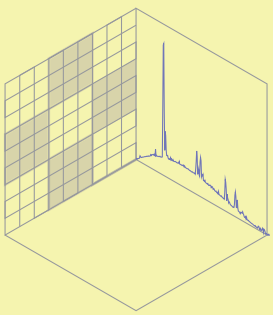
The elastic net:
Stability for sparsity methods

Stefan Schiffler



The elastic net: Stability for sparsity methods

von Stefan Schiffler

Dissertation
zur Erlangung des Grades eines Doktors der Naturwissenschaften
– Dr.rer.nat –

Vorgelegt im Fachbereich 3 (Mathematik & Informatik)
der Universität Bremen
im Februar 2010

Datum des Promotionskolloquiums: 24.06.2010

Gutachter:
Prof. Dr. Peter Maaß (Universität Bremen)
Prof. Dr. Dirk A. Lorenz (Technische Universität Braunschweig)

Acknowledgement

I would like to thank Prof. Dr. Peter Maaß for providing the possibility of writing this thesis in his pleasant working group as well as the chance to work on several interesting industrial projects, related to my research.

Moreover, I like to thank Prof. Dr. Dirk A. Lorenz and Prof. Dr. Peter Maaß for their always helpful supervision and persistent support. Furthermore, I like to thank Dr. Theodore Alexandrov, Dr. Christian Bredies, Dr. Bangti Jin, Prof. Dr. Dirk Lorenz and Dennis Trede for fruitful discussions, joint works and helpful suggestions. For careful proofreading and helpful comments, I am deeply grateful to Dr. Theodore Alexandrov, Dr. Bangti Jin, Sankar Natarajan, Sabine Schiffler and Dennis Trede.

Last but not least, I want to thank all my colleagues at the "Zentrum für Technomathematik" for the good time and the pleasant working atmosphere.

Finally, I acknowledge Bruker Daltonics for providing images and datasets.

Stefan Schiffler
Zentrum für Technomathematik
Universität Bremen

www.math.uni-bremen.de/~schiffi

Abstract

In recent years, methods for sparse approximation have gained considerable attention and have been successfully applied to numerous problems in various mathematical disciplines. This work starts by illustrating applications for sparse approximation to introduce the concept of sparsity. Afterwards, the mathematical framework and basic mathematical principles are introduced. Particularly ℓ^1 minimization, which is an important tool in the sparsity context, will be introduced as well as available algorithms. This forms a profound background to approach the problem of stability in ℓ^1 minimization for ill-conditioned linear equations. It turns out that a tool arising from statistics – the elastic net – promises to attenuate stability problems, while preserving the benefits of ℓ^1 minimization. The connection between ℓ^1 minimization and the elastic net is discussed. Analytical properties of the elastic net are stated and corresponding algorithms are developed. Numerical troubles of ℓ^1 minimization are demonstrated for sample problems as well as the influence of the elastic net. Finally, given all necessary tools, the discussion leads to the highlight of exact-recovery conditions for elastic-net minimization.

Contents

1 Introduction

In recent years, there has been a growing interest in sparsity leading to considerable success in many fields of mathematics, such as inverse problems or image processing. Actually, it has spawned a new field of research, named compressed sensing, which is revolutionizing the standard sampling theory. The idea of compressed sensing is to incorporate sparsity into a measurement process so as to produce reasonable results beyond the known sampling theorems. Before we discuss applications of sparsity in detail, we answer the question "What exactly does sparsity mean¿‘

1.1 What is sparsity?

To answer this question, we consider a linear operator A mapping an input vector x_0 onto a signal vector y. Such linear operators appear in many applications, e.g. as solution operators of differential equations in physics or as convolution operators in image processing. Usually, one is provided only a measurement y^δ of the signal vector y perturbed with noise. In this case, x_0 is unknown and shall be recovered by solving $Ax = y^\delta$. Due to the modeling background, one expects $Ax = y$ to be solvable (by x_0), in contrast to $Ax = y^\delta$.

Sparsity describes a property of x_0, where it is denoted to be a sparse solution of $Ax = y$ if there exists a basis, such that the coordinates with respect to this basis are almost all zero. Typically, one is given a fixed basis, such that the searched for solution x_0 is assumed to be a sparsest solution,

i.e. a solution with fewest non-zero coordinates. This assumption may be exploited to obtain better approximations for x_0 than obtained by classical approaches, such as the least squares method (if applicable) or Tikhonov regularization applied to $Ax = y^\delta$. Especially, sparsity can be used as a regularization.

To give a mathematical formulation incorporating the sparsity idea into an approximate solution, our first ingredient is a suitable basis $\{\varphi_i : i \in \mathbb{N}\}$ of the space of all possible input vectors for the model. This basis shall be such that it is expected to admit a sparse expansion of x_0, and it is typically arising naturally from the problem at hand. Since sparsity is a property involving the basis expansion, we need the synthesis operator S mapping a vector x onto the linear combination $\sum x_i\varphi_i$. With the help of S we define $K = AS$ and reformulate the problem into $Kx = y^\delta$. The second ingredient is the ℓ^0-norm, where $\|x\|_{\ell^0}$ is counting the non-zero coordinates of x. In fact, it is not a true norm, but nevertheless it is a common notation in the literature. Combining the ingredients, we end up with the formula

$$\min \|x\|_{\ell^0}, \quad \text{s.t. } \|Kx - y^\delta\| < \tau \tag{1.1}$$

for some $\tau > 0$. It is easy to notice that a solution of this problem is one of the sparsest vectors keeping the error bound τ on the residual $\|Kx - y^\delta\|$. If the basis was chosen carefully, one hopes that the solution x fulfills $x_0 \approx Sx$. Before we come to a mathematically more strict treatment and discuss the details of sparse approximation we motivate sparsity in different mathematical disciplines.

1.2 Applications of sparse approximation

This section introduces three different fields of mathematics, where the sparsity idea emerges and introduces sample problems. In the field of inverse problems, sparsity can be used as a regularization. In this field we discuss an application from mass spectrometry. The second field of mathematics is compressed sensing, where sparsity forms the backbone of the whole field. We introduce the standard example for compressed sensing, namely, the single pixel camera. The third field is image analysis, where sparsity emerges for example in noise-robust feature extraction. Some of the introduced sample problems are used later as numerical examples or form the background for further numerical experiments.

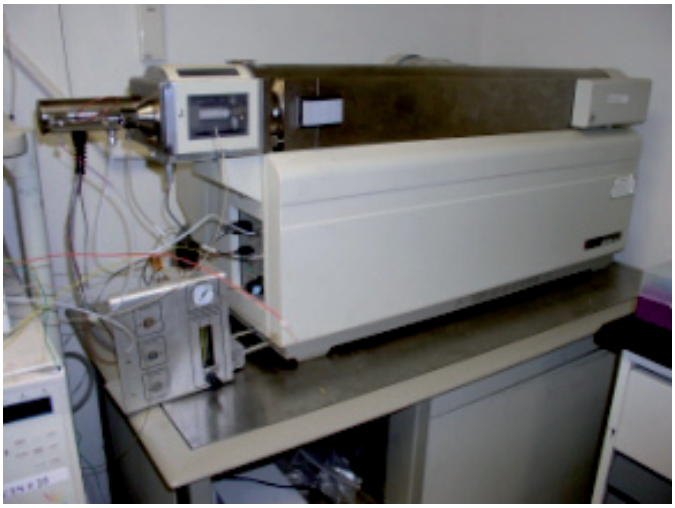


Figure 1.2.1: A mass spectrometer, Sciex 365.

1.2.1 Inverse problems

Now, we take a closer look on an application of the inverse-problems theory, which is also benefiting from the sparsity idea – peak picking in mass spectrometry (MS) [14, 57]. MS is a tool to measure the $\frac{m}{z}$ value, or mass-to-charge ratio, of molecules and is used in particular for medical purposes. For a given sample, it can count the number of molecules of a certain type, i.e. of a certain mass-to-charge ratio. These numbers are collected for several mass-to-charge ratios and charted in so called spectra. Considering the fact that on small scales all particles of the same type have exactly the same mass-to-charge ratio one expects that a spectrum consists of sharp peaks for each molecule type. This idealized concept is visualized in figure 1.2.2.

Assuming that figure 1.2.2 is a spectrum for healthy dermal tissue, the expectation is that for example cancerous tissue results in different particle distributions, i.e. the peaks have different heights and positions. This idea originates from the differences in metabolic processes of healthy and cancerous cells, which is the case for several tissue types. The ability to differ tissue types via MS is efficiently exploited in the new research field of MALDI imaging [10, 42], where a spatial structure of a slice of tissue, e.g. a brain, can be investigated (cf. figure 1.2.3).

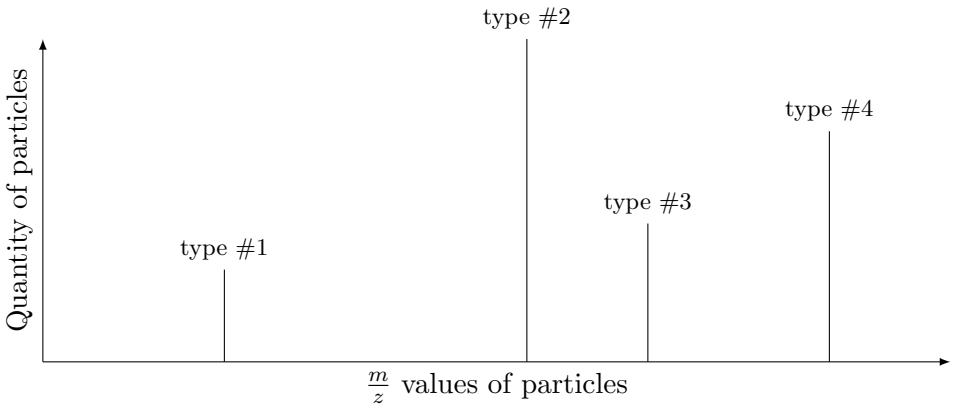


Figure 1.2.2: Idealized image of a mass spectrometry measurement, each particle type admits a sharp peak.

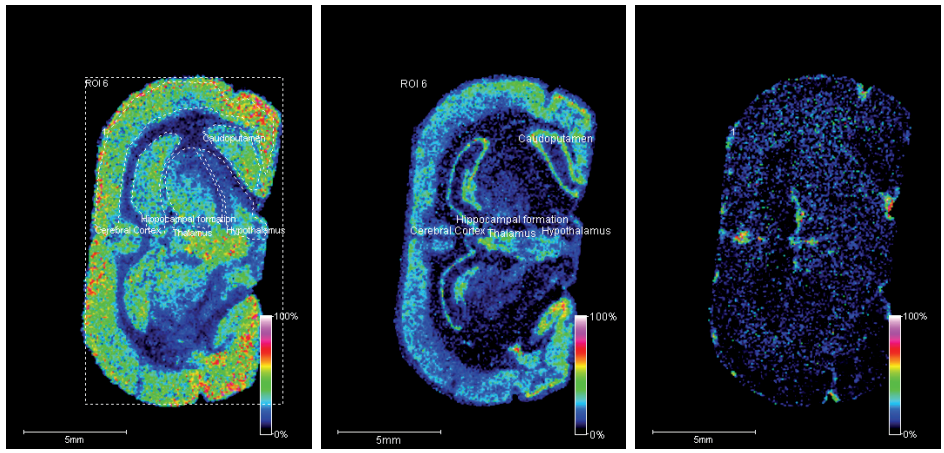


Figure 1.2.3: Segmentation of a rat brain in MALDI imaging is one application for mass spectrometry among many others. (Images by Theodore Alexandrov and Bruker Daltonics)

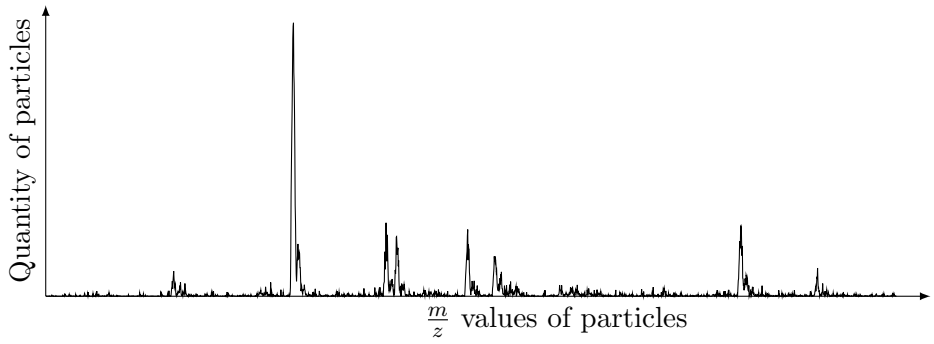


Figure 1.2.4: Section of real world mass spectrometry data. (Source: Bruker Daltonics)

After the measurement, the $\frac{m}{z}$ values are discretized. Hence, the output of a mass spectrometer is a vector with each coordinate corresponding to a certain $\frac{m}{z}$ value. At this point, it seems evident that a spectrum y can be modeled as a linear combination of discrete Dirac peaks (as used e.g. in [6, 35]) via

$$y = \sum x_i \delta(\cdot - i).$$

Compared to the resolution of the measurement device, the number of peaks is typically very low and hence this δ-peak model corresponds to a sparse representation of the spectrum, i.e. almost all coefficients x_i are zero. In fact, the measured spectra look more complicated, they are perturbed with noise and the peaks are blurry as can be seen in the mass spectrum shown in figure 1.2.4.

The question of peak picking in MS is "How to reconstruct the ideal peak spectra from the blurry and noisy measured ones?". As stated, we already have a basis which is expected to admit a sparse representation of an ideal spectrum – the basis consisting of Dirac peaks. The process of blurring can be modeled using a convolution operator A with a suitable kernel. This operator is linear and the kernel can be determined approximately by suitable measurements. In this context, the kernel is also referred to as point-spread function and can be regarded as a prototype of a measured blurry peak. The point-spread function may depend on the $\frac{m}{z}$ values as well as on the age or runtime of a measurement device. However, in models

it is often simplified to be a Gaussian convolution kernel. Thus, the problem of determining the coefficients x from a measured spectrum y^δ corresponds to solving the linear problem $Ax = y^\delta$.

Considering the noisy dataset, the implicit assumption is that we need a few big coefficients to represent the real peaks and many small coefficients to represent the noise. Hence, applying the sparse approximation framework should result in a good approximate solution x, representing the peaks but not the noise. So the benefit of sparse approximation is the possibility to simultaneously denoise and solve the problem of peak picking. This idea of exploiting sparsity based methods in mass spectrometry has been discussed for example in [6].

In practice, the conditioning of the convolution operator A is, depending on the kernel, arbitrarily poor and hence the inversion is not stable. Due to this fact, the problem is considered to be an ill-posed inverse problem.

Moreover, often one is not interested in quantities of single molecules but in combinations thereof, e.g. a protein of interest may result in a peak pattern. These patterns are known a priori from test measurements as well as theoretical fundamentals. Consequently, the Dirac-peak basis can be enhanced by adding patterns $\{c_i : i = 1 \dots n\}$. This means that the model for our idealized spectrum y changes into

$$y = \sum x_i \delta(\cdot - i) + \sum z_i c_i.$$

The sparsity idea exploits the fact that one needs one pattern to describe several peaks. Hence, for a sparse representation of y we will select a pattern if it appears in the data y instead of selecting several single peaks. Consequently, enhancing the Dirac-peak basis by patterns allows a more sparse representation of y.

In contrast to the Dirac-peak basis, we now need the synthesis operator S mapping the vector $((x_i)_i^T (z_i)_i^T)$ to the above sum. With $K = AS$, we now consider the problem $Kx = y^\delta$, for the noisy measurement y^δ, which no longer has a unique solution. In addition to the non-unique solution, the patterns for different molecules can overlap partly resulting in highly correlated columns of the matrix K in practice. Thus, K is not injective and is ill-conditioned. In the numerical experiments in later chapters, we see the impact of these problems on sparse-approximation algorithms and how one can resolve them.

1.2.2 Compression and compressed sensing

The field of compressed sensing is a rather new discipline, which has emerged in recent years and stands for a novel measurement paradigm applicable in many fields. It aims to go beyond standard sampling theory by introducing sparsity. A good introduction and motivation of compressed sensing can be found for example in [9, 17, 47].

A common procedure while measuring a quantity, e.g. taking a digital photo, is to perform many measurements (several millions for a usual digital camera) and producing a big dataset y. Afterwards this data is represented in a basis (e.g. wavelets) admitting a sparse representation of y. Only a few coefficients are stored such that a reconstruction up to a certain precision τ is possible. Such procedures are called lossy compression procedures. One of the most famous representatives is JPEG compression which uses such a procedure in combination with other advanced techniques.

Such a compression scheme is very wasteful in the sense that one takes millions of measurements and afterwards throws away information to store only a few thousand coefficients. At this point compressed sensing enters the scene and tries to replace this procedure by just taking a few thousand measurements and skip the compression step. The question is: "Is it possible to directly reduce the number of measurements to a few thousands while obtaining results of similar quality as using the traditional compression scheme with millions of measurements?"

The idea behind this question is that one tries to measure the correlation of y with other vectors, instead of pointwise measurements. We repeat this step a few thousand times and store the vectors correlating with in the rows of a matrix A (sensing matrix), which finally has millions of columns and a few thousand rows. Mathematically, we measure Ay. Since A is not injective, we can not recover the true signal y directly. Having a basis which allows a sparse representations of y, we denote the corresponding synthesis operator by S. To recover the signal y ,we can use the prior sparsity information and solve the equation $Ay = ASx$. Due to the dimensions of A, this equation still has infinitely many solutions but if A was chosen carefully it has only one sparse solution x_0. Hence, via $y = Sx_0$ the true signal y can be recovered. The research in compressed sensing deals with finding constraints on S and A such that we can recover x_0, or at least a solution x such that $y \approx Sx$.

