is more of a technical nature and basically assures that all locally optimal elements defined below are in the considered domain \mathcal{R} .

Definition 2.8 (Local Optimality) *A pair of regularization parameters (n_1, n_2) is called locally optimal if*

- $\psi_x(n_1, n_2)\rho(n_1, n_2) \leq 1$
- Either $\psi_x(n_1 - 1, n_2)\rho(n_1 - 1, n_2) > 1$ or $\psi_x(n_1, n_2 - 1)\rho(n_1, n_2 - 1) > 1$.

The set of all locally optimal regularization parameters is called \mathcal{N}_{opt} .

Remark 2.9 *As for locally optimal elements $\psi_x(n_1, n_2) \leq \rho(n_1, n_2)^{-1}$ it is sufficient to express rates in terms of $\rho(n_1, n_2)^{-1}$.*

It is hard to prove optimality results with respect to $d(x, x_{(n_1, n_2)}^\delta)$ because we do not have access to necessary lower bounds. Therefore in view of Corollary 2.6 we will consider the sum $\psi_x(n_1, n_2) + \rho(n_1, n_2)^{-1}$ and try to optimize with respect to this one, which is actually a standard procedure [9, 11].

Definition 2.10 (Optimality) An optimal element (o_1, o_2) shall be defined as

$$(o_1, o_2) = \underset{(n_1, n_2) \in \mathcal{R}}{\operatorname{argmin}} \{ \psi_x(n_1, n_2) + \rho(n_1, n_2)^{-1} \}.$$

If there are several arguments minimizing the above expression, we can choose an arbitrary one of them. Now we will show the optimality of the set \mathcal{N}_{opt} . In comparison to the one parameter case a major step is asserting the comparability of regularization parameters.

Lemma 2.11 The set \mathcal{N}_{opt} fulfills the following properties:

- $\mathcal{N}_{opt} \cap \mathcal{R} = \mathcal{N}_{opt}$
- $\mathcal{R} = \bigcup_{(n_1, n_2) \in \mathcal{N}_{opt}} \mathcal{C}_{(n_1, n_2)}^*$
- There exists a global constant c_{opt} such that for all pairs (m_1, m_2) we have

$$\psi_x(m_1, m_2) + \rho(m_1, m_2)^{-1} \geq c_{opt} \rho(n_1, n_2)^{-1}$$

for all $(n_1, n_2) \in \mathcal{N}_{opt} \cap \mathcal{C}_{(m_1, m_2)}^*$.

Proof

As for all $n_1 \leq N_1$ and $n_2 \leq N_2$ it holds that $\psi_x(n_1, N_2) \rho(n_1, N_2) < 1$ and $\psi_x(N_1, n_2) \rho(N_1, n_2) < 1$ (see Assumption 2.7) and on the other hand for every $(n_1, n_2) \in \mathcal{N}_{opt}$ there is no $(m_1, m_2) \leq (n_1, n_2)$ such that $\psi_x(m_1, m_2) \rho(m_1, m_2) \leq 1$ (see Definition 2.8) we immediately get $\mathcal{N}_{opt} \cap \mathcal{R} = \mathcal{N}_{opt}$.

For the second point first take $(m_1, m_2) \in \mathcal{R}$. Now there exists (normally more than one) a consecutive chain $\{(n_{k,1}, n_{k,2})\}$ of regularization parameters from $(1, 1) = (n_{1,1}, n_{1,2})$ to $(N_1, N_2) = (n_{N_1+N_2,1}, n_{N_1+N_2,2})$ which has no holes and contains (m_1, m_2) . I.e. we have $(n_{k,1}, n_{k,2}) \leq (n_{k+1,1}, n_{k+1,2})$ for all k and either $n_{k,1} = n_{k+1,1}$ and $n_{k,2} + 1 = n_{k+1,2}$ or $n_{k,1} + 1 = n_{k+1,1}$ and $n_{k,2} = n_{k+1,2}$.

