Lecture Notes

Inverse Problems

Winter 2007/08

Martin Burger
Institut für Numerische und Angewandte Mathematik
Westfälische Wilhelms-Universität Münster
martin.burger@uni-muenster.de

Contents

Chapter 1

Introduction

The goal of this lecture is to provide an overview of important techniques used for the analysis, regularization, and numerical solution of *inverse problems*. Generally speaking, inverse problems are concerned with finding causes for an *observed effect* or a *desired effect*. With respect to this aspect one usually divides into the terms

- *Identification* or *reconstruction*, if one looks for the cause for an observed effect.
- *Control* or *design*, if one looks for a possible cause of a desired effect.

Both problems are of course related, but there are also several mathematical consequences due to the different aims. E.g., in an identification problem a desirable property is uniqueness of a solution (*identifiability*), because there is probably a specific cause for the observed effect, which one would like to obtain. In a control or design problem, uniqueness is not really of importance, since non-uniqueness only means that the design goal can be reached by different strategies and hence, one has additional freedom (e.g. to incorporate further design goals).

A common property of a vast majority of inverse problems is their *ill-posedness*. In the sense of Hadamard, a mathematical problem (we can think of an equation or optimization problem) is well-posed if it satisfies the following properties:

1. **Existence:** For all (suitable) data, there exists a solution of the problem (in an appropriate sense).
2. **Uniqueness:** For all (suitable) data, the solution is unique.
3. **Stability:** The solution depends continuously on the data.

According to this definition, a problem is ill-posed if one of these three conditions is violated. However, in general we shall be concerned in particular with problems violating the third condition, i.e., the solution does not depend on the data in a stable way.

The prototype of inverse problem will be an equation of the form

$$F(x) = y, \tag{1.1}$$

with a function space setting to be specified below. For such an equation, the unknown is x and the data are usually the right-hand side y. If the stability condition is violated, the numerical solution of the inverse problem by standard methods is difficult and often yields

instability, even if the data are exact (since any numerical method has internal errors acting like noise). Therefore, special techniques, so-called *regularization methods* have to be used in order to obtain a stable approximation of the solution. Since the appropriate construction and analysis of regularization methods and subsequently (or simultaneously) of numerical schemes is the major issue in the solution of inverse problems, most of the lecture will be concerned with this task.

Finally, we notice that the nomenclature "inverse problem" somehow indicates the existence of a "direct problem" or "forward problem". This is not always true, but in most cases such a direct problem, which is well-posed, exists. For example, the direct problem could be to simply evaluate an integral operator of the first kind, and the corresponding inverse problem is to solve an integral equation involving this operator.

Inverse problems is a very active field of research in applied sciences, with a fast growing bibliography. Throughout the lecture notes we shall refer to various papers and monographs including further details on several aspects. As general references on inverse problems, and also as sources for contents in this lecture we refer to the monographs by Engl, Hanke, Neubauer [10], Kirsch [17], and Vogel [23], the latter focusing on computational methods.

Chapter 2

Examples of Inverse Problems

In the following we shall discuss some motivating examples of inverse problems, the first two of them being rather simple and therefore allowing a detailed insight into typical features of inverse problems. The other problems rather motivate the practical importance of inverse problems, but due to their complicated mathematical structure we will spend less on their analysis at this point.

2.1 Differentiation of Data

One of the simplest ill-posed problems is (numerical) differentiation of noisy functions, a task one faces in many applications. Assume that we want to compute the derivative of a function which includes additive noise, i.e., instead of the exact function f we are only given the function f^δ with

$$f^\delta(x) = f(x) + n^\delta(x), \qquad x \in [0,1]$$

and $f^\delta(0) = f(0) = 0$, $f^\delta(1) = f(1) = 0$, where $n^\delta(x)$ represents the data noise. In many typical measurement devices, the noise at each point x ($n^\delta(x)$) can be modeled as a normal distribution with mean zero and variance $\delta > 0$, being independent at different measurement points x_1 and x_2. From the law of large numbers one may expect that

$$\int_0^1 |n^\delta(x)|^2 \, dx \approx \delta^2,$$

i.e., one obtains some information of the noise. However, even if we know exactly that

$$\int_0^1 |n^\delta(x)|^2 \, dx = \delta^2$$

and δ is arbitrarily small, we cannot obtain any estimate on the derivative $\frac{df}{dx}$. In the worst case, the noise n^δ is not differentiable, so that one cannot even compute a derivative. However, even if we assume that the noise is differentiable (even analytic) the error in the derivative can be arbitrarily large. Take for example

$$n^\delta(x) = \sqrt{2}\delta \sin(2\pi k x)$$

for some $k \in \mathbb{N}$. Then, $\int_0^1 |n^\delta(x)|^2 \, dx = \delta^2$ and

$$\frac{df^\delta}{dx}(x) = \frac{df}{dx}(x) + \sqrt{2}2\pi\delta k \cos(2k\pi x)$$

Now note that k can be arbitrarily large and therefore δk can be arbitrarily large. Hence, the L^2-error

$$\left(\int_0^1 \left(\frac{df^\delta}{dx}(x) - \frac{df}{dx}(x) \right)^2 \, dx \right)^{1/2} = 2\pi\delta k$$

or the L^∞-error

$$\sup_{x\in[0,1]} \left| \frac{df^\delta}{dx}(x) - \frac{df}{dx}(x) \right| = \sqrt{2}2\pi\delta k$$

can be arbitrarily large. This statement holds true in general for ill-posed problems (and could actually be used as a definition):

> *Without regularization and without further information, the error between the exact and noisy solution can be arbitrarily large, even if the noise is arbitrarily small.*

How can additional information that helps to bound the error, look like ? Of course, one could assume that the noise is bounded in a stronger norm, e.g.,

$$\int_0^1 \left(\frac{dn^\delta}{dx}(x) \right)^2 \, dx \leq \delta^2.$$

In this case, we would obtain in a trivial way the error estimate

$$\left(\int_0^1 \left(\frac{df^\delta}{dx}(x) - \frac{df}{dx}(x) \right)^2 \, dx \right)^{1/2} \leq \delta,$$

but our result does not correspond to the practical applications, where we can hardly get an estimate for $\frac{dn^\delta}{dx}$. Thus, it seems not a good idea to assume stronger bounds on the noise.

A more realistic alternative is to assume further regularity of the solution f, e.g., $f \in C^2([0,1])$. The error is then still arbitrarily large for the original problem, but can be estimated if regularization is used. As a simple example we could smooth the data by solving

$$-\alpha \frac{d^2 f_\alpha}{dx^2}(x) + f_\alpha(x) = f^\delta(x), \qquad f_\alpha(0) = f_\alpha(1) = 0,$$

which is also equivalent to applying the associated Green operator (an integral operator with smooth kernel) to f^δ. We shall see later that this approach can be identified with so-called *Tikhonov regularization*. Note that due to the standard variational interpretation of elliptic differential operators, this smoothing is also equivalent to mimimizing the functional

$$H_\alpha(f_\alpha) = \int_0^1 (f_\alpha(x) - f^\delta(x))^2 \, dx + \alpha \int_0^1 \left(\frac{df_\alpha}{dx}(x) \right)^2 \, dx,$$

i.e., we perform a least-squares fit with a penalty term that enforces $\frac{df_\alpha}{dx}$ to be bounded. Then we have

$$-\alpha \frac{d^2}{dx^2}(f_\alpha(x) - f(x)) + (f_\alpha(x) - f(x)) = (f^\delta(x) - f(x)) + \alpha \frac{d^2 f}{dx^2}(x)$$

and multiplication by $f_\alpha(x) - f(x)$ and integration with respect to x yields

$$\int_0^1 \left[\alpha\left(\frac{df_\alpha}{dx} - \frac{df}{dx}\right)^2 + (f_\alpha(x) - f(x))^2\right] dx = \int_0^1 \left(f^\delta(x) - f(x) + \alpha\frac{d^2 f}{dx}\right)(f_\alpha(x) - f(x))\ dx$$

where we have used integration by parts for the first term. By applying the Cauchy-Schwarz inequality to the right-hand side we further obtain

$$\begin{aligned}\int_0^1 \left[\alpha\left(\frac{df_\alpha}{dx} - \frac{df}{dx}\right)^2 + \frac{1}{2}(f_\alpha(x) - f(x))^2\right] dx &\leq \int_0^1 (f^\delta(x) - f(x))^2\ dx + \alpha^2 C^2 \\ &\leq \delta^2 + \alpha^2 C^2,\end{aligned}$$

where $C = \|f\|_{C^2}$. Thus, we may conclude in particular

$$\int_0^1 \left(\frac{df_\alpha}{dx} - \frac{df}{dx}\right)^2 dx \leq \frac{\delta^2}{\alpha} + \alpha C^2,$$

i.e., we obtain a bound on the error in terms of α and δ. The obvious next question is the choice of the regularization parameter: *How to choose α such that the error in the solution is minimal ?* In general it will not be possible to really minimize the error, but with an estimate like the one above we can at least minimize the right-hand side, which happens for $\alpha = \frac{\delta}{C}$ and the error estimate takes the form

$$\int_0^1 \left(\frac{df_\alpha}{dx} - \frac{df}{dx}\right)^2 dx \leq 2\delta.$$

If we take the square root in this estimat to obtain the norm on the left-hand side, the error is $\sqrt{2\delta}$, i.e., of order $\sqrt{\delta}$ and hence, much larger than the data error δ. This is another typical effect for ill-posed problems: *Even with regularization, we can never achieve an error in the reconstruction which is as slow as the error in the data.* Note also that the error bound $\sqrt{2\delta}$ was only achieved for $f \in C^2([0,1])$. If we only assume that $f \in C^1([0,1])$, which seems actually much more natural for differentiating once, we would need to estimate alternatively

$$\begin{aligned}\int_0^1 \alpha\frac{d^2 f}{dx}(f_\alpha(x) - f(x))\ dx &= -\alpha\int_0^1 \alpha\frac{df}{dx}(\frac{df_\alpha}{dx}(x) - \frac{df}{dx}(x))\ dx \\ &\leq \frac{\alpha}{2}\int_0^1 \left(\frac{df_\alpha}{dx} - \frac{df}{dx}\right)^2 dx + \frac{\alpha}{2}\int_0^1 \left(\frac{df}{dx}\right)^2 dx.\end{aligned}$$

Since the second integral can be estimated by $\frac{\alpha}{2}C^2$ with $C = \|f\|_{C^1}$ the final estimate becomes

$$\int_0^1 \left(\frac{df_\alpha}{dx} - \frac{df}{dx}\right)^2 dx \leq 2\frac{\delta^2}{\alpha} + C^2,$$

and the right-hand side is larger than C no matter how we choose α. As we shall see later, one can show by different arguments that $\int_0^1 \left(\frac{df_\alpha}{dx} - \frac{df}{dx}\right)^2 dx \to 0$, but this convergence is

arbitrarily slow, another general statement for ill-posed problems: *Without additional smoothness assumptions on the exact solution, the convergence of regularized solutions is arbitrarily small.*

Above we have motivated inverse problems as the inversion of some kind of direct problem. For numerical differentiation, we have started with the inverse problem immediately. However, the direct problem can easily be obtained by integration. E.g., if $f(0) = 0$, then the direct problem is given by the integral equation of the first kind

$$f^\delta(x) = \int_0^x \frac{df}{dx}(y)\ dy.$$

This integral operator can be shown to be compact and we will see later that the inversion of a compact linear operator is always an ill-posed problem.

We finally mention that analogous reasoning can be applied to numerical differentiation of sampled data of a function f, e.g. by one-sided or central finite difference schemes. In this case, the difference scheme has the effect of a regularization method and the grid size h plays the role of a regularization parameter. A detailed analysis can be found in [10, 14].

2.2 Computerized Tomography: Radon Inversion

An inverse problem that became of high practical importance almost fourty years ago due to the invention of X–ray tomography is the inversion of the Radon transform of a function $f : \mathbb{R}^2 \to \mathbb{R}$. The Radon transform is defined as

$$\mathcal{R}f(s, w) = \int_{\mathbb{R}} f(sw + tw^\perp)\ dt, \qquad w \in \mathbb{R}^2, |w| = 1, s \in \mathbb{R}_+,$$

where $w^\perp$ denotes an orthogonal vector to w.

In X–ray tomography, the function f represents the spatially varying density in a domain $D \subset \mathbb{R}^2$, which might be the cross-section of a human body or of some material to be tested in a nondestructive way. The inverse problem consists in recovering the density f from X-ray measurements in a plane covering D. These X–rays travel along different lines, each ray is parametrized by its distance $s > 0$ from the origin, and by its unit normal vector $w \in \mathbb{R}^2$ ($|w| = 1$). The basic modeling assumption is that the decay $-\Delta I$ of the intensity of an X–ray beam along a small distance Δt is proportional to the intensity I itself, the density f, and to Δt. Hence,

$$\frac{\Delta I(sw + tw^\perp)}{\Delta t} = -I(sw + tw^\perp) f(sw + tw^\perp).$$

For $\Delta t \to 0$ we obtain the ordinary differential equation

$$\frac{dI(sw + tw^\perp)}{dt} = -I(sw + tw^\perp) f(sw + tw^\perp).$$

By integrating this differential equation from $t = 0$, the position of the emitter, to $t = L$, the position of the detector, we obtain

$$\ln I(sw + Lw^\perp) - \ln I(sw) = -\int_0^L f(sw + tw^\perp)\ dt.$$

The functions $I_L(s,w) := I(sw+Lw^\perp)$ and $I_0(s,w) := I(sw)$ can be measured at the emitters and detectors for all s and w (the positions of emitters and detectors can be changed), and since f can be extended to be zero for $t \notin (0,L)$ we can reformulate the inverse problem of computerized tomography as the inversion of the Radon-transform

$$\ln I_0(s,w) - \ln I_L(s,w) = \mathcal{R}f(s,w), \qquad w \in \mathbb{R}^2, |w| = 1, s \in \mathbb{R}_+.$$

An interesting special case consists of a radially symmetric density f and D being a circle. In this case it suffices to use a single direction w, e.g., $w_0 = (0,1)$, and moreover $f(sw + tw^\perp) = F(r)$ can be written as a function of the radius $r = \sqrt{s^2+t^2}$. Using a transformation to polar coordinates, the Radon transform can be rewritten as

$$\mathcal{R}f(s,w_0) = 2\int_s^\rho \frac{rF(r)}{\sqrt{r^2-s^2}}\, dr$$

with ρ sufficiently large such that $F(r) = 0$ for $r < \rho$. With the notation $g(s) = -\frac{1}{2}(\ln I_0(s,w_0) - \ln I_L(s,w_0))$, the Radon inversion in this special case can be written as the *Abel integral equation*

$$g(s) = \int_s^\rho \frac{rF(r)}{\sqrt{r^2-s^2}}\, dr, \qquad 0 < s \le \rho.$$

It is possible to find an explicit inversion formula for the Abel integral equation, which yields

$$F(r) = -\frac{2}{\pi}\int_r^\rho \frac{g'(s)}{\sqrt{s^2-r^2}}\, dr.$$

Note that in the inversion formula, the derivative g' appears and we have seen in the previous section that differentiation of data is ill-posed. The differentiation is compensated partly by the additional integration, but one can show that the inversion of the Abel integral equation is ill-posed. For the Radon inversion, an explicit (but more complicated) inversion formula exists, which also involves differentiation of data.

2.3 Image Denoising and Deblurring

Two basic problems in mathematical imaging are *image denoising* and *image deblurring*. In the case of denoising, the data are a noisy version of the original image u,

$$u^\delta(x) = u(x) + n^\delta(x), \qquad x \in \Omega \subset \mathbb{R}^2,$$

with the noise satisfying similar properties as in the first example, in particular a bound of the form

$$\int_\Omega (n^\delta(x))^2\, dx \le \delta^2.$$

The major goal is to compute an approximation of the original image u avoiding oversmoothing and keeping features of particular importance in the image such as e.g. edges. One might argue that denoising is not an ill-posed problem, since the data is the function u itself. The real ill-posedness in denoising is the fact that we want to obtain specific features of the image like edges, which are distorted by the noise. One can formally argue that features like edges are rather related to derivatives of u and hence, denoising is ill-posed in the same way as numerical differentiation.

In the case of deblurring, the image u is first transfered through a channel, which can be modeled via a linear integral operator, and the given output is

$$f^\delta(x) = \int_\Omega k(x,y)u(y)\ dy + n^\delta(x), \qquad x \in \Omega \subset \mathbb{R}^2.$$

The goal is the same as in image denoising, but it is much more difficult to reach, since some information is lost while the image ins transfered through the channel. A typical model for the *point-spread-function* k is a Gaussian, i.e.,

$$k(x,y) = c\exp\left(-\frac{|x-y|^2}{\sigma}\right).$$

With increasing σ, the Gauss kernel becomes broader and the averaging of the image is stronger than for small σ. Again, the direct problem involves an integral operator of the first kind with smooth kernel, which is in particular a compact linear operator and the inverse problem is clearly ill-posed.

2.4 Parameter Identification

By parameter identification one usually denotes the problem of reconstructing unknown coefficients in partial differential equations from indirect measurements of the solution. A simple example is the following model from *groundwater filtration*, which is modeled through the elliptic equation

$$-\text{div}\ (a\nabla u) = f, \qquad \text{in } \Omega \subset \mathbb{R}^d,$$

where u is the unkown, f a given source, and a the hydraulic permittivity. The direct problem consists in solving the partial differential equation for u given a and suitable boundary conditions on $\partial\Omega$. The inverse problem consists in reconstructing the unknown function a on Ω given a noisy measurement

$$u^\delta(x) = u(x) + n^\delta(x), \qquad x \in \Omega,$$

of the solution.

If the solution of the direct problem is unique for each parameter a, which is the case for the groundwater filtration problem with appropriate boundary conditions, then one can introduce the *parameter-to-solution map* $a \mapsto u_a$, where u_a is the solution of the direct problem given a specific a. Note that even if the direct problem is linear (for u), the inverse problem and the parameter-to-output map are usually nonlinear. E.g., in the groundwater filtration problem we have $u_{2a} = \frac{1}{2}u_a$, i.e., $u_{2a} \neq 2u_a$ and hence, the problem is not linear.

The uniqueness question for parameter identification problems is usually denoted as *identifiability*. In the case $\Omega = [0,1]$ with boundary conditions $\frac{du}{dx}(0) = 0$ and $u(1) = 0$ we can easily answer the question by integrating the equation with respect to x, which yields the formula

$$a(x)\frac{du}{dx}(x) = \int_0^x f(y)\ dy.$$

Hence, the parameter a is determined uniquely for every x such that $\frac{du}{dx}(x) \neq 0$. On the other hand, there are many realistic assumptions on f, which guarantee that $\frac{du}{dx} \neq 0$ almost everywhere. For example, if the antiderivative of f is positive in $(0,1)$, then the above formula

shows that $\frac{du}{dx} \neq 0$. Another possible assumption is $f(x) \neq 0$ for almost every x, then $\frac{du}{dx}$ cannot vanish on an open interval $I \subset [0,1]$, since otherwise

$$0 = \frac{d0}{dx} = \frac{d}{dx}\left(a(x)\frac{du}{dx}(x)\right) = f(x), \qquad x \in I$$

yields a contradiction. On the other hand, if $f \equiv 0$, then $u \equiv 0$ and $\frac{du}{dx} \equiv 0$ for any a and it is never possible to reconstruct the parameter. The choice of f or in reality the action leading to the source f is a matter of *design of experiments*, one could even ask the question what is the best source with respect to stable reconstruction of the parameter.

The solution formula

$$a(x) = \frac{\int_0^x f(y)\ dy}{\frac{du}{dx}(x)}$$

also shows that besides the usual *linear ill-posedness* arising from the fact that data (u) have to be differentiated, there is a *nonlinear ill-posedness* from the quotient, whose consequence is that errors at small values of $\frac{du}{dx}$ are amplified much stronger than errors at large values of $\frac{du}{dx}$. I.e., if $\frac{du}{dx}(x)$ is very small in an interval I, then we still have identifiability, but in practice we must expect very high errors due to the noise amplification.

Another interesting issue in parameter identification problems are stability estimates, which concerns the continuity of the inverse operator on special subsets. Note that for an ill-posed problem, the inverse operator (if it exists) is not continuous, but it is continuous on compact subsets of its domain. As an example we consider the compact subset

$$\mathcal{C}_{\gamma,M} = \{\ u \in C^2([0,1]) \mid\ , \|u\|_{C^2} \leq M, \frac{du}{dx} \geq \gamma > 0 \text{ in } [0,1]\ \}.$$

Let u_j be the solution of the forward problem for given parameter a_j, $j = 1, 2$. Then, from the above inversion formula we obtain

$$a_1(x) - a_2(x) = \frac{\int_0^x f(y)\ dy}{\frac{du_1}{dx}(x)\frac{du_2}{dx}(x)}\left(\frac{du_2}{dx}(x) - \frac{du_1}{dx}(x)\right)$$

Hence, we obtain

$$\int_0^1 (a_1(x) - a_2(x))^2\ dx \leq \frac{(\int_0^1 |f(y)|\ dy)^2}{\gamma^4}\int_0^1 \left(\frac{du_2}{dx}(x) - \frac{du_1}{dx}(x)\right)^2\ dx.$$

Using integration by parts and the Cauchy-Schwarz inequality we obtain

$$\begin{aligned}
&\int_0^1 \left(\frac{du_2}{dx}(x) - \frac{du_1}{dx}(x)\right)^2\ dx \\
&\quad = \int_0^1 (u_1(x) - u_2(x))\left(\frac{d^2u_2}{dx}^2(x) - \frac{d^2u_1}{dx}^2(x)\right)\ dx \\
&\quad \leq \sqrt{\int_0^1 (u_1(x) - u_2(x))^2\ dx}\sqrt{\int_0^1 \left(\frac{d^2u_2}{dx^2}(x) - \frac{d^2u_1}{dx^2}(x)\right)^2\ dx} \\
&\quad \leq 2M\|u_1 - u_2\|_{L^2}
\end{aligned}$$

Thus, for the difference $a_1 - a_2$ we obtain the estimate

$$\|a_1 - a_2\|_{L^2} \leq \frac{\|f\|_{L^1}}{\gamma^2}\sqrt{2M}\|u_1 - u_2\|_{L^2}^{1/2},$$

i.e., the inverse operator $G : u \in \mathcal{C}_\gamma \mapsto a$ is locally Hölder continuous with exponent $\frac{1}{2}$ in the L^2-norm. This result corresponds to the Hölder estimate we have derived for numerical differentiation above. The effect that the estimate is only a local one for the parameter identification problem, is a consequence of the nonlinearity. One clearly observes the influence of smoothness of the solution, for increasing M the constant in the Hölder estimate increases. Moreover, the nonlinear instability is reflected in the estimate by the term $\frac{1}{\gamma^2}$, i.e., the closer u gets to zero, the larger the constant becomes.

In practical applications, it is hardly the case that the solution of a partial differential equation can be measured on a whole domain, since one usually cannot place many detectors inside an object (e.g. a patient in medical applications or a microelectronic device). In such cases boundary measurements either on a space- or time-boundary are available. An example is the diffusion equation

$$\frac{\partial u}{\partial t} = \text{div } (a\nabla u) + f \qquad \text{in } \Omega \times (0, T),$$

with measurements at final time, i.e., $u(x, T)$, for $x \in \Omega$, or at the boundary, e.g., $\frac{\partial u}{\partial n}$ on $\partial\Omega \times (0, T)$. Of course, with such a measurement, the dimensionality of the data is much lower than the one of the unknown $a(x, t)$. Thus, in such cases one can only identify special parameters such as $a = a(x)$, which is however realistic since a might describe material properties that do not change in time.

2.5 Impedance Tomography

Impedance tomography can be considered as a parameter identification problem with boundary measurements. The technological setup is as follows: at the boundary of an object (represented by a domain $D \subset \mathbb{R}^d$), different electrical voltages are applied, and the arising electrical currents are measured. From these measurements one would like to reconstruct the conductivity as a function of space, which gives information about different materials inside the object.

The simplest mathematical model for this process is the solution of the elliptic partial differential equation

$$\text{div } (a\nabla u) = 0 \qquad \text{in } D,$$

where u is the electric potential and a is the conductivity, modeled as a function of the spatial location inside D. The applied voltages f are directly related to the electric potential u at the boundary, i.e.,

$$u = f \qquad \text{on } \partial D.$$

The measured currents over the boundary for a specific voltage f are given by

$$g_f = a\frac{\partial u}{\partial n} \qquad \text{on } \partial D.$$

Hence, if all possible voltages f (in the sense of all functions on ∂D in a certain class) are applied, and the corresponding currents are measured, the data consist of the *Dirichlet-to-Neumann map*

$$\Lambda_a : f \mapsto g_f,$$

which is a linear operator due to the linearity of the differential equation and boundary conditions for fixed a.

The inverse problem of impedance tomography (called *inverse conductivity problem*) consists in reconstructing the conductivity a as a function on D from a measurement of the Dirichlet-to-Neumann map Λ_a. Again, due to the appearance of a as a coefficient in the equation, the inverse problem is nonlinear, though the direct problem of computing the Dirichlet-Neumann map for given a is linear.

From the dimensionality of the data it is not clear whether one can reconstruct the conductivity uniquely, since the unknown is a function on D and the measurement is a linear operator on a class of functions on ∂D. The answer depends on the spatial dimension, for $d \geq 2$ it is indeed possible to identify the conductivity uniquely if the class of voltages f on ∂D is sufficiently large. For dimension $d = 1$, the answer is negative. Consider e.g. the domain $D = [0, 1]$ with boundary $\partial D = \{0, 1\}$. Then a function f on ∂D can be represented by two values, f_0 for $x = 0$, and f_1 for $x = 1$. Hence, the Dirichlet-to-Neumann map can be considered as a linear operator $\Lambda_a : \mathbb{R}^2 \to \mathbb{R}^2$. Since each such linear operator can be represented by a 2×2 matrix, the data consist only of 4 real numbers representing the matrix entries, Since the dimension of the data space ($\mathbb{R}^{2\times 2}$) is finite, but the dimension of the space for the unknown (e.g. $C^1(D)$) is infinite, the data cannot suffice to determine the conductivity uniquely.