The key point is that we can store the (few) measurements Ay directly, but if a digital camera would store the images in the compressed sensing scheme we would need to solve

$$\min \|x\|_0, \quad \text{s.t. } \|ASx - Ay\| < \tau$$

to view the images. In this way, compressed sensing transfers the computational effort from the measurement device to the analysis device and e.g. makes sensors more efficient. Another benefit of compressed sensing is that it enables us to reduce the number of necessary sensors (even a single pixel camera is possible, see [47]).

We pay our main attention to the fact that the sensing matrix A, due to its dimensions, yields a rank deficient matrix A^*A. As we will see, this causes troubles when solving sparse approximation problems.

1.2.3 Image processing

As an example for signal and image processing involving sparsity, we present an application from mechanics, namely, surface analysis of high-precision-turned surfaces. Turning describes a process where a piece of material (wood, metal, plastic etc.) is placed on a plate, not necessarily centered, which is rotating (cf. figure 1.2.5). Then a cutting tool is traversed along 2 axes of motion, while on each rotation of the workpiece the tool hits the workpiece and cuts a track into the material (see figure 1.2.6). It can be used for example to produce plain surfaces in high-precision turning for mirror-like surfaces. When hitting the workpiece, and while cutting, the tool starts to vibrate, which influences the quality of the surface, especially its optical characteristics.

Image processing comes into play when the created surfaces have to be evaluated by means of interferometric images (figure 1.2.6). The imaging process requires tilting the workpiece since one can not image the surface directly from above. For evaluation, this tilt needs to be "removed" from the images. Furthermore, we are not interested in the track information of the tool but in its vibrations. Hence, we want to remove tilt and track information to visualize the vibrations of the tool. We possibly need to remove more characteristics such as the shape of the turned object but for now we consider only plain surfaces. As a model for an image y of a plain surface, we take the following superposition

$$y = \text{tilt} + \text{tracks} + \text{vibrations}.$$

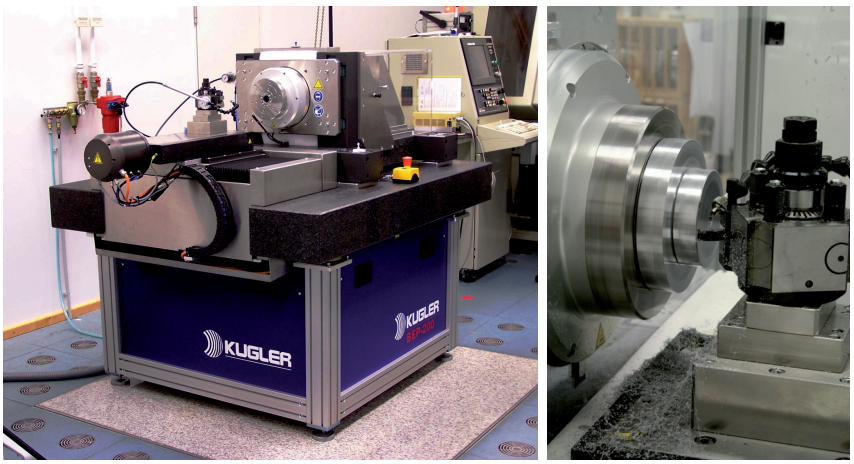


Figure 1.2.5: A turning device, on the right is a close up of the tool cutting the workpiece.

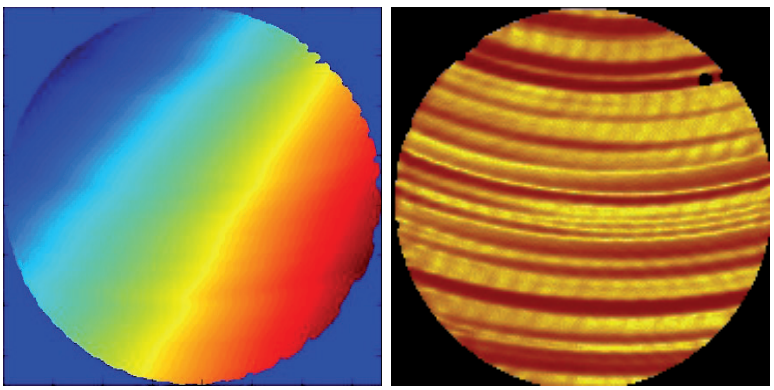


Figure 1.2.6: Interferometric image of a surface produced by turning and the same dataset without tilt where the tracks become visible.

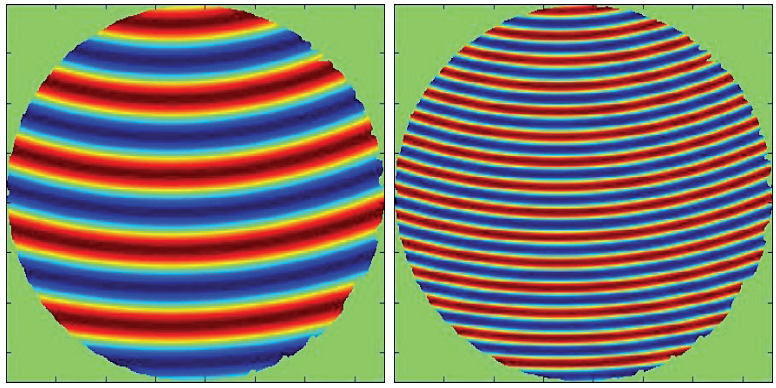


Figure 1.2.7: Visualization of two of the basis vectors used to represent track information.

In our case, the tilt is modeled as a polynomial of degree one. The tracks can be represented by taking sinusoids in radial direction of the circular tracks constantly propagated along the tracks (figure 1.2.7). Hence, our model for the surface becomes a linear combination of polynomials of degree one and sinusoids, where the collection of these functions is referred to as dictionary. Writing these dictionary functions as columns of a matrix K, we can represent our surface y as

$$y = Kx + \text{vibrations}.$$

This approach is a common procedure in surface analysis [36]. In contrast to the standard methods, as least squares, sparsity can be used to ensure that e.g. the tilt is fitted by the polynomial part and not by sinusoids. This is possible, because it takes one polynomial to model the tilt while it would take several sinusoids. If we do not restrict ourself to plain surfaces, and hence need to fit the shape of our turned object as well, sparsity becomes more important because it may anticipate to fit a feature such as tilt or tracks using dictionary functions designated to fit the shape. In this manner we are able to decompose the surface into certain parts e.g. tilt, tracks and vibrations (figure 1.2.8).

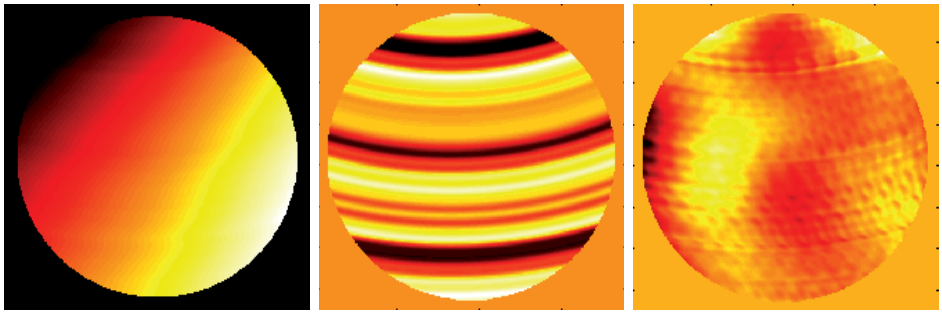


Figure 1.2.8: Left: Interferometric image of a surface. Middle: Extracted track information. Right: Residual containing vibrations and noise.

1.3 Overview

In the next chapter, we introduce some mathematical fundamentals as well as the strict mathematical framework which we use all over the rest of this work. This comprises assumptions on the linear operator under consideration and the corresponding spaces the operator maps between. Also different notions of differentiability are introduced.

Having all the fundamentals, we give an introduction to ℓ^1 minimization and its connection to sparse approximation. Afterwards, we give a short overview of available minimization algorithms followed by a motivation of elastic-net minimization. Research results on the elastic net are proved and the problem of parameter selection is discussed. Two active-set algorithms for elastic-net minimization are developed. New results about active-set choices and corresponding convergence are proved.

It follows a chapter on numerical properties of the elastic net compared to ℓ^1 minimization, demonstrating the differences in the numerical stability, but on the other side similarity in the results.

Finally, we address a special case of parameter-selection rules, namely, rules that allow to recover the support of the input vector, which we called x_0 in the beginning of the introduction. The applicability is demonstrated using a numerical example from mass spectrometry together with a recipe to choose the parameters a priori.

The work is closed with a conclusion and an overview of remaining, interesting questions which might be addressed in further research.

2 Notations, framework and basic principles

After motivating sparsity in the previous chapter, we now turn to a strict mathematical framework to see how the problem of sparse approximation can be solved and to have insights into some of its mathematical properties. Therefore, we firstly introduce the notations and preliminaries used all over the following work. Secondly, we introduce some basic principles from functional analysis, namely, Gâteaux differentiability, subdifferential calculus, and Newton differentiability which also plays a fundamental role.

2.1 Notations

We start with some general notations. By $\mathbb{R}_\infty$ we denote the extended real line $(-\infty, \infty]$, where we extend the usual ordering on $\mathbb{R}$ onto $\mathbb{R}_\infty$ by $-\infty < a \leq \infty$, for all $a \in \mathbb{R}_\infty$. For a subset I of $\mathbb{N}$ and a real number $p > 0$ we define the sequence spaces

$$\ell^p(I) = \{(a_i)_{i\in I} \subset \mathbb{R} : \sum_{i\in I} |a_i|^p < \infty\}.$$

These spaces are Banach spaces if equipped with the corresponding norms

$$\|a\|_{\ell^p(I)} = \big(\sum_{i\in I} |a_i|^p\big)^{\frac{1}{p}}.$$

Especially, the pair $(\ell^2, \|\cdot\|_{\ell^2})$ constitutes a Hilbert space together with the scalar product

$$\langle x, y\rangle_{\ell^2} = \sum_{i\in\mathbb{N}} x_i y_i.$$

As an abbreviation we write ℓ^p instead of $\ell^p(\mathbb{N})$. By $L(\mathcal{X}, \mathcal{Y})$ we denote the set of all bounded linear operators mapping between two spaces $\mathcal{X}$ and $\mathcal{Y}$. For a linear operator $T : \mathcal{X} \to \mathcal{Y}$ we denote its range by $R(T) = \{Tx : x \in X\} \subset Y$ and its pseudo inverse by $T^\dagger$.

2.2 Framework

The framework for a strict mathematical treatment requires some assumptions which need to be postulated in most of the theorems, lemmas and propositions in this work. In this section all those are collected while explaining the framework and shall be assumed all over the work. You can find the overview table in the last chapter.

All over this work, let $\mathcal{H}$ be a Hilbert space and $K \in L(\ell^2, \mathcal{H})$. Further, let $y \in \mathrm{R}(K) \subset \mathcal{H}$ with a noisy observation thereof which we call y^δ, defining the noise as $\eta := y^\delta - y$ and the noise level $\delta := \|\eta\|_{\mathcal{H}} < \infty$. Particularly, it follows that $y^\delta \in \mathcal{H}$. We assume that the linear equation $Kx = y$ has a sparse solution $x^\dagger$ which is sparse with respect to the standard basis $(e_i)_{i\in\mathbb{N}}$. We shall also need the definition of the ℓ^1 functional

$$\Psi_\alpha(x) := \tfrac{1}{2}\|Kx - y^\delta\|^2_{\mathcal{H}} + \alpha\|x\|_{\ell^1}, \quad (\ell^2 \to \mathbb{R}_\infty), \tag{2.1}$$

and the elastic-net functional

$$\Phi_{\alpha,\beta}(x) := \tfrac{1}{2}\|Kx - y^\delta\|^2_{\mathcal{H}} + \alpha\|x\|_{\ell^1} + \tfrac{1}{2}\beta\|x\|^2_{\ell^2}, \quad (\ell^2 \to \mathbb{R}_\infty) \tag{2.2}$$

for $\alpha, \beta \geq 0$. The dependency of the functionals Ψ_α and $\Phi_{\alpha,\beta}$ on the operator K and the measurement data y^δ is not denoted. The part $\frac{1}{2}\|Kx - y^\delta\|^2_{\mathcal{H}}$ is called the residual term while the remaining part is called the penalty term.

For a set $I \subset \mathbb{N}$ we denote the orthogonal projection onto the indices listed in I as $P_I : \ell^2 \to \ell^2(I)$. Note that in this setting the adjoint of P_I is the canonical embedding $P_I^* : \ell^2(I) \to \ell^2$

$$P_I^* x = \sum_{i\in I} x_i e_i.$$

In these terms we define $K_I := K(P_I^*) : \ell^2(I) \to \mathcal{H}$.

Remark 2.2.1. All statements that we make in this work do also hold if we have $D \subset \mathbb{N}$ and consider $K \in L(\ell^2(D), H)$. In this case the index set I for the definition of P_I, P_I^* and K_I has to be a subset of D. The minor role of the index set reflects in our abbreviation $\ell^2 = \ell^2(\mathbb{N})$.

Remark 2.2.2. The requirement that K is defined on ℓ^2 and that $x^\dagger$ is sparse with respect to the standard basis might appear as a strong requirement. In the following we draft that this is not the case.

Let $A \in L(\widetilde{\mathcal{H}}, \mathcal{H})$ a bounded linear operator between two Hilbert spaces and $D \subset \mathbb{N}$ an index set. Further, let $\{\varphi_i \in \widetilde{\mathcal{H}} : i \in D\}$ a collection of vectors (e.g. an orthonormal basis of $\widetilde{\mathcal{H}}$ if $\widetilde{\mathcal{H}}$ is separable), allowing a sparse representation of the searched for solution of $Ax = y$. We assume that the linear synthesis operator $S : \ell^2(D) \to \widetilde{\mathcal{H}}$,

$$Sx = \sum_{i \in D} x_i \varphi_i$$

is bounded (which is fulfilled for orthonormal bases of separable Hilbert spaces). Defining $K := AS \in L(\ell^2(D), \mathcal{H})$, instead of $Ax = y$ we can consider the equation $Kx = y$. In this fashion many problems can be transferred into our framework. In our context, a suitable set $\{\varphi_i \in \widetilde{\mathcal{H}} : i \in D\}$ is referred to as dictionary.

Remark 2.2.3. For the functionals Ψ_α and $\Phi_{\alpha,\beta}$ it is often desirable to fine tune the parameters. Hence, one replaces

$$\alpha\|x\|_{\ell^1} \rightsquigarrow \sum \alpha_i |x_i|, \quad \beta\|x\|_{\ell^2}^2 \rightsquigarrow \sum \beta_i |x_i|^2$$

where $(\alpha_i)_{i\in\mathbb{N}}$, $(\beta_i)_{i\in\mathbb{N}}$ are sequences of positive real numbers which are bounded from below by some constant $c > 0$. All stated theorems, lemmas, etc. in this work also hold for this modification but for notational simplicity are proved without it.

2.3 Gâteaux differentiability

This section introduces the concept of Gâteaux differentiability which plays a fundamental role for the development of algorithms in the elastic-net chapter. The principle can be found in full details in [50,58]. Here, we only recall the important properties.

Definition 2.3.1 (Gâteaux derivative)**.** Let $\mathcal{X}$,$\mathcal{Y}$ be normed spaces, $D \subset X$ an open set. A functional $f : D \to \mathcal{Y}$ is called Gâteaux differentiable at $x_0 \in D$ if there exists $T \in L(\mathcal{X}, \mathcal{Y})$ such that

$$\forall v \in \mathcal{X} : \quad \lim_{h \to 0} \frac{f(x_0 + hv) - f(x_0)}{h} = Tv.$$

In this case, we denote T as $Df(x_0)$ and call it Gâteaux derivative of f at x_0.

The next proposition shows that many properties of the usual derivative transfer to the concept of Gâteaux derivatives. It states a sum and a chain rule as well as its compatibility with scalar multiplication. The most important part is that it admits a necessary and sufficient condition for convex functionals to have a global minimizer as known for usual derivatives.

Proposition 2.3.2 (Properties of the Gâteaux derivative)**.** *Let $\mathcal{X}$,$\mathcal{Y}$ be Banach spaces and $f, g : \mathcal{X} \to \mathcal{Y}$ Gâteaux differentiable. Then the following properties hold:*

i) The function $f + g$ is Gâteaux differentiable and it holds the sum rule

$$D(f + g)(x_0) = Df(x_0) + Dg(x_0), \quad x_0 \in \mathcal{X}.$$

ii) For each scalar λ the function λf is Gâteaux differentiable with the Gâteaux derivative

$$D(\lambda f)(x_0) = \lambda Df(x_0), \quad x_0 \in \mathcal{X}.$$

iii) The function $f \circ g$ is Gâteaux differentiable and the following chain rule holds:

$$D(f \circ g) = Df(g(x_0)) \circ Dg(x_0), \quad x_0 \in \mathcal{X}.$$

iv) If f is convex and $x_0 \in \mathcal{X}$ such that $Df(x_0) = 0$, then x_0 is a global minimizer of f.

To get a better assessment for Gâteaux derivatives as well as to anticipate some work we state a few examples.