Because of Assumption 2.7 and the monotonicity of $\psi_x \rho$ there exists an element $(n_{k,1}, n_{k,2})$ and its successor $(n_{k+1,1}, n_{k+1,2})$ such that

$$\psi_x(n_{k,1}, n_{k,2}) \rho(n_{k,1}, n_{k,2}) > 1 \geq \psi_x(n_{k+1,1}, n_{k+1,2}) \rho(n_{k+1,1}, n_{k+1,2}).$$

By definition $(n_{k+1,1}, n_{k+1,2}) \in \mathcal{N}_{opt}$ and therefore $(m_1, m_2) \in \mathcal{C}_{(n_{k+1,1}, n_{k+1,2})}^* \subset \bigcup_{(n_1, n_2) \in \mathcal{N}_{opt}} \mathcal{C}_{(n_1, n_2)}^*$

Now we consider the third point. Consider first that $\psi_x(m_1, m_2) \rho(m_1, m_2) > 1$. Then by the second point there exists at least one $(n_1, n_2) \in \mathcal{N}_{opt}$ with $(m_1, m_2) \leq (n_1, n_2)$. Take one of them. Let (\bar{n}_1, \bar{n}_2) be one of the regularization parameters $(n_1 - 1, n_2)$ or $(n_1, n_2 - 1)$ (at least one has to exist) which fulfills $\psi_x(\bar{n}_1, \bar{n}_2) \rho(\bar{n}_1, \bar{n}_2) > 1$. Since $(m_1, m_2) \leq (n_1, n_2)$, $(\bar{n}_1, \bar{n}_2) \leq (n_1, n_2)$ and assumption 2.4 it holds that

$$\begin{aligned} \psi_x(m_1, m_2) + \rho(m_1, m_2)^{-1} &\geq \psi_x(m_1, m_2) \geq \psi_x(n_1, n_2) \geq c_1^{-1} \psi_x(\bar{n}_1, \bar{n}_2) \\ &\geq c_1^{-1} \rho(\bar{n}_1, \bar{n}_2)^{-1} \geq c_1^{-1} c_2^{-1} \rho(n_1, n_2)^{-1}. \end{aligned}$$

Otherwise consider $\psi_x(m_1, m_2)\rho(m_1, m_2) \leq 1$. Then by the second point there exists at least one $(n_1, n_2) \in \mathcal{N}_{opt}$ with $(m_1, m_2) \geq (n_1, n_2)$. Take one of them. Let (\bar{n}_1, \bar{n}_2) one of the regularization parameters $(n_1 - 1, n_2)$ respectively $(n_1, n_2 - 1)$ which fulfills $\psi_x(\bar{n}_1, \bar{n}_2)\rho(\bar{n}_1, \bar{n}_2) > 1$. It holds that

$$\psi_x(m_1, m_2) + \rho(m_1, m_2)^{-1} \geq \rho(m_1, m_2)^{-1} \geq \rho(n_1, n_2)^{-1}.$$

By choosing $c_{opt} = \min \{c_1^{-1}c_2^{-1}, 1\}$ we get the proposition. □

Remark 2.12 *The last statement in the above lemma can be interpreted in a way that whatever we do, the solutions corresponding to the set \mathcal{N}_{opt} are at most by a factor c_{opt} away from the other solutions with respect to comparable regularization parameters.*

However, it is possible that all but one of the regularization parameters are very far away from the optimal solution.

Corollary 2.13 *There exists a global constant c_{opt} such that for the optimal regularization parameter (o_1, o_2) there is a pair $(n_1, n_2) \in \mathcal{N}_{opt} \cap \mathcal{C}_{(o_1, o_2)}^*$ with*

$$c_{opt}\rho(n_1, n_2)^{-1} \leq \psi_x(o_1, o_2) + \rho(o_1, o_2)^{-1} \leq 2\rho(n_1, n_2)^{-1} \tag{6}$$

Proof

The first inequality was a part of the last theorem, the second is a direct consequence from Definitions 2.8 and 2.10. □

Now we will arrive at the key part of our algorithm. A motivation in the one parameter choice regime can be found in [9].