An interesting case in impedance tomography is the case of objects consisting only of two different materials and consequently of two different conductivity values, i.e.,

$$a(x) = \begin{cases} a_1 & \text{if } x \in \Omega \subset D \\ a_2 & \text{if } x \in D \backslash \Omega. \end{cases}$$

The subset Ω could for example represent the shape of some inclusion in the object. In such a case the interest is focused on identifying the shape Ω. Since the class of possible functions a is now strongly limited by introducing a-priori knowledge, one may argue that less measurements suffice in order to obtain uniqueness, at least for the shape Ω at given values a_1 and a_2. Indeed, one can show that the measurement of the Neumann value for a single Dirichlet value yields local uniqueness of the inverse problem.

2.6 Inverse Scattering

Inverse scattering problems are, generally speaking, inverse problems where one tries to recover information about an unknown object from measurements of waves (or fields) scattered by this object. Inverse scattering problems exist for all kinds of waves (e.g. acoustic and electromagnetic waves) and all kinds of models (e.g. wave equation, Helmholtz equation, Schrödinger equation, Maxwell equations). We consider the case of an acoustic scattering problem for the Helmholtz equation in the following.

The original mathematical model for the density of an acoustic wave is the wave equation

$$\frac{\partial^2 U}{\partial t} = \frac{1}{n^2} \Delta u \qquad \text{in } \mathbb{R}^d \times \mathbb{R}_+,$$

where $n = n(x)$ describes a spatially varying *acoustic profile* (reciprocal to the speed of sound), where n is scaled to equal one outside a compact domain ($n = 1$ may e.g. represent surrounding air or water). The region where $n(x) \neq 1$ represents the *scattering object*, the deviation of $n(x)$ from one provides information about the structure of the scatterer. If we only consider time harmonic waves of the form $U(x,t) = e^{ikt}u(x)$ for $k \in \mathbb{R}$, then the function u solves the Helmholtz equation

$$\Delta u + k^2 n^2 u = 0.$$

In inverse scattering, an incident wave u^i is sent in, which corresponds to the wave propagating in absence of the scatterer, i.e.,

$$\Delta u^i + k^2 u^i = 0.$$

The scattered wave, which is the difference between the really observed and the incident wave, i.e., with $u^s = u - u^i$ satisfies

$$\Delta u^s + k^2 u^s = k^2 f(u^i + u^s),$$

$f := 1 - n^2$. The inverse scattering problem consists in identifying the compactly supported function f from the knowledge of the incident wave u^i and a measurement of the scattered wave u^s. The scattered wave can only be measured far away from the scatterer, in many cases it is reasonable to assume that u^s can be measured at the sphere with radius $R >> 1$ including the scatterer, which is referred to as the *far-field pattern*.

A closer look at the dimensonality of the unknown (f) and the data ($u^s|_{r=R}$) shows that we have to identify a function on a d-dimensional domain ($d = 2, 3$), but the measurement is a function on a $d-1$-dimensional manifold. Hence, it seems obvious that a single measurement will not suffice to determine f uniquely. For this reason, one uses many different incident waves (varying the value of k) and measures the far-field pattern for all of them, which yields reasonable data for the inverse problem.

Due to the appearance of f as a coefficient in the Helmholtz equation, the acoustic inverse scattering problem is nonlinear. In several situations it is reasonable to assume that the scattered wave in a neighborhood of the scatterer is much smaller than the incident wave, i.e., the term $u^i + u^s$ on the right-hand side can be approximated well by u^i. Under this assumption one can use a linearized version of the inverse scattering problem via the equation

$$\Delta u^s + k^2 u^s = k^2 f u^i,$$

which is known as the *Born approximation*.

A related situation is *inverse obstacle scattering*, where the scattering appears at an obstacle (represented by a domain D), which is not penetrated by the wave. In this case, the Helmholtz equation is a model for wave propagation outside D, i.e.,

$$\Delta u + k^2 u = 0 \qquad \text{in } \mathbb{R}^d \backslash D,$$

coupled with a boundary condition of the form

$$\frac{\partial u}{\partial n} + \lambda u = 0 \qquad \text{on } \partial D.$$

The inverse obstacle scattering problem consists in identifying the shape D, and similary to the corresponding situation for electrical impedance tomography this can be achieved using less measurements (i.e., for only few values of k).

Chapter 3

Regularization of Linear Ill-Posed Problems

In this paper we shall derive the basic ideas of (linear) regularization methods for linear ill-posed problems. We motivate the main ideas for positive definite matrices and then discuss general regularization techniques for linear operator equations involving compact operators.

3.1 Ill-Conditioned Matrix Equations

We start with a linear matrix equation of the form

$$\mathbf{A}\mathbf{x} = \mathbf{y}, \tag{3.1}$$

with $\mathbf{A} \in \mathbb{R}^{n\times n}$ being a symmetric positive definite matrix. From the spectral theory of symmetric matrices it is well-known that there exist eigenvalues $0 \leq \lambda_1 \leq \ldots \leq \lambda_n$ and corresponding eigenvectors $\mathbf{u}_i \in \mathbb{R}^n$ ($\|\mathbf{u}_i\| = 1$) such that $\mathbf{A}$ has a representation of the form

$$\mathbf{A} = \sum_{i=1}^{n} \lambda_i \mathbf{u}_i \mathbf{u}_i^T.$$

It is well-known that the condition number (if $\lambda_1 \neq 0$) of $\mathbf{A}$ is given by the quotient of largest and smallest eigenvalue, i.e., $\kappa = \frac{\lambda_n}{\lambda_1}$. For the sake of simplicity and coherence with the analysis for the inifinite-dimensional case below, we shall assume that the scaling is such that $\lambda_n = 1$, then the condition number corresponds to $\kappa = \lambda_1^{-1}$.

The condition number is a measure for the stable solvability of the problem. Assume that we have noisy data y^δ instead of y, which satisfy

$$\|\mathbf{y}^\delta - \mathbf{y}\| \leq \delta \tag{3.2}$$

in the Euclidean norm on $\mathbb{R}^n$. Let $\mathbf{x}^\delta$ denote the solution with right-hand side $\mathbf{y}^\delta$. Then, from the spectral representation we obtain

$$\mathbf{x}^\delta - \mathbf{x} = \sum_{i=1}^{n} \lambda_i^{-1} \mathbf{u}_i \mathbf{u}_i^T (\mathbf{y}^\delta - \mathbf{y}).$$

Hence, we can estimate (using the orthogonality of eigenvectors)

$$\|\mathbf{x}^\delta - \mathbf{x}\|^2 = \sum_{i=1}^{n} \lambda_i^{-2} |\mathbf{u}_i^T(\mathbf{y}^\delta - \mathbf{y})|^2 \leq \lambda_1^{-2} \|\mathbf{y}^\delta - \mathbf{y}\|^2,$$

or, in other words

$$\|\mathbf{x}^\delta - \mathbf{x}\| \leq \kappa \|\mathbf{y}^\delta - \mathbf{y}\| \leq \kappa\delta.$$

The sharpness of this estimate can be seen immediately for $\mathbf{y}^\delta - \mathbf{y} = \delta\mathbf{u}_1$. One observes that with increasing condition number of the matrix $\mathbf{A}$, the noise amplification increases in the worst-case. For large κ one therefore speaks of an *ill-conditioned problem*. Note that a finite-dimensional linear problem is never ill-posed in the sense that the third condition in Hadamards definition is violated, but for κ large one certainly approximates this case.

We also observe that errors in low frequencies (i.e., corresponding to eigenvectors with large eigenvalues) are amplified less. In the particular, an error in the lowest frequency, i.e., $\mathbf{y}^\delta - \mathbf{y} = \delta\mathbf{u}_n$ is not amplified at all, we just obtain $\|\mathbf{x}^\delta - \mathbf{x}\| = \delta$ from the spectral representation. This is a typical effect for inverse problems: Not all possible versions of noise of the same size are equally bad, high-frequency noise corresponding to low eigenvalues is always worse then low-frequency noise. However, as discussed earlier, it is usually not possible to make any assumptions on the noise in practice, so that a regularization method should be able to deal with arbitrary noise.

So far, we have assumed that the minimal eigenvalue is positive. If this is not the case, i.e., the matrix has a non-trivial nullspace, one can decompose the vector space as

$$\mathbb{R}^n = \mathcal{N}(A) + \mathcal{R}(A), \tag{3.3}$$

where $\mathcal{R}$ denotes the range and $\mathcal{N}$ denotes the nullspace. If λ_m denotes the minimal nonzero eigenvalue, then the solution formula becomes

$$\mathbf{x} = \sum_{i=m}^{n} \lambda_i^{-1} \mathbf{u}_i \mathbf{u}_i^T \mathbf{y}$$

and the problem is only solvable if and only if $\mathbf{u}_i^T\mathbf{y} = 0$ for $i < m$. If the data are noisy ($\mathbf{y}^\delta$) we can use their projection $\mathbf{P}\mathbf{y}^\delta$ onto the range of $\mathbf{A}$ and obtain for the corresponding solution $\mathbf{x}^\delta$ with data $\mathbf{P}\mathbf{y}^\delta$ that

$$\mathbf{x}^\delta - \mathbf{x} = \sum_{i=m}^{n} \lambda_i^{-1} \mathbf{u}_i \mathbf{u}_i^T (\mathbf{P}\mathbf{y}^\delta - \mathbf{y}).$$

Since $\mathbf{u}_i^T\mathbf{P}\mathbf{y}^\delta = \mathbf{u}_i^T\mathbf{y}^\delta$ for $i \geq m$ we thus can estimate similary as above

$$\|\mathbf{x}^\delta - \mathbf{x}\| \leq \lambda_m \delta.$$

There is no error propagation in the nullspace components and the noise amplification is actually determined by minimal nonzero eigenvalue.

This property is a typical one for finite-dimensional operators (i.e., matrices), if A is an operator acting between arbitrary Hilbert spaces X and Y, and A^* denotes its adjoint

operator, then we only the range of the operators A and A^* respectively, is not necessarily closed. The decomposition becomes

$$X = \mathcal{N}(A) + \overline{\mathcal{R}(A^*)}, \qquad Y = \mathcal{N}(A^*) + \overline{\mathcal{R}(A)},$$

where $\overline{Z}$ denotes the closure of a subspace Z. In the case of a non-closed range, the projection operator $\mathbf{P}$ onto the range of $\mathcal{R}$ is not continuous, which creates the instability of the inverse problem. Another appearance of this effect is that there will be nonzero eigenvalues of $\mathbf{A}$ arbitrarily close to zero, so that the condition number is really infinite and the problem is ill-posed.

We now turn our attention to regularization methods. Since we have seen above that small eigenvalues of $\mathbf{A}$ cause most difficulties, it seems natural to approximate $\mathbf{A}$ by a family of matrices, whose smallest eigenvalues are shifted away from zero. A simple candidate is the matrix

$$\mathbf{A}_\alpha := \mathbf{A} + \alpha\mathbf{I}, \qquad \alpha > 0.$$

It is easy to see that the eigenvalues of $\mathbf{A}_\alpha$ are given by $\lambda_i + \alpha$ and the eigenvectors are the same as for $\mathbf{A}$. For $\mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$ and $\mathbf{x}_\alpha == \mathbf{A}_\alpha^{-1}\mathbf{y}$, we thus have

$$\mathbf{x} - \mathbf{x}_\alpha = \sum_{i=1}^{n} (\lambda_i^{-1} - (\lambda_i + \alpha)^{-1})\mathbf{u}_i\mathbf{u}_i^T\mathbf{y} = -\sum_{i=1}^{n} \frac{\alpha}{\lambda_i(\lambda_i + \alpha)}\mathbf{u}_i\mathbf{u}_i^T\mathbf{y}.$$

The *approximation error* of this regularization can be estimated by

$$E_a(\alpha) := \|\mathbf{x} - \mathbf{x}_\alpha\| \leq \frac{\alpha}{\lambda_1(\lambda_1 + \alpha)}\|y\|$$

and in particular, E_α decays to zero for $\alpha \to 0$. Besides the approximation error for exact data y, we can derive error estimates for the regularized solutions with noisy data, i.e., the for the difference of $\mathbf{x}_\alpha$ and $\mathbf{x}_\alpha^\delta = \mathbf{A}_\alpha^{-1}\mathbf{y}^\delta$. From the spectral representation we have

$$\mathbf{x}_\alpha^\delta - \mathbf{x}_\alpha = \sum_{i=1}^{n} (\lambda_i + \alpha)^{-1}\mathbf{u}_i\mathbf{u}_i^T(\mathbf{y}^\delta - \mathbf{y}).$$

The error can be estimated as

$$E_r(\alpha, \delta) := \|\mathbf{x}_\alpha^\delta - \mathbf{x}_\alpha\| \leq \frac{\delta}{\lambda_1 + \alpha}.$$

Finally, by the triangle inequality, we can estimate the error between the exact solution and the regularized solution for noisy data as

$$\|\mathbf{x} - \mathbf{x}_\alpha^\delta\| \leq E_a(\alpha) + E_r(\alpha, \delta) \leq \frac{\alpha}{\lambda_1(\lambda_1 + \alpha)}\|\mathbf{y}\| + \frac{\delta}{\lambda_\mathbf{1} + \alpha}.$$

Note that in practice one does not know the exact data $\mathbf{y}$ and hence, $\|\mathbf{y}\|$ is unknown, too. Using the bound for the noise, we can at least estimate $\|\mathbf{y}\| \leq \|\mathbf{y}^\delta\| + \delta$. Hence,

$$\|\mathbf{x} - \mathbf{x}_\alpha^\delta\| \leq \frac{\alpha}{\lambda_1(\lambda_1 + \alpha)}(\|\mathbf{y}^\delta\| + \delta) + \frac{\delta}{\lambda_1 + \alpha}$$

Note that the two error terms have different behaviour (for fixed δ), the second term is increasing for $\alpha \to 0$ and the first term is decreasing. Hence, it seems clear that there exists a positive $\alpha = \alpha(\delta, \mathbf{y}^\delta)$ which minimizes the total error. The choice of α in dependence of the noise level δ and the noisy data $\mathbf{y}^\delta$ is called parameter choice rule. The rule is *a-priori* if $\alpha = \alpha(\delta)$ and *a-posteriori* if $\alpha = \alpha(\delta, \mathbf{y}^\delta)$. Note that the a-priori choice will not depend on the specific data $\mathbf{y}^\delta$, but on the noise level only, i.e., it is fixed a-priori as a function on the real line. An a-posteriori choice will change the parameter choice for each specific $\mathbf{y}^\delta$. The parameter choice becomes even more important for really ill-posed problems, i.e., eigenvalues tending to zero.

We finally mention that all of the above analysis for symmetric positive semidefinite matrices can be extended to equations with a general matrix $\mathbf{A} \in \mathbb{R}^{n \times m}$ by considering the associated Gaussian normal equation

$$\mathbf{A}^T \mathbf{A} \mathbf{x} = \mathbf{A}^T \mathbf{y},$$

whose system matrix $\mathbf{A}^T \mathbf{A}$ is always symmetric positive semidefinite. We shall follow the analogous approach also for general linear inverse problems in the next sections.

3.2 Generalized Inverses

In the following we consider general linear operator equations of the form

$$Ax = y, , \tag{3.4}$$

where $A : X \to Y$ is a bounded linear operator acting between the Hilbert spaces X and Y. If the range of A is not the full image space (or not even dense), then (3.4) is not solvable for arbitrary right-hand side y. In this case it seems reasonable to look for x such that Ax has minimal distance to x. On the other hand, if A has a nontrivial nullspace, (3.4) may have multiple solutions and it seems reasonable to choose the one with minimal norm among them. This leads to the following definition:

Definition 3.1. An element $x \in X$ is called

(i) *least-squares solution* of (3.4) if

$$\|Ax - y\| = \inf\{ \|Az - y\| \mid z \in X \}. \tag{3.5}$$

(ii) *best-approximate solution* or *minimal-norm solution* of (3.4) if

$$\|x\| = \inf\{ \|z\| \mid z \text{ is least-squares solution of (3.4)} \}. \tag{3.6}$$

In general, a least-squares solution or minimal-norm solution need not exist for arbitrary y, since the range of A need not be closed, a fact we have seen in several examples above. It is easy to see that if a least-squares solution exists, then the minimal-norm solution is unique, because it is the minimizer of a strictly convex (quadratic) functional on a linear subspace.

For those y, where it exists, the minimal-norm solution can (at least in theory) be computed via the *Moore-Penrose generalized inverse*, which is defined as follows:

Definition 3.2. Let $A \in \mathcal{L}(X,Y)$ and let $\tilde{A} : \mathcal{N}(A)^\perp \to \mathcal{R}(A)$ denote its restriction. Then the Moore-Penrose generalized inverse $A^\dagger$ is defined as the unique linear extension of $\tilde{A}^{-1}$ to

$$\mathcal{D}(A^\dagger) := \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$$

with $\mathcal{N}(A^\dagger) = \mathcal{R}(A)^\perp$.

Note that the Moore-Penrose inverse $A^\dagger$ is well-defined: First of all, due to the restriction to $\mathcal{N}(A)^\perp$ and $\mathcal{R}(A)$, the operator $\tilde{A}$ is injective and surjective, and hence, $\tilde{A}^{-1}$ exists. As a consequence, $A^\dagger$ is well-defined on $\mathcal{R}(A)$. For arbitrary $y \in \mathcal{D}(A^\dagger)$ we can find unique $y_1 \in \mathcal{R}(A)$ and $y_2 \in \mathcal{R}(A)^\perp$ and from the linearity and $\mathcal{N}(A^\dagger) = \mathcal{R}(A)^\perp$ we finally obtain

$$A^\dagger y = A^\dagger y_1 + A^\dagger y_2 = \tilde{A}^{-1} y_1.$$

It can be shown that $A^\dagger$ is characterized by the "Moore-Penrose equations"

$$\begin{aligned} AA^\dagger A &= A \\ A^\dagger A A^\dagger &= A^\dagger \\ A^\dagger A &= I - P \\ AA^\dagger &= Q|_{\mathcal{D}(A^\dagger)}, \end{aligned}$$

where $P : X \to \mathcal{N}(A)$ and $Q : Y \to \overline{\mathcal{R}(A)}$ are the orthogonal projectors onto $\mathcal{N}(A)$ and $\overline{\mathcal{R}(A)}$, respectively.

We have announced above that minimal-norm solutions can be computed using the Moore-Penrose generalized inverse, which we make precise by the following result:

Theorem 3.3. *For each $y \in \mathcal{D}(A^\dagger)$, the equation (3.4) has a unique minimal-norm solution given by*

$$x^\dagger := A^\dagger y.$$

The set of all least-squares solutions is given by $\{x^\dagger\} + \mathcal{N}(A)$.

For non-symmetric matrices, it is well-known from linear algebra, that the Gaussian normal equation can be considered to obtain least-squares solutions. We now verify that this assertion is true in the general case:

Theorem 3.4. *For given $y \in \mathcal{D}(A^\dagger)$, $x \in X$ is a least-squares solution of (3.4) if and only if x satisfies the Gaussian normal equations*

$$A^*Ax = A^*y. \tag{3.7}$$

Proof. An element $x \in X$ is a least-squares solution if and only if Ax is the projection of y onto $\mathcal{R}(A)$, which is equivalent to $Ax - y \in \mathcal{R}(A)^\perp$. Since $\mathcal{R}(A)^\perp = \mathcal{N}(A^*)$, this is equivalent to (3.7). □

Since $A^\dagger y$ is the least-squares solution of minimal-norm, we obtain from Theorem 3.4 that $A^\dagger y$ is a solution of (3.7) with minimal norm, i.e.,

$$A^\dagger y = (A^*A)^\dagger A^* y.$$

This means that in order to approximate $A^\dagger y$ we may as well compute an approximation to the minimal-norm solution in (3.7), a fact we will heavily use in the construction of regularization methods.

3.3 Compact Linear Operators

In the following we consider the special case of A being a compact linear operator, compactness being defined as follows.

Definition 3.5. Let $A : X \to Y$ be a continuous linear operator between the Banach spaces X and Y. Then A is said to be compact if for any bounded set $B \subset X$ the image $A(B) \subset Y$ is pre-compact.

An alternative definition of compactness can be given in terms of sequences. If (x_n) is a bounded sequence in X, then compactness of A implies that (Ax_n) is contained in a precompact set and therefore has a convergent subsequence.

As we have seen already in the introductory examples, inverse problems with compact operators are an important case, actually most inverse problems involve compact operators. In other words one could argue that compactness of the operator A is a source of ill-posedness for the equation (3.4), which is confirmed by the following result:

Theorem 3.6. *Let $A : X \to Y$ be a compact linear operator between the infinite-dimensional Hilbert spaces X and Y, such that the dimension of $\mathcal{R}(A)$ is infinite. Then the problem* (3.4) *is ill-posed, i.e., $A^\dagger$ is discontinuous.*

Proof. Since X and $\mathcal{R}(A)$ are infinite-dimensional, also $\mathcal{N}(A)^\perp$ is infinite-dimensional (note that the dimension of $\mathcal{R}(A)$ always smaller or equal the dimension of $\mathcal{N}(A)^\perp$). Hence, we can find a sequence (x_n) with $x_n \in \mathcal{N}(A)^\perp$, $\|x_n\| = 1$, and

$$\langle x_n, x_k \rangle = 0 \qquad \text{for } k \neq n.$$

Since A is a compact operator, the sequence $(y_n) := (Ax_n)$ is compact and hence for each $\delta > 0$ we can find k, ℓ such that $\|y_k - y_\ell\| < \delta$, but

$$\|A^\dagger y_k - A^\dagger y_\ell\|^2 = \|x_k - x_\ell\|^2 = \|x_k\|^2 + \|x_k\|^2 - 2\langle x_k, x_\ell \rangle = 2.$$

Hence, $A^\dagger$ is unbounded. □

As for matrices, one can prove a spectral representation theorem for compact linear self-adjoint operators in Hilbert spaces. If A is compact, then the spectrum of A is given by $\{0\} \cup \{\lambda_n\}_{n=1}^\infty$, where λ_n are the (at most countably many) nonzero eigenvalues of A. Since A is self-adjoint, all eigenvalues are real, and hence, with a set of normalized eigenvectors x_n one has

$$Ax = \sum_{n=1}^{\infty} \lambda_n x_n \langle x, x_n \rangle \qquad \forall\, x \in X.$$

If A is not self-adjoint we can (as in the case of matrices) pass to the operators $B := A^*A$ and $C := AA^*$. Both B and C are compact, self-adjoint and even positive semidefinite, so that both admit a spectral representation with positive real eigenvalues of the form

$$Bx = \sum_{n=1}^{\infty} \sigma_n^2 u_n \langle x, u_n \rangle \qquad \forall\, x \in X,$$

and

$$Cx = \sum_{n=1}^{\infty} \tilde{\sigma}_n^2 v_n \langle x, v_n \rangle \qquad \forall\, y \in Y.$$

Note that for $\tilde{\sigma}_n^2$ being an eigenvalue of C with eigenvector v_n we have

$$\tilde{\sigma}_n^2 A^* v_n = A^* C v_n = A^* A A^* v_n = B A^* v_n$$

and hence, $\tilde{\sigma}_n^2$ is an eigenvalue of B (with eigenvector $A^* v_n$). Vice versa, it follows that an eigenvalue σ_n^2 of B is also an eigenvalue of C (with eigenvector Au_n). Thus, there is a one-to-one relation between eigenvalues and eigenvectors of B and C, and we can assume without restriction of generality that

$$\tilde{\sigma}_n = \sigma_n, \qquad v_n = \frac{Au_n}{\|Au_n\|}.$$

The above construction of eigensystems for B and C is the basis of the *singular value decomposition* of compact linear operators. We call (σ_n, u_n, v_n) singular system, and using this system we obtain

$$Ax = \sum_{n=1}^{\infty} \sigma_n \langle x, u_n \rangle v_n, \qquad \forall\ x \in X, \tag{3.8}$$

which is called singular value decomposition of the operator A. For the adjoint operator one can derive an analogous relation of the form

$$A^* y = \sum_{n=1}^{\infty} \sigma_n \langle y, v_n \rangle u_n, \qquad \forall\ y \in Y. \tag{3.9}$$

Note that the sums on the right-hand side of (3.8) and (**??**) converge due to the square integrability of the coefficients $\langle x, u_n \rangle$ (respectively $\langle y, v_n \rangle$), the orthogonality of singular vectors, and the boundedness of the singular values. E.g. for (3.8) the finite sums

$$\left\| \sum_{n=1}^{N} \sigma_n \langle x, u_n \rangle v_n \right\|^2 \leq \sum_{n=1}^{N} \sigma_n^2 \langle x, u_n \rangle^2 \leq \sigma_1^2 \sum_{n=1}^{N} \langle x, u_n \rangle^2 \leq \sigma_1^2 \|x\|^2$$

are uniformly bounded with respect to N, which allows to pass to the limit $N \to \infty$.