Example 2.3.3. For this example, we use the notations which are introduced in the previous section.

- Gâteaux derivative of a constant function: It is easy to see from the definition of Gâteaux differentiability that the Gâteaux derivative of a constant function is zero.

- Gâteaux derivative of K: The short calculation

$$\frac{K(x_0 + hv) - K(x_0)}{h} = \frac{hKv}{h} = Kv$$

shows that $DK(x_0) = K$ for every $x_0 \in \ell^2$.

- Gâteaux derivative of $\frac{1}{2}\|\cdot\|_{\mathcal{H}}^2 : \mathcal{H} \to \mathbb{R}$: We start the calculation by rewriting the norm in terms of the inner product and exploiting its linearity:

$$\tfrac{1}{2}(\|x_0 + hv\|_{\mathcal{H}}^2 - \|x_0\|_{\mathcal{H}}^2) = h\mathrm{Re}\langle x_0, v\rangle_{\mathcal{H}} + h^2\tfrac{1}{2}\|v\|_{\mathcal{H}}^2.$$

Plugging this result into definition 2.3.1 shows that the Gâteaux derivative at x_0 is given by $\mathrm{Re}\langle x_0, \cdot\rangle_{\mathcal{H}}$.

- Gâteaux derivative of $\frac{1}{2}\|K\cdot - y^\delta\|_{\mathcal{H}}^2 : \ell^2 \to \mathbb{R}$: Using the chain rule stated in proposition 2.3.2 and the previous examples, we find that the Gâteaux derivative at $x_0 \in \ell^2$ is given by

$$D(\tfrac{1}{2}\|Kx_0 - y^\delta\|_{\mathcal{H}}^2) = \mathrm{Re}\langle Kx_0 - y^\delta, \cdot\rangle_{\mathcal{H}} \circ K = \mathrm{Re}\langle K^*(Kx_0 - y^\delta), \cdot\rangle_{\ell^2}.$$

The "Re" function can be omitted since ℓ^2 is assumed to be a real Hilbert space. By the theorem of Riesz, we can identify

$$D(\tfrac{1}{2}\|Kx_0 - y^\delta\|_{\mathcal{H}}^2) = K^*(Kx_0 - y^\delta) \in \ell^2.$$

2.4 Subdifferential calculus

In the last section, we have, as an example, calculated the Gâteaux derivative of the residual part of the functionals Ψ_α and $\Phi_{\alpha,\beta}$. With a view on the penalty terms of the functionals, the problem arises that the term $\alpha\|x\|_{\ell^1}$ is not Gâteaux differentiable. In this section, we introduce a principle which enables us to overcome this problem by introducing the more general concept of subdifferentiability [44, 46, 50]. We restrict ourself to subgradient calculus for convex functions.

Recall. A function f, defined on a Banach space $\mathcal{X}$, mapping into $\mathbb{R}_\infty$ is called convex if

$$\forall x, y \in \mathcal{X},\ \forall t \in [0,1]: \quad f(tx + (1-t)y) \leq tf(x) + (1-t)f(y).$$

It is called strictly convex if the strict inequality holds for $x \neq y$ and for $t \in (0,1)$. The effective domain of f is defined by $\operatorname{dom}(f) := \{x \in \mathcal{X} : f(x) < \infty\}$.

Definition 2.4.1 (Subdifferential of a convex function)**.** Let $\mathcal{X}$ be a Banach space and $f : \mathcal{X} \to \mathbb{R}_\infty$ a convex function. Then the subdifferential of f at $x_0 \in \operatorname{dom}(f)$ is defined by

$$\partial f(x_0) := \{t \in \mathcal{X}' : f(x) - f(x_0) \geq t(x - x_0), x \in \mathcal{X}\}.$$

The function f is called subdifferentiable at x_0 if and only if the subdifferential is not empty.

Remark 2.4.2. If $\mathcal{X}$ is a Hilbert space, by the Riesz representation theorem the subgradient can be identified with

$$\partial f(x_0) = \{x^* \in \mathcal{X} : f(x) - f(x_0) \geq \langle x^*, x - x_0\rangle_{\mathcal{X}}, x \in \mathcal{X}\}.$$

Remark 2.4.3. The subgradient of f at x_0 is a singleton if f is Gâteaux differentiable and $x_0 \in \operatorname{int}\operatorname{dom}(f)$. In that case, we may write $Df(x_0) = \partial f(x_0)$ instead of $Df(x_0) \in \partial f(x_0)$ or $\{Df(x_0)\} = \partial f(x_0)$.

Remark 2.4.4. If $0 \in \partial f(x_0)$ it follows from the definition that f attains its global minimum at x_0 since

$$\forall x \in \mathcal{X}: \quad f(x) \geq f(x_0)$$

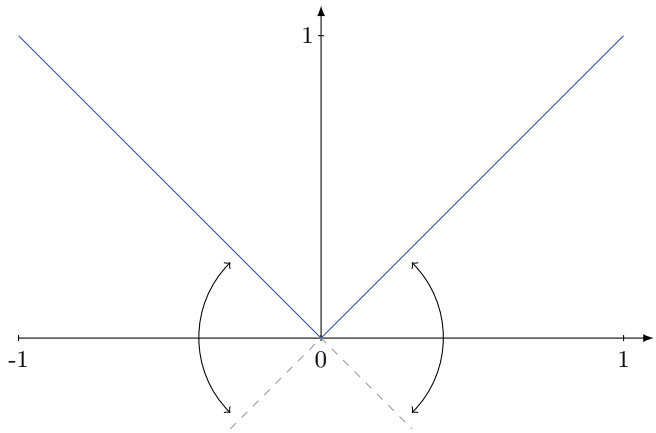


Figure 2.4.1: Visualization of the subgradient of the absolute value function.

Definition 2.4.5. We define the set valued Sign function as follows: $\mathrm{Sign}(x) : \mathbb{R} \to \{\{1\}, \{-1\}, [-1, 1]\}$

$$\mathrm{Sign}(x) = \begin{cases} \{\mathrm{sign}(x)\}, & x \neq 0, \\ [-1, 1], & x = 0. \end{cases}$$

Analogously, we define the function on ℓ^2 via $\mathrm{Sign}(x)_i = \mathrm{Sign}(x_i)$.

Example 2.4.6.

- Subgradient of $|\cdot| : \mathbb{R} \to \mathbb{R}$: In this case the condition for the subgradient reads $\partial|0| = \{x^* \in \mathbb{R} : |x| \geq x \cdot x^*, x \in \mathbb{R}\}$. By the Riesz representation theorem, it is easy to see that the subgradient is given by

  $$\partial|0| = [-1, 1],$$

  which is the set of all possible tangent slopes (figure 2.4.1). Further it is easy to see that for general $x \in \mathbb{R}$ there holds

  $$\partial|x| = \mathrm{Sign}(x).$$

- Subgradient of $\|\cdot\|_{\ell^1} : \ell^2 \to \mathbb{R}$: Using the Riesz representation theorem for ℓ^2, we start with the consideration of $\partial\|0\|_{\ell^1}$ which contains $x^* \in \ell^2$

if and only if

$$\forall x \in \ell^2 : \quad \sum |x_i| \geq \langle x, x^* \rangle_{\ell^2} = \sum x_i x_i^*.$$

Inserting the standard basis vectors and its scalar multiples, this formula can be reduced to the one dimensional case already covered above. Hence, by means of the Sign function, we can write for a general $x_0 \in \operatorname{dom}(\|\cdot\|_{\ell^1}) = \ell^1$

$$\partial \|x_0\|_{\ell^1} = \{x \in \ell^2 : x_i \in \operatorname{Sign}(x_{0i})\} = \operatorname{Sign}(x_0) \cap \ell^2.$$

Remark that $0 \in \operatorname{Sign}(x_0)$ implies that $0 \in \partial \|x_0\|_{\ell^1}$. We can see that Ψ_α and $\Phi_{\alpha,\beta}$ are subdifferentiable only for $x_0 \in \ell^0$, since otherwise $|\operatorname{Sign}(x)|_i = \pm 1$ for infinitely many indices $i \in \mathbb{N}$. For ℓ^2 consists of sequences tending to zero, the intersection is empty.

Proposition 2.4.7. *Let $\mathcal{X}$ be a Hilbert space and $f, g : \mathcal{X} \to \mathbb{R}$ convex. If there exists a point in $\operatorname{dom}(f) \cap \operatorname{dom}(g)$ at which f is continuous, then*

$$\partial(f + g) = \partial(f) + \partial(g).$$

Remark 2.4.8. Applying the previous proposition to Ψ_α and $\Phi_{\alpha,\beta}$ and combining example 2.3.3 and 2.4.6, we find that

$$\begin{aligned} \partial\Psi_\alpha(x) &= K^*(Kx - y^\delta) + \alpha \operatorname{Sign}(x) \cap \ell^2, \\ \partial\Phi_{\alpha,\beta}(x) &= K^*(Kx - y^\delta) + \alpha \operatorname{Sign}(x) \cap \ell^2 + \beta x \end{aligned}$$

each defined for $x \in \ell^0$. From remark 2.4.4 we also know that $0 \in \partial\Psi_\alpha(x)$ or $0 \in \partial\Phi_{\alpha,\beta}(x)$ respectively are sufficient conditions on x to be a minimizer of the corresponding functional. In fact, the conditions are not only sufficient but also necessary conditions.

In [46] we can find the following inversion rule for subgradients:

Proposition 2.4.9 (Inversion rule for subgradients)**.** *Let $\mathcal{X}$ be a Hilbert space and $f : \mathcal{X} \to \mathbb{R}$ lower semicontinuous and convex with $\operatorname{dom}(f) \neq \emptyset$. Then $(\partial f)^{-1}(x) = \operatorname{argmax}_{y \in \mathcal{X}} \langle x, y \rangle_{\mathcal{X}} - f(y)$.*

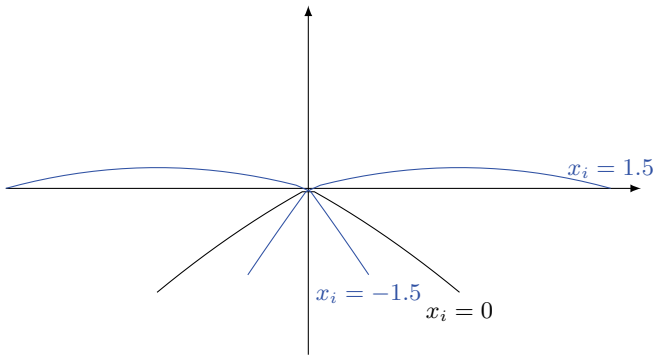


Figure 2.4.2: Graph of f_{x_i} for different values of x_i and $\alpha = \beta = 1$.

Example 2.4.10. In this example, we consider the latter part of the functional $\Phi_{\alpha,\beta}$, namely, $\alpha\|x\|_{\ell^1} + \frac{1}{2}\beta\|x\|_{\ell^2}^2$ as a functional on ℓ^2. We already know that $\partial(\alpha\|x\|_{\ell^1} + \frac{1}{2}\beta\|x\|_{\ell^2}^2) = \alpha\,\mathrm{Sign}(x) + \beta x$. Now, we calculate the inverse of the subgradient for $\beta > 0$ which is given by

$$x \longmapsto \operatorname*{argmax}_{y\in\ell^2} \langle y, x\rangle_{\ell^2} - \alpha\|y\|_{\ell^1} - \tfrac{1}{2}\beta\|y\|_{\ell^2}^2.$$

For a fixed vector $x \in \ell^2$, we can write the right-hand side as

$$\operatorname*{argmax}_{y\in\ell^2} \sum_i (x_i y_i - \alpha|y_i| - \tfrac{1}{2}\beta|y_i|^2).$$

We can maximize this sum component wise as a set of one dimensional optimization problems where y_i is the maximizing argument of the function (figure 2.4.2)

$$f_{x_i} : \mathbb{R} \to \mathbb{R}, \quad y \mapsto x_i y - \alpha|y| - \tfrac{\beta}{2}y^2.$$

At first, we notice that $f_{x_i}(0) = 0$ and that $y = 0$ maximizes f_{x_i} if $|x_i| \le \alpha$ since

$$|x_i| \le \alpha \implies |x_i y| = |x_i| \cdot |y| \le \alpha|y| \implies x_i y - \alpha|y| \le 0 \implies f_{x_i} \le 0.$$

For $y \neq 0$ the optimality condition for this function reads $x_i - \alpha\,\mathrm{sign}(y) - \beta y = 0$ which is solvable if and only if $|x_i| > \alpha$. Looking for $y > 0$ to

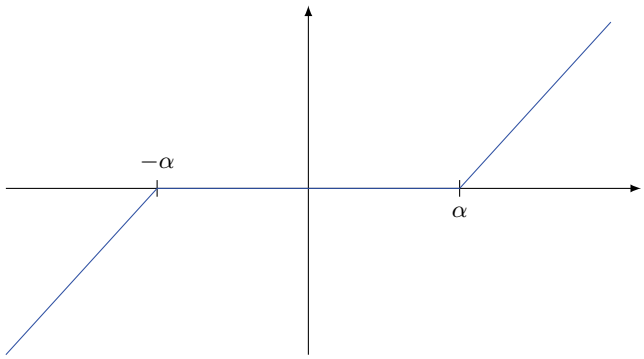


Figure 2.4.3: Graph of the 1D soft shrinkage function S_α.

solve the optimality condition, we find that x_i has to be greater than α and $y = \frac{x_i - \alpha}{\beta}$. Respectively, if we are looking for $y < 0$ we find that x_i has to be less than α and $y = \frac{x_i + \alpha}{\beta}$. All together, we have for $|x_i| > \alpha$ that either $y = \frac{x_i - \alpha\,\mathrm{sign}(x_i)}{\beta}$ or $y = 0$ maximizes f_{x_i}. Next, we consider $f_{x_i}(\frac{x_i - \alpha\,\mathrm{sign}(x_i)}{\beta})$ which reads

$$\frac{x_i^2 - \alpha|x_i|}{\beta} - \frac{\alpha|x_i - \alpha\,\mathrm{sign}(x_i)|}{\beta} - \frac{(x_i - \alpha\,\mathrm{sign}(x_i))^2}{2\beta}.$$

As first case, we take $x_i > \alpha$ where the above formula can be estimated as

$$\frac{x_i^2 - \alpha x_i}{\beta} - \frac{\alpha x_i - \alpha^2}{\beta} - \frac{(x_i - \alpha)^2}{2\beta} = \frac{(x_i - \alpha)^2}{\beta} - \frac{(x_i - \alpha)^2}{2\beta} = (x_i - \alpha)^2(\tfrac{1}{\beta} - \tfrac{1}{2\beta}) > 0.$$

Analogously, in the second case $x_i < -\alpha$:

$$\frac{x_i^2 + \alpha x_i}{\beta} + \frac{\alpha x_i + \alpha^2}{\beta} - \frac{(x_i + \alpha)^2}{2\beta} = \frac{(x_i + \alpha)^2}{\beta} - \frac{(x_i + \alpha)^2}{2\beta} = (x_i + \alpha)^2(\tfrac{1}{\beta} - \tfrac{1}{2\beta}) > 0.$$

Hence, we can exclude $y = 0$ as a candidate to maximize f_{x_i} with $|x_i| > \alpha$. All together, we find that the inverse of $\partial(\alpha\|\cdot\|_{\ell^1} + \frac{1}{2}\beta\|\cdot\|_{\ell^2}^2)$ is given by

$$(\partial(\alpha\|\cdot\|_{\ell^1} + \tfrac{1}{2}\beta\|\cdot\|_{\ell^2}^2)^{-1}(x))_i = \begin{cases} 0, & |x_i| \le \alpha \\ \frac{x_i - \alpha\,\mathrm{sign}(x_i)}{\beta}, & \text{otherwise.} \end{cases}, \quad i \in \mathbb{N}. \tag{2.3}$$

This function can also be expressed as $\frac{1}{\beta} S_\alpha(x)_i = \frac{1}{\beta}\max\{0, |x_i| - \alpha\} \cdot \mathrm{sign}(x_i)$ where S_α is known in the literature as soft shrinkage (figure 2.4.3).

2.5 Newton differentiability

As a third notion of differentiability, we need the Newton or slant differentiability as used in [12,31]. It is the key to semismooth Newton methods as we develop for the elastic net later on.

Definition 2.5.1 (Newton derivative)**.** Let $\mathcal{X}$,$\mathcal{Y}$ be Banach spaces, $D \subset \mathcal{X}$ an open subset. A mapping $F : \mathcal{X} \to \mathcal{Y}$ is said to be slantly or Newton differentiable at $x \in D$ if there exists a mapping $G : D \to L(\mathcal{X}, \mathcal{Y})$ such that the family $\{G(x+h\}$ is uniformly bounded in the operator norm for h sufficiently small and

$$\lim_{h \to 0} \frac{F(x+h) - F(x) - G(x+h)h}{\|h\|_{\mathcal{X}}} = 0.$$

The function G is called a slanting function or Newton derivative for F at x.

Remark 2.5.2. A slanting function is not necessarily unique.

Remark 2.5.3. In contrast to Gâteaux differentiability the limit is not directional and hence a Gâteaux differentiable function is not necessarily Newton differentiable. On the other hand, a Newton differentiable function does not need to be Gâteaux differentiable since for a slanting function G of F in x, $G(x)$ itself is not characterized in general by a limit.

Now that we defined the notion of Newton differentiability, we also need the following proposition proved in [12].