Definition 2.14 (Balancing Optimality) *Set*

$$b(m_1, m_2) = \max_{(m_1, m_2) \leq (l_1, l_2) \in \mathcal{R}} \frac{1}{4}d(x_{(m_1, m_2)}^\delta, x_{(l_1, l_2)}^\delta) \rho(l_1, l_2) \tag{7}$$

and

$$B(n_1, n_2) = \max_{(n_1, n_2) \leq (m_1, m_2) \in \mathcal{R}} b(m_1, m_2). \tag{8}$$

A pair of regularization parameters (n_1, n_2) is called locally balancing optimal w.r.t. $\kappa \geq 1$ if

- $B(n_1, n_2) \leq \kappa$
- *At least one of the following properties holds $B(n_1 - 1, n_2) > \kappa$ or $B(n_1, n_2 - 1) > \kappa$.*

The set of all balancing optimal regularization parameters is called \mathcal{B}_κ . The set $\widetilde{\mathcal{B}}_\kappa$ is defined as \mathcal{B}_κ with all elements removed for which there exists a smaller element in \mathcal{B}_κ , i.e.

$$\widetilde{\mathcal{B}}_\kappa = \{(n_1, n_2) \in \mathcal{B}_\kappa \text{ s.t. } \nexists (m_1, m_2) \in \mathcal{B}_\kappa \text{ fulfilling } (m_1, m_2) < (n_1, n_2)\} \tag{9}$$

Pictures of examples of the quantity B will be displayed in the numerics section (Figures 2 and 4).

Assumption 2.15 (Consistency) *Assume that $B(1, 1) > \kappa$.*

This will assure that \mathcal{B}_κ is non-empty. Next we will show in two steps the optimality of the balancing optimal regularization parameters. Again comparability is a key issue.

Lemma 2.16 *For all $(n_1, n_2) \in \mathcal{B}_\kappa$ it holds*

$$d(x, x_{(n_1, n_2)}^\delta) \leq 6\kappa\rho(m_1, m_2)^{-1} \quad (10)$$

for all $(m_1, m_2) \in \mathcal{N}_{opt} \cap \mathcal{C}_{(n_1, n_2)}^*$.

Proof

Consider $(m_1, m_2) \in \mathcal{N}_{opt} \cap \mathcal{C}_{(n_1, n_2)}^*$, where $(n_1, n_2) \in \mathcal{B}_\kappa$ and $(m_1, m_2) \leq (k_1, k_2) \leq (l_1, l_2) \in \mathcal{R}$. Since ψ_x and ρ are monotonically decreasing and $\kappa \geq 1$ it holds

$$\begin{aligned} d(x_{(k_1, k_2)}^\delta, x_{(l_1, l_2)}^\delta) &\leq \psi_x(k_1, k_2) + \rho(k_1, k_2)^{-1} + \psi_x(l_1, l_2) + \rho(l_1, l_2)^{-1} \\ &\leq 4\kappa\rho(l_1, l_2)^{-1}. \end{aligned}$$

Hence for all $(k_1, k_2) \geq (m_1, m_2)$ by (7) we have that $b(k_1, k_2) \leq \kappa$. From the definition (8) it follows that $B(n_1, n_2) \leq \kappa$. Because of the monotonicity of B we obtain that $(n_1, n_2) \leq (m_1, m_2)$ and therefore $\psi_x(n_1, n_2)\rho(n_1, n_2) \geq 1$.

Using Definition 2.8 we have that for every $(n_1, n_2) \in \mathcal{B}_\kappa$ it holds for every $(m_1, m_2) \in \mathcal{N}_{opt} \cap \mathcal{C}_{(n_1, n_2)}^*$ that $(n_1, n_2) \leq (m_1, m_2)$. Due to Lemma 2.11 (second statement) there is at least one such (m_1, m_2) . Now

$$\begin{aligned} d(x, x_{(n_1, n_2)}^\delta) &\leq d(x, x_{(m_1, m_2)}^\delta) + d(x_{(m_1, m_2)}^\delta, x_{(n_1, n_2)}^\delta) \\ &\leq 2\rho(m_1, m_2)^{-1} + 4\kappa\rho(m_1, m_2)^{-1} \\ &\leq (2 + 4\kappa)\rho(m_1, m_2)^{-1} \leq 6\kappa\rho(m_1, m_2)^{-1} \end{aligned}$$

which proves the proposition. \square

Remark 2.17 *The constant κ is in principle just necessary if we deal with a non-standard noise model (see [9]) or just have a very rough access to ρ (i.e. overestimating it with respect to reality).*

Lemma 2.18 *For every $(m_1, m_2) \in \mathcal{N}_{opt}$ there exists $(n_1, n_2) \in \mathcal{B}_\kappa \cap \mathcal{C}_{(m_1, m_2)}^*$.*

Proof

Let $(m_1, m_2) \in \mathcal{N}_{opt}$. In the previous proof we have shown that $B(m_1, m_2) \leq \kappa$. Now we construct a sequence $\{(n_{k,1}, n_{k,2})\}$ in the following way.