We can now derive a representation of the generalized inverse $A^\dagger$ in terms of the singular system. Note that $A^\dagger = (A^* A)^\dagger A^*$ and hence, for $x^\dagger = A^\dagger y$ we have

$$\sum_{n=1}^{\infty} \sigma_n^2 \langle x^\dagger, u_n \rangle u_n = A^* A x^\dagger = A^* y = \sum_{j=1}^{n} \sigma_n \langle y, v_n \rangle u_n,$$

and hence, by comparison of the respective linearly independent components we obtain

$$\langle x^\dagger, u_n \rangle = \frac{1}{\sigma_n} \langle y, v_n \rangle u_n.$$

As a direct consequence we obtain the singular value decomposition of the generalized inverse $A^\dagger$ via

$$x^\dagger = A^\dagger y = \sum_{n=1}^{\infty} \frac{1}{\sigma_n} \langle y, v_n \rangle u_n. \tag{3.10}$$

Again we have to check if the sum on the right-hand side of (3.10) converges. Opposed to the case of A and A^* this is not always true for $A^\dagger$ and clearly reflects the unboundedness of the generalized inverse. The convergence criteria for the sum, namely

$$\|A^\dagger y\|^2 \ = \ \sum_{n=1}^{\infty} \frac{\langle y, v_n \rangle^2}{\sigma_n^2} \quad < \infty \tag{3.11}$$

is called *Picard criterion*. This criterion can be interpreted as a smoothness condition on the data y, since its coefficients with respect to the singular vectors v_n must decay faster than the singular values. The unboundedness of the generalized inverse for compact operators can again be seen from (3.11), since for the normalized sequence v_n we have $\|A^\dagger v_n\| = \frac{1}{\sigma_n} \to \infty$. One observes again that errors in high frequencies, i.e., in the coefficients $\langle y, v_n \rangle$ corresponding to singular vectors with large n (and small σ_n) are amplified much stronger than those for low frequencies (large σ_n).

The way how high frequency errors are amplified depends on the operator A, more precisely on the decay speed of its singular values. The faster this decay happens, the more severe the Picard criterion (3.11) becomes. This motivates the following distinction into mildly and severely ill-posed problems:

- *Mildly ill-posed linear inverse problems* are problems of the form (3.4) with an operator A, whose singular values decay at most with polynomial speed, i.e., there exist $\gamma, C > 0$ and such that $\sigma_n \geq Cn^{-\gamma}$ for all n.
- *Severely ill-posed linear inverse problems* are problems of the form (3.4) with an operator A, whose singular values decay faster than with polynomial speed, i.e., for all $\gamma, C > 0$ one has $\sigma_n \leq Cn^{-\gamma}$ for n sufficiently large.

Example 3.7. As a simple example for the singular value decomposition, we consider again numerical differentiation, i.e., the operator $A : L^2([0,1]) \to L^2([0,1])$ is given by

$$(Ax)(t) := \int_0^t x(s)\ ds.$$

From the theory of integral operators it follows that A is compact, since its integral kernel is square-integrable (cf. [9]). For $y \in L^2([0,1])$ we have

$$\langle Ax, y \rangle = \int_0^1 \int_0^t x(s)\ y(t)\ ds\ dt = \int_0^1 \int_s^1 y(t)\ dt\ x(s)\ ds = \langle x, A^* y \rangle$$

and therefore, the adjoint is given by

$$(A^* y)(t) := \int_t^1 x(s)\ ds.$$

Now assume that $\lambda \neq 0$ is an eigenvalue of A^*A with eigenvector u. Then,

$$\int_t^1 \int_0^s u(\tau)\ d\tau\ ds = (A^*Au)(t) = \lambda u(t).$$

In particular, we have $u(1) = 0$ and u has a derivative given by

$$u'(t) = -\frac{1}{\lambda} \int_0^t u(\tau)\ d\tau.$$

Now we can conclude $u'(0) = 0$ and

$$u''(t) = -\frac{1}{\lambda} u(t).$$

With $\sigma = \sqrt{\lambda}$ it thus follows that

$$u(t) = c_1 \sin \sigma t + c_2 \cos \sigma t,$$

with constants c_1, c_2. The only nontrivial (normalized) solutions satisfying the boundary conditions $u(1) = 0$ and $u'(0) = 0$ are

$$\sigma_n = \frac{2}{(2n-1)\pi}, \qquad u_n(t) = \sqrt{2} \cos \frac{t}{\sigma_n}.$$

Hence, the singular values of the operator A decay like n^{-1} and the problem is mildly ill-posed.

For the eigensystem of AA^* we obtain

$$\int_0^t \int_s^1 v(\tau)\ d\tau\ ds = (A^*Av)(t) = \lambda v(t),$$

and thus, v satisfies

$$v''(t) = -\frac{1}{\lambda} v(t), \qquad v(0) = 0, v'(1) = 0$$

and we obtain the singular vectors

$$v_n(t) = \sqrt{2} \sin \frac{t}{\sigma_n}.$$

The Picard criterion in this case becomes

$$\sum_{n=1}^{\infty} \frac{(2n-1)^2\pi^2}{2} \left(\int_0^1 y(t)\ \sin \frac{t}{\sigma_n}\ dt \right)^2 < \infty,$$

which is just the condition for the differentiability of the Fourier series by differentiating its components.

3.4 Regularization Methods

In this section, we shall introduce the notion of regularization in a rigorous way and discuss some basic properties. In general, we shall describe a linear regularization method by a family of continuous linear operators $R_\alpha : Y \to X$, for $\alpha \in I \subset (0, \alpha_0)$, where the index set I includes at least one sequence $\alpha_n \to 0$. Of course, the regularization operator should converge to the generalized inverse in some sense as $\alpha \to 0$. This means that, as $\alpha \to 0$, we need the convergence $R_\alpha y \to A^\dagger y$ for $y \in \mathcal{D}(A^\dagger)$. If $y \in Y \setminus \mathcal{D}(A^\dagger)$, we have to expect that $\|R_\alpha y\| \to \infty$ due to the unboundedness of the generalized inverse.

When dealing with noisy data y^δ satisfying

$$\|y - y^\delta\| \le \delta, \tag{3.12}$$

one has to choose the regularization parameter α in dependence of the noise level δ and, possibly, in dependence of the noisy data y^δ. We shall call the specific strategy of choosing $\alpha = \alpha(\delta, y^\delta)$ *parameter choice rule*. Clearly, for $y \in \mathcal{D}(A^\dagger)$ we would like to obtain $R_{\alpha(\delta, y^\delta)} y^\delta \to A^\dagger y$ as $\delta \to 0$. This desired convergence property leads to the following definition:

Definition 3.8. A family $\{R_\alpha\}_{\alpha\in I}$ of continuous linear operators is called regularization (or regularization operator) for $A^\dagger$, if for all $y \in \mathcal{D}(A^\dagger)$ there exists a parameter choice rule $\alpha : \mathbb{R}^+ \times Y \to I$ such that

$$\limsup_{\delta\to 0}\{\|R_{\alpha(\delta,y^\delta)}y^\delta - A^\dagger y\| \mid y^\delta \in Y, \|y - y^\delta\| \le \delta\} = 0 \tag{3.13}$$

and

$$\limsup_{\delta\to 0}\{\alpha(\delta, y^\delta) \mid y^\delta \in Y, \|y - y^\delta\| \le \delta\} = 0. \tag{3.14}$$

For a specific $y \in \mathcal{D}(A^\dagger)$, the pair (R_α, α) is called *(convergent) regularization method* of (3.4) if (3.13) and (3.14) hold.

As announced before, we will distinguish the parameter choice via the dependence on the noisy data y^δ, the rule $\alpha = \alpha(\delta, y^\delta)$ is called

- *a-priori* if α does not depend on y^δ. We shall write $\alpha = \alpha(\delta)$ in this case.
- *a-posteriori* otherwise.

In practice, there might be a temptation to choose α in dependence of the known noisy data y^δ, but independent of the noise level δ. The following result due to Bakushinskii [2] shows that such an approach cannot result in a convergent regularization method for an ill-posed problems, or, in other words, such a strategy can only work for well-posed problems, which could also be solved without regularization:

Theorem 3.9. *Let $A : X \to Y$ be a bounded linear operator and let $\{R_\alpha\}$ be a regularization for $A^\dagger$, such that the regularization method converges for every $y \in \mathcal{D}(A^\dagger)$ and such that the parameter choice α depends on y^δ only (and not on δ). Then $A^\dagger$ can be extended to a continuous operator from X to Y.*

Proof. For $\alpha = \alpha(y^\delta)$, (3.13) implies

$$\limsup_{\delta\to 0}\{\|R_{\alpha(y^\delta)}y^\delta - A^\dagger y\| \mid y^\delta \in Y, \|y - y^\delta\| \le \delta\} = 0$$

and in particular $R_{\alpha(y)}y = A^\dagger y$ for all $y \in \mathcal{D}(A^\dagger)$. For any sequence $\{y_n\} \subset \mathcal{D}(A^\dagger)$ converging to y, we obtain

$$A^\dagger y_n = R_{\alpha(y_n)}y_n \to R_{\alpha(y)}y = A^\dagger y,$$

and hence, $A^\dagger$ is continuous on $\mathcal{D}(A^\dagger)$. Since $\mathcal{D}(A^\dagger)$ is dense, there exists a unique continuous extension of $A^\dagger$ to Y. □

This result rules out *error-free* parameter choices $\alpha = \alpha(\delta)$ as convergent regularization methods. Of course, it does not mean that such strategies (which are actually used as heuristic approaches in practice) do not behave well for finite δ, but at least it indicates that the results have to be considered with care in such cases.

Once we have clarified the concepts of regularization methods and parameter choice rules, the obvious questions how to construct such regularizations and how to perform parameter choice, arises. This questions will be considered in the next section, but before we derive some basic properties to be satisfied by regularization methods.

Proposition 3.10. *Let $A : X \to Y$ be a continuous linear operator and $R_\alpha : Y \to X$ be a family of continuous operators, $\alpha \in \mathbb{R}_+$. Then, the family $\{R_\alpha\}$ is a regularization of $A^\dagger$ if*

$$R_\alpha \to A^\dagger \quad \text{pointwise on } \mathcal{D}(A^\dagger) \quad \text{as } \alpha \to 0.$$

In particular, in this case there exists an a-priori parameter choice rule α such that (R_α, α) is a convergent regularization method for (3.4).

Proof. Let $y \in \mathcal{D}(A^\dagger)$ be arbitrary, but fixed. Due to the pointwise convergence, we can find a monotone function $\sigma : \mathbb{R}^+ \to \mathbb{R}_+$ such that for every $\epsilon > 0$

$$\|R_{\sigma(\epsilon)} y - A^\dagger y\| \leq \frac{\epsilon}{2}.$$

The operator $R_{\sigma(\epsilon)}$ is continuous for fixed ϵ and hence, there exists $\rho(\epsilon) \in I$ such that

$$\|R_{\sigma(\epsilon)} z - R_{\sigma(\epsilon)} y\| \leq \frac{\epsilon}{2} \qquad \text{if } \|z - y\| \leq \rho(\epsilon).$$

Without restriction of generality, we can assume that ρ is monotone increasing, continuous, and $\lim_{\epsilon \to 0} \rho(\epsilon) = 0$. Hence, there exists a well-defined inverse function ρ^{-1} on the range of ρ with the same properties. Now we extend ρ^{-1} to a continuous, strictly monotone function on $\mathbb{R}^+$ and define the parameter choice rule as

$$\alpha : \mathbb{R}^+ \to \mathbb{R}^+, \qquad \delta \mapsto \sigma(\rho^{-1}(\delta)).$$

By our construction, we have for $\delta := \rho(\epsilon)$ that

$$\|R_{\alpha(\delta)} y^\delta - A^\dagger y\| \leq \|R_{\alpha(\delta)} y^\delta - R_{\alpha(\delta)} y\| + \|R_{\alpha(\delta)} y - A^\dagger y\| \leq \epsilon,$$

if $\|y - y^\delta\| \leq \delta$, because we have $\alpha(\delta) = \sigma(\epsilon)$. Hence, (R_α, α) is a convergent regularization method for (3.4). □

We now know that any family of continuous operators that converges pointwise to the generalized inverse defines a regularization method. Vice versa, we can conclude from (3.14) that

$$\lim_{\delta \to 0} R_{\alpha(\delta, y^\delta)} y = A^\dagger y, \qquad \forall\, y \in \mathcal{D}(A^\dagger),$$

and thus, if α is continuous in δ, this implies

$$\lim_{\sigma \to 0} R_\sigma y = A^\dagger y,$$

i.e., a convergent regularization method with continuous parameter choice rule implies pointwise convergence of the regularization operators.

Now we turn our attention to the behaviour of the regularization operators on $Y \setminus \mathcal{D}(A^\dagger)$. Since the generalized inverse is not defined on this set, we cannot expected that R_α remains bounded on this set as $\alpha \to 0$. This is indeed confirmed by the next result:

Proposition 3.11. *Let $A : X \to Y$ be a continuous linear operator and $R_\alpha : Y \to X$ be a family of continuous linear regularization operators. Then, $x_\alpha := R_\alpha y$ converges to $A^\dagger y$ as $\alpha \to 0$ for $y \in \mathcal{D}(A^\dagger)$. Moreover, if*

$$\sup_{\alpha > 0} \|A R_\alpha\| < \infty, \tag{3.15}$$

then $\|x_\alpha\| \to \infty$ for $y \notin \mathcal{D}(A^\dagger)$.

Proof. The convergence of x_α for $y \in \mathcal{D}(A^\dagger)$ has been verified above. Now, let $y \notin \mathcal{D}(A^\dagger)$ and assume there exists a sequence $\alpha_n \to 0$ such that $\|x_{\alpha_n}\|$ is uniformly bounded. Then there exists a weakly convergent subsequence (again denoted by x_{α_n}) with some limit $x \in X$, and since continuous linear operators are weakly continuous, too, we have $Ax_{\alpha_n} \to Ax$. On the other hand, since AR_α are uniformly bounded operators, we conclude $Ax_{\alpha_n} = AR_{\alpha_n} y \to Qy$. Hence, $Ax = Qy$ and consequently $y \in \mathcal{D}(A^\dagger)$, a contradiction. □

We finally consider the properties of a-priori parameter choice rules:

Proposition 3.12. *Let $A : X \to Y$ be a continuous linear operator and $R_\alpha : Y \to X$ be a family of continuous linear regularization operators, with a-priori parameter choice rule $\alpha = \alpha(\delta)$. Then, (R_α, α) is a convergent regularization method if and only if*

$$\lim_{\delta \to 0} \alpha(\delta) = 0, \qquad \lim_{\delta \to 0} \delta \|R_{\alpha(\delta)}\| = 0. \tag{3.16}$$

holds.

Proof. If (3.16) hold, then for all $y^\delta \in Y$ with $\|y - y^\delta\| \le \delta$,

$$\begin{aligned} \|R_{\alpha(\delta)} y^\delta - A^\dagger y\| &\le \|x_{\alpha(\delta)} - A^\dagger y\| + \|x_{\alpha(\delta)} - R_{\alpha(\delta)} y^\delta\| \\ &\le \|x_{\alpha(\delta)} - A^\dagger y\| + \delta \|R_{\alpha(\delta)}\|. \end{aligned}$$

Because of (3.16) and since regularization operators converge pointwise, the right-hand side tends to zero as $\delta \to 0$, i.e., (R_α, α) is a convergent regularization method.

Assume vice versa that (R_α, α) is a convergent regularization method. Assume that there exists a sequence $\delta_n \to 0$ such that $\delta_n \|R_{\alpha(\delta_n)}\| \ge C > 0$ for some constant C. Then we can find a sequence z_n with $\|z_n\| = 1$ such that $\delta_n \|R_{\alpha(\delta_n)} z_n\| \ge \frac{C}{2}$. Then, for any $y \in \mathcal{D}(A^\dagger)$ and $y_n := y + \delta_n z_n$ we obtain $\|y - y_n\| \le \delta_n$, but

$$(R_{\alpha(\delta_n)} y_n - A^\dagger y) = (R_{\alpha(\delta_n)} y - A^\dagger y) + \delta_n R_{\alpha(\delta_n)} z_n$$

does not converge, since the second term is unbounded. Hence, for δ_n sufficiently small, (3.14) is not satisfied. □

3.5 Construction of Regularization Methods

Now we turn our attention to the construction of regularization methods for linear ill-posed problems. We have seen above that the first two points defining a well-posed problem, can always be enforced by considering the generalized inverse $A^\dagger$. A violation of the third point, i.e. instability, arises if the spectrum of the operator A is not bounded away from zero. Thus, it seems natural to construct regularizing approximations via modifying the smallest singular values.

With the singular value decomposition of the generalized inverse, we can easily realize such a modification of small singular values and construct regularization operators of the form

$$R_\alpha y := \sum_{n=1}^{\infty} g_\alpha(\sigma_n) \langle y, v_n \rangle u_n, \qquad y \in Y, \tag{3.17}$$

with some function $g_\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ such that $g_\alpha(\sigma) \to \frac{1}{\sigma}$ for $\sigma > 0$ as $\alpha \to 0$. Such an operator R_α is a regularization operator if

$$g_\alpha(\sigma) \leq C_\alpha < \infty, \qquad \forall\, \sigma \in \mathbb{R}_+. \tag{3.18}$$

If (3.18) is satisfied, we can estimate

$$\|R_\alpha y\|^2 = \sum_{n=1}^{\infty} (g_\alpha(\sigma_n))^2 |\langle y, v_n\rangle|^2 \leq C_\alpha^2 \sum_{n=1}^{\infty} |\langle y, v_n\rangle|^2 \leq C_\alpha^2 \|y\|^2, \textit{i.e.},$$

C_α is a bound for the norm of R_α.

From the pointwise convergence of g_α we immediately conclude the pointwise convergence of R_α to $A^\dagger$. Moreover, condition (3.16) on the choice of the regularization parameter can be replaced by

$$\lim_{\delta \to 0} \delta C_{\alpha(\delta)} = 0. \tag{3.19}$$

since $\|R_\alpha\| \leq C_\alpha$. Hence, if (3.19) is satisfied, we know that (R_α, α) with R_α defined by (3.17) is a convergent regularization method.

In order to make this construction more concrete, we consider some particular examples.

Example 3.13 (Truncated Singular Value Decomposition). The main idea of truncated singular value decomposition is to ignore all singular values below a certain threshold value, which we can identify with the regularization parameter α. Thus, the function g_α is given by

$$g_\alpha(\sigma) := \begin{cases} \frac{1}{\sigma} & \text{if } \sigma \geq \alpha \\ 0 & \text{if } \sigma < \alpha \end{cases} \tag{3.20}$$

We obviously obtain $C_\alpha = \frac{1}{\alpha}$ and hence, truncated singular value decomposition is a convergent regularization method if $\frac{\delta}{\alpha} \to 0$.

The representation of the regularized solution is given by

$$x_\alpha := R_\alpha y = \sum_{\sigma_n \geq \alpha} \frac{1}{\sigma_n} \langle y, v_n\rangle u_n, \qquad y \in Y, \tag{3.21}$$

which explains the name truncated singular value decomposition, since all terms in the sum corresponding to small singular values are truncated. Note that since 0 is the only accumulation point of the singular values of a compact operator, the sum in (3.21) is always finite for $\alpha > 0$. In particular, only a finite number of singular values and singular vectors has to be computed in order to realize this method. On the other hand, for α being small, the number of singular values that need to be computed can increase strongly.

Example 3.14 (Lavrentiev Regularization). The main idea of Lavrentiev Regularization is to shift all singular values by α, i.e., $g_\alpha(\sigma) = \frac{1}{\sigma+\alpha}$ and

$$x_\alpha := R_\alpha y = \sum_{n=1}^{\infty} \frac{1}{\sigma_n + \alpha} \langle y, v_n\rangle u_n, \qquad y \in Y. \tag{3.22}$$

In this case, the sum is really infinite and the full singular system is needed in order to compute the solution. However, if A is a positive semidefinite operator (and thus $\lambda_n = \sigma_n$, $u_n = v_n$), one obtains

$$(A + \alpha I) x_\alpha = \sum_{n=1}^{\infty} \frac{1}{\sigma_n + \alpha} \langle x_\alpha, u_n\rangle u_n = \sum_{n=1}^{\infty} \langle y, u_n\rangle u_n = y.$$

Hence, the regularized solution can also be obtained in this case without any knowledge of the singular system as the solution of the linear equation

$$(A + \alpha I)x_\alpha = y.$$

Since $\frac{1}{\sigma+\alpha} \leq \frac{1}{\alpha}$, we again obtain $C_\alpha = \frac{1}{\alpha}$ and the condition to obtain a convergent regularization method is again $\frac{\delta}{\alpha} \to 0$.

Example 3.15 (Tikhonov Regularization). For Tikhonov Regularization, the function g_α is given by $g_\alpha(\sigma) = \frac{\sigma}{\sigma^2+\alpha}$ and the regularized solution is

$$x_\alpha := R_\alpha y = \sum_{n=1}^{\infty} \frac{\sigma_n}{\sigma_n^2 + \alpha} \langle y, v_n \rangle u_n, \qquad y \in Y. \tag{3.23}$$

We can estimate $\sigma^2 + \alpha \geq 2\sigma\sqrt{\alpha}$ and hence, $g_\alpha(\sigma) \leq C_\alpha := \frac{1}{2\sqrt{\alpha}}$. Thus, the condition for a convergent regularization method in this case becomes $\frac{\delta}{\sqrt{\alpha}} \to 0$.

As in the case of Lavrentiev regularization, the we can compute x_α defined by (3.23) without knowledge of the singular system, but now for arbitrary linear operators A. It is easy to see that

$$(A^*A + \alpha I)x_\alpha = A^*y$$

and hence, we can solve a well-posed linear system to obtain x_α. From this representation one also observes that Tikhonov regularization is just Lavrentiev regularization applied to the Gaussian normal equation. We will discuss Tikhonov regularization and related approaches in detail in the next chapter.

Example 3.16 (Asymptotic Regularization). Asymptotic regularization is usually constructed from the solution u of the initial value problem

$$\begin{aligned} u'(t) &= -A^*(Au(t) - y) \qquad t \in \mathbb{R}_+ \\ u(0) &= 0, \end{aligned}$$

as $x_\alpha = u(\frac{1}{\alpha})$. By representing u in terms of the singular vectors u_n as

$$u(t) = \sum_{n=1}^{\infty} \gamma_n(t) u_n$$

with $\gamma_n(0) = 0$, we obtain from the singular value decomposition

$$\gamma_n'(t) = -\sigma_n^2 \gamma_n(t) + \sigma_n \langle v_n, y \rangle.$$

This ordinary differential equation can be solved analytically as

$$\gamma_n(t) = (1 - \exp(-\sigma_n^2 t)) \frac{1}{\sigma_n} \langle v_n, y \rangle.$$

Hence, the regularized solution is given by

$$x_\alpha = \sum_{n=1}^{\infty} (1 - \exp(-\frac{\sigma_n^2}{\alpha})) \frac{1}{\sigma_n} \langle y, v_n \rangle u_n, \qquad y \in Y. \tag{3.24}$$

and $g_\alpha(\sigma) = (1 - \exp(-\frac{\sigma 2}{\alpha}))\frac{1}{\sigma}$.

When we consider the error of the regularization in the case of noisy data y^δ, we have (with the notation $x_\alpha^\delta := R_\alpha y^\delta$)

$$x^\dagger - x_\alpha^\delta = (x^\dagger - x_\alpha) + (x_\alpha - x_\alpha^\delta).$$

The first term $(x^\dagger - x_\alpha)$ is the approximation error of the regularization method, which is independent of the noise. The second term $(x_\alpha - x_\alpha^\delta)$ corresponds to the propagation of data noise in the regularized case. Through the triangle inequality we can estimate

$$\|x^\dagger - x_\alpha^\delta\| \leq \|x^\dagger - x_\alpha\| + \|x_\alpha - x_\alpha^\delta\|. \tag{3.25}$$

Thus, in order to estimate the error between the regularized solution and the exact solution, we can estimate these two error terms separately. Such an estimation can also yield some guideline for the parameter choice, namely by choosing α such that the terms on the right-hand side are balanced.