Proposition 2.5.4. *Let $\mathcal{X}$,$\mathcal{Y}$ be Banach spaces. A mapping $F : \mathcal{X} \to \mathcal{Y}$ is Newton differentiable at x if and only if F is Lipschitz continuous at x.*

Example 2.5.5. The shrinkage function $S_\alpha : \mathbb{R} \to \mathbb{R}$ is Newton differentiable at any point $x \in \mathbb{R}$ (since it clearly is Lipschitz continuous). According to [28], a Newton derivative is given by

$$G(x) = \begin{cases} 0 : \mathbb{R} \to \mathbb{R}, & |x| \leq \alpha \\ id : \mathbb{R} \to \mathbb{R}, & |x| > \alpha \end{cases}$$

as a Newton derivative. Accordingly, the shrinkage function $S_\alpha : \ell^2 \to \ell^2$ is Newton differentiable at any point $x \in \ell^2$ with

$$G(x) = \begin{pmatrix} \mathrm{id}_{\{i \in \mathbb{N} : |x_i| > \alpha\}} & 0 \\ 0 & 0 \end{pmatrix} : \ell^2 \to \ell^2$$

as a Newton derivative (cf. [28]).

3 A short trip to ℓ^1 minimization

Remeber our framework, i.e. we consider a Hilbert space $\mathcal{H}$ and a bounded linear operator $K \in L(\ell^2, \mathcal{H})$. Having $y \in R(K)$ we assume that $Kx = y$ has a sparse solution $x^\dagger$. Further, y^δ is a noisy observation of y.

In this chapter, we show how the problem of sparse approximation can be transfered into a so called ℓ^1-minimization problem, i.e. minimizing the functional

$$\Psi_\alpha(x) = \tfrac{1}{2}\|Kx - y^\delta\|_{\mathcal{H}}^2 + \alpha\|x\|_{\ell^1}.$$

We discuss analytical properties of ℓ^1 minimization, e.g. an optimality condition and the problem of selecting α. After the theoretical treatise, we turn to the practical question of solving the minimization problem. In this context, we present two classes of algorithms, iterative thresholding methods and active-set methods, as well as several representatives.

3.1 From sparse approximation to ℓ^1 minimization

In the introduction, we have seen the mathematical characterization of sparse approximation via

$$\min \|x\|_{\ell^0}, \quad \text{s.t. } \|Kx - y^\delta\|_{\mathcal{H}} < \tau. \tag{3.1}$$

Unfortunately, solving this problem is computationally unfeasible since it is NP hard [43]. Thus, there are several strategies to approximate solutions,

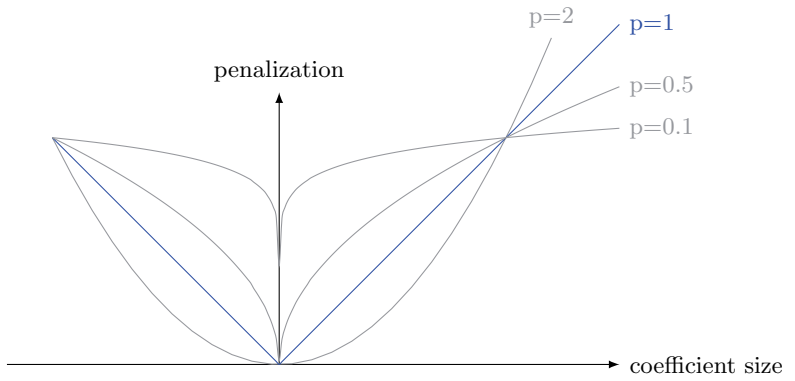


Figure 3.1.1: How does $\|\cdot\|_{\ell^p}$ influences the minimal value (3.1) with respect to the size of the coefficients x_i.

e.g. matching pursuit [16, 41, 51] or basis pursuit [11]. A good overview is given in [8, 54].

The interesting thing, using $\|\cdot\|_{\ell^0}$ to measure the coefficients is that every non-zero coefficient affects the minimal value of problem (3.1) by 1, regardless of the coefficients value. In contrast, using the $\|\cdot\|_{\ell^2}$ would increase the impact of big coefficients while reducing those of small coefficients. Roughly speaking, using the $\|\cdot\|_{\ell^0}$ leads to minimizers with as few non-zero entries as possible, while the $\|\cdot\|_{\ell^2}$ leads to minimizers containing many tiny coefficients, instead of a few big ones. Using the ℓ^p norms $\|\cdot\|_{\ell^p}$ for $p \in (0, 2)$ one can somehow interpolate this behavior (figure 3.1.1). For replacing $p = 0$ by $p \in (0, 1)$ we refer to [25, 60] and for $p \in [1, 2)$ we recommend the reference [13].

For our purposes, we follow a widely used variant via convex relaxation using $\|\cdot\|_{\ell^1}$, which is in the above sense the closest convex norm to $\|\cdot\|_{\ell^0}$, leading to the minimization problem

$$\min \|x\|_{\ell^1}, \quad \text{s.t. } \|Kx - y^\delta\|_{\mathcal{H}} < \tau. \tag{3.2}$$

This is similar to the lasso problem [21, 48, 61]. We want to mention that for the noise free case, i.e. $\delta = 0$, one would use $\tau = 0$ together with the non-strict inequality. In that case, one can show that under some conditions the minimizers of (3.1) and (3.2) coincide (cf. [18, 19]).

However, using the Lagrangian multipliers there exists $\alpha > 0$ such that the solution of problem (3.2) is equivalent to the minimization of the convex optimization problem (cf. [40, 48])

$$\min \tfrac{1}{2}\|Kx - y^\delta\|_{\mathcal{H}}^2 + \alpha\|x\|_{\ell^1}.$$

This problem is referred to as ℓ^1-minimization problem.

Lemma 3.1.1 (Optimality condition & Sparsity property)**.** *A necessary and sufficient condition on a minimizer of Ψ_α is given by*

$$-K^*(Kx - y^\delta) \in \alpha \operatorname{Sign}(x). \tag{3.3}$$

Moreover, each minimizer x^* of Ψ_α is sparse, i.e. has only finitely many non-zero entries.

Proof. By remark 2.4.8, we know that (3.3) is a sufficient condition on a minimizer of Ψ_α. Moreover, we showed in example 2.4.6 that every x meeting the optimality condition is sparse i.e. an element of ℓ^0. To the other end, let x^* be a minimizer of Ψ_α, then $x^* \in \ell^0$ since otherwise we could project the problem onto the components where $x_i^* \neq 0$ and calculate the optimality condition for Ψ_α – so assume w.l.o.g. $x_i \neq 0$, $i \in \mathbb{N}$ – which would be

$$|-K^*(Kx^* - y^\delta)| = \alpha.$$

This is a contradiction to K^* maps into ℓ^2. Hence, the subgradient is not empty for the minimizer of Ψ_α and also admits a necessary condition.

Note that we do not need $\operatorname{Sign}(x) \cap \ell^2$ on the right hand side since K^* maps into ℓ^2. □

Remark 3.1.2. If we consider the case of a unitary operator K, i.e. $K^*K = \mathrm{id}$, we find from the above optimality condition (3.3) that

$$K^*y^\delta \in \alpha \operatorname{Sign}(x) + x.$$

Applying the outcome of example 2.4.10 on the inversion of subgradients, we deduce

$$x = S_\alpha(K^*y^\delta)$$

with the soft-shrinkage function S_α. Hence, the solution can be calculated explicitly. This fact can be used for example for denoising where $K = \mathrm{id}$.

A question naturally arising in the context of ℓ^1 minimization is how to choose the parameter α. In fact, this is, to our best knowledge, an open question if one is seeking a minimizer with respect to sparsity motivated criteria. For example, if one is seeking a five-sparse minimizer. At least, it is possible to tighten the interval of meaningful choices for α by finding out for which value of α a minimizer is given by 0. From the above optimality condition, we can find that 0 is a minimizer if and only if

$$(K^* y^\delta)_i \in [-\alpha, \alpha], \quad i \in \mathbb{N} \tag{3.4}$$

and hence a suitable α should be not outside the interval $[0, \max_{i\in\mathbb{N}} |K^* y^\delta|_i]$. In general, for K not injective the minimizer of Ψ_α is not unique but if we assume that for $\alpha \geq \max_{i\in\mathbb{N}} |K^* y^\delta|_i$ there exist an additional minimizer x^* we find

$$\begin{aligned}
&\tfrac{1}{2}\|Kx^* - y^\delta\|_{\mathcal{H}}^2 + \alpha\|x^*\|_{\ell^1} = \tfrac{1}{2}\|y^\delta\|_{\mathcal{H}}^2 \\
&\implies \tfrac{1}{2}\|Kx^*\|_{\mathcal{H}}^2 - \langle Kx^*, y^\delta\rangle + \alpha\|x^*\|_{\ell^1} = 0 \\
&\implies \tfrac{1}{2}\|Kx^*\|_{\mathcal{H}}^2 + \alpha\|x^*\|_{\ell^1} = \langle Kx^*, y^\delta\rangle = \langle x^*, K^* y^\delta\rangle \overset{(3.4)}{\leq} \alpha\|x^*\|_{\ell^1} \\
&\implies \tfrac{1}{2}\|Kx^*\|_{\mathcal{H}}^2 = 0.
\end{aligned}$$

Hence, we have for the reconstruction $Kx^* = 0$ and reinserting this result into the first equation we obtain $x^* = 0$. This proves the upper bound for the interval of suitable choices for α, moreover, it proved the next lemma.

Lemma 3.1.3. *The unique minimizer of Ψ_α is given by 0 if and only if $\alpha \geq \max_{i\in\mathbb{N}} |K^* y^\delta|_i$.*

Stronger parameter-choice rules can be established if one is not focused on the sparsity property of Ψ_α itself but on the regularizing properties to inverse problems. Since this is not the main interest of this work we refer to [22, 26, 38, 45].

We address the question of how to select α if one can hope for exact recovery of $x^\dagger$, in the sense that the support of the minimizer of Ψ_α coincides with the support of $x^\dagger$, later for the more general elastic-net case. For now, we refer to [24, 39, 52, 53].

3.2 An overview of minimization algorithms

Essentially, there are two classes of minimization algorithms for the functional Ψ_α, the active-set type algorithms and iterative thresholding methods. A prototype of thresholding methods is the iterated soft thresholding (IST) [13]. From the performance point of view this algorithm is often not reasonable but there are many spin offs dealing with those performance deficiencies, e.g. fixed-point continuation (FPC) [29]. Iterative thresholding algorithms often are proven to have a linear convergence behavior but the rate may be arbitrarily poor.

On the other side, the active-set methods, e.g. the semismooth Newton method (SSN) [28] or the feature-sign-search algorithm (FSS) [37], are proven only local convergence or the convergence speed scales exponentially with the dimension of the problem. Anyway, in practice these algorithms often outperform iterative thresholding methods.

At this point, we shall refer to a few more well known algorithms for ℓ^1 minimization to give a broader but still incomplete picture of the algorithmic landscape: Gradient projection for sparse reconstruction (GPSR) [23], iterated hard thresholding (IHT) [7], fast iterative soft-thresholding algorithm (FISTA) [3], sparse reconstruction by separable approximation (SPARSA) [59], l1-ls [34] and the homotopy method [20]. Some of them might occur in the numerical examples later on. A good comparison of some of them can be found in [40].

In this work we mainly focus on algorithms of active set type, namely, the SSN and the FSS. As a reference for iterative thresholding methods we take the IST. The SSN is a locally super linearly convergent method which indeed often suffers from convergence problems. We consider it mainly for analytical purposes since it has a nice and straight forward analysis. For practical purposes, we use the FSS which is similar to the SSN but with a more modest active-set strategy. It benefits from its global convergence properties and while the proved rate of convergence is exponentially bad in practice it often converges rapidly.

The IST is globally convergent with a linear rate which may be arbitrarily bad. It is the prototype of most iterated shrinkage methods and often used for verification.

4 The elastic net

In several applications, ℓ^1 minimization yields good results, while in our experience many algorithms have numerical troubles when applied to rank-deficient or ill-conditioned operators. Many algorithms take longer (more iterations) to minimize the ℓ^1 functional as the operator becomes worse conditioned until convergence can no longer be achieved. Sometimes, one is satisfied with approximate solutions to the ℓ^1-minimization problem which can often be achieved by relaxing the stopping criterion of the used algorithm, e.g. by demanding that the step size is below a threshold. Usually, in this case there are no qualitative analytical informations on the "solutions" available yet, depending on the used stopping criteria. Even if there were such information the question arises what to do if the result is not satisfactory. Hence, we find it worthwhile to establish a way, abolishing the numerical problems, such that qualitative results can be achieved. Furthermore, we try to open a possibility to obtain the exact ℓ^1 minimizer anyway. In this aim, we focus mostly on the stabilization of active-set methods.

In this chapter we present a generalization of the ℓ^1 functional, the elastic net. We prove similar analytical properties for the elastic net as for ℓ^1 minimization and show how to obtain algorithms for the minimization of the elastic-net functional from existing ℓ^1 algorithms. On top, we derive two active-set algorithms explicitly for the elastic net showing that we can derive algorithms for the elastic net which do not transfer from ℓ^1 minimization. We see later on that the elastic net is capable of rectifying or at least attenuating the numerical problems of ℓ^1 minimization.

4.1 Relations to ℓ^1 minimization

Our way to the elastic net started with the observation of numerical troubles in ℓ^1-minimization algorithms. Especially, we worked on the SSN and FSS active-set methods for ℓ^1 minimization. In every iteration these methods basically have to invert a submatrix of

$$K^*K.$$

For ill-conditioned operators, the numerical problems are obvious and a known regularization method can help to attenuate the troubles, namely, Tikhonov regularization, i.e. for small $\beta > 0$ replace the inversion by the inversion of

$$K^*K + \beta \,\mathrm{id}\,.$$

Our experiments have shown that this regularization in fact leads to reasonable results. Driven by the question of how these modifications affect the output of the algorithm, we ended up with the answer that it minimizes a functional, which later turned out to be the elastic net, i.e.

$$\Phi_{\alpha,\beta}(x) = \tfrac{1}{2}\|Kx - y^\delta\|_{\mathcal{H}}^2 + \alpha\|x\|_{\ell^1} + \tfrac{1}{2}\beta\|x\|_{\ell^2}^2.$$

The elastic net $\Phi_{\alpha,\beta}$ is a tool developed in statistics [15, 62]. It is obvious that the elastic net is a generalization of the ℓ^1 functional since they coincide for $\beta = 0$. A fortiori it is surprising that also the ℓ^1 functional can be interpreted as a generalization of the elastic net in the following sense. We can rephrase the elastic-net functional as

$$\Phi_{\alpha,\beta}(x) = \frac{1}{2}\left\| \begin{pmatrix} K \\ \sqrt{\beta}\,\mathrm{id} \end{pmatrix} x - \begin{pmatrix} y^\delta \\ 0 \end{pmatrix} \right\|_{\mathcal{H}\times\ell^2}^2 + \alpha\|x\|_{\ell^1} \tag{4.1}$$

which is a restriction of the ℓ^1 functional to special operators and datasets. This rephrased version shows that many properties of the ℓ^1 functional transfer immediately to the elastic net. Anyway, we can prove some statements which are not obvious from this redraft and for convenience we proof some results which do transfer but are not proven in the last chapter.

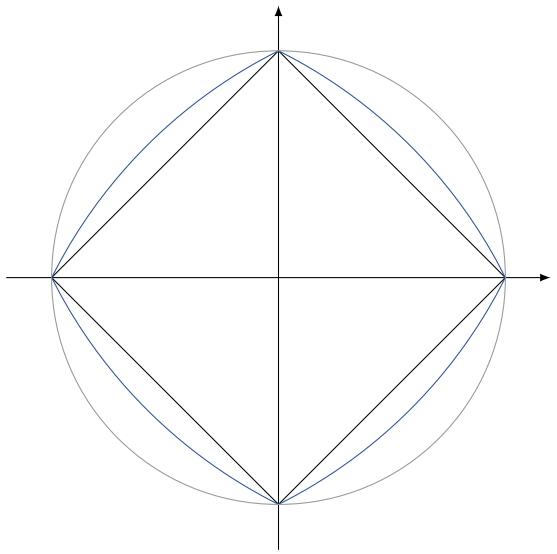


Figure 4.1.1: Contour of ℓ^1 (black), ℓ^2 (gray) and elastic net (blue) penalty term in $\mathbb{R}^2$.

4.2 Properties of the elastic-net functional and its minimizers

A property of the elastic-net functional $\Phi_{\alpha,\beta}$, which we exploit extensively, is its convexity. That $\Phi_{\alpha,\beta}$ is convex could also be seen from redrafted elastic-net functional (4.1) since it is known that Ψ_α is convex. Moreover, since $\begin{pmatrix} K \\ \sqrt{\beta}\,\mathrm{id} \end{pmatrix}$ is always injective for $\beta > 0$ we have that $\Phi_{\alpha,\beta}$ is strictly convex for $\beta > 0$. This fact is summarized in the following lemma and is proved for the sake of completeness.