Set $(m_1, m_2) = (n_{1,1}, n_{1,2})$ and repeat the following argument: Assume that $B(n_{k,1}, n_{k,2}) \leq \kappa$. Then two cases can happen. In the first $B(n_{k,1} - 1, n_{k,2}) > \kappa$ or $B(n_{k,1}, n_{k,2} - 1) > \kappa$, i.e. $(n_{k,1}, n_{k,2}) \in \mathcal{B}_\kappa$. Otherwise set $(n_{k+1,1}, n_{k+1,2})$ to either $(n_{k,1} - 1, n_{k,2})$ or $(n_{k,1}, n_{k,2} - 1)$, whatever is possible.

This procedure has to stop because $B(1, 1) > \kappa$. Set (n_1, n_2) to the last element of the chain constructed beforehand. \square

Corollary 2.19 For every $(m_1, m_2) \in \mathcal{N}_{opt}$ there exists $(n_1, n_2) \in \widetilde{\mathcal{B}}_\kappa \cap \mathcal{C}_{(m_1, m_2)}^*$.

Corollary 2.20 There is an $(n_1, n_2) \in \widetilde{\mathcal{B}}_\kappa$ and a global constant c such that

$$d(x, x_{(n_1, n_2)}^\delta) \leq c(\psi_x(o_1, o_2) + \rho(o_1, o_2)^{-1}). \quad (11)$$

Proof

Immediate consequence from Corollary 2.13, estimate (10) and Corollary 2.19. □

Remark 2.21 Looking into the recent article [10] we see that Assumptions 2.4, 2.5, 2.7 are stronger than we actually need if we just want to prove the balancing principle.

It would be sufficient that there exists the set \mathcal{N}_{opt} and for all $(n_1, n_2) \in \mathcal{N}_{opt} \cap \mathcal{R}$ there exists a descending function ρ such that

$$d(x, x_{(n_1, n_2)}^\delta) \leq 2\rho(n_1, n_2)^{-1}.$$

The corresponding proofs will stay the same, however one will lose the intuitive side of the algorithm and some interesting bits depending on ψ_x . Therefore we have presented the above version.

The problem we face right now is to choose an appropriate $(n_1, n_2) \in \widetilde{\mathcal{B}}_\kappa$. However, without using additional information it is impossible to do this.

In reality choosing $(n_1, n_2) \in \widetilde{\mathcal{B}}_\kappa$ as

$$(n_1^*, n_2^*) = \underset{(n_1, n_2) \in \widetilde{\mathcal{B}}_\kappa}{\operatorname{argmin}} \{ \rho(n_1, n_2)^{-1} \} \quad (12)$$

proved to be successful and is in principle sensible if the real noise behavior is not too far away from its worst case description ρ . However, it is important to keep in mind that this is purely heuristic and could yield the worst of the best possible rates.

3. Choosing ρ

In order to apply the proposed procedure (see Definition 2.14) we need to get hold of ρ . One possibility which is rarely observed in practice is that we know ρ by some means in advance. Right now we will propose a data driven way to observe this quantity in practice.

Consider the K independent input data sets $y^{k, \delta}$ with corresponding regularized solutions $x_{(n_1, n_2)}^{k, \delta}$. Then we have

$$d(x_{(n_1, n_2)}^{k, \delta}, x_{(n_1, n_2)}^{l, \delta}) \leq d(x_{(n_1, n_2)}^0, x_{(n_1, n_2)}^{l, \delta}) + d(x_{(n_1, n_2)}^{k, \delta}, x_{(n_1, n_2)}^0) \leq 2\rho(n_1, n_2)^{-1}.$$

Under strong assumptions about the measurements one can show in a Hilbert space setting that the above inequality gets an equality (without the factor of 2) in expectation [12]. Therefore we set ρ to the empirical expectation of the pairwise distances:

$$\rho(n_1, n_2)^{-1} = \frac{4}{K(K+1)} \sum_{1 \leq k < l \leq K} d(x_{(n_1, n_2)}^{k, \delta}, x_{(n_1, n_2)}^{l, \delta}). \quad (13)$$

However we would like to point out that again this is just a heuristics which works surprisingly well in practice. For a more detailed account on possible noise models and possibilities to obtain ρ we refer to [13].