We start with an estimate of the approximation error, which is independent of the noise level δ:

Theorem 3.17. *Let $g_\alpha : \mathbb{R}^+ \to \mathbb{R}$ be a piecewise continuous function satisfying the assumptions above and*

$$\sup_{\alpha,\sigma} (\sigma g_\alpha(\sigma)) \leq \gamma$$

for some constant $\gamma > 0$. Moreover, let the regularization operator be defined by (3.17)*. Then, for all $y \in \mathcal{D}(A^\dagger)$,*

$$R_\alpha y \to A^\dagger y \qquad as \qquad \alpha \to 0.$$

Proof. From the singular value decomposition we have

$$R_\alpha y - A^\dagger y = \sum_{n=1}^{\infty} \left(g_\alpha(\sigma_n) - \frac{1}{\sigma_n} \right) \langle y, v_n \rangle u_n = \sum_{n=1}^{\infty} (\sigma_n g_\alpha(\sigma_n) - 1) \langle x^\dagger, u_n \rangle u_n.$$

From the assumptions on g_α we obtain

$$|(\sigma_n g_\alpha(\sigma_n) - 1)\langle x^\dagger, u_n \rangle| \leq (\gamma + 1)\|x^\dagger\|,$$

and hence, we may estimate

$$\begin{aligned} \limsup_\alpha \|R_\alpha y - A^\dagger y\|^2 &\leq \limsup_\alpha \sum_{n=1}^{\infty} (\sigma_n g_\alpha(\sigma_n) - 1)^2 \langle x^\dagger, u_n \rangle^2 \\ &\leq \sum_{n=1}^{\infty} \Big(\lim_\alpha (\sigma_n g_\alpha(\sigma_n)) - 1 \Big)^2 \langle x^\dagger, u_n \rangle^2. \end{aligned}$$

From the pointwise convergence $\sigma g_\alpha(\sigma) \to 1$, we deduce that $\lim_\alpha(\sigma_n g_\alpha(\sigma_n)) - 1 = 0$ and hence, $\|R_\alpha y - A^\dagger y\| \to 0$ as $\alpha \to 0$. □

From the proof we again observe the arbitrarily slow convergence of the regularized solutions. In the particular case $x^\dagger = u_n$ we have

$$\lim_{\alpha \to 0} \|R_\alpha y - A^\dagger y\| = \lim_{\alpha \to 0} |(\sigma_n g_\alpha(\sigma_n)) - 1|.$$

The function $t \mapsto tg_\alpha(\sigma)$ converges pointwise to the function

$$g(t) = \begin{cases} 0 & \text{if } t > 0 \\ 1 & \text{if } t = 0. \end{cases}$$

Due to the discontinuity at zero, the convergence of $tg_\alpha(t) - 1$ to zero is becoming slower and slower as t decreases to zero. Since we can always find an arbitrarily small singular value σ_n and the minimal norm solution $x^\dagger = u_n$, the convergence of regularized solutions is arbitrarily slow. On the other hand, we observe from the proof that there is a possibly faster convergence if the components $\langle x^\dagger, u_n\rangle$ decay sufficiently fast compared to the eigenvalues. E.g., if we know that $|\langle x^\dagger, u_n\rangle| \le c\sigma_n^\mu$ for some constant $c > 0$ and $\mu > 0$, then we have

$$\limsup_\alpha \|R_\alpha y - A^\dagger y\|^2 \le \limsup_\alpha c^2 \sum_{n=1}^\infty (\sigma_n g_\alpha(\sigma_n) - 1)^2 \sigma_n^{2\mu} \le c^2 \sum_{n=1}^\infty \lim_\alpha \left(\sigma_n^{1+\mu} g_\alpha(\sigma_n) - \sigma_n^\mu\right)^2.$$

Thus, one has to consider the limit of the function $t \mapsto |t^{1+\mu} g_\alpha(t) - t^\mu|$ as $t \to \infty$ instead, which is usually much faster. E.g., for truncated singular value decomposition, we obtain

$$|t^{1+\mu} g_\alpha(t) - t^\mu| = \begin{cases} 0 & \text{if } t \ge \alpha \\ t^\mu & \text{if } t < \alpha. \end{cases}$$

If the singular values of the operator decay sufficiently fast (which is the typical case for ill-posed problems), e.g., $\sum_{n=1}^\infty \sigma_n^\mu < \infty$, we obtain

$$\|R_\alpha y - A^\dagger y\|^2 \le c^2 \sum_{\sigma_n < \alpha} \sigma_n^{2\mu} \le c^2 \alpha^\mu \sum_{n=1}^\infty \sigma_n^\mu,$$

and thus, $\|R_\alpha y - A^\dagger y\|$ is of order $\alpha^{\mu/2}$. I.e., we somehow need smoothness of the solution (in terms of the smoothing properties of the operator) in order to obtain a convergence rate in terms of α. We shall pursue this idea by considering so-called *source conditions* below.

Now we consider the propagation of the data error through the regularization:

Theorem 3.18. *Let g_α and γ be as in Theorem 3.17, and let $x_\alpha := R_\alpha y$, $x_\alpha^\delta := R_\alpha y^\delta$. Then,*

$$\|Ax_\alpha - Ax_\alpha^\delta\| \le \gamma\delta, \tag{3.26}$$

and

$$\|x_\alpha - x_\alpha^\delta\| \le C_\alpha \delta, \tag{3.27}$$

hold.

Proof. From the singular value decomposition we can directly estimate

$$\begin{aligned} \|Ax_\alpha - Ax_\alpha^\delta\|^2 &\le \sum_{n=1}^\infty \sigma_n^2 g_\alpha(\sigma_n)^2 |\langle y - y^\delta, v_n\rangle|^2 \\ &\le \gamma^2 \sum_{n=1}^\infty |\langle y - y^\delta, v_n\rangle|^2 = \gamma^2 \|y - y^\delta\|^2 \le (\gamma\delta)^2, \end{aligned}$$

which yields (3.26). In the same way we estimate

$$\begin{aligned}\|x_\alpha - x_\alpha^\delta\|^2 &\leq \sum_{n=1}^{\infty} g_\alpha(\sigma_n)^2 |\langle y - y^\delta, v_n\rangle|^2 \\ &\leq \sum_{n=1}^{\infty} |\langle y - y^\delta, v_n\rangle|^2 = C_\alpha^2 \|y - y^\delta\|^2 \leq (C_\alpha \delta)^2,\end{aligned}$$

implying (3.27). □

Note that (3.27) estimates the norm of R_α by C_α. It is clear that C_α increases with $\alpha \to 0$ and hence, we need to choose $\alpha = \alpha(\delta, x^\delta)$ such that

$$C_{\alpha(\delta,y^\delta)}\delta \to 0 \qquad \text{as } \delta \to 0, \tag{3.28}$$

for convergence, which is equivalent to (3.16).

Combining the assertions of Theorem 3.17 and Theorem 3.18, we obtain the following result for the convergence of the regularized solutions.

Corollary 3.19. *Let the assumptions of Theorem 3.17 and Theorem 3.18 hold, as well as* (3.28)*. Then,* $x_{\alpha(\delta,y^\delta)} \to x^\dagger$ *as* $\delta \to 0$*.*

3.6 Convergence Rates

In the following we investigate the possible convergence rates of regularized solutions to inverse problems, which can be obtained under additional smoothness assumptions on the minimal norm solution $x^\dagger$ (and consequently on the exact data y). Classical conditions, so-called *source-conditions* are of the form

$$\exists\, w \in X : x^\dagger = (A^*A)^\mu w. \tag{3.29}$$

The power $\mu > 0$ of the operator A^*A can be defined via spectral theory, i.e.,

$$(A^*A)^\mu w = \sum_{n=1}^{\infty} \sigma_n^{2\mu} \langle w, u_n\rangle u_n.$$

This corresponds to our preliminary analysis above, since in this case we have $\langle x^\dagger, u_n\rangle = \sigma_n^{2\mu}\langle w, u_n\rangle$, i.e., the coefficients of $x^\dagger$ with respect to the n-th singular vector decay faster than $\sigma_n^{2\mu}$.

The rate to be obtained depends on the regularization scheme, i.e., on the specific choice of the function g_α. We assume that

$$t^\mu |t g_\alpha(t) - 1| \leq \omega_\mu(\alpha), \qquad \forall\ t > 0$$

holds, in the typical case $\omega_\mu(\alpha) = \alpha^\mu$, as e.g. seen above for truncated singular value decomposition.

Under this condition we can modify the estimate in the proof of 3.17 to

$$\begin{aligned}\|R_\alpha y - A^\dagger y\|^2 &\leq \sum_{n=1}^{\infty} (\sigma_n g_\alpha(\sigma_n) - 1)^2 \langle x^\dagger, u_n\rangle^2 \\ &= \sum_{n=1}^{\infty} (\sigma_n g_\alpha(\sigma_n) - 1)^2 \sigma_n^{2\mu} \langle w, u_n\rangle^2 \\ &\leq \omega_\mu(\alpha)^2 \|w\|^2,\end{aligned}$$

and hence,

$$\|x_\alpha - x^\dagger\| \leq \omega_\mu(\alpha)\|w\|. \tag{3.30}$$

Hence, by combining this estimate with (3.27), we obtain

$$\|x^\delta_{\alpha(\delta)} - x^\dagger\| \leq \omega_\mu(\alpha)\|w\| + C_\alpha \delta. \tag{3.31}$$

The minimum on the right-hand side is obtained if α is chosen such that $\frac{\omega_\mu(\alpha)}{C_\alpha}\|w\| = \delta$. In the typical case $\omega_\mu(\alpha) = \alpha^\mu$ and $C_\alpha = \alpha^{-1/2}$ this implies $\alpha(\delta) = \left(\frac{\delta}{\|w\|}\right)^{2/(2\mu+1)}$ and finally yields the estimate

$$\|x^\delta_{\alpha(\delta)} - x^\dagger\| \leq 2\delta^{2\mu/(2\mu+1)}\|w\|^{1/(2\mu+1)}. \tag{3.32}$$

Note that, no matter how large μ is, the rate of convergence $\delta^{2\mu/(2\mu+1)}$ is always of lower order than the order δ in the data noise. It is again a consequence of the ill-posedness of the problem that the error in the solution cannot be decreased to the same order as the error in the data, i.e., some information is always lost in the reconstruction. It can be shown (cf. [10]), that an error of order $\delta^{2\mu/(2\mu+1)}$ is the minimal error that can in general be obtained under a condition like (3.29), and hence, the regularization schemes are of *optimal order* in this case.

We also consider the error in the output in this case, i.e.,

$$\begin{aligned}\|Ax_\alpha - y\|^2 &\leq \sum_{n=1}^{\infty} \sigma_n^2 (\sigma_n g_\alpha(\sigma_n) - 1)^2 \langle x^\dagger, u_n\rangle^2 \\ &= \sum_{n=1}^{\infty} (\sigma_n g_\alpha(\sigma_n) - 1)^2 \sigma_n^{2(\mu+1)} \langle w, u_n\rangle^2 \\ &\leq \omega_{\mu+1}(\alpha)^2 \|w\|^2.\end{aligned}$$

Thus, in the above case of $w_\mu(\alpha) = \alpha^\mu$ and $\alpha(\delta) = \left(\frac{\delta}{\|w\|}\right)^{2/(2\mu+1)}$, we obtain the estimate

$$\|Ax_\alpha - y\| \leq \delta$$

and together with (3.26)

$$\|Ax^\delta_\alpha - y\| \leq (1+\gamma)\delta. \tag{3.33}$$

Hence, the error in the output is always of the same order as the noise level δ.

We mention that for most standard regularization methods, there exists a $\mu_0 > 0$ such that $\omega_\mu(\alpha) = c\alpha^\mu$ for $\mu \leq \mu_0$ and $\omega_\mu(\alpha) = c\alpha^{\mu_0}$ for $\mu > \mu_0$. This implies that no rate better than $\delta^{2\mu_0/(2\mu_0+1)}$ can ever be achieved with such a method. The number μ_0 is called

qualification of the method, e.g., Tikhonov regularization has qualification $\mu_0 = 0$. We refer to [10, Chapter 4.2] for a more detailed discussion of saturation.

The observation that the output error is always of the order δ motivates a simple, but widely used a-posteriori stopping rule, the *discrepancy principle*. One the one hand we have seen so far that the output error is of order δ, and on the other hand it is not necessary to look for a regularized solution such that

$$\|Ax_\alpha^\delta - y^\delta\| < \delta,$$

since also the exact solution $x^\dagger$ may have error δ in the output. Hence, the discrepancy principle is defined as the parameter choice

$$\alpha(\delta, y^\delta) = \sup\{\alpha > 0 \mid \|Ax_\alpha^\delta - y^\delta\| \leq \tau\delta\} \tag{3.34}$$

with

$$\tau > \sup_{\alpha>0, t\in[0,\|[\|A]} g_\alpha(t).$$

Note that this condition is satisfied in particular for $\tau = \gamma + 1$. It can be shown that an a-posteriori choice of the regularization parameter via the discrepancy principle yields indeed a convergent regularization method of optimal order (cf. [10, Chapter 4.3]). We will meet the discrepancy principle below in several regularization methods for nonlinear inverse problems.

Chapter 4

Tikhonov-type Regularization

In this section we investigate Tikhonov regularization and related schemes more closely. In general, we shall now consider a nonlinear operator equation of the form

$$F(x) = y, \tag{4.1}$$

where $F : X \to Y$ is a continuous nonlinear operator. The extension of the regularization method to the nonlinear case is not obvious, since one can neither carry out a singular value decomposition nor define an adjoint of a nonlinear operator. The generalization to the nonlinear case therefore needs a reformulation of Tikhonov regularization, which we shall discuss in the following.

4.1 Tikhonov Regularization

We start again from the Tikhonov regularization of a linear operator equation, which is determined by the solution of the equation

$$(A^*A + \alpha I)x_\alpha^\delta = A^*y^\delta.$$

It is easy to verify that this linear equation is the first order optimality condition of the quadratic optimization problem

$$J_\alpha(x) := \|Ax - y^\delta\|^2 + \alpha\|x\|^2 \to \min_{x\in X}. \tag{4.2}$$

Note that J_α is strictly convex, which follows from

$$J''(x)(\varphi,\varphi) = 2\|A\varphi\|^2 + 2\alpha\|\varphi\|^2 > 0$$

and hence, x_α^δ is the unique global minimizer of the functional J_α.

The characterization of the regularized solution as a minimizer of the functional (4.2) offers the possibility of an immediate generalization to the nonlinear case, since we can define a regularized solution via

$$x_\alpha^\delta \in \arg\min_{x\in X}\left[\|F(x) - y^\delta\|^2 + \alpha\|x - x^*\|^2\right]. \tag{4.3}$$

Here, $x^* \in X$ is a given prior, which might represent a-priori knowledge about the solution. Note that in the nonlinear case, the value $x^* = 0$ plays no special role, so we can in principle consider any prior x^*. Consequently, we must also adapt our definition of generalization solution to the nonlinear case:

Definition 4.1. We shall call $\overline{x} \in X$ *least-squares solution* of (4.1) if

$$\|F(\overline{x}) - y\| = \inf_{x \in X} \|F(x) - y\|.$$

A least-squares solution $x^\dagger$ is called x^*-minimum norm solution, if

$$\|x^\dagger - x^*\| = \inf\{ \|x - x^*\| \mid x \text{ is least squares solution of (4.1)} \}.$$

For general nonlinear operators, we cannot expect the functional

$$J_\alpha(x) := \|F(x) - y^\delta\|^2 + \alpha\|x - x^*\|^2 \tag{4.4}$$

to be convex and hence, minimizers need not be unique. Moreover, there may exist global as well as local minimizers, but we will only consider global minimizers as regularized solutions.

We start with differentiability properties of the functional J_α:

Proposition 4.2. *If F is Frechet-differentiable, then the functional $J_\alpha : X \to \mathbb{R}$ is is Frechet-differentiable with derivative*

$$J'_\alpha(x)\varphi = 2\langle F(x) - y, F'(x)\varphi\rangle + 2\alpha\langle x - x^*, \varphi\rangle. \tag{4.5}$$

Moreover, if F is twice Frechet-differentiable, then the functional $J_\alpha : X \to \mathbb{R}$ is Frechet-differentiable with second derivative

$$J''_\alpha(x)(\varphi_1, \varphi_2) = 2\langle F'(x)\varphi_1, F'(x)\varphi_2\rangle + 2\langle F(x) - y, F''(x)(\varphi_1, \varphi_2)\rangle + 2\alpha\langle\varphi_1, \varphi_2\rangle. \tag{4.6}$$

Again, we can use the first-order optimality condition to verify that a regularized solution satisfies

$$F'(x_\alpha^\delta)^*(F(x_\alpha^\delta) - y^\delta) + \alpha(x_\alpha^\delta - x^*) = 0, \tag{4.7}$$

the nonlinear analogue of the original equation. On the other hand, not every solution of (4.7) is necessarily a regularized solution, since it could as well be a local minimum, saddle-point, or even maximum of J_α.

So far, we have not yet considered the problem of existence of regularized solutions, which is not obvious in the nonlinear case. In order to prove existence, we need an additional condition, namely *weak sequential closedness* of the operator F:

$$F(x) = y \qquad \text{if } F(x_n) \rightharpoonup y, x_n \rightharpoonup x. \tag{4.8}$$

This assumption is no severe restriction for inverse problems, in particular every compact nonlinear operator is weakly sequentially closed.

Theorem 4.3. *Let $F : X \to Y$ be a continuous operator satisfying* (4.8)*. Then, there exists a minimizer $x_\alpha^\delta \in X$ of the functional J_α defined by* (4.4)*.*

Proof. We first consider the level sets $L_M := \{ x \in X \mid J_\alpha(x) \le M \}$. Since $J_\alpha(x^*) = \|F(x^*) - y^\delta\|^2 < \infty$, the set L_M is nonempty for M sufficiently large. Moreover, $x \in L_M$ implies $\alpha\|x - x^*\|^2 \le M$ and, due to the triangle inequality

$$\|x\| \le \|x_*\| + \sqrt{\frac{M}{\alpha}} =: R,$$

i.e., L_M is contained in a ball with radius R. Since balls in X are compact with respect to the weak topology, the sets L_M are weakly pre-compact.

Since J_α is bounded below by zero, its infimum is finite and thus, there exists a minimizing sequence x_n. Since $x_n \in L_M$ for n sufficiently large, we can extract a weakly convergent subsequence (again denoted by x_n) with some limit $\overline{x} \in X$. Moreover, the sequence $F(x_n)$ is bounded due to

$$\|F(x_n) - y^\delta\|^2 \leq M$$

and hence, there exists a weakly convergent subsequence (again denoted by the subscript n) $F(x_n) \to z \in Y$. Because of the weak sequential closedness, we conclude $z = F(\overline{x})$, and thus,

$$J_\alpha(\overline{x}) = \lim_{n\to\infty} J_\alpha(x_n) = \inf_{x\in X} J_\alpha(x),$$

i.e. $x_\alpha^\delta = \overline{x}$ is a minimizer of J_α. □

We now turn our attention to the stability properties of Tikhonov regularization for (4.1). In the linear case, we have derived a Lipschitz estimate for the regularization operators, which is not possible in the general nonlinear case. In the nonlinear case, we only obtain a weak stability in a set-valued sense:

Proposition 4.4. *Let $F : X \to Y$ be a continuous operator satisfying* (4.8)*. Moreover, let $y_n \in Y$ be a sequence such that $y_n \to y^\delta$ and let x_n be a corresponding sequence of minimizers of J_α with y^δ replaced by y_n. Then x_n has a weakly convergent subsequence and every weak accumulation point is a minimizer of J_α.*

Proof. Due to Theorem 4.3 we can find a sequence of minimizers x_n corresponding to the data y_n. Since

$$\|x_n - x_*^2\| \leq \frac{1}{\alpha}\|F(x_n) - y_n\|^2 + \|x_n - x^*\|^2 \leq \frac{1}{\alpha}\|F(x^*) - y_n\|^2$$

and since y_n converges to y^δ, x_n is contained in a ball with radius independent of n. Due to weak compactness we can extract a convergent subsequence. Now let x be a weak accumulation point of x_n, without restriction of generality we assume that $x_n \rightharpoonup x$. Since $\|F(x_n) - y_n\| \leq \|F(x^*) - y_n\|$ we also conclude boundedness of $F(x_n)$ and consequently existence of a weak subsequence with limit z, and the weak sequential closedness implies $z = F(x)$. Finally, from the weak lower semicontinuity of the square of the norm in Hilbert spaces we conclude

$$\begin{aligned} J_\alpha(x) &= \|F(x) - y^\delta\|^2 + \alpha\|x - x^*\|^2 \leq \liminf_n \|F(x_n) - y_n\|^2 + \alpha\|x_n - x^*\|^2 \\ &\leq \liminf_n \|F(x_\alpha^\delta) - y_n\|^2 + \alpha\|x_\alpha^\delta - x^*\|^2 \\ &= \|F(x_\alpha^\delta) - y_\delta\|^2 + \alpha\|x_\alpha^\delta - x^*\|^2 = J_\alpha(x_\alpha^\delta). \end{aligned}$$

Since x_α^δ is a minimizer of J_α, x must be a minimizer, too. □

Proposition 4.4 ensures that Tikhonov regularization has indeed a regularizing effect, i.e., the approximate problems are well-posed for $\alpha > 0$. The obvious next question is convergence of the regularization method with suitable choice of α in dependence on δ. Similar to stability, this convergence appears in a set-valued sense:

Theorem 4.5. *Let $y \in Y$ such that there exists a x^*-minimum norm solution $x^\dagger \in X$ with $F(x^\dagger) = y$. Let y^δ be noisy data satisfying (3.12) and let x_α^δ be a regularized solution satisfying (4.3). If $\alpha = \alpha(\delta, y^\delta)$ is chosen such that*

$$\alpha \to 0, \qquad \frac{\delta^2}{\alpha} \to 0 \qquad \text{as } \delta \to 0, \tag{4.9}$$

then there exists a strongly convergent subsequence $x_{\alpha_n}^{\delta_n}$ (with $\delta_n \to 0$) and the limit of each convergent subsequence is a x^*-minimum norm solution of (4.1).

Proof. Since x_α^δ is a minimizer of J_α, we conclude

$$\begin{aligned} \|x_\alpha^\delta - x^*\|^2 &\leq \frac{1}{\alpha}\|F(x_\alpha^\delta) - y^\delta\|^2 + \|x_\alpha^\delta - x^*\|^2 = \frac{1}{\alpha} J_\alpha(x_\alpha^\delta) \\ &\leq \frac{1}{\alpha} J_\alpha(\overline{x}) = \frac{1}{\alpha}\|F(x^\dagger) - y^\delta\|^2 + \|x^\dagger - x^*\|^2 \\ &\leq \frac{\delta^2}{\alpha} + \|x^\dagger - x^*\|^2. \end{aligned}$$

Since $\frac{\delta^2}{\alpha} \to 0$, it is bounded in particular, and hence, $\|x_\alpha^\delta - x^*\|$ is uniformly bounded with respect to δ, which allows to extract a weakly convergent subsequence. For $x_{\alpha_n}^{\delta_n}$ being a weakly convergent subsequence with limit $\overline{x}$, the above estimate yields

$$\|\overline{x} - x^*\|^2 \leq \limsup_n \|x_{\alpha_n}^{\delta_n} - x^*\|^2 \leq \limsup_n \frac{\delta_n^2}{\alpha_n} + \|x^\dagger - x^*\|^2 = \|x^\dagger - x^*\|^2$$

and

$$\|F(\overline{x} - y\|^2 \leq \limsup_n \|F(x_{\alpha_n}^{\delta_n}) - y^{\delta_n}\|^2 \leq \limsup_n (\delta_n^2 + \alpha_n \|x^\dagger - x^*\|^2) = 0.$$

Hence, $\overline{x}$ satisfies $F(\overline{x}) = y$ and, by the definition of the minimum norm solution

$$\|\overline{x} - x^*\| \leq \|x^\dagger - x^*\| = \inf\{\ \|x - x^*\| \mid x \text{ is least squares solution of (4.1)}\ \},$$

which implies that $\overline{x}$ is a minimum norm solution of (4.1).

It remains to verify strong convergence of $x_{\alpha_n}^{\delta_n}$. For this sake we expand

$$\|x_{\alpha_n}^{\delta_n} - \overline{x}\|^2 = \|x_{\alpha_n}^{\delta_n} - x^*\|^2 + \|\overline{x} - x^*\|^2 - 2\langle x_{\alpha_n}^{\delta_n} - x^*, \overline{x} - x^*\rangle.$$

Due to the weak convergence we know that

$$-2\langle x_{\alpha_n}^{\delta_n} - x^*, \overline{x} - x^*\rangle \to -2\|\overline{x} - x^*\|^2.$$

Moreover, we have concluded above that

$$\limsup_n \|x_{\alpha_n}^{\delta_n} - x^*\|^2 \leq \|\overline{x} - x^*\|^2,$$

and thus,

$$\limsup_n \|x_{\alpha_n}^{\delta_n} - \overline{x}\|^2 \leq \|\overline{x} - x^*\|^2 + \|\overline{x} - x^*\|^2 - 2\|\overline{x} - x^*\|^2 = 0,$$

which implies the strong convergence. □

Note that the convergence proof applies only to situations where the data y are attainable, i.e., $F(x^\dagger) = y$. If $\|F(x^\dagger) - y\| > 0$ a slightly different proof under changed conditions on the parameter choice has to be carried out, we refer to [4] for further details.

We finally consider convergence rates in the nonlinear setting. For the sake of simplicity we restrict our attention to the case corresponding to $\mu = \frac{1}{2}$ for the linear problem. In this case, the source condition $x^\dagger = (A^*A)^{1/2}w$ is equivalent to $x^\dagger = A^*p$ for

$$p = \sum_{n=1}^{\infty} \langle w, u_n \rangle v_n \quad \in Y.$$

The condition $x^\dagger = A^*p$ is easier to interpret from an optimization point of view. A mininum norm solution $x^\dagger$ is determined as a minimizer of the constrained problem

$$\frac{1}{2}\|x\|^2 \to \min_{x \in X}, \text{ subject to } Ax = y,$$

and it is natural to consider the associated Lagrangian

$$L(x; p) := \frac{1}{2}\|x\|^2 - \langle Ax, y \rangle.$$

It is easy to see that for $(x^\dagger, p)$ being a stationary point of the Lagrangian, $x^\dagger$ is a solution of the above constrained problem, i.e., a minimum norm solution. In the case of an ill-posed operator equation, the converse does not hold, since the constraint operator A^* is not surjective. Hence, the existence of a Lagrange multiplier is an additional smoothness condition on the exact solution $x^\dagger$. Since we always have $\frac{\partial}{\partial p} L(x^\dagger; p) = Ax^\dagger - y = 0$, it is clear that a stationary point p exists if and only if

$$0 = \frac{\partial}{\partial x} L(x^\dagger; p) = x - A^*p,$$

i.e., if and only if the source condition is satisfied.

Again, the optimization viewpoint allows an immediate generalization to the nonlinear case, where the Lagrangian is given by

$$L(x; p) = \frac{1}{2}\|x - x^*\|^2 - \langle F(x) - y, p \rangle.$$

Thus, the source condition becomes

$$0 = \frac{\partial}{\partial x} L(x^\dagger; p) = x^\dagger - x^* - F'(x^\dagger)^* p,$$

i.e.,

$$\exists\, p \in Y : \quad x^\dagger - x^* = F'(x^\dagger)^* p. \tag{4.10}$$

In order to prove a convergence rate we also assume that F' is Lipschitz continuous with module L and that the smallness condition

$$L\|p\| < 1 \tag{4.11}$$

holds.