Lemma 4.2.1 (Convexity)**.** *The elastic-net functional is convex, if $\beta > 0$ or K injective also strictly convex.*

Proof. For this proof we exploit the linearity of K, the triangle inequality

and the convexity of the function $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$:

$$\begin{aligned}
&\Phi_{\alpha,\beta}(\lambda x + (1-\lambda)z) \\
&\quad = \tfrac{1}{2}\|K(\lambda x + (1-\lambda)z) - y^\delta\|_{\mathcal{H}}^2 + \alpha\|\lambda x + (1-\lambda)z\|_{\ell^1} \\
&\qquad + \tfrac{1}{2}\beta\|\lambda x + (1-\lambda)z\|_{\ell^2}^2 \\
&\quad = \tfrac{1}{2}\|\lambda Kx + (1-\lambda)Kz - \lambda y^\delta - (1-\lambda)y^\delta)\|_{\mathcal{H}}^2 + \alpha\|\lambda x + (1-\lambda)z\|_{\ell^1} \\
&\qquad + \tfrac{1}{2}\beta\|\lambda x + (1-\lambda)z\|_{\ell^2}^2 \\
&\quad \leq \tfrac{1}{2}(\lambda\|Kx - y^\delta\|_{\mathcal{H}}^2 + (1-\lambda)\|Kz - y^\delta)\|_{\mathcal{H}}^2 + \alpha(\lambda\|x\|_{\ell^1} + (1-\lambda)\|z\|_{\ell^1}) \\
&\qquad + \tfrac{1}{2}\beta(\lambda\|x\|_{\ell^2}^2 + (1-\lambda)\|z\|_{\ell^2}^2) \\
&\quad = \lambda\Phi_{\alpha,\beta}(x) + (1-\lambda)\Phi_{\alpha,\beta}(z).
\end{aligned}$$

The strict convexity for $\beta > 0$ follows from the last inequality where we used the triangle inequality and the strict convexity of the square function on the second penalty term. The functional might be not strictly convex for $\beta = 0$ because in the same inequality there might be $Kx = Kz$. Hence, for K injective the functional is again strictly convex. □

From the convexity of the elastic-net functional $\Phi_{\alpha,\beta}$, it follows a result on the uniqueness of its minimizers. It is known from the rewritten formula (4.1) for $\Phi_{\alpha,\beta}$ using known facts about ℓ^1 minimization. Again, we state the proof for the sake of completeness.

Lemma 4.2.2 (Uniqueness of the minimizer)**.** *Let $\beta > 0$ or K injective. The minimizer of $\Phi_{\alpha,\beta}$ is unique.*

Proof. Let x_1, x_2 be minimizers of $\Phi_{\alpha,\beta}$. Assuming $x_1 \neq x_2$ and using the strict convexity of $\Phi_{\alpha,\beta}$ we conclude

$$\Phi_{\alpha,\beta}(\tfrac{1}{2}(x_1 + x_2)) < \tfrac{1}{2}(\Phi_{\alpha,\beta}(x_1) + \Phi_{\alpha,\beta}(x_2)) = \Phi_{\alpha,\beta}(x_1).$$

The last equality follows from the fact that x_1 and x_2 are minimizers and hence $\Phi_{\alpha,\beta}(x_1) = \Phi_{\alpha,\beta}(x_2)$. Thus, the equation establishes a contradiction to the minimizing property of x_1 and hence proofs the claim. □

Lemma 4.2.1 and lemma 4.2.2 become important in the algorithmic section of this chapter. They are a first hint that the elastic-net functional can free us from the requirement that K needs to be injective. This is important for problems involving rank-deficient operators.

The next result is an optimality condition for the elastic-net functional $\Phi_{\alpha,\beta}$, i.e. a necessary and sufficient condition on a vector $x \in \ell^2$ to be a minimizer of the elastic-net functional $\Phi_{\alpha,\beta}$. Again, this condition transfers from ℓ^1 minimization via equation (4.1).

Lemma 4.2.3 (Optimality condition & sparsity property)**.** *An element $x^* \in \ell^2$ is a minimizer of $\Phi_{\alpha,\beta}$ if and only if*

$$-K^*(Kx^* - y^\delta) - \beta x^* \in \alpha \operatorname{Sign}(x^*). \tag{4.2}$$

Moreover, if $\alpha > 0$ a minimizer x^* of $\Phi_{\alpha,\beta}$ is sparse, i.e. $\operatorname{supp}(x^*)$ *is a finite set.*

In the situation of $\beta > 0$, we are able to formulate a variant of the optimality condition for the elastic-net functional $\Phi_{\alpha,\beta}$ which does not hold for the ℓ^1 functional, i.e. $\beta = 0$. It uses the shrinkage operator S_α introduced for the inversion formula for subgradients in section 2.4 resulting in a fixed-point equation.

Lemma 4.2.4 (Optimality condition via shrinkage)**.** *Let $\beta > 0$. An element $x^* \in \ell^2$ is a minimizer of $\Phi_{\alpha,\beta}$ if and only if it fulfills the fixed point equation*

$$S_\alpha(-K^*(Kx^* - y^\delta)) = \beta x^*. \tag{4.3}$$

Proof. The optimality condition $-K^*(Kx^*-y^\delta)-\beta x^* \in \alpha \operatorname{Sign}(x^*)$ (cf (4.2)) can be rewritten equivalently using the inversion formula for subgradients (2.3) from example 2.4.10:

$$\begin{aligned} &- K^*(Kx^* - y^\delta) - \beta x^* \in \alpha \operatorname{Sign}(x^*) \\ &\iff -K^*(Kx^* - y^\delta) \in \alpha \operatorname{Sign}(x^*) + \beta x^* \\ &\iff S_\alpha(-K^*(Kx^* - y^\delta)) = \beta x^*. \end{aligned}$$

□

Remark 4.2.5. As for ℓ^1 minimization, if we consider the case of a unitary operator K, i.e. $K^*K = \mathrm{id}$, we can calculate the minimizer directly. This

can be understood by a similar calculation as in the lemma above:

$$\begin{aligned} & -K^*(Kx^* - y^\delta) - \beta x^* \in \alpha \operatorname{Sign}(x^*) \\ \iff & K^* y^\delta \in \alpha \operatorname{Sign}(x^*) + (\beta + 1)x^* \\ \iff & S_\alpha(K^* y^\delta) = (\beta + 1)x^*. \end{aligned}$$

Here, we again used example 2.4.10 on the inversion of subgradients.

As mentioned in the introduction of this chapter, we aim to regularize ℓ^1-minimization methods such that the results are still a good approximation for the ℓ^1 minimizer. Hence, we desire a stability result of the elastic-net minimizer with respect to the choice of β. In this context the following two interesting properties have been proved in [33].

Theorem 4.2.6 (Stability of the minimizer w.r.t. α and β)**.** *Let $\alpha, \beta > 0$ and $(\alpha_n)_{n\in\mathbb{N}}$,$(\beta_n)_{n\in\mathbb{N}} \subset \mathbb{R}^+$ sequences with $\alpha_n \to \alpha$, $\beta_n \to \beta$. Further, denote by x_n the minimizer of Φ_{α_n,β_n} and by x^* the minimizer of $\Phi_{\alpha,\beta}$ respectively. There holds $x_n \to x^*$.*

Proposition 4.2.7. *Let the sequences $(\alpha_n)_{n\in\mathbb{N}}$,$(\beta_n)_{n\in\mathbb{N}} \subset \mathbb{R}^+$ satisfy that for some $\gamma \geq 0$ and $\alpha > 0$ there holds*

$$\lim_{n\to\infty} \beta_n = 0 \quad \textit{and} \quad \lim_{n\to\infty} \frac{\alpha_n - \alpha}{\beta_n} = \gamma.$$

The minimizer of Φ_{α_n,β_n} converges to the unique minimum $\gamma\|\cdot\|_{\ell^1} + \frac{1}{2}\|\cdot\|^2_{\ell^2}$ element of all minimizers to the functional $\Phi_{\alpha,0}$.

To complete these statements, we close the gap in between, namely, the stability for $\alpha \to 0$. The proof is similar to the proofs of theorem 4.2.6 and proposition 4.2.7 as stated in [33].

Proposition 4.2.8. *Let the sequences $(\alpha_n)_{n\in\mathbb{N}}$,$(\beta_n)_{n\in\mathbb{N}} \subset \mathbb{R}^+$ satisfy that for some $\beta > 0$ there holds*

$$\lim_{n\to\infty} \alpha_n = 0 \quad \textit{and} \quad \lim_{n\to\infty} \beta_n = \beta.$$

The minimizer of Φ_{α_n,β_n} converges to the unique minimizer of $\Phi_{0,\beta}$.

Proof. The minimizing property of $x_n^* := \operatorname{argmin} \Phi_{\alpha_n,\beta_n}(x)$ implies that the sequences $(\|Kx_n^* - y^\delta\|)_{n\in\mathbb{N}}$, $(\|x_n^*\|_{\ell^1})_{n\in\mathbb{N}}$, and $(\|x_n^*\|_{\ell^2})_{n\in\mathbb{N}}$ are uniformly bounded. In particular, there exists a subsequence of $(x_n^*)_{n\in\mathbb{N}}$, also denoted by $(x_n^*)_{n\in\mathbb{N}}$ converging weakly to some $\widetilde{x} \in \ell^2$.

By the weak continuity of K and the weak lower-semicontinuity of norms, we have

$$\|K\widetilde{x} - y^\delta\| \le \liminf_{n\to\infty} \|Kx_n^* - y^\delta\|, \quad \text{and} \quad \|\widetilde{x}\|_{\ell^2} \le \liminf_{n\to\infty} \|x_n^*\|_{\ell^2}. \tag{4.4}$$

Consequently, we have

$$\begin{aligned}
\Phi_{0,\beta}(\widetilde{x}) &= \frac{1}{2}\|K\widetilde{x} - y^\delta\|^2 + \frac{\beta}{2}\|\widetilde{x}\|_{\ell^2}^2 \\
&\le \frac{1}{2}\liminf_{n\to\infty} \|Kx_n^* - y\|^2 + \liminf_{n\to\infty} \alpha_n\|x_n^*\|_{\ell^1} + \liminf_{n\to\infty} \frac{\beta_n}{2}\|x_n^*\|_{\ell^2}^2 \\
&\le \liminf_{n\to\infty} \left\{ \frac{1}{2}\|Kx_n^* - y\|^2 + \alpha_n\|x_n^*\|_{\ell^1} + \frac{\beta_n}{2}\|x_n^*\|_{\ell^2}^2 \right\} \\
&= \liminf_{n\to\infty} \Phi_{\alpha_n,\beta_n}(x_n^*).
\end{aligned}$$

Next, we let $x^* := \operatorname{argmin} \Phi_{0,\beta}(x)$ and want to show that $\Phi_{0,\beta}(x^*) \ge \limsup_{n\to\infty} \Phi_{\alpha_n,\beta_n}(x_n^*)$. To this end, we observe

$$\begin{aligned}
\limsup_{n\to\infty} \Phi_{\alpha_n,\beta_n}(x_n^*) &\le \limsup_{n\to\infty} \Phi_{\alpha_n,\beta_n}(x^*) \\
&= \lim_{n\to\infty} \Phi_{\alpha_n,\beta_n}(x^*) = \Phi_{0,\beta}(x^*)
\end{aligned}$$

by the minimizing property of x_n^*. Consequently,

$$\limsup_{n\to\infty} \Phi_{\alpha_n,\beta_n}(x_n^*) \le \Phi_{0,\beta}(x^*) \le \Phi_{0,\beta}(\widetilde{x}) \le \liminf_{n\to\infty} \Phi_{\alpha_n,\beta_n}(x_n^*).$$

Therefore, $\widetilde{x}$ is a minimizer of $\Phi_{0,\beta}$, and the uniqueness of its minimizer implies $\widetilde{x} = x^*$. Hence, the whole sequence $(x_n^*)_{n\in\mathbb{N}}$ converges weakly to x^*. Next, we show that the functional value $\|x_n^*\|_{\ell^2} \to \|x^*\|_{\ell^2}$, for which it suffices by (4.4) to show that

$$\limsup_{n\to\infty} \|x_n^*\|_{\ell^2} \le \|x^*\|_{\ell^2}.$$

Assume that this does not hold. Then there exists a constant c such that $c := \limsup_{n\to\infty} \|x_n^*\|_{\ell^2}^2 > \|x^*\|_{\ell^2}^2$, and a subsequence of $(x_n^*)_{n\in\mathbb{N}}$, denoted by $(x_n^*)_{n\in\mathbb{N}}$ again, such that

$$x_n^* \to x^* \text{ weakly} \quad \text{and} \quad \|x_n^*\|_{\ell^2}^2 \to c.$$

By the continuity of $\Phi_{0,\beta}(x^*)$ in (α, β), we have

$$\begin{aligned} & \lim_{n\to\infty} \left\{ \frac{1}{2}\|Kx_n^* - y^\delta\|^2 + \alpha_n \|x_n^*\|_{\ell^1} \right\} \\ = \; & \Phi_{0,\beta}(x^*) - \lim_{n\to\infty} \frac{\beta_n}{2}\|x_n^*\|_{\ell^2}^2 \\ = \; & \frac{1}{2}\|Kx^* - y^\delta\|^2 + \frac{\beta}{2}\left(\|x^*\|_{\ell^2}^2 - c\right) \\ < \; & \frac{1}{2}\|Kx^* - y^\delta\|^2. \end{aligned}$$

Here, the last inequality follows from the definition of c. This is in contradiction with the lower-semicontinuity result in equation (4.4). Therefore, we have

$$\limsup_{n\to\infty} \|x_n^*\|_{\ell^2} \leq \|x^*\|_{\ell^2}.$$

This, together with equation (4.4) implies that $\|x_n^*\|_{\ell^2} \to \|x^*\|_{\ell^2}$, from which the desired convergence in ℓ^2 follows directly. □

4.3 Parameter-choice rules

For our purposes, the elastic-net functional is only a method to improve the computability of sparse solutions, such as ℓ^1 solutions, for ill-conditioned or rank-deficient operators. This motivation advises to choose β as small as possible to keep its influence negligible. This choice is rectified by the continuity results in the last section. For some applications, the operator at hand is no longer ill conditioned if restricted to the support of the true minimizer. Hence, the ill condition only affects the ℓ^1 algorithms if the starting guess is too far from the solution. In such cases, the elastic net might be used to calculate an ℓ^1 solution by subsequently reducing the value of β.

As calculated for ℓ^1 minimization in section 3.1, equation (3.4), one can easily see from the optimality condition for $\Phi_{\alpha,\beta}$ that the interval for meaningful values of α can be localized to be at most $[0, \max_{i\in\mathbb{N}} |K^* y^\delta|_i]$. Unfortunately, a similar approach can not be applied to β, even for fixed α. The case of parameter choices for exact recovery is treated separately in chapter 6

However, as for ℓ^1 minimization parameter-selection rules can be established if one is interested in the regularizing properties of the elastic-net functional to inverse problems (cf. [33]).

4.4 Active-set algorithms and properties

We started this chapter by sketching our way to the elastic net, regularizing ℓ^1-minimization algorithms. This might suggest that ℓ^1 algorithms transfer directly as can also be seen from the rewritten elastic-net functional (4.1). However, considering the elastic net directly might yield variants of these algorithms which can not be obtained by applying ℓ^1 algorithms to (4.1). For example, applying the SSN to the reformulated elastic-net functional indeed differs from only regularizing the inversion of K in the SSN since it affects the active-set choice. It is surprising that this change to the active set can be omitted, which we show in the following where we derive the SSN for the elastic net introducing new active-set strategies. After the treatise of the SSN, we consider a variant of the FSS applied to the rewritten elastic-net functional (4.1). The results of this section can be found as the author's contribution to [33].

4.4.1 Regularized semismooth Newton method (RSSN)

We now derive an algorithm for the elastic-net functional $\Phi_{\alpha,\beta}$ based on the semismooth Newton method [31, 56], which in turn coincides with a regularization of the SSN [28] and hence we call it RSSN.

As shown in section 4.1 the SSN algorithm can be adapted easily by rephrasing $\Phi_{\alpha,\beta}$. However, we deduce the algorithm in full detail to show new active-set choices. The first new choice is presented in the deduction, it can not be derived directly from the ℓ^1 algorithm applied to the rephrased elastic-net functional. Another one is derived in a remark afterwards together with the known active-set choice from SSN via rephrasing $\Phi_{\alpha,\beta}$.

From lemma 4.2.4 we know that x is an optimal point of $\Phi_{\alpha,\beta}$ if and only if

$$F(x) := \beta x - S_\alpha(-K^*(Kx - y^\delta)) = 0.$$

As pointed out by the name of this algorithm we want to solve this fixed point problem via Newton's method. S_α is not differentiable in the classical sense and thus a generalized notion of differentiability is required for applying Newton's method. We shall use the notion of Newton-derivative (also known as slant derivative) [12, 31] as introduced in section 2.5 in the basic principles chapter. As stated in example 2.5.5 the shrinkage operator S_α is Newton differentiable. More precisely, we have the next result proved in [28]:

Lemma 4.4.1. *A Newton derivative of S_α is given by*

$$G(x) = \begin{pmatrix} \mathrm{id}_{\{i\in\mathbb{N}:|x_i|>\alpha\}} & 0 \\ 0 & 0 \end{pmatrix}$$

and for any linear operator $T : \ell^2 \to \ell^2$ and any $b \in \ell^2$ a Newton derivative of $S_\alpha(Tx + b)$ is given by $G(Tx + b)T$.

Applying this lemma, a Newton derivative of F is given by $D(x) = \beta\,\mathrm{id} + G(-K^*(Kx - y^\delta))K^*K$. Given a set $A \subset \mathbb{N}$, we split the operator K^*K as

$$K^*K = \begin{pmatrix} M_A & M_{AA^c} \\ M_{A^cA} & M_{A^c} \end{pmatrix}.$$

Upon letting $A_x := \{i \in \mathbb{N} : |K^*(Kx - y^\delta)|_i > \alpha\}$, we have

$$D(x) = \begin{pmatrix} \beta\,\mathrm{id}_{A_x} + M_{A_x} & M_{A_xA_x^c} \\ 0 & \beta\,\mathrm{id}_{A_x^c} \end{pmatrix}.$$

Obviously, A_x is a finite set since K^* maps into ℓ^2 and hence the sequence $K^*(Kx - y^\delta) \in \ell^2$ tends to zero for all $x \in \ell^2$.