4. Numerics

In order to show that our method works we will consider two examples. The first is an artificially perturbed matrix inversion where we have control over all quantities and the second is originating from crack reconstructing by electrostatic imaging methods.

4.1. Matrix inversion

We want to solve a system of linear equations

$$Ax = y^\delta, \quad (14)$$

where A is a 200×200 diagonal matrix, with elements $a_{ii} = i^{-3}$ and y^δ being the perturbed data. The noise is taken to be Gaussian with mean zero and standard deviation σ equal to 0.00003% of $\|y\|$. The solution x is chosen to be a random vector with decay behavior like $n^{-3/2}$.

We want that the regularization operator depends in a way from the regularization parameters such that there is no possibility to solve the problem separately. Therefore we choose to regularize in the following way:

$$x_{\alpha,\beta}^\delta = (A + \alpha\beta^{0.6}I + \beta\alpha^{0.3}C)^{-1}y^\delta.$$

Here I is an identity matrix and C is a diagonal matrix with elements $c_{kk} = k^{-1}$.

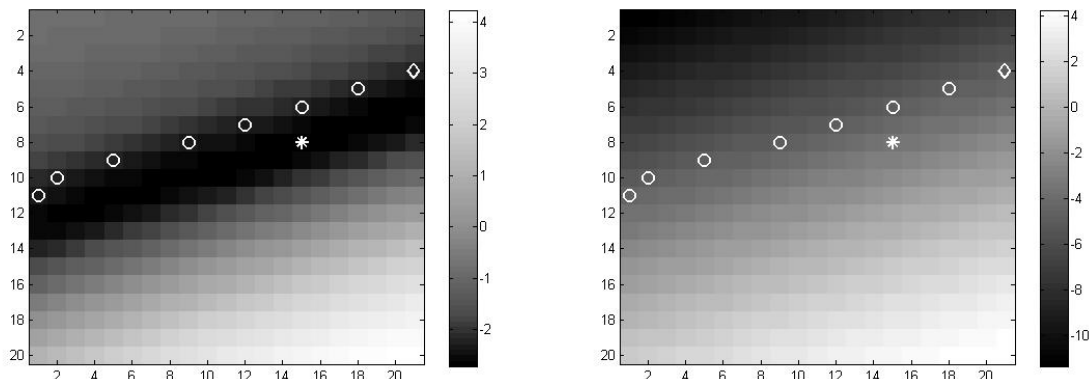


Figure 1. Left: exact noise behavior. Right: propagated noise behavior.

In order to find the regularization parameters by the balancing principle we sample them exponentially according to

$$\alpha = 100 \cdot 2^{-n_1}, \quad \beta = 0.01 \cdot 2^{-n_2}, \quad n_1 \in \{1, \dots, 20\}, \quad n_2 \in \{1, \dots, 21\}.$$