Theorem 4.6. *In addition to the assumptions of Theorem 4.5 assume that* (4.10) *and* (4.11) *hold. Moreover, if* $\alpha(\delta, y^\delta) \sim \delta$*, there exists* $\delta_0 > 0$*, such that for all* $\delta < \delta_0$ *the estimate*

$$\|x_\alpha^\delta - x^\dagger\| \leq c\sqrt{\delta} \tag{4.12}$$

holds for some constant $c > 0$*.*

Proof. Since x_α^δ minimizes J_α we obtain

$$\|F(x_\alpha^\delta) - y^\delta\|^2 + \alpha\|x_\alpha^\delta - x^*\|^2 \leq \delta^2 + \alpha\|x^\dagger - x^*\|^2$$

and after a simple calculation, this inequality can be rewritten as

$$\|F(x_\alpha^\delta) - y^\delta\|^2 + \alpha\|x_\alpha^\delta - x^\dagger\|^2 \leq \delta^2 + 2\alpha\langle x^\dagger - x^*, x^\dagger - x_\alpha^\delta\rangle. \tag{4.13}$$

Now we insert the source condition (4.10) into the last term on the right-hand side to obtain

$$-2\alpha\langle x^\dagger - x^*, x^\dagger - x_\alpha^\delta\rangle = 2\alpha\langle p, F'(x^\dagger)(x_\alpha^\delta - x^\dagger)\rangle.$$

From a Taylor-expansion we obtain that

$$F'(x^\dagger)(x_\alpha^\delta - x^\dagger)\rangle = F(x_\alpha^\delta) - F(x^\dagger) + r_\alpha^\delta,$$

and due to the Lipschitz continuity of F' we have

$$\|r_\alpha^\delta\| \leq \frac{L}{2}\|x_\alpha^\delta - x^\dagger\|^2.$$

Hence, we may estimate

$$2\alpha\|\langle p, F'(x^\dagger)(x_\alpha^\delta - x^\dagger)\rangle\| \leq 2\alpha\|p\|\|F(x_\alpha^\delta) - F(x^\dagger)\| + \alpha L\|p\|\|x_\alpha^\delta - x^\dagger\|^2.$$

Combining this estimate with (4.13) we deduce

$$\|F(x_\alpha^\delta) - y^\delta)\|^2 + \alpha(1 - L\|p\|)\|x_\alpha^\delta - x^\dagger\|^2 \leq \delta^2 + 2\alpha\|p\|\|F(x_\alpha^\delta) - y\| \leq \delta^2 + 2\alpha\|p\|(\|F(x_\alpha^\delta) - y^\delta\| + \delta),$$

or, rewritten

$$\frac{1}{\alpha}(\|F(x_\alpha^\delta) - y^\delta)\| - \alpha\|p\|)^2 + (1 - L\|p\|)\|x_\alpha^\delta - x^\dagger\|^2 \leq \frac{\delta^2}{\alpha} + \alpha\|p\|^2 + 2\delta\|p\|.$$

If $c_1\delta \leq \alpha \leq c_2\delta$, we obtain (using nonnegativity of the first term on the left-hand side)

$$(1 - L\|p\|)\|x_\alpha^\delta - x^\dagger\|^2 \leq \frac{\delta}{c_1} + c_2\delta\|p\|^2 + 2\delta\|p\|,$$

and hence, (4.12) holds with

$$c = \sqrt{\frac{1 + c_1c_2\|p\|^2 + 2c_1\|p\|}{c_1(1 - L\|p\|)}}.$$

□

We finally mention that a general source condition can be generalized to

$$\exists\, p \in Y: \quad x^\dagger - x^* = (F'(x^\dagger)^*F'(x^\dagger))^\mu p, \tag{4.14}$$

and analogous convergence rate results to the linear case can be shown for $\mu > \frac{1}{2}$.

4.2 Construction of Tikhonov-type Regularization Methods

The idea of Tikhonov regularization can easily be generalized with respect to the choice of a regularization functional. With a nonnegative functional $J_R : X \to \mathbb{R}$ one could consider a "Tikhonov-type" regularization via

$$x_\alpha^\delta \in \arg\min_{x \in X} \left[\|F(x) - y^\delta\|^2 + \alpha J_R(x) \right]. \tag{4.15}$$

Such a method is a regularization if the regularization functional J_R has a suitable properties. In particular, if J_R is weakly lower semicontinuous in some topology $\mathcal{T}$, i.e.,

$$J_R(x) \leq \liminf_n J_R(x_n) \qquad \forall\ x_n \to_{\mathcal{T}} x,$$

and if the sub level sets of J_R are precompact in the topology $\mathcal{T}$, the results on existence, stability, and convergence for Tikhonov regularization can be carried over to (4.15) with minor modifications of the proofs, when convergence is considered with respect to the topology $\mathcal{T}$. Since the topology $\mathcal{T}$ need not correspond to the strong or weak topology in a Hilbert space, one can carry out regularization via (4.15) also if X is a metric space. We shall meet this situation for two Banach spaces in the Sections below and for a metric space of shapes in the last chapter.

From this generalization one observes that the main regularizing effect of Tikhonov regularization comes from the fact that the sub level sets of the functional

$$J_\alpha(x) = \|F(x) - y^\delta\|^2 + \alpha J_R(x)$$

are precompact in the topology $\mathcal{T}$, i.e., the regularization acts by compactification. In the case of a Hilbert space, the natural choice for the topology $\mathcal{T}$ is the weak topology, the fact that one finally even obtains strong convergence is a particularity. In a similar setup for Banach spaces one cannot expect strong convergence, as we shall see for total variation regularization below.

4.3 Maximum-Entropy Regularization

Maximum entropy regularization is a method of particular interest for the reconstruction of *probability density functions*, i.e., functions in the space

$$PDF(\Omega) := \{x \in L^1(\Omega) \mid x \geq 0, \int_\Omega x(t)\ dt = 1\}.$$

The (negative) entropy borrowed from physics and information theory is defined as the functional

$$E(x) := \int_\Omega x(t) \log x(t)\ dt, \qquad \forall x \in L^1(\Omega), x \geq 0, \int_\Omega x(t)\ dt = 1. \tag{4.16}$$

For a continuous operator $F : L^1(\Omega) \to Y$ with Y being some Hilbert space, we can consider the regularized problem

$$\|F(x) - y\|^2 + \alpha E(x) \to \min_{x \in PDF(\Omega)}. \tag{4.17}$$

The convergence analysis of maximum entropy regularization (cf. [8, 11]) can be related to the one for Tihonov regularization in Hilbert space with a simple trick: one can find a monotone function $\psi : \mathbb{R}^+ \to \mathbb{R}$ such that

$$E(x) = \int_\Omega x(t) \log x(t)\ dt = \int_\Omega \psi(x(t))^2\ dx.$$

Hence, with the operator

$$\Psi : L^2(\Omega) \to L^1(\Omega), \quad z \mapsto \psi^{-1}(z),$$

can be rewritten as

$$\|F(\Psi(z)) - y\|^2 + \alpha \int_\Omega z(t)^2\ dt \to \min_{z \in L^2(\Omega)}.$$

With suitable assumptions on the solution and the admissible set, one can verify that the new nonlinear operator $F \circ \Phi : L^2(\Omega) \to Y$ satisfies all needed properties for Tikhonov regularization and thus, the convergence (rate) analysis can be carried over, we refer to [11] for further details.

If a prior $x^* \in PDF(\Omega)$ is available, then one often uses the *relative entropy* (or Kullback-Leibler divergence)

$$E^*(x) := \int_\Omega x(t) \log \frac{x(t)}{x^*(t)}\ dt, \qquad \forall x \in L^1(\Omega), x \geq 0, \int_\Omega x(t)\ dt = 1, \tag{4.18}$$

the convergence analysis in this case is similar.

4.4 Total Variation Regularization

Total variation regularization is an approach originally introduced for image restoration (cf. [21]) with the aim of preserving edges in the image, i.e., discontinuities in the solution. Formally the total variation functional can be defined as

$$|u|_{TV} = \int_\Omega |\nabla u|\ dt, \qquad u \in C^1(\Omega).$$

A more rigorous definition is based on the dual form

$$|u|_{TV} := \sup_{\mathbf{g} \in C_0^\infty(\Omega)^d} \int_\Omega u \text{ div } \mathbf{g}\ dt. \tag{4.19}$$

The general definition of the space of functions of bounded variation $BV(\Omega)$ is

$$BV(\Omega) := \{\ u \in L^1(\Omega) \mid |u|_{TV} < \infty\ \}.$$

With this definition, the space $BV(\Omega)$ includes also discontinuous functions. Consider e.g. $\Omega = [-1, 1]$ and, for $R < 1$,

$$u^R(x) = \begin{cases} 1 & \text{if } |x| \leq R \\ 0 & \text{else.} \end{cases}$$

Then,

$$\int_\Omega u \text{ div } g\ dt = \int_{-R}^{R} \frac{dg}{dt}\ dt = g(R) - g(-R).$$

For $\|g\|_\infty \leq 1$, we have $g(R)-g(-R) \leq 2$ and it is easy to construct a function $g \in C_0^\infty([-1,1])$ such that $g(R) = 1, g(-R) = -1$ and $\|g\|_\infty \leq 1$. Hence,

$$|u|_{TV} := \sup_{g \in C_0^\infty([-1,1])} [g(R) - g(-R)] = 2.$$

In general, for a function u being equal to 1 in $D \subset\subset \Omega$ and $u = 0$ else, the total variation $|u|_{TV}$ equals the surface area (or curve length) of ∂D.

Total variation regularization is defined via the minimization problem

$$\|F(u) - y\|^2 + \alpha |u|_{TV} \to \min_{u \in BV(\Omega)}.$$

The convergence analysis (cf. [1]) is based on the compact embedding $BV(\Omega) \hookrightarrow L^p(\Omega)$, where $p > 1$ depends on the spatial dimension d. One can use this property to deduce that sub level sets of the regularized functional are compact in the strong topology of $L^p(\Omega)$, and if F is weakly sequentially closed in this topology, one can carry out an analogous convergence proof as for Tikhonov regularization.

In order to obtain further insight, we consider the formal optimality condition in the case $F = Id : BV(\Omega) \to L^2(\Omega)$, i.e., the classical case of denoising considered in [21]. By differentiating formally, we have

$$u - y^\delta = \alpha \text{div} \left(\frac{\nabla u}{|\nabla u|} \right).$$

If u is a continuously differentiable function, the term div $\left(\frac{\nabla u}{|\nabla u|}\right)$ is equal to the mean curvature of the level sets $\{u = \sigma\}$, $\sigma \in \mathbb{R}$. Hence, the optimality condition is a condition on the smoothness of the level sets only, there is no condition on the size of $|\nabla u|$.

Again by formal arguments, we can derive a dual problem for total variation minimization. Consider again the denoising case $F = Id : BV(\Omega) \to L^2(\Omega)$, then the minimization problem to solve is

$$\inf_u \left[\int_\Omega (u - y^\delta)^2 \, dt + \alpha |u|_{TV} \right] = \inf_u \sup_{\mathbf{g}} \left[\int_\Omega (u - y^\delta)^2 \, dt + 2\alpha \int_\Omega u \text{ div } \mathbf{g} \, dt. \right]$$

Under the assumption that we can exchange the inf and sup (in a suitable function space setting), we obtain

$$\sup_{\mathbf{g}} \inf_u \left[\int_\Omega (u - y^\delta)^2 \, dt + 2\alpha \int_\Omega u \text{ div } \mathbf{g} \, dt. \right]$$

The minimization over u is a strictly convex problem and its unique minimizer can be computed from the first-order optimality conditon as

$$u = y^\delta - \alpha \text{ div } \mathbf{g}.$$

Hence, after eliminating u we end up with the maximization problem

$$\sup_{\mathbf{g}} \left[\alpha^2 \int_\Omega (\text{ div } \mathbf{g})^2 \, dt + 2\alpha \int_\Omega (y^\delta - \alpha \text{ div } \mathbf{g}) \text{ div } \mathbf{g} \, dt \right]$$

Since we can add constant terms without changing the maximizer, the problem is equivalent to

$$- \int_\Omega (\alpha \text{ div } \mathbf{g} - y^\delta)^2 \, dx \to \max_{|\mathbf{g}|_\infty \leq \mathbf{1}}.$$

Now let $p := \alpha g$, then we end up with (using the fact that maximization of a functional is equivalent to minimization of the negative functional)

$$\int_\Omega (\ \mathrm{div}\ p - y^\delta)^2\ dt \to \min_{|p|_\infty \le \alpha}. \tag{4.20}$$

The minimization problem (4.20) is the dual problem of the total variation regularization, if we have computed a solution p, then the primal solution can be computed as $u = y^\delta - \ \mathrm{div}\ p$.

Motivated from the dual problem we can also consider the dual space of BV, namely

$$BV^* := \{\ q = \ \mathrm{div}\ p \mid p \in L^\infty(\Omega)\ \}$$

with the dual norm

$$\|q\|_{BV^*} := \inf\{\|p\|_\infty \mid q = \ \mathrm{div}\ p\}.$$

Note that (4.20) has the structure of a projection, namely it projects y^δ to the ball of radius α in the dual space BV^*.

The dual problem also allows further insight into the so-called stair-casing phenomenon, i.e., the fact that the total variation regularization favours piecewise constant regularized solutions. Consider for simplicity the one-dimensional case and let $\frac{df}{dx} = y^\delta$. Then, with $q = p - f$, the dual problem can be rewritten as

$$\int_\Omega \left(\frac{dq}{dt}\right)^2\ dt \to \min \qquad \text{subject to } -\alpha \le q + f \le +\alpha. \tag{4.21}$$

Consider formally the associated Lagrangian

$$\mathcal{L}(q;\lambda,\mu) = \int_\Omega \left[\left(\frac{dq}{dt}\right)^2 + \lambda(q + f - \alpha) - \mu(q + f + \alpha)\right]\ dt$$

for positive functions λ and μ. Then the optimality condition becomes

$$-2\frac{d^2q}{dt^2} + \lambda - \mu = 0$$

and moreover, the constraints

$$\lambda \ge 0, \quad \mu \ge 0, \quad -\alpha \le q + f \le \alpha$$

and the complementarity conditions

$$\lambda(q + f - \alpha) = 0, \quad \mu(q + f + \alpha) = 0$$

hold. Thus, we have three cases:

1. $q(t) + f(t) = \alpha$, which implies that $\mu(t) = 0$ and $2\frac{d^2q}{dt^2} = \lambda(t) \ge 0$.
2. $q(t) + f(t) = -\alpha$, which implies that $\lambda(t) = 0$ and $2\frac{d^2q}{dt^2} = -\mu(t) \le 0$.
3. $q(t) + f(t) \notin \{\alpha, -\alpha\}$, which implies $\lambda(t) = \mu(t) = 0$, and hence, $\frac{d^2q}{dt^2}(t) = 0$.

Note that the third case shows that q is linear in regions where $q + f \notin \{\alpha, -\alpha\}$ and hence, $u = -\frac{dq}{dt}$ is constant.

If y^δ is a piecewise constant function, then f is piecewise linear and thus, the cases 1. and 2. imply that $q = \pm\alpha - f$ is piecewise linear in the respective subregions, and thus, $u = -\frac{dq}{dt}$ is piecewise constant. This means that, together with case 3., u must (locally) either be piecewise constant or exactly equal to the image f. If f is very noise (ans includes fine scale oscillations), then it is very unlikely that u follows f, since the total variation of the noise is usually high. Hence, one has to expect a piecewise constant solution, which is confirmed by various numerical results. Consider for example the special case

$$y^\delta(x) = \begin{cases} 1 & \text{if } |x| \le R \\ 0 & \text{else.} \end{cases}$$

with $0 < \alpha < R < 1$. By a simple integration we obtain the anti-derivative

$$f(x) = \begin{cases} -R & \text{if } x \le -R \\ x & \text{if } -R \le x \le R \\ R & \text{if } x \ge R \end{cases}$$

Thus, the dual problem becomes

$$\int_\Omega \left(\frac{dq}{dt}\right)^2 \, dt \to \min$$

subject to

$$\begin{array}{ll} -\alpha \le q - R \le \alpha & \text{if } x \le -R \\ -\alpha \le q + x \le \alpha & \text{if } -R \le x \le R \\ -\alpha \le q + R \le \alpha & \text{if } x \ge R \end{array} .$$

Now let

$$\overline{q}(x) = \begin{cases} R - \alpha & \text{if } x \le -R \\ \frac{\alpha - R}{R} x & \text{if } -R \le x \le R \\ \alpha - R & \text{if } x \ge R \end{cases} ,$$

Then $\overline{q}$ satisfies the constraints and

$$\int_\Omega \left(\frac{dq}{dt}\right)^2 \, dt = \frac{2(R-\alpha)^2}{R}.$$

Moreover, for arbitrary q satisfying the constraints we obtain

$$(R - \alpha) - (\alpha - R) \le q(-R) - q(R) = -\int_{-R}^{R} \frac{dq}{dt} \, dt$$

Hence, from the Cauchy-Schwarz inequality we deduce

$$\int_\Omega \left(\frac{dq}{dt}\right)^2 \, dt \ge \int_{-R}^{R} \left(\frac{dq}{dt}\right)^2 \, dt \ge \frac{1}{2R} \left(\int_{-R}^{R} \left(\frac{dq}{dt}\right) \, dt\right)^2 \ge \frac{2(R-\alpha)^2}{2R}.$$

This shows that $\overline{q}$ is a minimizer of the dual problem and consequently, the regularized solution is given by

$$u(x) = -\frac{d\overline{q}}{dx} = \begin{cases} 0 & \text{if } x \le -R \\ 1 - \frac{\alpha}{R} & \text{if } -R \le x \le R \\ 0 & \text{if } x \ge R \end{cases}$$

Thus, the regularized solution is discontinuous and even has the same discontinuity set as the data, but shrinks the height (in a monotone way with α). It is easy to see by analogous reasoning that for $R < \alpha$, the regularized solution is $u \equiv 0$, i.e., α marks a critical size below which features in the solution will be eliminated. Such small features are usually due to noise, so the regularization eliminates really parts that one would consider to be noise, while it maintains important discontinuities. A nice detailed discussion of the properties of solutions in total variation regularization can be found in the book by Meyer [19].

Chapter 5

Iterative Regularization

In the following we discuss the basic principles of iterative regularization. We shall carry out a detailed analysis for the possibly simplest method, namely Landweber iteration, which allows to present the main ideas and properties avoiding technical details. We will then discuss some speficic properties of more advanced iteration methods like conjugate gradients and Newton-type methods.

5.1 Landweber Iteration

We start again from the linear operator equation (3.4), respectively from the associated Gaussian normal equation, which can be rewritten equivalently as

$$x = x - \tau A^*(Ax - y)$$

for any $\tau \in \mathbb{R}$. This form motivates a simple fixed-point iteration, namely

$$x^{k+1} = x^k - \tau A^*(Ax^k - y) = (I - \tau A^* A)x^k + \tau A^* y, \tag{5.1}$$

which is called *Landweber iteration*. In the linear case, the standard choice of the initial value is $x^0 = 0$.

Note that the Landweber iteration is equivalently a gradient method for the least-squares problem

$$\frac{1}{2}\|Ax - y\|^2 \to \min_{x \in X},$$

and from well-known results on gradient methods one can conclude that the least-squares functional is decreasing during the iteration if τ is sufficiently small.

Exact Data

We can again employ the singular value decomposition to obtain an equivalent form of (5.1) as

$$\sum_{n=1}^{\infty} \langle x^{k+1}, u_n \rangle u_n = \sum_{n=1}^{\infty} (1 - \tau\sigma_n^2)\langle x^k, u_n \rangle u_n + \tau\sigma_n \langle y, v_n \rangle u_n.$$

Due to linear independence, this yields the recursion

$$\langle x^{k+1}, u_n \rangle = (1 - \tau\sigma_n^2)\langle x^k, u_n \rangle + \tau\sigma_n \langle y, v_n \rangle$$

for the n-th coefficient. This recursion can be solved to obtain

$$\begin{aligned}\langle x^k, u_n\rangle &= \tau\sigma_n\langle y, v_n\rangle \sum_{j=1}^{k}(1-\tau\sigma_n^2)^{k-j} = \frac{\left(1-(1-\tau\sigma_n^2)^{k-1}\right)}{\tau\sigma_n^2}\tau\sigma_n\langle y, v_n\rangle \\ &= \left(1-(1-\tau\sigma_n^2)^{k-1}\right)\frac{1}{\sigma_n}\langle y, v_n\rangle.\end{aligned}$$

If $|1-\tau\sigma_n^2| < 1$, then $(1-\tau\sigma_n^2)^{k-1} \to 0$ and hence,

$$\langle x^k, u_n\rangle \to \frac{1}{\sigma_n}\langle y, v_n\rangle = \langle x^\dagger, u_n\rangle.$$

Hence, we need that $0 < \tau\sigma_n^2 < 2$ for all n in order obtain convergence of the coefficients in the singular value expansion. Since $\sigma_1 = \max_n \sigma_n = \|A\|$, this means we have to choose $0 < \tau < \frac{2}{\|A\|^2}$.

Note that under this condition on τ we can also guarantee the decay of the least-squares functional, since

$$\begin{aligned}\|Ax^{k+1}-y\|^2 &= \|Ax^k-y\|^2 + \tau^2\|AA^*(Ax^k-y)\|^2 - 2\tau\langle Ax^k-y, AA^*(Ax^k-y)\rangle \\ &= \|Ax^k-y\|^2 + \tau\left(\tau\|AA^*(Ax^k-y)\|^2 - \|A^*(Ax^k-y)\|^2\right) \\ &\leq \|Ax^k-y\|^2 + \tau\|A^*(Ax^k-y)\|^2 \underbrace{\left(\tau\|A\|^2-1\right)}_{\leq 0} \leq \|Ax^k-y\|^2.\end{aligned}$$

If we interpret $\alpha := \frac{1}{k}$ as the regularization parameter, then we have

$$g_\alpha(\sigma) = \left(1-(1-\tau\sigma^2)^{1/(\alpha-1)}\right)\frac{1}{\sigma}.$$

Clearly, for $\tau\sigma^2 < 1$, $g_\alpha(\sigma)$ converges to σ as $\alpha = \frac{1}{k} \to 0$, and hence, from the results of Chapter 3 we conclude that $x^k \to x^\dagger$ as $k \to \infty$. We have also seen that the speed of convergence is arbitrarily slow unless the exact data satisfy additional smoothness assumptions.

In order to obtain a speed of convergence, we assume that the solution satisfies the source condition $x^\dagger = A^*p$ for $p \in Y$. In this case we have

$$\langle x^k - x^\dagger, u_n\rangle = (1-\tau\sigma_n^2)^{k-1}\langle x^\dagger, u_n\rangle = \sigma_n(1-\tau\sigma_n^2)^{k-1}\langle p, v_n\rangle.$$

The positive function $r(\sigma) := \sigma(1-\tau\sigma^2)^{k-1}$ has a unique maximum in the interval $(0, \sqrt{\frac{2}{\tau}})$ at $\overline{\sigma} = \frac{1}{\sqrt{\tau(2k-1)}}$ with and hence,

$$|\langle x^k - x^\dagger, u_n\rangle| \leq r(\overline{\sigma})\|p\| \leq \frac{1}{\sqrt{\tau(2k-1)}}\|p\|.$$

This implies that the error decays as

$$\|x^k - x^\dagger\| = \mathcal{O}(\frac{1}{\sqrt{k}}).$$

Noisy Data

In the case of noisy data, we can of course carry out the same iteration procedure, in order to clarify the dependence of the iterates upon the noise we shall write x_δ^k in the following

$$x_\delta^{k+1} = x_\delta^k - \tau A^*(Ax_\delta^k - y^\delta) = (I - \tau A^*A)x_\delta^k + \tau A^* y^\delta. \tag{5.2}$$

Since one only has to apply the continuous operators A and A^*, this method is well-defined and one could in principle iterate until $k \to \infty$. However, we have interpreted $\alpha = \frac{1}{k}$ as the regularization parameter, and since α should be positive (i.e., k finite) in the case of noisy data, one should terminate the iteration procedure by a stopping rule involving the noise and noise level. This means we only carry out (5.2) as long as $k < k_*(\delta, y^\delta)$. The regularized solution is then just $x_\delta^{k_*}$.

Once again, we use the singular value decomposition to write the error as

$$\begin{aligned}\langle x_\delta^k - x^\dagger, u_n\rangle &= \left(1 - (1-\tau\sigma_n^2)^{k-1}\right)\frac{1}{\sigma_n}\langle y^\delta, v_n\rangle - \frac{1}{\sigma_n}\langle y, v_n\rangle \\ &= \left(1 - (1-\tau\sigma_n^2)^{k-1}\right)\frac{1}{\sigma_n}\langle y^\delta - y, v_n\rangle + (1-\tau\sigma_n^2)^{k-1}\frac{1}{\sigma_n}\langle y, v_n\rangle.\end{aligned}$$

The second term in the sum decays exponentially to zero as $k \to \infty$, we have

$$(1-\tau\sigma_n^2)^{k-1}\frac{1}{\sigma_n}\langle y, v_n\rangle = (1-\tau\sigma_n^2)^{k-1}\langle x^\dagger, v_n\rangle.$$

For $k > 1$, the absolute value of the first term can be estimated by

$$\left(1 - (1-\tau\sigma_n^2)^{k-1}\right)\frac{1}{\sigma_n}|\langle y^\delta - y, v_n\rangle| = \tau\sigma_n \sum_{j=0}^{k-2}(1-\tau\sigma_n^2)^j |\langle y^\delta - y, v_n\rangle| \le \tau\sigma_n k\delta.$$

Now we can take a closer look at the error between the exact and regularized solutions, whose components in the singular value decomposition can be estimated using the above arguments as

$$|\langle x_\delta^{k_*(\delta)} - x^\dagger, u_n\rangle| \le \tau\sigma_n k_*\delta + (1-\tau\sigma_n^2)^{k-1}\|x^\dagger\|.$$

If, as $\delta \to 0$, we choose the stopping index such that $k_*(\delta) \to \infty$ and $k_*(\delta)\delta \to 0$, then all components converge to zero and hence, $x_\delta^{k_*(\delta)} \to x^\dagger$, i.e., the Landweber iteration is a convergent regularization method.