Lemma 4.4.2. *Let $\beta > 0$. For every $x \in \ell^2$, $D(x)$ is invertible and $\|D(x)^{-1}\|$ is uniformly bounded.*

Proof. Consider the equation $D(x)f = g$ or block wise

$$\begin{pmatrix} \beta \operatorname{id}_{A_x} + M_{A_x} & M_{A_x A_x^c} \\ 0 & \beta \operatorname{id}_{A_x^c} \end{pmatrix} \begin{pmatrix} f|_{A_x} \\ f|_{A_x^c} \end{pmatrix} = \begin{pmatrix} g|_{A_x} \\ g|_{A_x^c} \end{pmatrix},$$

i.e.

$$\beta f|_{A_x^c} = g|_{A_x^c} \quad \text{and} \qquad (\beta \operatorname{id}_{A_x} + M_{A_x}) f|_{A_x} = g|_{A_x} - M_{A_x A_x^c} f|_{A_x^c}.$$

Therefore, the invertibility of $D(x)$ only depends on the invertibility of $\beta \operatorname{id}_{A_x} + M_{A_x}$.

Since A_x is finite the operator $M_{A_x} = K_{A_x}^* K_{A_x}$ maps $\ell^2(A_x) \to \ell^2(A_x)$, and thus between finite dimensional spaces. Hence, it is given by a matrix which is self-adjoint and positive semidefinite. Therefore, the eigenvalues of $\beta \operatorname{id}_{A_x} + M_{A_x}$ are contained in the interval $[\beta, \infty)$. Consequently, the matrix $\beta \operatorname{id}_{A_x} + M_{A_x}$ is invertible, and $\|(\beta \operatorname{id}_{A_x} + M_{A_x})^{-1}\|_{\ell^2(A_x)} \le \beta^{-1}$. Now, the assertion follows from

$$\begin{aligned} &\left\| \begin{pmatrix} \beta \operatorname{id}_{A_x} + M_{A_x} & M_{A_x A_x^c} \\ 0 & \beta \operatorname{id}_{A_x^c} \end{pmatrix}^{-1} g \right\|_{\ell^2} \\ &= \left\| \begin{pmatrix} (\beta \operatorname{id}_{A_x} + M_{A_x})^{-1} & -\beta^{-1} (\beta \operatorname{id}_{A_x} + M_{A_x})^{-1} M_{A_x A_x^c} \\ 0 & \beta^{-1} \operatorname{id}_{A_x^c} \end{pmatrix} g \right\|_{\ell^2} \\ &\le \beta^{-1} (\|g_{A_x}\|_{\ell^2} + \beta^{-1} \|M_{A_x A_x^c} g_{A_x^c}\|_{\ell^2(A_x)} + \|g_{A_x^c}\|_{\ell^2}) \\ &\le \beta^{-1} (2 + \beta^{-1} \|K^* K\|_{\ell^2, \ell^2}) \cdot \|g\|_{\ell^2}, \end{aligned}$$

where we have used the inequality $\|M_{A_x A_x^c}\|_{\ell^2(A_x^c), \ell^2(A_x)} \le \|K^* K\|_{\ell^2, \ell^2}$. □

This lemma verifies the computability of Newton iterations. For notational simplicity, in the following calculation we write M as an abbreviation

for $M_{A_{x^k}}$ and M_c for $M_{A_{x^k}A^c_{x^k}}$ and analogously for id.

$$
\begin{aligned}
&x^k - x^{k+1} = D(x^k)^{-1}F(x^k) \\
&= \begin{pmatrix} (\beta\,\mathrm{id}+M)^{-1} & -\beta^{-1}(\beta\,\mathrm{id}+M)^{-1}M_c \\ 0 & \beta^{-1}\,\mathrm{id}_c \end{pmatrix} \begin{pmatrix} (\beta x^k + K^*(Kx^k - y) \pm \alpha)|_{A_{x^k}} \\ \beta x^k|_{A^c_{x^k}} \end{pmatrix} \\
&= \begin{pmatrix} (\beta\,\mathrm{id}+M)^{-1}(\beta x^k|_{A_{x^k}} + (K^*Kx^k)|_{A_{x^k}} - (K^*y)|_{A_{x^k}} \pm \alpha - M_c x^k|_{A^c_{x^k}}) \\ x^k|_{A^c_{x^k}} \end{pmatrix} \\
&= \begin{pmatrix} (\beta\,\mathrm{id}+M)^{-1}(\beta x^k|_{A_{x^k}} + Mx^k|_{A_{x^k}} - (K^*y)|_{A_{x^k}} \pm \alpha) \\ x^k|_{A^c_{x^k}} \end{pmatrix} \\
&= \begin{pmatrix} x^k|_{A_{x^k}} - (\beta\,\mathrm{id}+M)^{-1}((K^*y)|_{A_{x^k}} \pm \alpha) \\ x^k|_{A^c_{x^k}} \end{pmatrix}.
\end{aligned}
$$

And finally

$$
x^{k+1} = \begin{pmatrix} (\beta\,\mathrm{id}+M)^{-1}((K^*y)|_{A_{x^k}} \pm \alpha) \\ 0 \end{pmatrix}.
$$

In particular, this shows that the next iterate depends on the previous one only via the active set – this shows the importance of the active-set choice.

We are ready to state the complete algorithm.

Step 1 Initialize: $k = 0$, $x^0 = 0$

Step 2 Choose active set $A_{x^k} = \{i \in \mathbb{N} : |K^*(Kx^k - y^\delta)|_i > \alpha\}$ and calculate

$$
s_i^k = \begin{cases} 1, & [-K^*(Kx^k - y^\delta)]_i > \alpha \\ -1, & [-K^*(Kx^k - y^\delta)]_i < -\alpha \\ 0, & \text{else} \end{cases}.
$$

Step 3 Update for the next iterate x^{k+1}

$$
\begin{aligned}
x^{k+1}|_{A_{x^k}} &= (\beta\,\mathrm{id}_{A_{x^k}} + M_{A_{x^k}})^{-1}(K^*y^\delta - \alpha s^k)|_{A_{x^k}} \\
x^{k+1}|_{A^c_{x^k}} &= 0
\end{aligned}
\tag{4.5}
$$

Step 4 Check stop criteria. Return x^{k+1} or set $k \leftarrow k+1$ and continue at step 2.

Remark 4.4.3. We shall note that the active-set choice is the same for the elastic-net version as well as for the ℓ^1-minimization version of this algorithm. On the other hand using the ℓ^1-minimization version to minimize the elastic-net functional, via the reformulation of $\Phi_{\alpha,\beta}$ in the previous section, would change the active-set strategy. In that case the choice would be

$$A_{x^k} = \{i \in \mathbb{N} : |K^*(Kx^k - y^\delta) + \beta x^k|_i > \alpha\}.$$

A natural stopping criterion for the algorithm is the change of the active set. If it does not change for two consecutive iterations, then a minimizer has been attained. The next result shows the super-linear local convergence of the algorithm.

Theorem 4.4.4. *Let $\alpha, \beta > 0$. The RSSN converges locally super linearly.*

Proof. Let x^* the minimizer of $\Phi_{\alpha,\beta}$. Using $F(x^*) = 0$ we have

$$\begin{aligned}\|x^{k+1} - x^*\|_{\ell^2} &= \|x^k - D(x^k)^{-1}F(x^k) - x^*\|_{\ell^2} \\ &= \|x^k - D(x^k)^{-1}F(x^k) - x^* + D(x^k)^{-1}F(x^*)\|_{\ell^2} \\ &= \|D(x^k)^{-1}\|_{\ell^2,\ell^2}\|D(x^k)(x^k - x^*) - F(x^k) + F(x^*)\|_{\ell^2}.\end{aligned}$$

The definition of Newton derivatives implies

$$\lim_{x \to x^*} \frac{\|F(x) - F(x^*) - D(x)(x - x^*)\|_{\ell^2}}{\|x - x^*\|_{\ell^2}} = 0$$

Now, let $\varepsilon > 0$ arbitrary and $\|x^k - x^*\|$ sufficiently small. Utilizing the uniform boundedness of $\|D(x)^{-1}\|$ as shown in lemma 4.4.2, we have

$$\begin{aligned}&\|D(x^k)^{-1}\|_{\ell^2,\ell^2}\|D(x^k)(x^k - x^*) - F(x^k) + F(x^*)\|_{\ell^2} \\ &< \|D(x^k)^{-1}\|_{\ell^2,\ell^2} \cdot \varepsilon\|x^k - x^*\|_{\ell^2}\end{aligned}$$

which shows the desired super-linear local convergence. □

Remark 4.4.5. Several comments on the algorithm are in order. Firstly, this algorithm differs from the classical SSN [28] only in the regularization of the equation in step 3. Secondly, we note that the proposed method is

different from the standard regularized Newton method (also known as the Levenberg-Marquardt method) via

$$x^{k+1} = x^k - (D(x^k) + \eta \,\mathrm{id})^{-1} F(x^k),$$

for some $\eta > 0$, in that the latter regularizes globally whereas the former regularizes only on the active set. Thirdly, there are several equivalent formulations of the minimization problem. For instance, multiplying the optimality condition (4.2) by $\gamma > 0$ and adding x gives

$$x - \gamma K^*(Kx - y^\delta) - \gamma\beta x \in x + \gamma\alpha \,\mathrm{Sign}(x), \tag{4.6}$$

and also an alternative characterization of a minimizer of $\Psi_{\alpha,\beta}$: $x - S_{\gamma\alpha}(x - \gamma K^*(Kx - y^\delta) - \gamma\beta x) = 0$. It leads to a similar algorithm but with a different active set, i.e.

$$A_x^1 = \{i \in \mathbb{N} : |x - \gamma K^*(Kx - y^\delta) - \gamma\beta x|_i > \gamma\alpha\}.$$

Another choice of the active set derives by rewriting equation (4.6) as $x - \gamma K^*(Kx - y^\delta) \in (1 + \gamma\beta)x + \gamma\alpha \,\mathrm{Sign}(x)$. This gives $(1 + \gamma\beta)x - S_{\gamma\alpha}(x - \gamma K^*(Kx - y^\delta)) = 0$, and also a third choice of the active set

$$A_x^2 = \{i \in \mathbb{N} : |x - \gamma K^*(Kx - y^\delta)|_i > \gamma\alpha\}.$$

These different choices may affect the convergence behavior of the respective algorithms.

4.4.2 Regularized feature-sign-search method (RFSS)

The main drawback of the RSSN is its potential lack of global convergence. Globalization may be achieved by alternative selection strategies for the active set. The RFSS is one such example with active-set selection strategy. It derives from the FSS [37] as the RSSN from the SSN. In this section we describe the RFSS algorithm in detail and show the next convergence result. We consider only finite-dimensional problems: $K : \mathbb{R}^s \to \mathbb{R}^m$, $y^\delta \in \mathbb{R}^m$ and $\underline{s} = \{1, 2, \ldots, s\}$.

The following Theorem is the main result of this section and we prove it along with explaining the algorithm.

Theorem 4.4.6. *The RFSS converges globally in finitely many steps, moreover every iteration strictly decreases the value of the functional* $\Phi_{\alpha,\beta}$.

For the algorithm, and hence for the proof of this theorem, we shall need the notion of consistency, which plays a fundamental role in the RFSS.

Definition 4.4.7. Let $A \subset \underline{s}$, $x = (x_i)_{i\in\underline{s}} \in \mathbb{R}^s$ and $\theta = (\theta_i)_{i\in\underline{s}} \in \{-1, 0, 1\}^s$. The triple (A, x, θ) is called consistent if

$$\begin{aligned} i \in A &\implies \operatorname{sign}(x_i) = \theta_i \neq 0, \\ i \in A^c &\implies x_i = \theta_i = 0. \end{aligned}$$

Assisted with a consistent triple (A, x, θ) we can split the optimality condition (4.2) into

$$(-K^*(Kx - y^\delta) - \beta x)_i = \alpha\theta_i, \quad i \in A, \tag{4.7}$$

$$|K^*(Kx - y^\delta)|_i \leq \alpha, \quad i \in A^c. \tag{4.8}$$

Remark 4.4.8. The formulas (4.7) and (4.8) correspond to the optimality condition for the following auxiliary functional

$$\Phi_{\alpha,\beta,\theta}(x) = \frac{1}{2}\|Kx - y^\delta\|^2 + \alpha\langle x, \theta\rangle_{\ell^2} + \frac{\beta}{2}\|x\|^2_{\ell^2}. \tag{4.9}$$

By the definition of consistency, $\Phi_{\alpha,\beta}(x) = \Phi_{\alpha,\beta,\theta}(x)$ if $\operatorname{sign}(x)_i = \theta_i$ for all non-zero components of x. In any case we have

$$\Phi_{\alpha,\beta,\theta}(x) \leq \Phi_{\alpha,\beta}(x).$$

We are ready to describe the complete FSS algorithm in five steps. The description provides a constructive proof of Theorem 4.4.6.

Step 1 Initialize: $k = 1$, $A_0 = \emptyset$, $x^0 = 0$ and $\theta^0 = 0$. Any consistent triple (A_0, x^0, θ^0) is valid for initialization. Then check the optimality condition (4.2) and take one of the actions

(i) return the solution if fulfilled;

(ii) continue with Step 2 if (4.8) is not fulfilled;

(iii) continue with Step 3 otherwise.

Step 2 At this step, the following premises hold: The optimality condition (4.8) is not fulfilled and the triple $(A_{k-1}, x^{k-1}, \theta^{k-1})$ is consistent.

This step performs a greedy scheme by selecting the index i_0^k violating condition (4.8) the most, i.e.

$$i_0^k \in \operatorname*{argmax}_{i \in A_{k-1}^c} |K^*(Kx^{k-1} - y^\delta)|_i - \alpha,$$

then update the active set by $A_k = A_{k-1} \cup \{i_0^k\}$, update θ^k by

$$\theta_i^k = \begin{cases} \theta_i^{k-1}, & i \neq i_0^k \\ -\operatorname{sign}((K^*(Kx^{k-1} - y^\delta))_{i_0^k}), & i = i_0^k \end{cases}$$

and continue with Step 3.

Step 3 Calculate the next iterate x^k such that (4.7) is fulfilled, i.e. x^k is optimal for $\Phi_{\alpha,\beta,\theta^k}$, by

$$x^k|_{A_k} = (\beta \operatorname{id} + M_{A_k})^{-1}(K^* y^\delta - \alpha\theta^k)|_{A_k} \quad \text{and} \quad x^k|_{A_k^c} = 0.$$

Observe that the update coincides with that in the RSSN. If the triple (A_k, x^k, θ^k) is consistent, continue with Step 5, and otherwise continue with Step 4. For the former, we deduce from Remark 4.4.8 that

$$\Phi_{\alpha,\beta}(x^k) = \Phi_{\alpha,\beta,\theta^k}(x^k) < \Phi_{\alpha,\beta,\theta^k}(x^{k-1}) \leq \Phi_{\alpha,\beta}(x^{k-1}).$$

Step 4 This step handles inconsistent (A_k, x^k, θ^k). We consider two different cases separately.

Case 1 The preceding step of Step 3 is Step 4, i.e. (A_k, x^{k-1}, θ^k) is consistent. Therefore, there must be at least one index such that the signs of x^k and x^{k-1} differ. Let λ_0 the smallest $\lambda \in (0, 1)$ such that

$$\exists i_0 \in A_k : \quad (\lambda x^k + (1-\lambda)x^{k-1})_{i_0} = 0,$$

and denote $x_{\lambda_0} = \lambda_0 x^k + (1-\lambda_0)x^{k-1}$. The convexity of $\Phi_{\alpha,\beta,\theta^k}$ implies

$$\begin{aligned} \Phi_{\alpha,\beta}(x_{\lambda_0}) &= \Phi_{\alpha,\beta,\theta^k}(x_{\lambda_0}) \\ &\leq \lambda_0 \Phi_{\alpha,\beta,\theta^k}(x^k) + (1-\lambda_0)\Phi_{\alpha,\beta,\theta^k}(x^{k-1}) \\ &< \lambda_0 \Phi_{\alpha,\beta,\theta^k}(x^{k-1}) + (1-\lambda_0)\Phi_{\alpha,\beta,\theta^k}(x^{k-1}) \\ &= \Phi_{\alpha,\beta,\theta^k}(x^{k-1}) = \Phi_{\alpha,\beta}(x^{k-1}), \end{aligned}$$

by the minimizing property of x^k for $\Phi_{\alpha,\beta,\theta^k}$. Now, we update (A_k, x^k, θ^k) by

$$x^k \leftarrow x_{\lambda_0}, \quad A_k \leftarrow \{i \in \underline{s} : x_i^k \neq 0\}, \quad \theta^k \leftarrow \operatorname{sign}(x^k)$$

and check equation (4.7). If fulfilled continue with step 5, otherwise increase k by one and continue with step 3.