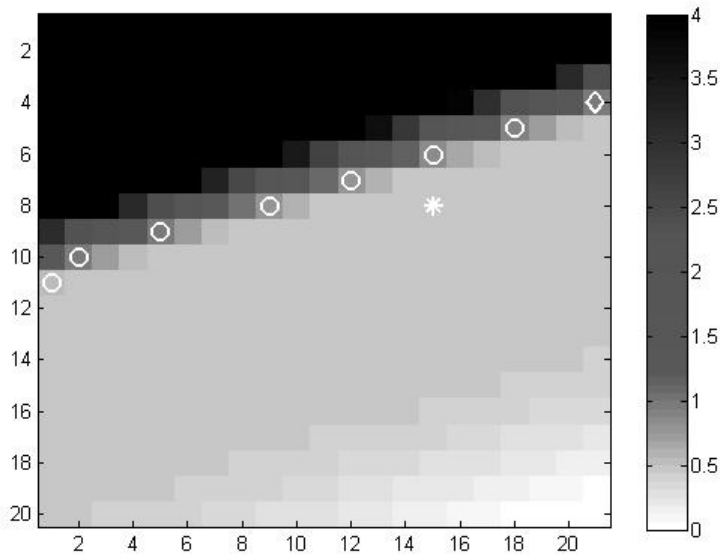


Figure 2. Balancing functional

We make use of the formula (13) for ρ to estimate the noise behavior. In Figure 1. exact and estimated measurement noise w.r.t. to the regularization are displayed. They coincide very well for all $(n_1, n_2) > \mathcal{N}_{opt}$. On the x-axis we find n_1 and on the y-axis n_2 ; the values of the displayed functions are color coded.

The corresponding balancing functional is shown in Fig. 2. Since we have access to the exact solution of (14) we can find a pair of regularization parameters, which is a minimizer to $d(x_{\alpha,\beta}^\delta, x)$ and we will denote it by a star in each figure. The pairs of regularization parameters suggested by the balancing principle with $\kappa = 1$ are marked by circles, one of them chosen by the heuristics (12) is denoted by a diamond.

We observed that the absolute error in the optimal point is 0.0658 and the maximum error in one of the chosen points is 0.1099, i.e. we just loose a factor of 1.6702.

4.2. Crack determination

Further, we have tested this method on a crack determination problem, which occurs in nondestructive testing and evaluation.

Let D be a bounded domain in \mathbb{R}^2 with a smooth boundary Γ_1 . The direct problem for a domain with a perfectly conducting crack is finding a solution $u \in H^1(D)$ for a given function $f \in H^{1/2}(\Gamma_1)$, such that

$$\Delta u = 0 \quad \text{in } D \setminus \Gamma_c$$

satisfying the Dirichlet boundary conditions

$$u = 0 \quad \text{on } \Gamma_c, \quad u = f \quad \text{on } \Gamma_1.$$

The inverse problem is determining the crack Γ_c from an imposed voltage f on the outer boundary Γ_1 and measured Neumann data (i.e. the resulting currents)

$$g = \frac{\partial u}{\partial \nu} \quad \text{on } \Gamma_1.$$

To solve the inverse problem we use an algorithm [3], which is based on nonlinear integral equations arising from the reciprocity gap principle. We arrive at a two-by-two system of integral equations for the pair of the unknown crack Γ_c and the jump of the normal derivative $h := [\partial u / \partial \nu]_{\pm}$ across Γ_c .

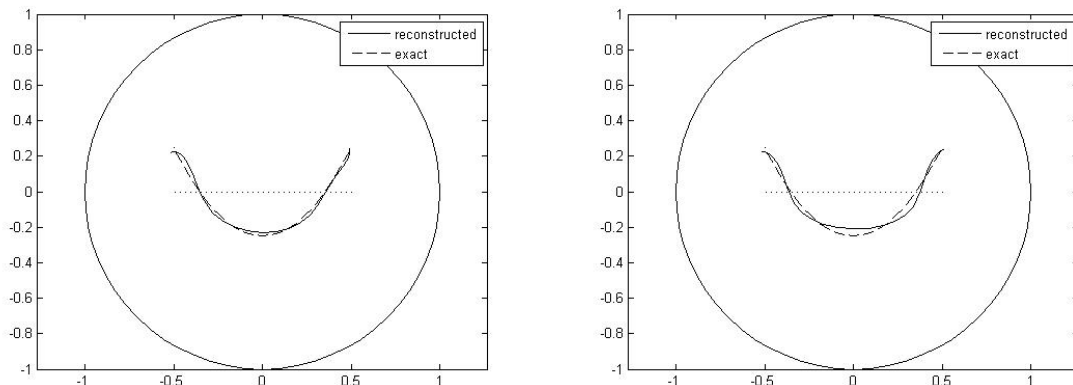


Figure 3. Reconstruction of the crack. Left: optimal regularization. Right: balancing optimal regularization.

For the solution of the integral equations we apply a Newton iteration via linearization of the full system. Due to the ill-posedness of the problem Tikhonov regularization [14, 4] is incorporated with a regularization parameter α for the density h and a regularization parameter β corresponding to the crack Γ_c . For more details we refer to [2].