For iterative methods such as the Landweber iteration it is easy to use a-posteriori stopping rules such as the *discrepancy principle*,

$$k_*(\delta, y^\delta) := \inf\{k \in \mathbb{N} \mid \|Ax_k^\delta - y^\delta\| < \eta\delta\},$$

with $\eta \ge \frac{2}{2-\tau\|A\|}$. Roughly speaking, this means that we stop the iteration the first time the error reaches the same size as the noise level. Thus, in order to implement this stopping rule one just has to monitor the residual $Ax_k^\delta - y^\delta$ (which is computed anyway during the Landweber iteration) and compare its norm with the noise level. In order to understand how

the discrepancy principle works, we again look at the error during the iteration

$$\begin{aligned}
\|x_\delta^{k+1} - x^\dagger\|^2 - \|x_\delta^k - x^\dagger\|^2 &= \tau^2\|A^*(Ax_\delta^k - y^\delta)\|^2 - 2\tau\langle x_\delta^k - x^\dagger, A^*(Ax_\delta^k - y^\delta)\rangle \\
&= \tau^2\|A^*(Ax_\delta^k - y^\delta)\|^2 - 2\tau\langle Ax_\delta^k - y, Ax_\delta^k - y^\delta\rangle \\
&\leq \tau^2\|A\|\|Ax_\delta^k - y^\delta\|^2 - 2\tau\|Ax_\delta^k - y^\delta\|^2 - 2\tau\langle y^\delta - y, Ax_\delta^k - y^\delta\rangle \\
&\leq -\tau\|Ax_\delta^k - y^\delta\|\left((2 - \tau\|A\|)\|Ax_\delta^k - y^\delta\| - 2\delta\right) \\
&\leq -\frac{\tau}{\eta}\|Ax_\delta^k - y^\delta\|\left(\|Ax_\delta^k - y^\delta\| - \eta\delta\right).
\end{aligned}$$

Note that as long as $k < k_*$ one can guarantee that the right-hand side is negative and hence,

$$\|x_\delta^{k+1} - x^\dagger\|^2 < \|x_\delta^k - x^\dagger\|^2,$$

i.e. the error is decreased at least until the stopping index is reached. This is another good motivation for the use of the discrepancy principle as a stopping criterion for the Landweber iteration. One can indeed show (cf. [10]) that the discrepancy principle is a convergent regularization method. In the case of $x^\dagger$ satisfying the source condition $x^\dagger = A^*p$ for $p \in Y$, one can even show

$$\|x_\delta^{k_*} - x^\dagger\| = \mathcal{O}(\sqrt{\delta}),$$

an analogous result as for the continuous regularization methods.

Nonlinear Problems

We now discuss the generalization of the Landweber iteration for nonlinear problems of the form (4.1). The key observation is that the Landweber iteration in the nonlinear case is just a gradient descent method for the associated least-squares functional $\frac{1}{2}\|Ax - y\|^2$. Hence, we consider the least-squares functional

$$J(x) := \frac{1}{2}\|F(x) - y^\delta\|^2$$

for the nonlinear problem. Its derivative is given by

$$J'(x) = F'(x)^*(F(x) - y)$$

and hence, the Landweber iteration in the nonlinear case can be obtained as

$$x_\delta^{k+1} = x_\delta^k - \tau F'(x_\delta^k)^*(F(x_\delta^k) - y). \tag{5.3}$$

In the nonlinear case the choice of the initial value is of particular importance, x^0 plays the same role as the prior x^* in Tikhonov regularization and can (in the case of multiple solutions) also determine the particular solution to which the algorithm converges.

A detailed convergence analysis of the nonlinear Landweber iteration can be found in [10], here we shall take closer look on the behaviour of the error. Similar to the linear case we have

$$\begin{aligned}
&\|x_\delta^{k+1} - x^\dagger\|^2 - \|x_\delta^k - x^\dagger\|^2 \\
&\quad= \tau^2\|F'(x_\delta^k)^*(F(x_\delta^k) - y^\delta)\|^2 - 2\tau\langle x_\delta^k - x^\dagger, F'(x_\delta^k)^*(F(x_\delta^k) - y^\delta)\rangle \\
&\quad\leq \tau^2\|F'(x_\delta^k)\|^2(F(x_\delta^k) - y^\delta)^2 - 2\tau\langle F(x_\delta^k) - y, F(x_\delta^k) - y^\delta\rangle \\
&\qquad+2\tau\langle F(x_\delta^k) + F'(x_\delta^k)(x_\delta^k - x^\dagger) - F(x^\dagger), F(x_\delta^k) - y^\delta\rangle
\end{aligned}$$

Note that the first two terms are analogous to those appearing in the linear case, while the additional third term reflects the nonlinearity of the problem. Note that $F(x_\delta^k)+F'(x_\delta^k)(x_\delta^k - x^\dagger) - F(x^\dagger)$ is a Taylor-expansion of first order, and therefore, one can expect an estimate (locally around the solution) of the form

$$\|F(x_\delta^k) + F'(x_\delta^k)(x_\delta^k - x^\dagger) - F(x^\dagger)\| \leq c\|x_\delta^k - x^\dagger\|^2.$$

As we have seen several times before, the error $\|x_\delta^k - x^\dagger\|^2$ can be much larger than $\|F(x_\delta^k) - F(x^\dagger)\|$ and consequently such an estimate cannot help to obtain the decrease of the error from the above estimate. For the convergence analysis it turns out that a condition of the form

$$\|F(x_\delta^k) + F'(x_\delta^k)(x_\delta^k - x^\dagger) - F(x^\dagger)\| \leq c\|F(x_\delta^k) - F(x^\dagger)\| \tag{5.4}$$

with $c < \frac{1}{2}$ is sufficient, at least locally around the solution. Condition 5.4 restricts the nonlinearity of the operator F, it is called *tangential cone condition*. Such a nonlinearity condition somehow replaces the continuous invertibility of $F'(x^\dagger)$, which is a fundamental ingredient in the convergence analysis of iterative methods in the case of well-posed problems. Conditions on the nonlinearity are not just artificial inventions, it can be shown that a slightly weaker condition than (5.4) is actually necessary for the convergence of the Landweber iteration (cf. [22]).

5.2 Conjugate Gradient Methods

The conjugate gradient method is probably the most popular iteration scheme for linear equations involving symmetric positive definite linear operators. Therefore it seems tempting to consider the conjugate gradient method as an iterative regularization method. Without assuming that A is symmetric and positive definite, one can apply the conjugate gradient method only to the normal equation (CGNE)

$$A^*Ax = A^*y. \tag{5.5}$$

The most important property of conjugate gradient methods is that the residual is minimized in a Krylov subspace (shifted by x^0 in the case of non-homogeneous initial values), i.e.

$$\|Ax^k - y\| = \min\{\|Ax - y\| \mid x - x^0 \in \mathcal{K}_k(A^*(y - Ax^0), A^*A)\},$$

with

$$\mathcal{K}_k(z, B) := \{\ B^j z \mid j = 0, 1, 2, \ldots, k\ \}.$$

Note that the minimization of $\|Ax^k - y\|$ is an ill-posed problem in general, but it is regularized by a restriction to the compact finite-dimensional subspace $\mathcal{K}_k(A^*(y - Ax^0), A^*A)$.

The full algorithm reads as follows:

- Initialize x^0, $d^0 = y - Ax^0$, $p^1 = s^0 = A^*d^0$.

- For $k = 1, 2, \ldots$ and while $s^{k-1} \neq 0$ compute

$$\begin{aligned} q^k &= Ap^k \\ \alpha^k &= \|s^{k-1}\|^2 / \|q^k\|^2 \\ x^k &= x^{k-1} + \alpha^k p^k, \\ d^k &= d^{k-1} - \alpha^k q_k, \\ s^k &= A^* d^k, \\ \beta^k &= \|s^k\|^2 / \|s^{k-1}\|^2, \\ p^{k+1} &= s^k + \beta^k p^k. \end{aligned}$$

From the iteration procedure one observes that the conjugate gradient method is a nonlinear iteration scheme, which is the most fundamental difference to all regularization methods considered above. This means that the conjugate gradient method (even for linear ill-posed problems) is a nonlinear regularization method. The convergence analysis can therefore not be based on general results obtain for linear regularization methods as above, but has to be carried out by other means. Fortunately, the nonlinearity is not too strong to destroy the possibility of using the singular value decomposition for most parts of the analysis (cf. [10, 13]).

5.3 Newton-type Methods

The basic idea of the Newton method for a nonlinear equation like (4.1) is a local linearization. One step of the Newton method would involve the solution of the linear equation

$$F'(x^k)(x^{k+1} - x^k) = -(F(x^k) - y). \tag{5.6}$$

Since $F'(x^k)$ is not a regular linear operator in the case of an ill-posed problem, the equation for the update in the Newton method is a linear ill-posed problem itself and consequently x^{k+1} might not be well-defined.

A common approach to the construction of Newton-type methods for nonlinear ill-posed problems is to regularize (5.6) using a regularization method for linear ill-posed problems. For example we can apply linear Tikhonov-regularization (interpreting $x^{k+1} - x^k$ as the unknown), which yields

$$(F'(x^k)^* F'(x^k) + \alpha_k I)(x^{k+1} - x^k) = -F'(x^k)^*(F(x^k) - y), \tag{5.7}$$

the so-called *Levenberg-Marquardt method*. As an alternative one can also apply Tikhonov-regularization with a different prior, the most common form is to use the initial value x^0 as a prior throughout the whole iteration, which yields

$$(F'(x^k)^* F'(x^k) + \alpha_k I)(x^{k+1} - x^k) = -F'(x^k)^*(F(x^k) - y) + \alpha_k(x^0 - x^k), \tag{5.8}$$

called *iteratively regularized Gauss-Newton method*. The additional term on the right-hand side improves the stability of the method in some cases.

Note that in both methods α_k is not the regularization parameter of the iterative method, which is again the stopping index. The parameters α_k can be chosen to decay to zero as $k \to \infty$ in order to avoid overdamping of the Newton-type method. In the case of the iteratively regularized Gauss-Newton method, $\alpha_k \to 0$ is even necessary for the convergence

of the method, since for $\alpha_k \to \alpha_\infty > 0$ one would approximate the solution of a Tikhonov-regularized problem instead of the solution of (4.1).

The convergence analysis of Newton-type method is rather involved and shall therefore be omitted here, we only mention that similar (sometimes weaker) nonlinearity conditions as for the nonlinear Landweber iteration are used.

We finally mention that any suitable linear regularization method can be applied to the Newton step, in particular also linear iterative methods such as Landweber iteration or conjugate gradient methods, leading to methods called *Newton-Landweber* or *Newton-CG*. The above decay of α_k to zero corresponds to an increase of the inner iteration number in such cases.

5.4 Iterative Methods as Time-Discrete Flows

We finally discuss an interpretation of iterative regularization methods as time discretizations of a gradient flow. We start again from the Landweber iteration, which can be rewritten as

$$\frac{x^{k+1} - x^k}{\tau} = -F'(x^k)(F(x^k) - y).$$

If we interpret τ as a step size parameters and $x^k = x(k\tau)$ as time steps of some flow, the Landweber iteration corresponds to an explicit time discretization (forward Euler) of the flow

$$\frac{dx}{dt}(t) = -F'(x(t))^*(F(x(t)) - y), \tag{5.9}$$

i.e., asymptotical regularization.

From this correspondance to the flow (5.9) it seems natural to try other time discretizations. The implicit time discretization (backward Euler) yields the nonlinear equation

$$\frac{x^{k+1} - x^k}{\tau} = -F'(x^{k+1})(F(x^{k+1}) - y),$$

which is the optimality condition of the optimization problem

$$\|F(x) - y\|^2 + \frac{1}{\tau}\|x^{k+1} - x^k\|^2 \to \min_{x \in X},$$

well-known from Tikhonov regularization. The corresponding iterative procedure is therefore called *iterated Tikhonov regularization*. If we perform a semi-implicit time discretization, i.e., approximating $F'(x)^*$ explicitely and $F(x)$ by a first-order Taylor expansion around the last time step x^k we end up with the Levenberg-Marquardt method. The iteratively regularized Gauss-Newton method corresponds to a non-consistent semi-implicit time discretization.

From this motivation it is not surprising that general Runge-Kutta methods (even non-consistent ones) applied to the flow (5.9) yield convergent iterative regularization methods, as recently shown by Rieder [20].

Chapter 6

Discretization

In the following we turn our attention to the discretization of inverse problems. The motivation for considering discretization can be twofold. First of all one has to discretize the problem in any case in order to perform a computational scheme. On the other hand, the discretization acts as a regularization itself. We have already seen an example of regularization by discretization in Chapter 3, namely the truncated singular value decomposition. In this case the discretization is caused by the fact that we only use the singular values larger than some threshold, which is always a finite number for an ill-posed problem. Here we focus on a discussion of regularization by projection and then discuss some aspects related to the numerical solution of regularized problems.

6.1 Regularization by Projection

We start by considering a simple least-squares approximation strategy for the linear problem (3.4), by choosing a sequence of finite dimensional subspaces $X_1 \subset X_2 \subset \ldots \subset X$, whose union is dense is X. Now we consider the problem of computing a least-squares solution of minimal norm in the finite-dimensional space X_k. Such a solution x_k is in particular a solution of

$$\|Ax - y\|^2 \to \min_{x \in X_k} . \tag{6.1}$$

Thus, for all $h \in X_k$ we have

$$0 \leq \|A(x_k + h) - y\|^2 - \|Ax_k - y\|^2 = \|Ah\|^2 + 2\langle Ah, Ax_k - y\rangle.$$

Since h is arbitrary (and in particular its norm can be arbitrarily small), this identity can be satisfied only if

$$\langle h, A^*(Ax_k - y)\rangle = \langle Ah, Ax_k - y\rangle = 0$$

for all $h \in X_h$. With the projection operator $P_k : X \to X_k$ we can rewrite the identity as

$$P_k A^*(Ax_k - y) = P_k^* A^*(AP_k x_k - y) = A_k^*(A_k x_k - y) = 0,$$

where $A_k := AP_k$. We look for the discrete least-squares solution of minimal norm, and it is easy (by arguments similar to those of Section 3) to show that that $x_k = A_k^\dagger y$. Note that the range of the operator A_k is equal to the finite-dimensional space $A(X_k)$ and hence $\mathcal{R}(A_k)$ is closed, which implies the existence of a continuous inverse. For general choices of subspaces

X_k, this simple least-squares projection is not a convergent regularization method, since the norm of x_k can increase with k (i.e., $\|x_k\|$ is not uniformly bounded, cf. [10]). In general, it seems favourable to choose "smooth" basis functions in X_k, e.g., such that $X_k \subset \mathcal{R}(A^*)$, we refer to [10] for a detailed discussion.

So far we have discussed a projection method in the space X, but there is no reason why one should not investigate projection in the image space Y. For this sake we choose a sequence of finite-dimensional subspaces $Y_1 \subset Y_2 \subset \ldots \subset Y$, whose union is dense is Y. The so-called dual projection method seeks a minimum-norm solution $x_k \in X$ of

$$\langle Ax_k - y, z\rangle = 0 \qquad \forall\ z \in Y_k, \tag{6.2}$$

or equivalently, with the projector $Q_k : Y \to Y_k$,

$$Q_k A x_k = Q_k y. \tag{6.3}$$

The convergence analysis in the case of exact data is provided by the following result:

Theorem 6.1. *Let $y \in \mathcal{D}(A^\dagger)$, then there exists a unique solution x_k of minimal norm of the equation* (6.2)*. Moreover, $x_k := P_k x^\dagger$, where P_k is the orthogonal projector onto the subspace $X_k := T^* Y_k$. Consequently, if $\overline{\bigcup_k Y_k} = Y$, then $x_k \to x^\dagger$ as $k \to \infty$.*

Proof. Let $x_k := P_k x^\dagger$, then $\langle x_k - x^\dagger, A^* z\rangle = 0$ for all $z \in Y_k$ due to the properties of orthogonal projections. Because of

$$\langle x_k - x^\dagger, A^* z\rangle = \langle Ax_k - Ax^\dagger, z\rangle = \langle Ax_k - y, z\rangle,$$

x_k is a solution of (6.2). Let $x \in X$ be any other solution of (6.2), then

$$0 = \langle Ax - y, z\rangle = \langle AP_k x - y, z\rangle + \langle x - P_k x, A^* z\rangle$$

for all $y \in Y_k$. Since $x - P_k x$ is orthogonal to X_k, the second term vanishes, and hence $P_k x$ is also a solution of (6.2) and $\|P_k x\| \le \|x\|$. Thus, we may assume that $x \in X_k$. Then

$$0 = \langle Ax - y, z\rangle - \langle Ax_k - y, z\rangle = \langle x - x_k, A^* z\rangle$$

for all $z \in Y_k$. Hence, $x - x_k \in X_k$ is orthogonal to X_k, which implies $x = x_k$, and consequently x_k is the solution of minimal norm. □

We now turn our attention to the convergence analysis in the case of noisy data. So far, we have not introduced an explicit regularization parameter for regularization by discretization, it is hidden in the dimension of the subspace X_k. The actual regularization parameter is μ_k, the smallest singular value of the operator $A_k := Q_k A$. If the subspaces are chosen such that $Y_k \subset \mathcal{N}(A)^\perp$, then we always have $\mu_k > 0$. This allows to perform the following stability analysis:

Theorem 6.2. *Let $y \in \mathcal{D}(A^\dagger)$ and let y^δ be perturbed data satisfying $\|y - y^\delta\| \le \delta$. Moreover, let x_k^δ denote the minimum norm solution of* (6.2)*. If $\frac{\delta}{\mu_k} \to 0$ as $\delta \to 0$ and $k \to \infty$ ($\mu_k \to 0$), then $x_k^\delta \to x^\dagger$.*

Proof. As usual we use the triangle inequality to obtain

$$\|x_k^\delta - x^\dagger\| \leq \|x_k^\delta - x_k\| + \|P_k x^\dagger - x^\dagger\|.$$

Theorem 6.1 guarantees the convergence of the second term $\|P_k x^\dagger - x^\dagger\|$ to zero. The first term can be estimated via

$$\|x_k^\delta - x_k\| = \|A_k^\dagger Q_k(y^\delta - y)\| \leq \|A_k^\dagger\| \|Q_k(y^\delta - y)\| \leq \frac{\|y^\delta - y\|}{\mu_k} = \frac{\delta}{\mu_k},$$

and thus, it also converges to zero if $\frac{\delta}{\mu_k} \to 0$. □

The optimal choice of the subspaces Y_k is given as the span of the singular vectors $\{v_1, \ldots, v_k\}$, in this case we obtain the truncated singular value decomposition. Finally, we mention that the method of least-squares collocation for integral equations is a prominent example of the dual projection method.

6.2 Discretization of Regularized Problems

If one wants to solve inverse problems numerically, one always has to perform some discretization. The discretization can be a regularization itself, as considered in the previous section, or one can also consider a discretization of previously regularized problems. Here we consider the discretization of a problem with Tikhonov-regularization as an example.

The standard Galerkin approximation of Tikhonov-regularization consists in minimizing

$$\|Ax - y^\delta\|^2 + \alpha\|x\|^2 \to \min_{x \in X_k}$$

for a finite-dimensional subspace $X_k \subset X$. As usual for Galerkin approximations, the regularized solution $x_{\alpha,k}^\delta \in X_k$ is characterized by the variational equations

$$\langle Ax_{\alpha,k}^\delta - y, Az\rangle + \alpha\langle x_{\alpha,k}^\delta, z\rangle = 0 \qquad \forall\ z \in X_k$$

or, equivalently, as

$$x_{\alpha,k}^\delta = (A_k^* A_k + \alpha I)^{-1} A_k^* y^\delta,$$

where $A_k := AP_k$, P_k being the projector onto X_k.

The convergence of $x_{\alpha,k}^\delta$ could be carried out in two steps: First of all, convergence of $k \to \infty$ would yield convergence of $x_{\alpha,k}^\delta$ to x_α^δ due to the well-posedness of the regularized problem. As a second step one can then use the previous results on the convergence as $\alpha, \delta \to 0$. In practice, one is rather choosing α and k simultaneously, so that one is interested in conditions on α and k in dependence on δ. A general statement is that the discretization size k should not converge to infinity too slow compared to the convergence of α to zero, since otherwise the regularizing effect by projection dominates the Tikhonov regularization and the Galerkin approximation is not a regularization in general. We refer to [12] for a detailed discussion. Similary, one could use a dual projection method (as discussed in the previous section) for the regularized problem and analogous reasoning is possible (cf. [10] for a detailed discussion).

6.3 Numerical Solution of Discretized Problems

After discretizing the problem (either directly or after regularization), we end up with a system of the form

$$Kz = b, \qquad K \in \mathbb{R}^{n\times n}, b \in \mathbb{R}^n \tag{6.4}$$

to be solved for a coefficient vector $z \in \mathbb{R}^n$. For standard discretizations the matrix K is a discretization of an operator like A^*A or $A^*A + \alpha I$ and therefore one may assume that K is symmetric and positive definite. The scaling can be chosen such that the largest eigenvalue of K is related to the norm of A, i.e., of order one, and the smallest eigenvalue is related to the regularization parameter α (penalty parameter for Tikhonov regularization, truncation level for TSVD, ...). Thus, the condition number of K is proportional to $\frac{1}{\alpha}$ and since we are interested in situations $\alpha \to 0$ (as $\delta \to 0$), this condition number will be rather high in general.

Due to the ill-conditioning of the problem, standard Gauss elimination or factorization should not be the method of choice for the solution of (6.4). For problems of moderate size a possibility for a robust direct solution is a generalized Cholesky factorization in the form

$$K = LDL^T,$$

where L is a lower triangular matrix with diagonal entries $L_{ii} = 1$ and D is a diagonal matrix. In the generalized Cholesky factorization, the ill-conditioning only affects the matrix D, which can be inverted reasonably due to its diagonal structure.

For problems of large scale, a direct solution of (6.4) is not possible due to restrictions of memory and computation time. The alternative is an iterative solution, usually with the conjugate gradient method being the method of choice. Clearly, the high condition number and complicated eigenvalue pattern of the matrix K will cause an undesirably high number of CG iterations, and therefore one should find an appropriate preconditioner B and apply the CG method to

$$B^{-1}Kx = B^{-1}b.$$

As usual, the preconditioner should satisfy two main properties: B^{-1} should be easy to evaluate and B should approximate K. On the other hand, if B approximates K too well, then it will be ill-conditioned itself. Therefore it seems unlikely to find B such that all eigenvalues of $B^{-1}K$ cluster around 1, one expects at least two clusters, one around one and the other around the value of the regularization parameter.

Consider for example the case of Tikhonov regularization for a severly ill-posed problem. Under appropriate approximation we may expect that also the eigenvalues of the discretization of A^*A decay to zero exponentially and therefore only few of them will be large compared to the regularization parameter α. Since K corresponds to the discretization of $A^*A + \alpha I$ one may therefore expect an eigenvalue cluster at α and some larger eigenvalues. A simple idea for constructing a preconditioner is to act only on the large eigenvalues: Let (λ_i, u_i) be the eigensystem of K and let $m << n$ be the maximal index such that λ_i is significantly larger than α. Then we could choose

$$B = \sum_{j=1}^{m} \lambda_j u_j u_j^T + \sum_{j>m} u_j u_j^T.$$

Note that $\sum_{j<m} u_j u_j^T$ is just the projector on the subspace orthogonal to the span of $\{u_1, \ldots, u_m\}$ and we need not know the u_j to construct it. Since the spectral decomposition of K is

$$K = \sum_{j=1}^{m} \lambda_j u_j u_j^T + \sum_{j>m} \lambda_j u_j u_j^T$$

we obtain

$$B^{-1}K = \sum_{j=1}^{m} u_j u_j^T + \sum_{j>m} \lambda_j u_j u_j^T,$$

i.e., the preconditioned matrix has m eigenvalues equal to one and $n-m$ eigenvalues clustering at α. Once we have computed the eigenvalues and eigenvectors up to $j = m$, the evaluation of B^{-1} is cheap. The main computational effort is the construction of the first m eigenvalues and eigenvectors, but the exponential decay helps in this respect, since it ensures that m is very large. In this sense the preconditioning of severely ill-posed problems is even easier than the preconditioning of mildly ill-posed problems. Appropriate preconditioning techniques are often very problem-specific and are still subject of intense research, examples can be found in [5, 15, 16].