Case 2 The preceding step of Step 3 is Step 2, i.e. $|K^*(Kx^{k-1} - y^\delta)|_{i_0^k} \geq \alpha$ and $x_{i_0^k}^{k-1} = 0$. The choice of $\theta_{i_0^k}^k$ implies

$$\operatorname{sign}(\nabla\Phi_{\alpha,\beta,\theta^k}(x^{k-1}))_{i_0^k} = \operatorname{sign}((K^*(Kx^{k-1} - y^\delta))_{i_0^k} + \alpha\theta_{i_0^k}^k) = -\theta_{i_0^k}^k,$$

and moreover,

$$\nabla\Phi_{\alpha,\beta,\theta^k}(x^{k-1})|_{A_{k-1}} = 0.$$

Hence, the Taylor expansion of $\Phi_{\alpha,\beta,\theta^k}$ at x^{k-1} yields that for $\tilde{x}$ near to x^{k-1} with

$$\begin{aligned} 0 &> \Phi_{\alpha,\beta,\theta^k}(\tilde{x}) - \Phi_{\alpha,\beta,\theta^k}(x^{k-1}) \\ &= \nabla\Phi_{\alpha,\beta,\theta^k}(x^{k-1})_{i_0^k}(\tilde{x} - x^{k-1})_{i_0^k} + R \\ &= \nabla\Phi_{\alpha,\beta,\theta^k}(x^{k-1})_{i_0^k}\tilde{x}_{i_0^k} + R, \end{aligned} \tag{4.10}$$

which consequently implies

$$0 > -\theta_{i_0^k}^k \tilde{x}_{i_0^k} \implies \theta_{i_0^k}^k = \operatorname{sign} \tilde{x}_{i_0^k}.$$

The minimizing property in Step 3, implies that $\Phi_{\alpha,\beta,\theta^k}(x^k) < \Phi_{\alpha,\beta,\theta^k}(x^{k-1})$, which further implies together with the convexity of $\Phi_{\alpha,\beta,\theta^k}$ that there exists a $\tilde{x}$ near to x^{k-1} on the line from x^{k-1} to x^k such that $\Phi_{\alpha,\beta,\theta^k}(\tilde{x}) < \Phi_{\alpha,\beta,\theta^k}(x^{k-1})$. Consequently,

$$\operatorname{sign}(x_{i_0^k}^k) = \theta_{i_0^k}^k.$$

Thus there can be a sign change for x^{k-1} to x^k only on a component other than i_0^k. Now, continue analogously to Case 1.

Step 5 In this case, the following premises are fulfill: (A_k, x^k, θ^k) is consistent and the optimality condition (4.7) is fulfilled. Check (4.8). If fulfilled, stop, otherwise continue with Step 2.

From the strictly reducing property of the algorithm we know that every possible active set is attained at most once. This guarantees the convergence in finitely many steps. We observe that the reduction properties hold also for infinite-dimensional problems.

Before embarking on numerical results, let us briefly comment on the algorithmic part. In the RFSS, the active set changes only in a modest way, which enables efficient implementations of Step 3. The symmetric matrix in Step 3 changes only by one row and one column in almost every iteration. Therefore, it is advisable to use the Cholesky factorization to solve the equation because of its straightforward update and reduction in computational efforts.

5 Numerical examples

Before we start numerical calculations we shortly note the design parameters of our ecosystem. All calculations are performed on an AMD Athlon 64 X2 dual core processor 3800+ equipped with a 64 bit Linux running Matlab release 2008a. There are some Matlab toolboxes, namely the Regularization Tools version 4.1 [30] and SPARCO version 1.2 [4], which are often used and cited in scientific articles. They form a common testing environment for inverse problems and sparsity related problems and are also used in the following. Further, we utilize implementations of different ℓ^1 and elastic-net minimization algorithms:

- **IST, IHT:** Implemented by Dirk Lorenz, September 30, 2008.
- **RSSN, RFSS:** Implemented by Stefan Schiffler, December 14, 2009.
- **SSN, FSS:** We use RFSS with $\beta = 0$ (RSSN resp.).
- **GPSR:** GPSR Basic version 5.0 by Mario Figueiredo, Robert Nowak, Stephen Wright, December 4, 2007.

Having a different stopping criterion and different tolerance, the first problem is how to compare whether an algorithm converged or not. Therefore, we set all algorithms to deliver the most accurate result possible by choosing their tolerances as small as possible. After the algorithm has stopped or reached an upper bound for the number of iterations we check for convergence by means of the optimality condition (3.3) to ensure a fair

comparison between all used algorithms. An algorithm is considered to be convergent if for a given tolerance $\varepsilon > 0$ the output x fulfills

$$\begin{aligned} &|K^*(Kx - y) + \beta x + \alpha \operatorname{sign}(x)|_i < \varepsilon, && x_i \neq 0 \\ &|K^*(Kx - y) + \beta x|_i - \alpha < \varepsilon, && x_i = 0 \end{aligned} \tag{5.1}$$

for every component x_i of x.

We do not compare the residuals of the minimizers with the true signals since we do not want to show the qualitative properties of ℓ^1 minimization, for this information we refer to [18, 26, 27, 38, 55]. Instead we want to find out how to obtain a highly accurate ℓ^1 minimizer in the above sense, i.e. a minimizer fulfilling the equations 5.1 for small values of ε.

5.1 Basis learning in mass spectrometry

The problem initiating our interest to stabilize algorithms for ℓ^1 minimization came from mass spectrometry. We have published the results in [1]. For our example problem in mass spectrometry, we are not interested in the distribution of proteins or molecules but only in the decision whether a spectrum has a certain property or not. For example, we are facing two possible classes of tissue, cancerous and healthy, and we want to decide to which of these classes a new given sample corresponds. The idea traced in [1] is to determine prototype spectra for each of the classes followed by determination of the characteristic peaks. Due to the size of real world spectra, here we use simulated data with spectra of length 110 (figure 5.1.1). In our setting we use 50 spectra of 2 classes each with 20% noise and spuriously added peaks. The determination of prototype spectra can be done by applying basis learning procedures, in this case the one proposed in [37]. It starts with an initial guess for prototype spectra written to the columns of a matrix K. Afterwards, it iteratively minimizes the ℓ^1 functional Ψ_α for each given spectrum, resulting in coefficients x for each spectrum, and then optimizes the prototype spectra using a Newton search. The output of this method is a collection of optimized prototype spectra. For a closer discussion on this method and the quality of the output we refer to [1].

The basis learning comprises ℓ^1 minimization, choosing the parameter α by hand to be $\alpha = 10$. While first results where promising (figure 5.1.2) we were facing the fact that ℓ^1 minimization struggled with rank-deficient matrices or such with highly correlated columns. We ran the procedure on

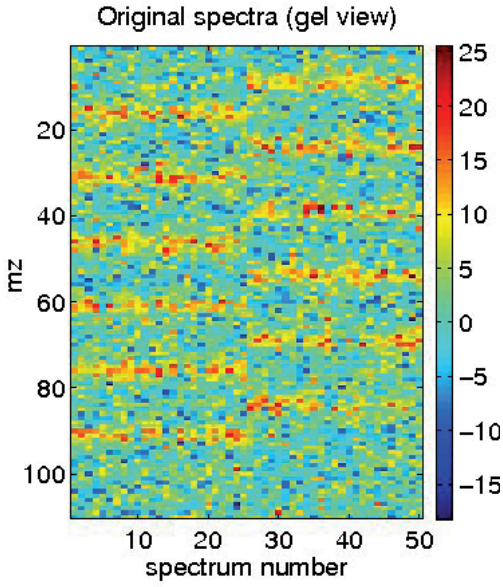


Figure 5.1.1: 50 simulated mass spectra of length 110 (each column is one spectrum) representing 2 classes, ill and healthy.

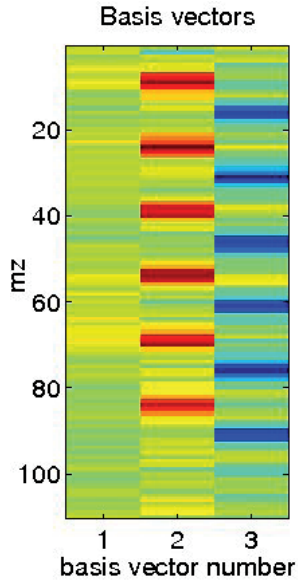


Figure 5.1.2: Automatically learned prototype spectra (each column corresponds to one spectrum prototype).

Algorithm	$\varepsilon = 10^{-10}$	$\varepsilon = 10^{-5}$
FSS	52%	52%
RFSS($\beta = 10^{-6}$)	33%	33%

Table 5.1: Comparison of FSS and RFSS for automated basis learning. We state the failure rates in 100 experiments for different tolerances ε for accepting the optimality condition (5.1).

100 simulated noisy datasets and obtained that the FSS failed to converge for 52 cases. Here, the algorithm is considered to be failed to converge if condition (5.1) is not fulfilled with a tolerance of $\varepsilon = 10^{-10}$ after a maximal number of 10000 iterations. We shall mention that the number of iterations has never been reached, instead the algorithm always stopped after less than 500 iterations for one of the three reasons:

1. it converged,
2. it produced NaN results,
3. it hang in a loop inserting an index into the active set and immediately removing the same index.

The problems can be ascribed to numerical problems in solving the linear equation for the next iterate, i.e. in step 3 of the RFSS for $\beta = 0$. In step 3, the RFSS needs to invert a submatrix of K^*K which turns out to be numerically instable for this example. Hence, the resulting next iterate contains NaN entries or does not fulfill the stopping criterion of the RFSS despite of the correct active set.

As described in the introduction to the elastic net, we used a regularization technique to attenuate the numerical troubles, i.e. choosing $\beta > 0$. Repeating the above experiment, using the elastic net with $\beta = 10^{-6}$, resulted in a higher success rate. The results are summarized in table 5.1. We can see that the tolerance for accepting the minimizer does not play an important role for the result since increasing it to $\varepsilon = 10^{-5}$ does not affect the result.

While these results were obtained only for the FSS, we compare different algorithms for several ill-conditioned problems in the next sections. It was the starting point for exploring numerical difficulties for ℓ^1 algorithms.

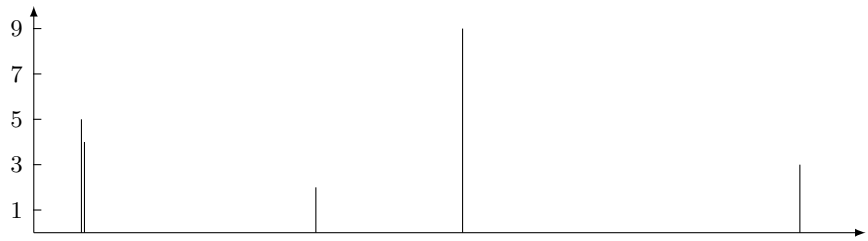


Figure 5.2.1: Source signal y

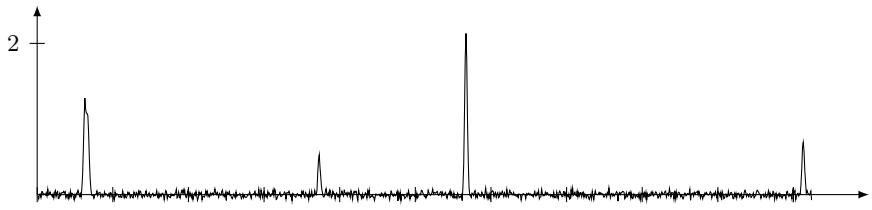


Figure 5.2.2: Blurry and noisy measurement y^δ

5.2 Rank-deficient operators

To demonstrate difficulties of ℓ^1-minimization algorithms with problems comprising rank-deficient operators, we take an example already mentioned in the introduction for mass spectrometry: Extraction of peak clusters from mass spectra. Since real-world mass spectra have an enormous length and hence lead to huge problem size and extensive computation time we restrict ourself to a simulation. We simulate signals of length 1024 as a linear combination of cluster prototypes. Our test signal x_0 contains δ-peaks at positions 63, 67, 373, 567, 1013 with the corresponding peak heights of 5, 4, 2, 9, 3 (figure 5.2.1). This signal is usually unknown and it is the objective to extract it from measurements. The measurement results are typically blurred which we model by convolution with a Gaussian kernel K (using "opConvolution1d“ from the SPARCO toolbox with $\sigma = 2$ sampled on 9 positions). Finally, we add noise with a noise level of 20% resulting in the final measurement y^δ (figure 5.2.2). The negative values in this simulation appear also in real-world datasets after a baseline correction. Since we have the prior knowledge that the signal we are looking for consists of a

linear combination of peak-patterns, we setup a dictionary as follows: We use a basis of δ peaks and add all possible patterns of up to four peaks out of the five peaks used to build our sample data but each peak normalized to unit height. The corresponding synthesis operator is called S. As described in the introduction, using the blurred dictionary elements as columns of the operator K, we should be able to extract, or approximate the true signal, by solving $Kx = y^\delta$ and obtaining $x_0 = Sx$. Since K is rank-deficient the solution can not be expected to be unique. For the problem depends on the noise, all numerical results are obtained as the mean over 100 runs. Corresponding to the parameter choice section, we should select $\alpha \in [0, \max_{i\in\mathbb{N}} |K^* y^\delta|_i]$. For convenience, we use the following choices for α, where $m := \max_{i\in\mathbb{N}} |K^* y^\delta|_i$: $\frac{m}{4}$, $\frac{m}{12}$, $\frac{m}{20}$. For applying the elastic net, we choose $\beta = 10^{-15}$.

Table 5.2 summarizes the results in terms of how often in the 100 experiments the algorithms converged. Here, we used 10000 as the upper bound for the number of iterations and we used tolerances of $\varepsilon \in \{10^{-15}, 10^{-10}, 10^{-8}\}$ for the convergence condition (5.1). We can summarize for the IST that it always stops, i.e. the update during the fixed-point iteration is zero, but the optimality condition is usually fulfilled only for $\varepsilon > 10^{-8}$. Changing the stopping criterion to the optimality condition does not make sense, because the update of the iteration is 0. The IHT shows a convergence like behavior but even after 500000 iterations the optimality condition was not fulfilled. Considering the GPSR the demand of high accuracy seems to influence the stopping behavior, also it does not stop it often fulfills our convergence criteria for a tolerance of about $\varepsilon = 10^{-10}$. Last but not least the FSS gives the most satisfactory results since it converges more often than all other ℓ^1 methods and simultaneously fulfills the convergence criteria up to the smallest tolerance. Still, the FSS does not converge for a major part of the experiments. For big values of α the ill conditioning of the operator seems not to penetrate the linear equation in the update step of the FSS, since it is solved only on small active sets. Increasing α, and hence the active sets, this equation becomes more problematic. Correspondingly, the number of accepted minimizers decreases for increased α values. Applying the elastic net leads to better results in the sense that convergence can be achieved independent from the choice of α. The RFSS always converged for $\beta = 10^{-15}$, satisfying the smallest tolerance on the convergence condition (5.1), within a few iterations.

					Converged		
		Stopped	Iterations	Time (ms)	$\varepsilon = 10^{-15}$	$\varepsilon = 10^{-10}$	$\varepsilon = 10^{-8}$
$\alpha = \frac{m}{4}$	IST	85	2978	1044	0	0	0
	IHT	0	-	-	0	0	0
	GPSR	4	184.5	82.3	37	73	100
	FSS	100	5.1	1.3	98	98	98
	RFSS	100	5.3	1.4	100	100	100
$\alpha = \frac{m}{12}$	IST	100	1474.5	517.3	0	0	2
	IHT	0	-	-	0	0	0
	GPSR	0	-	-	3	35	100
	FSS	100	12	4.2	75	75	75
	RFSS	100	27.1	8.4	100	100	100
$\alpha = \frac{m}{20}$	IST	100	2143.1	751.2	0	0	36
	IHT	0	-	-	0	0	0
	GPSR	0	-	-	4	41	100
	FSS	100	13.1	3.9	50	50	50
	RFSS	100	27.6	8.6	100	100	100

Table 5.2: Comparison of ℓ^1 algorithms vs. RFSS. We state how often the algorithms stopped within 10000 iterations for 100 experiments, the mean number of iterations and time (in milliseconds) until stop and how often our convergence criterion has been fulfilled with the specified tolerance ε.

The elastic-net minimizers always fulfilled the convergence condition for ℓ^1 minimization, i.e. (5.1) with $\beta = 0$, with the tolerance level $\varepsilon = 10^{-10}$. Choosing $\beta = 10^{-16}$ we also found that the elastic-net minimizers always fulfilled the convergence condition for ℓ^1 minimization also for the tolerance level $\varepsilon = 10^{-15}$, still having the same characteristics as in table 5.2. That means that the elastic-net minimizer is a more accurate minimizer of the ℓ^1 functional than the minimizers calculated directly by ℓ^1 algorithms. This effect can be assigned to the ill conditioning of K.

We can conclude from this experiment that the elastic net is able to overcome the numerical troubles resulting in as accurate ℓ^1 minimizers as the tested ℓ^1 algorithms itself, sometimes even more accurate.

Next, we analyze the impact of β to the results in terms of sparsity. Therefore, we use a fixed $\alpha = \frac{m}{4}$ (resp. $\alpha = \frac{m}{8}$) and increase β successively. Figure 5.2.3 shows that for a small value of β the sparsity of the minimizer is not affected compared to the ℓ^1 minimizer. In fact, the support of the minimizers coincide.

5.3 Sudoku

This application is based on the article [2]. Beside its sportive purpose the Sudoku is an example for compressed sensing. Before explaining the relation between the Sudoku game and compressed sensing we state the Sudoku problem.

Let $n \in \mathbb{N}$, then a Sudoku is a matrix of size $n^2 \times n^2$ (cf. figure 5.3.1) containing the numbers $\{1 \dots n^2\}$ such that

- in every row each number is contained exactly once,
- in every column each number is contained exactly once,
- dividing the matrix into boxes of size $n \times n$ every box contains each number exactly once.