We have tested the proposed scheme on a parabolic crack in a circular domain. In order to generate reference for the noise data with sufficient accuracy we generated ten different data sets with respect to the same noise level $\delta = 3\%$. For this problem we use a smoothed version of ρ , defined by formulae (13). As regularization parameters we use the set $\{(4^{-n_1-5}, 2^{-n_2-5})\}_{n_1 \in \{1, \dots, 15\}, n_2 \in \{1, \dots, 20\}}$. Due to the high unknown noise components we chose $\kappa = 2$ to be on the safe side.

In Fig. 4 the balancing functional is presented. The absolute error for optimal pair of regularization parameters is 0.1220 and the error for the pair suggested by (12) (marked by diamond) is 0.1856. The maximum absolute error for one of the pairs suggested by balancing principle is 3.2206.

5. Conclusion

We were able to show that the balancing principle for choosing the regularization parameter in multi-parameter problems works both theoretically and numerically. An

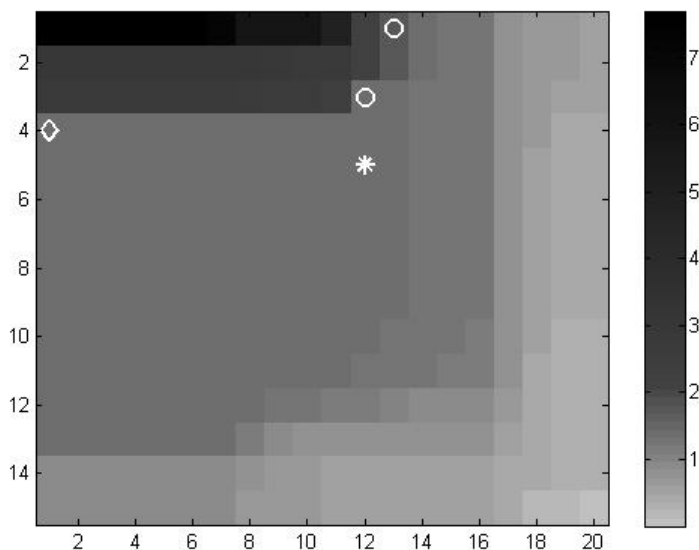


Figure 4. Balancing functional

extension to more than two regularization parameters is tedious but possible.

However, in practical situations we are limited by two problems. First of all we need to sample very many regularization parameters which is time consuming. Furthermore for the much higher number of samples noise seems to play a bigger role. Therefore it appears to be advisable to keep the number of regularization parameter as small as possible.

Up till now we have no possibility to choose a regularization out of the range $\widetilde{\mathcal{B}}_\kappa$ or even \mathcal{N}_{opt} . The only thing we can do is a reliable reduction to a very small number of sensible parameters which still might be an interesting feature in practice.

Acknowledgments

The support by the “Graduiertenkolleg 1023: Identification in mathematical models”, University of Göttingen is gratefully acknowledged. Furthermore the first author would like to thank Prof. Pereverzev (RICAM Linz) for valuable discussions.

References

- [1] Justen L, Ramlau R. A non-iterative regularization approach to blind deconvolution. *Inverse Problems*. 2006;22(3):771–800.
- [2] Ivanyshyn O, Kress R. Nonlinear integral equations for solving inverse boundary value problems for inclusions and cracks. *Jour Integral Equations and Appl*. 2006;p. (to appear).
- [3] Kress R, Rundell W. Nonlinear integral equations and the iterative solution for an inverse boundary value problem. *Inverse Problems*. 2005;21:1207 – 1223.
- [4] Engl H, Hanke M, Neubauer A. *Regularization of Inverse Problems*. Dordrecht, Boston, London: Kluwer Academic Publisher; 1996.