Chapter 7

Parameter Identification

In the following we discuss parameter identification problems in further detail, in particular their numerical solution. In the introductory examples of parameter identification we have observed a particular property of such problems, namely two types of unknowns, the parameter a and the state u. Parameter and state are linked via an equation, which we formally write as

$$e(u;a) = 0, \qquad e : X \times Q \to Z, \tag{7.1}$$

for Hilbert spaces X, Q, and Z. The equation (7.1), which usually represents a system of partial differential equations, is called *state equation*. The solution of the state equation for given a can be interpreted as the direct problem. In typical examples it is reasonable to assume that e is continuously Frechet differentiable and $\frac{\partial e}{\partial u}(u;a) : X \to Z$ is a continuous linear operator with continuous inverse. Hence, by the implicit function theorem we can conclude that (7.1) has a unique solution $u = u(a)$. It is therefore possible to introduce a well-defined operator

$$\Phi : Q \to X, \qquad a \mapsto u(a) \text{ solving (7.1)}.$$

Φ is called *parameter-to-solution map*. The data are related to the state in most examples via a linear observation operator $B : X \to Y$, such that $y = Bu$. The observation operator could either be the identity (distributed measurement), a restriction operator to part of the domain (partial distributed measurements), a trace operator to boundary values of the solution (boundary measurements), or a trace operator to final values of a solution in a time-dependent problem. By employing the parameter-to-solution map, we can also define a nonlinear operator $F := B \circ \Phi : Q \to Y$ and formulate the parameter-identification problem in a standard way as the nonlinear operator equation

$$F(a) = y. \tag{7.2}$$

If the operator F is injective, then the parameter a is identifiable. In the case of noisy data, one has several possibilities of a least-squares formulation for the problem. The most frequently used one is the *output least-squares formulation*

$$\|F(a) - y^\delta\|^2 \to \min_{a \in Q} \tag{7.3}$$

This formulation is equivalent to the constrained problem

$$\|Bu - y^\delta\|^2 \to \min_{(u;a) \in X \times Q} \qquad \text{subject to } e(u;a) = 0. \tag{7.4}$$

If we interpret the parameter identification problem as the coupled system

$$Bu = y, \qquad e(u;a) = 0 \tag{7.5}$$

instead, we could also consider the *total least-squares formulation*

$$\|Bu - y^\delta\|^2 + \|e(u;a)\|^2 \to \min_{(u,a)\in X\times Q}. \tag{7.6}$$

The nonlinear operator used in the total least-squares approach is defined on the product space as $\tilde{F} := (B, e) : X \times Q \to Y \times Z$ and it corresponds to the equation

$$\tilde{F}(u,a) = (Bu, e(u;a)) = (y^\delta, 0). \tag{7.7}$$

7.1 Derivatives and the Adjoint Method

For typical numerical solution methods one always needs to compute derivatives of the nonlinear operator F or $\tilde{F}$ and the associated least-squares functionals, respectively. Using the chain rule and the linearity of the observation operator we obtain that

$$F'(a) = B \circ \Phi'(a),$$

and since $e(\Phi(a);a) = 0$ we have

$$\frac{\partial e}{\partial u}(\Phi(a);a)\Phi'(a) + \frac{\partial e}{\partial a}(\Phi(a);a) = 0.$$

i.e., since we have assumed that $\frac{\partial e}{\partial u}$ is regular,

$$\Phi'(a) = -\frac{\partial e}{\partial u}(\Phi(a);a)^{-1}\frac{\partial e}{\partial a}(\Phi(a);a).$$

Altogether, the derivative of the operator F is given by

$$F'(a) = -B \circ \frac{\partial e}{\partial u}(\Phi(a);a)^{-1} \circ \frac{\partial e}{\partial a}(\Phi(a);a).$$

Hence, in order to evaluate the directional derivative $F'(a)h$, we have to solve the linearized problem

$$\frac{\partial e}{\partial u}(\Phi(a);a)[\Phi'(a)h] + \frac{\partial e}{\partial a}(\Phi(a);a)h = 0, \tag{7.8}$$

and then apply the observation operator to the solution. Note that the linearized problem (7.8) is a system of (linear) differential equations. Consider for example $X = H_0^1(\Omega)$, $Z = H^{-1}(\Omega)$ and

$$e(u;a) := -\operatorname{div}(a\nabla u) - f, \qquad B = Id : H_0^1(\Omega) \to L^2(\Omega), \tag{7.9}$$

then the derivatives are given by

$$\frac{\partial e}{\partial u}(u;a)v = -\operatorname{div}(a\nabla v), \qquad \frac{\partial e}{\partial a}(u;a)v = -\operatorname{div}(h\nabla u).$$

Hence, the linearized problem is the solution of the linear partial differential equation

$$-\operatorname{div}(a\nabla v) = \operatorname{div}(h\nabla u),$$

and its solution equals $\Phi'(a)h = v$. Thus, in order to compute a directional derivative, one has to solve a linear partial differential equation. In order to compute the full Frechet derivative $F'(a)$ one has to solve an equation for each h. With the formula for $F'(a)$ it is easy to compute the derivative of the output least-squares functional

$$J_O(a) := \|F(a) - y^\delta\|^2, \tag{7.10}$$

as

$$\begin{aligned} J_O'(a)h &= 2\langle F(a) - y^\delta, F'(a)h\rangle = 2\langle F'(a)^*(F(a) - y^\delta), h\rangle \\ &= -2\langle \frac{\partial e}{\partial a}(\Phi(a); a)^*(\frac{\partial e}{\partial u}(\Phi(a); a)^*)^{-1}B^*(F(a) - y^\delta), h\rangle. \end{aligned}$$

Hence,

$$J_O'(a)h = -2\frac{\partial e}{\partial a}(\Phi(a); a)^*(\frac{\partial e}{\partial u}(\Phi(a); a)^*)^{-1}B^*(F(a) - y^\delta).$$

Using this formula involving the adjoints of the derivatives of e, we can directly compute the gradient of the functional J_O as $J_O'(a) = -2\frac{\partial e}{\partial a}(\Phi(a); a)^* w$, where w is the solution of the adjoint equation

$$\frac{\partial e}{\partial u}(\Phi(a); a)^* w = B^*(F(a) - y^\delta). \tag{7.11}$$

In example (7.9) we can compute the adjoint via

$$\begin{aligned} \langle \frac{\partial e}{\partial u}(u; a)v, w\rangle &= -\int_\Omega \text{div } (a\nabla v)\ w\ dx = \int_\Omega a\nabla v \cdot \nabla w\ dx \\ &= -\int_\Omega \text{div } (a\nabla w)\ v\ dx = \langle v, \frac{\partial e}{\partial u}(u; a)^* w\rangle. \end{aligned}$$

Thus, the adjoint equation is the linear partial differential equation

$$-\text{ div } (a\nabla w) = u - y^\delta. \tag{7.12}$$

For complicated parameter identification problems, the direct computation of the adjoint is rather involved. An attractive alternative is a computation via the derivatives of the Lagrangian

$$L(u, a, w) := \|Bu - y^\delta\|^2 + \langle e(u; a), w\rangle. \tag{7.13}$$

It is easy to see that

$$\begin{aligned} \frac{\partial L}{\partial u}(u, a, w) &= 2B^*(Bu - y^\delta) + \frac{\partial e}{\partial u}(u; a)^* w \\ \frac{\partial L}{\partial a}(u, a, w) &= \frac{\partial e}{\partial a}(u; a)^* w \\ \frac{\partial L}{\partial w}(u, a, w) &= e(u; a). \end{aligned}$$

Thus, for given $a \in Q$, the solution $u \in X$ of $\frac{\partial L}{\partial w}(u, a, w) = 0$ equals $\Phi(a)$. Let, for given u and a, w be the solution of $\frac{\partial L}{\partial u}(u, a, w) = 0$, then

$$\begin{aligned} \frac{\partial L}{\partial a}(u, a, w) &= \frac{\partial e}{\partial a}(\Phi(a); a)^* w \\ &= -2\frac{\partial e}{\partial a}(\Phi(a); a)^*(\frac{\partial e}{\partial u}(\Phi(a); a)^*)^{-1}B^*(B\Phi(a) - y^\delta) \\ &= F'(a)^*(F(a) - y^\delta). \end{aligned}$$

Hence, we can compute the derivative of the least-squares functional directly from the Lagrangian by subsequently solving the equations $\frac{\partial L}{\partial w} = 0$, $\frac{\partial L}{\partial u} = 0$ and evaluating $\frac{\partial L}{\partial a}$.

In an analogous way we can compute derivatives of the operator $\tilde{F}$ as

$$\begin{aligned} \frac{\partial \tilde{F}}{\partial u}(u;a) &= (B, \frac{\partial e}{\partial u}(u;a)), \\ \frac{\partial \tilde{F}}{\partial a}(u;a) &= (0, \frac{\partial e}{\partial a}(u;a)). \end{aligned}$$

The derivative of the total least-squares functional

$$J_T(u,a) = \|\tilde{F}(u,a) - (y^\delta, 0)\|^2 = \|Bu - y^\delta\|^2 + \|e(u;a)\|^2$$

is given by

$$J_T'(u,a)(v,h) = 2\langle Bv, Bu - y^\delta\rangle + 2\langle \frac{\partial e}{\partial u}(u,a)v, e(u,a)\rangle + 2\langle \frac{\partial e}{\partial a}(u;a)v, e(u;a)\rangle.$$

The terms involved in the computation of the derivative J_T' are again the same as appearing in the derivative of J_O' .

7.2 Regularization

Under usual assumptions, one has to expect that a parameter identification problem is ill-posed (and most parameter identification problems are actually ill-posed). Therefore it is a natural first step to investigate the regularization of parameter identification problems. For this sake one needs to understand on which variable the regularization should act. From the viewpoint of (7.2) and (7.3) it seems clear that any regularization method for nonlinear ill-posed problems can be applied directly, with regularization acting on the only variable $a \in Q$. For the formulation (7.6) or (7.7) it is not obvious whether one should also incorporate regularization on u. However, it can be shown that such an additional regularization is not necessary due to the inherent well-posedness of the problem (respectively equation (7.1)) with respect to the state u.

Tikhonov Regularization

We start with the investigation of Tikhonov regularization. From (7.3), we arrive at the regularized problem

$$\|F(a) - y^\delta\|^2 + \alpha\|a - a^*\|^2 \to \min_{a\in Q} \tag{7.14}$$

or, equivalently,

$$\|Bu - y^\delta\|^2 + \alpha\|a - a^*\|^2 \to \min_{(u;a)\in X\times Q} \qquad \text{subject to } e(u;a) = 0 \tag{7.15}$$

The condition of weak sequential closedness of the operator F needed for the analysis of Tikhonov regularization is equivalent to the weak sequential closedness of the parameter-to-solution map Φ, because the continuous linear observation operator will preserve this property.

The Tikhonov regularization of the total least-squares formulation (7.6) is

$$\|Bu - y^\delta\|^2 + \|e(u;a)\|^2 + \alpha\|a - a^*\|^2 \to \min_{(u;a)\in X\times Q}. \tag{7.16}$$

The condition of weak sequential closedness of the operator $\tilde{F} := (B, e) : X \times Q \to Y \times Z$ is equivalent to weak sequential closedness of the equation operator e. A possible advantage of the output least-squares formulation is a natural way of dealing with perturbations in the equation. If, instead of $e(u; a)$, a perturbation $e(u; a) + f^\delta$ with $\|f^\delta\| \leq \delta$ is given, we can analyze convergence in the same way as for standard regularization.

In order to gain some insight into the structure of the regularized problem, we consider the example (7.9). For simplicity we consider $a - a^* \in H_0^1(\Omega)$ (which is indeed a regularization for $d = 1$) with the norm

$$\|b\|_{H_0^1} := \sqrt{\int_\Omega |\nabla b|^2 dt}.$$

The output least-squares formulation is equivalent to

$$\begin{aligned} \textstyle\int_\Omega (u - y^\delta)^2 dt + \alpha \int_\Omega |\nabla(a - a^*|^2 dt \to \min_{(u;a) \in H_0^1(\Omega) \times H_0^1(\Omega)} \\ \text{subject to} \qquad -\operatorname{div}(a \nabla u) = f \quad \text{in } \Omega. \end{aligned}$$

Every global minimizer of the Tikhonov functional is also a saddle-point of the Lagrangian

$$\mathcal{L}_\alpha(u, a, w) = \int_\Omega (u - y^\delta)^2 dt + \alpha \int_\Omega |\nabla(a - a^*|^2 dt + \int_\Omega (a \nabla u \cdot \nabla w - f w) dt, \tag{7.17}$$

where we have used Gauss' Theorem to convert the state equation to its weak form. Thus, the optimality condition becomes

$$\begin{aligned} 0 = \frac{\partial \mathcal{L}_\alpha}{\partial a}(u, a, w) &= -2\alpha \operatorname{div}(\nabla(a - a^*)) + \nabla u \cdot \nabla w, \\ 0 = \frac{\partial \mathcal{L}_\alpha}{\partial u}(u, a, w) &= -\operatorname{div}(a \nabla w) + 2(u - y^\delta), \\ 0 = \frac{\partial \mathcal{L}_\alpha}{\partial u}(u, a, w) &= -\operatorname{div}(a \nabla u) - f. \end{aligned}$$

Thus, the regularized solution can (at least in principle) be computed as the solution of a system of partial differential equations.

Total Variation Regularization

In several applications, the unknown parameter can be modeled as a piecewise constant function, but with unkown function values and unknown discontinuity sets. An example is the reconstruction of material parameters on domains that consist of a mixture of different materials (and each material is characterized by a specific scalar value). Under these conditions it is natural to use total variation regularization for the parameter identification problem, i.e., to minimize,

$$\|F(a) - y^\delta\|^2 + \alpha TV(a) \to \min_{a \in Q \subset BV(\Omega)}. \tag{7.18}$$

As we have seen above, the total variation functional favours piecewise constant solutions and the discontinuity set of the exact parameter is approximated well by the regularized solution.

Iterative Regularization by the Landweber Method

The simplest iterative regularization method, namely Landweber iteration, is given in the abstract setting as

$$a^{k+1} = a^k - \tau^k F'(a^k)^*(F(a^k) - y^\delta).$$

In terms of the funcional J_O and the associated Lagrangian we can rewrite the iteration as

$$a^{k+1} = a^k - \tau^k J_O'(a^k) = a^k - \tau^k \frac{\partial \mathcal{L}}{\partial a}(u^k, a^k, w^k)$$

for a suitable damping parameter $\tau^k > 0$, where $u^k = \Phi(a^k)$ is determined as the solution of

$$\frac{\partial \mathcal{L}}{\partial w}(u^k, a^k, w^k) = e(u^k; a^k)$$

and subsequently w^k as the solution of

$$\frac{\partial \mathcal{L}}{\partial u}(u^k, a^k, w^k) = 2B^*(Bu^k - y^\delta) + \frac{\partial e}{\partial u}(u^k; a^k)^* w^k = 0$$

Hence, the computation of one iteration step of the Landweber iteration consists of three parts: First of all, given a^k the state equation is solved to compute u^k, then the adjoint equation is solved to compute w^k and finally, $\frac{\partial \mathcal{L}}{\partial a}(u^k, a^k, w^k) = \frac{\partial e}{\partial a}(u^k, a^k)^* w^k$ is evaluated to determine the update in the iteration procedure. We again take a closer look at the iteration procedure for (7.9). The Lagrangian is given by

$$L(u, a, w) = \int_\Omega (u - y^\delta)^2 dt + \int_\Omega (a \nabla u \cdot \nabla v - vf)\ dx$$

and hence, in order to compute the update we have to solve the partial differential equations

$$\begin{aligned} 0 &= \frac{\partial \mathcal{L}}{\partial w}(u^k, a^k, w^k) = -\operatorname{div}(a^k \nabla u^k) - f \\ 0 &= \frac{\partial \mathcal{L}}{\partial u}(u^k, a^k, w^k) = -\operatorname{div}(a^k \nabla w^k) + 2(u^k - y^\delta). \end{aligned}$$

The update formula has to be carried out in the Hilbert space $H_0^1(\Omega)$, i.e., in weak form we have

$$\langle a^{k+1} - a^k, \varphi \rangle = -\frac{\tau^k}{2} J_O'(a^k)\varphi, \qquad \forall\ \varphi \in H_0^1(\Omega).$$

If we choose the same scalar product as above, then

$$\langle a^{k+1} - a^k, \varphi \rangle == \int_\Omega \nabla(a^{k+1} - a^k) \cdot \nabla \varphi dt = -\int_\Omega \operatorname{div}\ \nabla(a^{k+1} - a^k) dt$$

Hence, the update involves the solution of another partial differential equation of the form

$$-\operatorname{div}\ \nabla(a^{k+1} - ak) = -\tau^k \nabla u^k \cdot \nabla w^k$$

Note that once the gradient is known, it is also easy to use a quasi-Newton approach such as BFGS with little extra effort in order to obtain faster convergence.

7.2.1 Iterative Regularization by the Levenberg-Marquardt Method

The last regularization approach we discuss is the Levenberg-Marquardt method, where the iterates are computed from

$$(F'(a^k)^*F'^(a^k) + \alpha^k I)(a^{k+1} - a^k) = -F'(a^k)^*(F(a^k) - y^\delta),$$

which is equivalent to the minimization problem

$$J^k(a) := \|F(a^k) - y^\delta + F'(a^k)(a - ak)\|^2 + \alpha^k\|a - a^k\|^2 \to \min_{a\in Q}$$

This minimization is equivalent to

$$\begin{aligned} &\|Bu^k - y^\delta + Bv\|^2 + \alpha^k\|a - a^k\|^2 \min_{(u,a)\in X\times Q} \\ &\text{subject to } \frac{\partial e}{\partial u}(u^k, a^k)v + \frac{\partial e}{\partial a}(u^k, a^k)(a - a^k) = 0, \end{aligned}$$

where $u^k = \Phi(a^k)$ The optimality condition for this constrained problem are given by the system

$$\begin{aligned} 0 &= 2\alpha^k(a^{k+1} - a^k) + \frac{\partial e}{\partial a}(u^k, a^k)^* w^k \\ 0 &= 2B^*(Bu^k - y^\delta + Bv^k) + \frac{\partial e}{\partial a}(u^k, a^k)^* w^k \\ 0 &= \frac{\partial e}{\partial u}(u^k, a^k)v^k + \frac{\partial e}{\partial a}(u^k, a^k)(a^{k+1} - a^k) \end{aligned}$$

to be solved for $a^{k+1} \in Q$, $v^k \in X$, $w^k \in Z$. Hence, the realization of the Levenberg-Marquardt method enforces the solution of a linear system of differential equations, which is close to the linearization of the optimality conditions for Tikhonov regularization.

7.3 Large Scale Problems

We finally discuss the solution of large scale problems such as the examples of electrical impedance tomography and inverse scattering discussed before. In theory, one assumes to measure the full Dirichlet-to-Neumann map or the full far-field pattern, but in practice one clearly can measure only a finite number of evaluations of the maps. E.g., in impedance tomography, it is reasonable to measure $\Lambda_a(f_j)$ for $j = 1, \ldots, N$, and N being a very large number. This means we have to solve N state equations

$$\operatorname{div}(a\nabla u^j) = 0$$

with boundary values $u^j = f_j$. The general form corresponding to such a case is a state $u = (u^1, \ldots, u^N)$ with state equation

$$e(u; a) = (e_1(u^1; a), \ldots, e_N(u^N; a)) = 0 \tag{7.19}$$

and observation operator

$$Bu = (B_1u^1, \ldots, B_Nu^N). \tag{7.20}$$

The derivative of the associated output least-squares functional in this case can be computed again by the adjoint method, but since u^j only appears in the j-th equation we obtain a very peculiar structure. It is easy to see that

$$F'(a)^*(F(a) - y^\delta) = \sum_{j=1}^{N} \frac{\partial e_j}{\partial a}(u^j; a)^* w^j$$

where the adjoint state is the solution of

$$\frac{\partial e_j}{\partial u}(u^j; a)^* w^j + B_j^*(B_j u^j - y_j^\delta) = 0.$$

and the state is just determined from $e_j(u^j; a) = 0$.

The special structure of the derivative can be used to compute gradients with reasonable memory consumption. Note that if N is large and the discretization is reasonably fine, the degrees of freedom for the state variables u^j and the adjoint states w^j may produce a very high number of unknowns to be saved.Therefore, it seems advantageous not to compute and store all of them at the same time, but to compute them in a sequential way (or separately distributed on several processors). Such a computation is easy from the above form of the gradient, we start with $g_0 := 0$ and then use the recursion

$$g_j := g_{j-1} + \frac{\partial e_j}{\partial a}(u^j; a)^* w^j, \qquad j = 1, \dots, N$$

with states u^j and adjoint states w^j as above. In this way we only need the memory for u^1 and w^1, which can later be used for u^j and w^j subsequently.

With this way of computing the gradient it is straight-forward to realize the Landweber iteration, with the setting $a^{k,0} = a^k$ we compute

$$a^{k,j} = a^{k,j-1} - \tau^j \frac{\partial e_j}{\partial a}(u^{k,j}; a^k)^* w^{k,j}, \qquad j = 1, \dots, N$$

to obtain the new iterate $a^{k+1} = a^{k,N}$. Here $u^{k,j}$ and $w^{k,j}$ are the solutions of

$$e_j(u^{k,j}; a^k) = 0, \qquad \frac{\partial e_j}{\partial u}(u^{k,j}; a^k)^* w^{k,j} + B_j^*(B_j u^{k,j} - y_j^\delta) = 0.$$

Instead of the additive splitting in the computation of the update a^{k+1} one could also use a multiplicative splitting, i.e.,

$$a^{k,j} = a^{k,j-1} - \tau^j \frac{\partial e_j}{\partial a}(u^j; a^{k,j-1})^* w^j, \qquad j = 1, \dots, N$$

now with $u^{k,j}$ and $w^{k,j}$ being the solutions of

$$e(u^{k,j}; a^{k,j-1}) = 0, \qquad \frac{\partial e_j}{\partial u}(u^{k,j}; a^{k,j-1})^* w^{k,j} + B_j^*(B_j u^{k,j} - y_j^\delta) = 0.$$

This approach is called *Landweber-Kaczmarz method* (cf. [18]), for some practical problems one observes even better convergence properties for this method than for the simple Landweber iteration. We mention that the relation between Landweber and Landweber-Kaczmarz is of the same type as between Jacobi and Gauss-Seidel iteration for linear systems. The Kaczmarz-type approach also offers the possibility to perform Newton-type methods with reasonable memory consumption, for example one can perform a Levenberg-Marquardt type approach by freezing $u^{k,m}$ for $m \neq j$ and coupling the iterations in a cyclic way (cf. [6]).

Chapter 8

Shape Reconstruction Problems

In this section we shall deal with the solution of parameter identification problems, where the unknown variable is a shape or geometry in $\mathbb{R}^d$. Shapes can be considered as sets with regular boundary and therefore we may perform standard set operations like unions or intersections. However, there is no way to make a class of shapes into a linear space in general, but only with severe restrictions. An obvious way of solving a problem in a linear space instead of a problem on a class of shapes is to use parametrization (e.g. as piecewise graphs, by polar coordinates, or locally around a given shape). Since the parametrization is usually represented by a function on a fixed set, one can just minimize over all such functions in an appropriate Hilbert or Banach space. This allows to use standard methods as discussed above, but strongly limits the class of admissible shapes.

8.1 Shape Sensitivity Analysis

The main idea of shape sensitivity analysis is to consider "natural deformations" of shapes and inspect the corresponding variations of the objective functional. The general setup in the following is the minimization of

$$J(\Omega) \to \min_{\Omega \in \mathcal{K}},$$

where $\mathcal{K}$ is a suitable class of compact subsets of $\mathbb{R}^d$, with regular boundary.

There are two different ways of deriving shape sensitivities (both leading to the same result), namely via "direct deformations" or via the "speed method". We shall follow the latter, since this approach fits very well to the level set method, which we will discuss below as a possible solution method for shape optimization problems.

Before considering shapes we illustrate the idea of the speed method when applied to Gateaux-derivatives in linear spaces. In order to compute the directional derivative of a functional $J : \mathcal{U} \to \mathbb{R}$, we have so far considered the variation between the values of J at $\overline{u} \in \mathcal{U}$ and at its local deformation $\overline{u} + tv$. Alternatively, we could define $u(t) = \overline{u} + tv$ by

$$\frac{du}{dt} = v, \quad u(0) = \overline{u},$$

which is an initial value problem for an ordinary differential equation in $\mathcal{U}$. Using the chain rule, we can then compute

$$\frac{d}{dt} J(u(t)) = J'(u(t)) \frac{du}{dt} = J'(u(t))v.$$

In particular,

$$\frac{d}{dt}J(u(t))\Big|_{t=0} = J'(\overline{u})v,$$

i.e., we obtain the directional derivative at $\overline{u}$ by evaluating the time derivative of $J(u(t))$ at time $t = 0$.

In a similar way, we can define derivatives of shapes. Let $V : \mathbb{R}^d \to \mathbb{R}^d$ be a given velocity field and define $x(t)$ via

$$\frac{dx}{dt}(t) = V(x(t)), \quad x(0) = \overline{x}, \tag{8.1}$$

for each $\overline{x} \in \mathbb{R}^d$. We can then define the shape sensitivity

$$dJ(\overline{\Omega}; V) := \left(\frac{d}{dt}J(\Omega(t))\right)\Big|_{t=0},$$

where

$$\Omega(t) = \{x(t) \mid x(0) \in \overline{\Omega}\}.$$

Note that the main difference to derivatives in linear spaces is that the deformation defined by the ODE (8.1) is nonlinear, since V depends on x itself.