- [5] Bauer F, Pereverzev S. An utilization of a rough approximation of a noise covariance within the framework of multi-parameter regularization. *International Journal of Tomography & Statistics*. 2006;4:(1–12).
- [6] Aliev B. Two-parameter regularization method for finding L -pseudosolutions. *Vestnik Moskov Univ Ser XV Vychisl Mat Kibernet*. 1986;(2):45–50.
- [7] Brezinski C, Redivo-Zaglia M, Rodriguez G, Seatzu S. Multi-parameter regularization techniques for ill-conditioned linear systems. *Numer Math*. 2003;94(2):203–228.
- [8] Mathé P. The Lepskij principle revisited. *Inverse Problems*. 2006;(22):L11–L15.
- [9] Bauer F, Pereverzev S. Regularization without preliminary knowledge of smoothness and error behaviour. *European J Appl Math*. 2005;16(3):303–317.
- [10] Bauer F, Hohage T. A Lepskij-type stopping rule for regularized Newton methods. *Inverse Problems*. 2005;21(6):1975–1991.
- [11] Mathé P, Pereverzev S. Geometry of Linear Ill-Posed Problems in Variable Hilbert Spaces. *Inverse Problems*. 2003;19(3):789–803.
- [12] Bauer F. An Alternate Approach to the Oblique Derivative Problem in Potential Theory. University of Kaiserslautern, Department of Mathematics, Geomathematics Group; 2004.
- [13] Bauer F, Mathé P, Pereverzev S. Local Solutions to Inverse Problems in Geodesy: The Impact of the Noise Covariance Structure upon the Accuracy of Estimation. *Journal of Geodesy*. to appear;.
- [14] Colton D, Kress R. Inverse acoustic and electromagnetic scattering theory. vol. 93 of Applied Mathematical Sciences. 2nd ed. Berlin: Springer-Verlag; 1998.

Institut für Numerische und Angewandte Mathematik
Universität Göttingen
Lotzestr. 16-18
D - 37083 Göttingen

Telefon: 0551/394512

Telefax: 0551/393944

Email: trapp@math.uni-goettingen.de URL: <http://www.num.math.uni-goettingen.de>

Verzeichnis der erschienenen Preprints:

- | | | |
|---------|---|---|
| 2006-01 | R. Schaback | Limit Problems for Interpolation by Analytic Radial Basis Function |
| 2006-02 | N. Bissantz, T. Hohage, A. Munk, F. Ruymgaart | Convergence rates of general regularization methods for statistical inverse problems and applications |
| 2006-03 | J. Brimberg, H. Juel, A. Schöbel | Locating a circle on a sphere |
| 2006-04 | J. Brimberg, H. Juel, A. Schöbel | Locating a circle on the plane using the minimax criterion |
| 2006-05 | L. Ling, R. Opfer, R. Schaback | Results on Meshless Collocation Techniques |
| 2006-06 | G. Lube, T. Knopp et.al. | Domain decomposition methods for indoor air flow simulation |
| 2006-07 | T. Apel, T. Knopp, G. Lube | Stabilized finite element methods with anisotropic mesh refinement for the Oseen problem |
| 2006-08 | R. Schaback | Recovery of Functions From Weak Data Using Unsymmetric Meshless Kernel-Based Methods |
| 2006-09 | H.W. Hannacher, S. Horn, A. Schöbel | Stop Location Design in Public Transportation Networks: Covering and Accessibility Objective |
| 2006-10 | Q.T. Le Gia, F.J. Narcowich, J.D. Ward, H. Wendland | Continuous and Discrete Least-Squares Approximation by Radial Basis Functions on Spheres |
| 2006-11 | R. Schaback | Unsymmetric Meshless Methods for Operator Equations |
| 2006-12 | P. Giesl, H. Wendland | Meshless Collocation: Error Estimates with Application to Dynamical Systems |
| 2006-13 | H. Wendland | On the stability of meshless symmetric collocation for boundary value problems |
| 2006-14 | G. Lube | Stabilized FEM for incompressible flow. Critical review and new trends |

- 2006-15 J. Puerto, A. Schöbel, S. A Class of Infinite Potential Games
Schwarze
- 2006-16 S. Hein, T. Hohage, W. Koch, Acoustic Resonances in High Lift Configuration
J. Schöberl
- 2006-17 T. Hohage, M.L. Rapún, F.J. Detecting corrosion using thermal
Sayas measurements
- 2006-18 H. Wendland Spatial coupling in aeroelasticity by meshless
kernel-based methods
- 2006-19 F. Bauer, O. Ivanyshyn Optimal Regularization with Two Interdepen-
dent Regularization Parameters