We start with some examples. Let $g : \mathbb{R}^d \to \mathbb{R}$ be a continuously differentiable function and define

$$J(\Omega) := \int_\Omega g(x)\ dx.$$

Then, by change of variables

$$\begin{aligned} J(\Omega(t)) &= \int_{\Omega(t)} g(x)\ dx \\ &= \int_\Omega g(x_y(t))|M_y|\ dy \end{aligned}$$

where $x_y(t)$ is defined by

$$\frac{dx_y}{dt}(t) = V(x_y, t), \quad x_y(0) = y \in \Omega$$

and $M_y = \det \frac{\partial x_y}{\partial y}$. Hence, the time derivative can be computed as

$$\frac{d}{dt}J(\Omega(t)) = \int_\Omega \overline{v}\nabla g(x_y)\frac{\partial x_y}{\partial t}|M_y|\ dy + \int_\Omega g(x_y)\frac{\frac{\partial M_y}{\partial t}M_y}{|M_y|}dy.$$

For the derivative of the determinant we have

$$\begin{aligned} \frac{\partial M_y}{\partial t} &= \frac{\partial}{\partial t}\left(\sum_{(i_1,\dots i_d)\in\Pi(d)}(-1)^{i_1+\dots+i_d}\prod_{k=1}^{d}\frac{\partial(x_y)_k}{\partial y_{i_k}}\right) \\ &= \left(\sum_{(i_k)\in\Pi(d)}(-1)^{\sum i_k}\sum_j \frac{\partial^2(x_y)_j}{\partial y_{i_j}\partial t}\prod_{l\neq j}^{d}\frac{\partial(x_y)_l}{\partial y_{i_l}}\right) \\ &= \sum_{(i_k)\in\Pi(d)}(-1)^{\sum i_k}\sum_j \frac{\partial V_j}{\partial y_{i_j}}\prod_{l\neq j}^{d}\frac{\partial(x_y)_l}{\partial y_{i_l}} \end{aligned}$$

For $t = 0$, we have $\frac{\partial x_y}{\partial y} = I, M_y = 1$, and this implies

$$\frac{\partial M_y}{\partial t} = \sum_j \frac{\partial V_j}{\partial y_j} = div(V)$$

As a consequence, we have

$$\begin{aligned}
\frac{d}{dt} J(\Omega(t))\Big|_{t=0} &= \int_\Omega \Big(\nabla g(x_y) \frac{\partial x_y}{\partial t} \Big) \Big|_{t=0} dy + \int_\Omega \Big(g(x_y)\ div\ V(x_y) \Big) \Big|_{t=0} dy \\
&= \int_\Omega \Big(\nabla g(y) V(y) + g(y)\ div\ V(y) \Big) dy \\
&= \int_\Omega div \Big(g(y) V(y) \Big) dy \\
&= \int_{\partial\Omega} g(y) V(y).n\ ds,
\end{aligned}$$

where n denotes the unit outer normal on $\partial\Omega$. I.e., the shape sensitivity is a linear functional fo V concentrated on $\partial\Omega$. Another key oberservation is that the shape sensitivity $J'(\Omega)V := \frac{d}{dt} J(\Omega(t))|_{t=0}$ depends on $V.n|_{\partial\Omega}$ only, while it is completely independent of the values for V inside Ω and of its tangential component. Consequently, we may directly consider variations of $\partial\Omega$ with a velocity $V = V_n.n$, where V_n is a scalar speed function. The shape sensitivity then becomes

$$J'(\Omega)V_n = \int_{\partial\Omega} g.V_n\ ds.$$

The statement that the shape sensitivity is a linear functional of $V.n$ only holds for very general classes of objective functionals, it is usually known as the "Hadamard-Zolésio Structure Theorem". The independence of the shape sensitivity on tangential components is clear from geometric intuition, since those components correspond to a change of parametrization only. The independence on values of V in the interior of Ω seems obvious, too, since they do not change the domain of integration in the objective functional.

In most typical applications of shape optimization, the objective functional depends on a state variable u that satisfies a partial differential equation related to Ω. This relation can arise in several ways, e.g.

1. u solves a partial differential equation in a domain $\Omega \subset\subset D$, and $\partial\Omega$ is the discontinuity set for some of the parameters. A simple example is the optimal design of two conductive materials, where the conductivity a takes two different values, i.e.,

   $$a(x) = \begin{cases} a_1 & x \in \Omega \\ a_2 & x \in D \backslash \Omega. \end{cases}$$

   A typical shape optimization problem consists in the optimization of some functional $J(\Omega) = \tilde{J}(u_\Omega)$, where u_Ω solves

   $$-div\ (a \nabla u_\Omega) = 0.$$

2. u solves a partial differential equation in Ω and satisfies a boundary condition on $\partial\Omega$.

3. u solves a partial differential equation on the surface of $\partial\Omega$.

The general structure fo such problems is

$$J(\Omega) = \tilde{J}(u_\Omega, \Omega) \to \min_\Omega$$

subject to

$$e(u_\Omega, \Omega) = 0,$$

where e denotes the partial differential equation. In this case we have to use the chain rule and an implicit function theorem to compute the shape sensitivity. Let $\Omega(t)$ be as above and let $u(t)$ denote the solution of

$$e(u(t), \Omega(t)) = 0$$

with $\Omega(t)$ given. Then the shape sensitivity of J is given by

$$\begin{aligned} J'(\Omega)V &= \frac{d}{dt} J(\Omega(t))\Big|_{t=0} \\ &= \frac{d}{dt}\Big(\tilde{J}(u(t), \Omega(t))\Big)\Big|_{t=0} \\ &= \frac{\partial \tilde{J}}{\partial u}(u(0), \Omega(0))u'(0) + \frac{\partial \tilde{J}}{\partial \Omega}(u(0), \Omega(0))V. \end{aligned}$$

Here $\frac{\partial \tilde{J}}{\partial u}$ denotes the (Gateaux-)derivative of $\tilde{J}$ with respect to u (for Ω fixed) and $\frac{\partial \tilde{J}}{\partial \Omega}$ denotes the shape sensitivity of $\tilde{J}$ with respect to Ω (for u fixed). Due to the chain rule we obtain for $u'(0) = \frac{d}{dt} u(t)|_{t=0}$ the equation

$$0 = \frac{d}{dt} e(u(t), \Omega(t)) = \frac{\partial e}{\partial u}(u(t), \Omega(t))u'(t) + \frac{\partial e}{\partial \Omega}(u(t), \Omega(t))V.$$

Here, $\frac{\partial e}{\partial \Omega}(u, \Omega(t))V = \frac{d}{dt} e(u, \Omega(t))$, for u fixed, i.e., it means a generalization of shape sensitivities from functionals to operators. The function $u' = u'(0)$ is usually called "shape derivative".

We shall discuss the computation of shape derivatives for two examples. First, consider the maximization of current for a conductive material. The objective is given by

$$J(\Omega) = -\int_\Gamma a \frac{\partial u_\Omega}{\partial n}\, ds,$$

where $\Gamma \subset D, \Omega \subset\subset D$ and u solves

$$-div\ (a\nabla u) = f, \quad \text{in } D$$

with homogeneous boundary values $u = 0$ on ∂D. Here, f is a given function and a is defined as above, i.e.

$$a(x) = \begin{cases} a_1 & x \in \Omega \\ a_2 & x \in D \backslash \Omega. \end{cases}$$

The shape sensitivity is then given by (note that $\Omega \subset\subset D$ and thus $a = a_2$ on ∂D)

$$J'(\Omega)V = -\int_\Gamma a_2 \frac{\partial u'}{\partial n} ds,$$

where u' is the shape derivative corresponding to the above state equation. In order to compute the shape derivative u', we consider the state equation in its weak form, i.e. we seek $u \in H_0^1(D)$ satisfying

$$\int_D a\nabla u\nabla v \; dx = \int_D fv \; dx \quad \forall v \in H_0^1(D)$$

We can write the left-hand side as

$$< v, e(u,\Omega) >= \int_D a_2\nabla u\nabla v \; dx + \int_\Omega (a_1 - a_2)\nabla u\nabla v \; dx.$$

The derivative with respect to u is given by

$$\frac{\partial e}{\partial u}(u,\Omega)u' = \int_D a_2\nabla u'\nabla v \; dx + \int_\Omega (a_1 - a_2)\nabla u'\nabla v \; dx = \int_D a\nabla u'\nabla v \; dx.$$

In order to compute the derivative with respect to Ω, we can use the above results on shape sensitivities for the functional $\int_\Omega g \; dx$, now with $g = (a_1 - a_2)\nabla u.\nabla v$. Thus,

$$\frac{\partial e}{\partial \Omega}(u,\Omega)V = \int_{\partial\Omega} \Big((a_1 - a_2)\nabla u.\nabla v\Big)V.n \; ds \quad \forall v \in H_0^1(D).$$

As for standard optimal design problems, we can also employ the adjoint method to compute the shape sensitivity. For this sake, let $u^* \in H_0^1(D)$ be the unique weak solution of

$$\int_\Gamma a_2 \frac{\partial w}{\partial n} dx = \int_D a\nabla w\nabla u^* \; dx \quad \forall w \in H_0^1(D).$$

Then we obtain

$$-\int_\Gamma a_2 \frac{\partial u'}{\partial n} ds = -\int_D a\nabla u'\nabla u^* \; dx = \int_{\partial\Omega} \Big((a_1 - a_2)\nabla u.\nabla u^*\Big)V.n \; ds,$$

i.e., the shape sensitivity is again a functional of $V.n$ concentrated on $\partial\Omega$.

Our second example is the shape derivative for a state equation with Dirichlet boundary condition, i.e.

$$\begin{aligned} \Delta u &= f \quad \text{in } \Omega \\ u &= 0 \quad \text{on } \partial\Omega. \end{aligned}$$

It is easy to show that

$$\Delta u' = 0 \quad \text{in } \Omega.$$

For the boundary condition, let $y \in \partial\Omega$ and let $\frac{dx}{dt}(t) = V(x(t)), x(0) = y$. Then $u(x(t)) = 0$ for all t and thus

$$\frac{d}{dt}u(x(t)) = u'(x(t)) + \nabla u(x(t)).V(x(t)) = 0.$$

Hence, u' satisfies

$$u' = -\nabla u.V \quad \text{on } \partial\Omega.$$

We finally notice that second derivatives, so-called shape Hessians can be computed by applying the same technique as for shape sensitivities to $J'(\Omega)V$, now with a second velocity W.

8.2 Level Set Methods

Level set methods recently received growing attention in shape optimization due to their capabilities of solving shape optimization problems without parametrizations. The main idea of the level set method is to represent a shape as

$$\Omega(t) = \{\phi(.,t) < 0\},$$

where $\phi : \mathbb{R}^d \times \mathbb{R}^+ \to \mathbb{R}$ is a suitable continuous function, ideally the signed distance function to $\partial\Omega$ (i.e., equal to the distance between x and $\partial\Omega$ if $x \in \mathbb{R}^d \backslash \Omega$, and equal to the negative distance if $x \in \Omega$). For an appropriate ϕ we have that

$$\partial\Omega(t) = \{\phi(.,t) = 0\}.$$

Now consider the motion of points in $\Omega(t)$ by $\frac{dx}{dt} = V(x)$. Then we obtain from the chain rule for $x(t) \in \partial\Omega(t)$

$$0 = \frac{d}{dt}\phi(x(t),t) = \frac{\partial\phi}{\partial t} + V.\nabla\phi = 0,$$

i.e., ϕ can be determined by solving a transport equation. As we have seen above, the most interesting case is the one of a motion in normal direction on $\partial\Omega(t)$, i.e., $V = V_n.n$. In order to use such a velocity in the level set method, we have to express the normal in terms of the level set function ϕ. Assume that $\{\tilde{x}(s,t)|s \in (-\epsilon,\epsilon)\}$ is an arc on $\partial\Omega(t)$, locally parametrized by s around $x(t) = \tilde{x}(0,t)$. Then

$$0 = \frac{d}{ds}\phi(\tilde{x}(s,t),t) = \nabla\phi(\tilde{x}(s,t),t)\frac{\partial\tilde{x}}{\partial s}.$$

Since $\frac{\partial\tilde{x}}{\partial s}$ can be any tangential direction, we obtain that $\nabla\phi$ is a normal direction, and one obtains the unit normal as

$$n(s,t) = \frac{\nabla\phi}{|\nabla\phi|}(\tilde{x}(s,t),t).$$

Using these formulas together with the transport equation for ϕ, we obtain the Hamilton-Jacobi equation

$$\frac{\partial\phi}{\partial t} + V_n|\nabla\phi| = 0 \tag{8.2}$$

for ϕ. One can show that the motion of $\Omega(t)$ is determined by

$$\Omega(t) = \{\phi(.,t) < 0\}$$

if ϕ is a solution of (8.2) in $\mathbb{R}^d \times \mathbb{R}^+$ where V_n is an arbitrary extension from $\{\phi(.,0) < 0\}$ to $\mathbb{R}^d$.

For further details and applications of the level set method we refer to the monograph by Osher and Fedkiw.

8.3 Computing Shape Sensitivities by Level Set Methods

Using the level set method, we can formally compute shape sensitivities in a simple way. Consider again the functional

$$J(\Omega) = \int_\Omega g(x)\ dx$$

and let $\partial\Omega(t)$ move with normal speed V_n. Then we obtain

$$\begin{aligned} J(\Omega(t)) &= \int_{\{\phi(.,t)<0\}} g(x)\ dx \\ &= \int_{\mathbb{R}^d} H(-\phi(x,t))g(x)\ dx, \end{aligned}$$

where H denotes the Heaviside function

$$H(p) = \begin{cases} 1 & \text{if } p > 0 \\ 0 & \text{else.} \end{cases}$$

Since the derivative of the Heaviside function is the Dirac-delta-distribution, we obtain formally

$$\begin{aligned} \frac{d}{dt}J(\Omega(t)) &= \int_{\mathbb{R}^d} -H'(-\phi(x,t))\frac{\partial\phi}{\partial t}(x,t)\ g(x)\ dx \\ &= \int_{\mathbb{R}^d} \delta(\phi(x,t))|\nabla\phi(x,t)|\ V_n\ g(x)\ dx \end{aligned}$$

Now we apply the co-area formula, i.e.

$$\int_{\mathbb{R}^d} A(\phi(x))\ B(x)\ |\nabla\phi(x)|\ dx = \int_{\mathbb{R}} A(p) \int_{\{\phi=p\}} B(x)\ ds(x)\ dp.$$

This implies

$$\begin{aligned} \frac{d}{dt}J(\Omega(t))\Big|_{t=0} &= \int_{\mathbb{R}^d} \delta(\phi(x,0))\ g(x)\ V_n(x)\ |\nabla\phi(x,0)|\ dx \\ &= \int_{\mathbb{R}} \delta(p) \int_{\{\phi=p\}} g(x)\ V_n(x)\ ds\ dp \\ &= \int_{\{\phi=0\}} g(x)\ V_n(x)\ ds(x) \\ &= \int_{\partial\Omega} g\ V_n\ ds, \end{aligned}$$

i.e., we recover the above formula for the shape sensitivity.

In a similar way we can compute the shape sensitivity of the functional

$$J(\Omega) = \int_{\partial\Omega} g\ ds$$

For this sake we use again the δ-distribution and the coarea formula to deduce

$$\begin{aligned} J(\Omega(t)) &= \int_{\{\phi(.,t)=0\}} g(x)\ ds(x) \\ &= \int_{\mathbb{R}} \delta(p) \int_{\{\phi(.,t)=p\}} g(x)\ ds(x)\ dp \\ &= \int_{\mathbb{R}^d} \delta(\phi(x,t))\ g(x)\ |\nabla\phi(x,t)|\ dx \end{aligned}$$

Thus, we can try to compute the time derivative as

$$\begin{aligned}
\frac{d}{dt}J(\Omega(t)) &= \int_{\mathbb{R}^d} g\left(\delta'(\phi)|\nabla\phi|\phi_t + \delta(\phi)\frac{\nabla\phi\nabla\phi_t}{|\nabla\phi|}\right)\,dx \\
&= \int_{\mathbb{R}^d} g\left(\frac{\nabla\delta(\phi)\nabla\phi}{|\nabla\phi|}\phi_t + \delta(\phi)\frac{\nabla\phi\nabla\phi_t}{|\nabla\phi|}\right)\,dx \\
&= \int_{\mathbb{R}^d} \delta(\phi)\left(-div\,\left(g\frac{\nabla\phi}{|\nabla\phi|}\phi_t\right) + g\frac{\nabla\phi\nabla\phi_t}{|\nabla\phi|}\right)\,dx \\
&= -\int_{\mathbb{R}^d} \delta(\phi)\left(\frac{\nabla g.\nabla\phi}{|\nabla\phi|}.\phi_t + g\;div\,\left(\frac{\nabla\phi}{|\nabla\phi|}\right)\phi_t\right)\,dx \\
&= \int_{\mathbb{R}^d} \delta(\phi)|\nabla\phi|\;V_n\,\left(\nabla g\frac{\nabla\phi}{|\nabla\phi|} + g\;div\,\left(\frac{\nabla\phi}{|\nabla\phi|}\right)\right)\,dx \\
&= \int_{\{\phi=0\}} V_n\left(\nabla g\frac{\nabla\phi}{|\nabla\phi|} + g\;div\left(\frac{\nabla\phi}{|\nabla\phi|}\right)\right)\,ds
\end{aligned}$$

One observes that on $\partial\Omega = \{\phi = 0\}$ we have

$$u = \frac{\nabla\phi}{|\nabla\phi|}, \quad \kappa = div\; n = div\left(\frac{\nabla\phi}{|\nabla\phi|}\right),$$

where n is the unit normal and κ is the mean curvature. Thus,

$$J'(\Omega)V_n = \int_\Gamma V_n\left(\frac{\partial g}{\partial n} + g\;\kappa\right)\,ds.$$

We finally notice that the above strategy of removing the term $\delta'(\phi)$ by rewriting

$$\delta'(\phi)|\nabla\phi| = \nabla\delta(\phi)\frac{\nabla\phi}{|\nabla\phi|}$$

and applying Gauss' Theorem can be used for general functionals (e.g. for second derivatives of the functional J above). In this way, we always obtain a term of the form

$$-\delta(\phi)\;div\left(\frac{\nabla\phi}{|\nabla\phi|}\right),$$

i.e., the mean curvature on $\{\phi = 0\} = \partial\Omega$. In particular, we can rewrite all derivatives as surface integrals on $\partial\Omega$, involving only natural geometric quantities like the normal n or the curvature κ, its normal derivative $\frac{\partial\kappa}{\partial n}$, etc. It is a good advice to check all quantities that one obtains by computing shape sensitivities in this way with respect to their geometric meaning. If some terms do not have a geometric interpretation, then most likely the calculation was wrong.

8.4 Numerical Solution

In order to obtain computational methods for shape optimization problems we can again employ the level set method. In principle, we can apply any of the optimization methods discussed in chapter 4, once we know how to compute derivatives. The major difference is the

way we update the design variable. In the setting of chapter 4, we have computed a search direction s to obtain

$$u_{k+1} = u_j + \tau_k s.$$

Obviously, we cannot use the same strategy in shape optimization, since a formula like

$$\Omega_{k+1} = \Omega_k + \tau_k s$$

does not make sense for shapes Ω_k. However, there is a natural update offered by the speed method. First we notice that the update for a design variable u in a Hilbert space can be rewritten as

$$u_{k+1} = u(\tau_k), \quad \frac{du}{dt} = s, \quad u(0) = u_k.$$

As in the context of shape derivatives, the corresponding speed method for shapes gives

$$\Omega_{k+1} = \left\{ x(t_k) \;\middle|\; \frac{dx}{dt} = s,\ x(0) \in \Omega_k \right\}.$$

Since the motion depends only on the normal velocity on $\partial\Omega$, we can define the update also via the level set method as

$$\begin{aligned} \Omega_{k+1} &= \{\phi(.,\tau_k) < 0\} \\ \frac{\partial \phi}{\partial t} + s_n |\nabla \phi| &= 0 \quad \text{in } (0,\tau_k) \\ \{\phi(.,0\} &= \Omega_k, \end{aligned}$$

where s_n is the normal component of the update s. Hence, the iterative method is characterized by choosing a normal update. Below, we shall detail some possible ways for choosing this update.

We start with a gradient-type method. One observes that for optimization in Hilbert spaces, the gradient method is characterized by choosing the update s via

$$< s, v >= -J'(u)v \quad \forall v \in \mathcal{U}.$$

We can now write an analogous formula for the update s_n, namely

$$< s_n, V_n >= -J'(\Omega)V_n \quad \forall V_n \in \mathcal{U},$$

where $\mathcal{U}$ is a suitable Hilbert space for which we have several possibilities. We start with the simple choice $\mathcal{U} = L^2(\partial\Omega)$, i.e.,

$$< S_n, V_n >= \int_{\partial\Omega} S_n V_n \; ds.$$

As we have seen above, one can usually write the shape sensitivity in the form

$$J'(\Omega)V_n = \int_{\partial\Omega} h.V_n \; ds$$

(with $h = g$ for $J(\Omega) = \int_\Omega g \; dx$, and $h = \frac{\partial g}{\partial n} + g\kappa$ for $J(\Omega) = \int_{\partial\Omega} g \; ds$). Thus, the equation for S_n becomes

$$\begin{aligned} \int_{\partial\Omega} S_n V_n \; ds =< S_n, V_n > &= -J'(\Omega)V_n \\ &= -\int_{\partial\Omega} h \; V_n \; ds \quad \forall V_n \in L^2(\partial\Omega) \end{aligned}$$

which is equivalent to choosing $S_n = -h$.

Another interesting Hilbert space is $H^1(\partial\Omega)$. The scalar product in this space is given by

$$\begin{aligned} < S_n, V_n > &= \int_{\partial\Omega} (\nabla_s S_n \nabla_s \, V_n + S_n V_n) \, ds \\ &= \int_{\partial\Omega} V_n(-\Delta_s S_n + S_n) \, ds, \end{aligned}$$

where Δ_s denotes the gradient with respect to the surface variable S on $\partial\Omega$ and Δ_s is the surface Laplacian. Consequently, the update S_n can be computed by solving the Laplace-Beltrami equation

$$-\Delta_s S_n - S_n = h$$

on $\partial\Omega$ (note that we do not need a boundary condition, since the boundary of the surface $\partial\Omega$ is empty).

In general, we can write a Hilbert space scalar product as

$$< S_n, V_n >= \int_{\partial\Omega} (AS_n)V_n \, ds,$$

where A is a positive definite operator. Thus, we may choose any search direction of the form

$$S_n = -A^{-1}h,$$

where A is a positive definite operator. Since

$$J'(\Omega)S_n = - < S_n, S_n >= -\|S_n\|^2,$$

this yields a descent direction and we can use line search techniques to find a reasonable τ_k.

In a similar way to gradient methods we can derive Newton-type methods, for which S_n is choosen solving

$$J''(\Omega)(S_n, V_n) = -J'(\Omega)V_n, \quad \forall V_n \in \mathcal{U}.$$

Bibliography

[1] R.Acar, C.R.Vogel, *Analysis of total variation penalty methods*, Inverse Problems **10** (1994), 1217–1229.

[2] A.B.Bakushinskii, *Remarks on choosing a regularization parameter using the quasi-optimality and ratio criterion*, USSR Comp. Math. Math. Phys. **24** (1984), 181-182.

[3] H.T.Banks, K.Kunisch, *Estimation Techniques for Distributed Parameter Systems* (Birkhäuser, Boston, 1989).

[4] A.Binder, H.W. Engl, C.W. Groetsch, A. Neubauer, O. Scherzer, *Weakly closed nonlinear operators and parameter identification in parabolic equations by Tikhonov regularization*, Appl. Anal. **55** (1994), 215-234.

[5] J.H.Bramble, J.E.Pasciak, P.S.Vassilevski, *Computational scales of Sobolev norms with application to preconditioning*, Math. Comp. **69** (2000), 463–480.

[6] M.Burger, B.Kaltenbacher, *Regularizing Newton-Kaczmarz methods for nonlinear ill-posed problems*, SFB-Report 04-17 (University Linz, 2004), and submitted.

[7] D.Colton, R.Kress, *Inverse Acoustic and Electromagnetic Scattering Theory* (2nd ed., Springer, Berlin, 1998).

[8] P.P.B.Eggermont, *Maximum entropy regularization for Fredholm integral equations of the first kind*, SIAM J. Math. Anal. **24** (1993), 1557-1576.

[9] H.W. Engl, *Integralgleichungen* (Springer, Vienna, 1997).

[10] H.W. Engl, M. Hanke and A. Neubauer, *Regularization of Inverse Problems* (Kluwer, Dordrecht, 1996).

[11] H.W.Engl, G.Landl, *Convergence rates for maximum entropy regularization*, SIAM J. Numer. Anal. **30** (1993), 1509-1536.

[12] C.W. Groetsch, *Theory of Tikhonov regularization for Fredholm equations of the first kind* (Pitman, London, 1984).

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

[14] M.Hanke, O.Scherzer, *Inverse problems light: Numerical differentiation* Amer. Math. Monthly **108** (2001), 512-521.

[15] M.Hanke, C.R.Vogel, *Two-level preconditioners for regularized inverse problems. I. Theory*, Numer. Math. **83** (1999), 385–402.

[16] T.Hohage, *On the numerical solution of a three-dimensional inverse medium scattering problem*, Inverse Problems **17** (2001), 1743–1763.

[17] A.Kirsch, *An Introduction to the Mathematical Theory of Inverse Problems* (Springer, Berlin, 1996).

[18] R.Kowar, O.Scherzer, *Convergence analysis of a Landweber-Kaczmarz method for solving nonlinear ill-posed problems*, in: S.Romanov, S.I.Kabanikhin, Y.E.Anikonov, A.L.Bukhgeim, eds., *Ill-Posed and Inverse Problems* (VSP Publishers, Zeist, 2002).

[19] Y.Meyer, *Oscillating Patterns in Image Processing and Nonlinear Evolution Equations*, AMS, Providence, 2001.

[20] A.Rieder, *Runge-Kutta integrators yield optimal regularization schemes*, Inverse Problems **21** (2005), 453-471.

[21] L.I.Rudin, S.Osher, E.Fatemi, *Nonlinear total variation based noise removal algorithms*, Physica D **60** (1992), 259–268.

[22] O.Scherzer, *Convergence criteria of iterative methods based on Landweber iteration for solving nonlinear problems*, J. Math. Anal. Appl. **1995**, 911-933.

[23] C.Vogel, *Computational Methods for Inverse Problems* (SIAM, Philadelphia, 2002).