We assume that we are given only a few pre-specified matrix entries (clues) such that there exists a unique Sudoku with these entries. We now want to rewrite the problem of finding the full Sudoku as a linear equation. The first step is a binary representation of the numbers $\{1 \dots n^2\}$ by means of the corresponding standard basis vector $i \rightsquigarrow e_i \in \mathbb{R}^{n^2}$. Applying this encoding, we can write a Sudoku matrix row wise as a vector in $\{0,1\}^{n^6}$ where empty

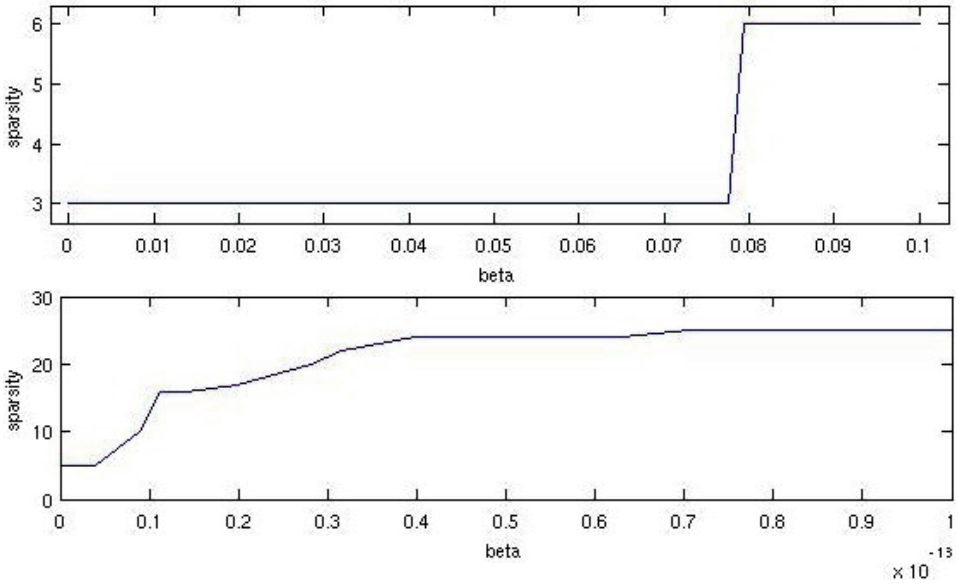


Figure 5.2.3: Impact of β on the sparsity of the minimizer. Top: For $\alpha = \frac{m}{4}$. Bottom: For $\alpha = \frac{m}{8}$

		3			9		8	1
			2				6	
5				1		7		
8	9							
		5	6		1	2		
							3	7
		9		2				8
	7				4			
2	5		8			6		

Figure 5.3.1: A sample Sudoku game.

cells are encoded as a 0 vector of length n^2. Equipped with this encoding of Sudoku matrices as a vector we can think on how to incorporate the Sudoku constraints into a linear equation.

We start with the statement "The first row has to contain the number 1 exactly once". If we take an empty Sudoku matrix and insert the number 1 in every cell of the first row and then encode it, an encoded Sudoku x should fulfill

$$(\underbrace{\underbrace{10\dots0}_{\text{length } n^2}\dots\underbrace{10\dots0}_{\text{length } n^2}}_{n^2 \text{ times}}\ \underbrace{0\dots0}_{\text{length } n^4\cdot(n^2-1)})x = 1.$$

Here, each block of length n^2 corresponds to the encoded number in a cell of the Sudoku matrix by means of the standard basis vectors. Since we have the number 1 everywhere in the first row, the first n^2 blocks consist of $e_1 \in \mathbb{R}^{n^2}$. If the correlations with x was not 1, this would mean that x has the number 1 not exactly once in the first row, which is a contradiction to the assumption that x is a Sudoku. Analogously, for the statement "The first row has to contain the number 2 exactly once" we have

$$(\underbrace{010\dots0}_{\text{length } n^2}\dots\underbrace{010\dots0}_{\text{length } n^2}\ \underbrace{0\dots0}_{\text{length } n^4\cdot(n^2-1)})x = 1.$$

Hence, the requirement "The first row has to contain every number exactly once" can be expressed vie the equations

$$(\underbrace{\mathrm{id}_{n^2\times n^2}\dots\mathrm{id}_{n^2\times n^2}}_{n^2 \text{ times}}\ 0_{n^2\times n^4(n^2-1)})x = 1_{n^2\times 1}.$$

Consequently, the whole row constraint "Every row has to contain every number exactly once" corresponds to

$$\begin{pmatrix} \mathrm{id}_{n^2\times n^2}\dots\mathrm{id}_{n^2\times n^2} & & 0 \\ & \ddots & \\ 0 & & \mathrm{id}_{n^2\times n^2}\dots\mathrm{id}_{n^2\times n^2} \end{pmatrix} x = 1_{n^4\times 1}.$$

Remember that every row of this matrix corresponds to an encoded Sudoku matrix where the corresponding number is contained all over one row while

all other entries are zero. This principle can be also applied to the column constraints "Every column has to contain every number exactly once":

$$\begin{pmatrix} \mathrm{id}_{n^2\times n^2} & 0_{n^2\times n^4-n^2} & \dots & \mathrm{id}_{n^2\times n^2} & 0 \\ \ddots & \ddots & & \ddots & \\ 0 & \mathrm{id}_{n^2\times n^2} & 0_{n^2\times n^4-n^2} & \dots & \mathrm{id}_{n^2\times n^2} \end{pmatrix} x = 1_{n^4\times 1}$$

and similarly to the box constraints "Every box has to contain every number exactly once". Note that a solution to the equation systems may contain other components than $\{0, 1\}$, hence the equation system is not equivalent to solving a Sudoku.

For the equation system consisting of the collection of all those constraints it has been shown in [2] that

- every encoded Sudoku is a solution (since together with the Sudoku constraints it fulfills the equations),
- every solution has at least n^4 non-zero components (since the constraints ensure that every Sudoku cell (i.e. every n^2 block of x) has to contain at least one number),
- every solution with n^4 non-zero entries is an encoded Sudoku (Since the equations are fulfilled with n^4 non-zero entries and every Sudoku cell has to contain a number we have that every Sudoku cell contains exactly one number. Finally the equations ensure that the Sudoku constraints are fulfilled).

For our initial problem we need to incorporate the clues to our equations system using the same principle as before: For every clue take an empty Sudoku matrix, insert a clue and encode it. The correlation with x has to be 1. By our assumptions, adding these clues to the equation system ensures that there exists exactly one solution with n^4 non-zero entries. Combining all constraints into the operator K, solving the Sudoku means finding the sparsest solution of the equation system

$$Kx = 1.$$

We shall mention that this type of matrices also appear in compressed sensing, since the right hand side of the equation can be considered as a

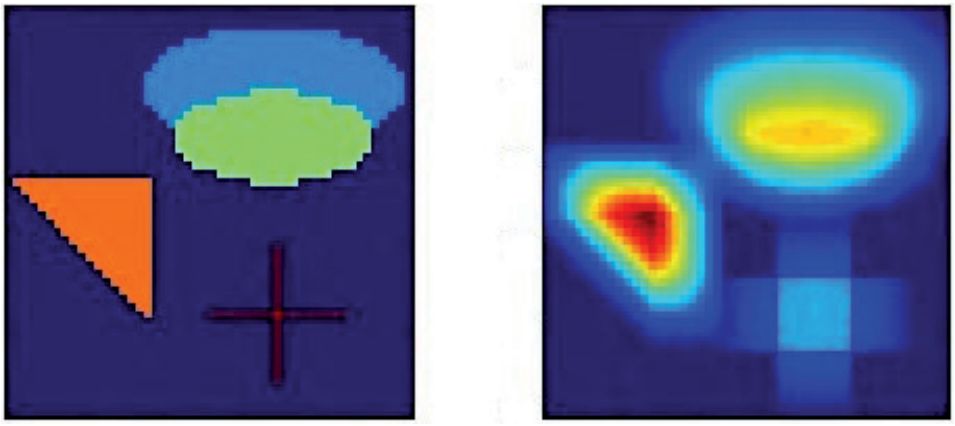


Figure 5.4.1: Left: Blur sample image (x_0) from the Regularization Tools toolbox. Right: A blurry measurement simulation.

measurement of the given Sudoku correlated with vectors. In this sense, K is the sensing matrix.

Note that in [2] they add the additional and redundant constraints that every cell in the Sudoku must contain a value. In that case ℓ^1 as well as elastic-net algorithms do work. In our case, i.e. without these additional constraints, using the Sudoku in figure 5.3.1, ℓ^1 minimization fails independent of the tested algorithms. In contrast, the elastic net with $\beta = 10^{-13}$ does converge resulting in the correct solution.

5.4 Ill-conditioned operators

In this section, we take a closer look on ill-conditioned operators where we use the blur example from the Regularization Tools toolbox. We choose a sample size of 50 which creates a 50×50 pixel sample image x_0 with about 26% non-zero entries (Figure 5.4.1) and a blur operator A. We simulate a measurement y^δ via the equation $y^\delta = Ax_0 + noise$ where the noise level is 5% (Figure 5.4.1), i.e. denoting the noise by η we have $\|\eta\| = 0.05 \cdot \|y\|$.

We use a dictionary build of the pixel basis and a unit-normed row-wise heavy-side basis, i.e. for every row the images having pixels 1 to j, $j = 1 \dots 50$ set to 1 form a basis vector (after normalization). In this

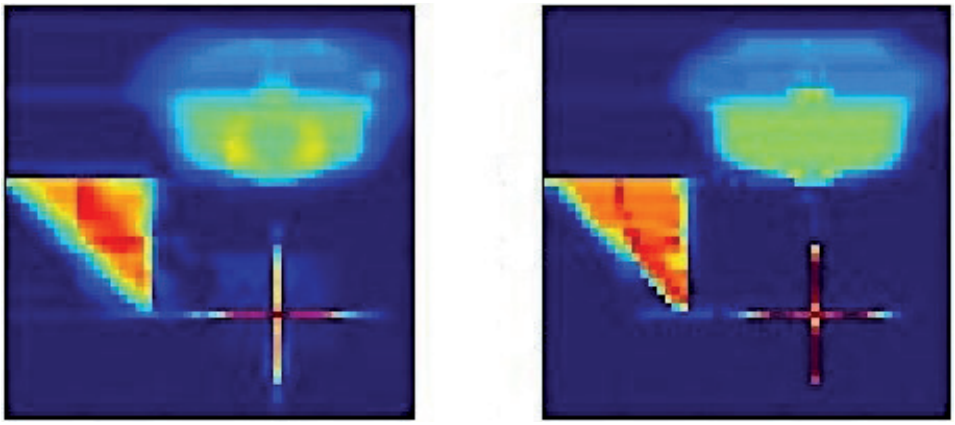


Figure 5.4.2: Different reconstructions of the blur example using ℓ^1 methods stopped after 10000 iterations.

dictionary every row of the image matrix can be represented in at most 4 basis vectors.

Choosing $\alpha = 10^{-4}$ by hand and applying ℓ^1 minimization comprises troubles as in the former experiments. The FSS does not converge, the IHT, IST and GPSR converge very slowly, after 10000 iterations our convergence criteria (5.1) is violated for $\varepsilon = 10^{-4}$ (Figure 5.4.2). Using RFSS with a value of $\beta = 10^{-18}$ leads to convergence within 611 iterations. This elastic-net minimizer also fulfills the convergence condition for the ℓ^1 functional, i.e. $\beta = 0$, with a tolerance of $\varepsilon = 10^{-14}$. Using the following path-following strategy [32], we were able to finally obtain an ℓ^1 minimizer within a total of 612 iterations which had an accuracy of 10^{-15} w.r.t. ℓ^1 minimization.

Step 1: Select α, β

Step 2: Minimize $\Phi_{\alpha,\beta}$

Step 3: Stop if the convergence criteria (5.1) for ℓ^1 minimization (i.e. with $\beta = 0$) is fulfilled for a pre-specified tolerance ε, otherwise update $\beta \leftarrow \frac{\beta}{2}$ and continue at step 2.

This algorithm is motivated by theorem 4.2.6 on the stability of elastic-net minimizers with respect to α and β, as well as the propositions thereafter.

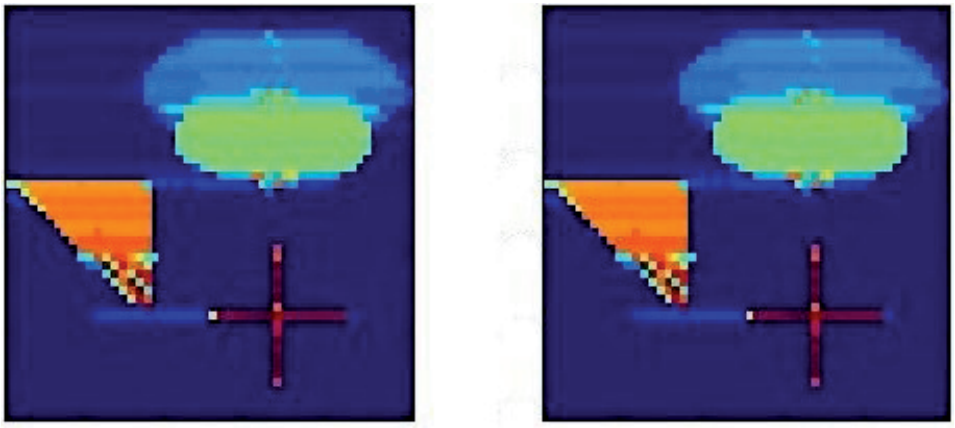


Figure 5.4.3: The elastic-net minimizer (left) of the blur example and the ℓ^1 minimizer (right) obtained via path following.

Figure 5.4.3 displays the output of the elastic-net minimizer and the ℓ^1 minimizer obtained via path following.

6 Exact-recovery conditions

Recall that we consider a Hilbert space $\mathcal{H}$ and a bounded linear operator $K \in L(\ell^2, \mathcal{H})$. Having $y \in R(K)$ we assume that $Kx = y$ has a sparse solution $x^\dagger$ and we denote $I = \operatorname{supp}(x^\dagger)$, $P_I^* : \ell^2(I) \to \ell^2$ the embedding and $K_I = KP_I^*$. y^δ is a noisy observation of y with $\eta = y^\delta - y$ and $\delta = \|\eta\|_{\mathcal{H}}$. In this chapter, we require K to be injective, where this requirement can be replaced by the assumption that only $K_I = K(P_I^*)$ is injective. However, this replacement does not contribute to the understanding of the key-results of this chapter but complicates the proofs and hence is not considered.

For ℓ^1 minimization, Tropp [52, 53] established parameter-choice rules which – if they can be fulfilled – guarantee the exact recovery of the support, i.e. the support of the minimizer of the ℓ^1 functional

$$\Psi_\alpha = \tfrac{1}{2}\|Kx - y^\delta\|_{\mathcal{H}}^2 + \alpha\|x\|_{\ell^1}$$

coincides with those of $x^\dagger$. In this chapter Tropp's results are transferred to the elastic-net functional

$$\Phi_{\alpha,\beta} = \tfrac{1}{2}\|Kx - y^\delta\|_{\mathcal{H}}^2 + \alpha\|x\|_{\ell^1} + \tfrac{1}{2}\beta\|x\|_{\ell^2}^2.$$

We shall mention that Tropp listed this transfer as "valuable to study" in the discussion section of his paper "Just Relax" [52]. After the transfer, we show how the established parameter-choice rules can be applied to concrete problems giving a parameter-choice recipe. In the end of this chapter, we demonstrate the applicability of the parameter-choice rules for an example problem from mass spectrometry. The results of this chapter can also be found in [39, 49].

6.1 Parameter-choice rules

In equation (4.1) we have redrafted the elastic-net functional $\Phi_{\alpha,\beta}$ as an ℓ^1 functional via

$$\Phi_{\alpha,\beta}(x) = \frac{1}{2}\left\| \begin{pmatrix} K \\ \sqrt{\beta}\,\mathrm{id} \end{pmatrix} x - \begin{pmatrix} y^\delta \\ 0 \end{pmatrix} \right\|^2_{\mathcal{H}\times\ell^2} + \alpha\|x\|_{\ell^1}.$$

This section states the parameter-choice results of Tropp [52, 53] for exact recovery for this redrafted functional and modifies them for, what we consider, better applicability. The respective problems are discussed at due position.

We start with the first result of Tropp which reads in our setting applied to the redrafted the elastic-net functional:

Proposition 6.1.1. *[53, Lemma 2] Let $\mathcal{H}$ be a Hilbert space, $K \in L(\ell^2, \mathcal{H})$ injective, $y \in R(K)$ and $x^\dagger$ the sparse solution of $Kx = y$. Further, let y^δ be a noisy observation of y and $\alpha, \beta \geq 0$. We denote $I = \mathrm{supp}(x^\dagger)$ and $x_I^{\alpha,\beta} = \mathrm{argmin}_{\mathrm{supp}(x)\subset I}\, \Phi_{\alpha,\beta}(x)$. The implicit parameter rule*

$$\alpha > \sup_{i\in I^c} \left|\left\langle \begin{pmatrix} K \\ \sqrt{\beta}\,\mathrm{id} \end{pmatrix} x_I^{\alpha,\beta} - \begin{pmatrix} y^\delta \\ 0 \end{pmatrix}, \begin{pmatrix} K \\ \sqrt{\beta}\,\mathrm{id} \end{pmatrix} e_i \right\rangle_{\mathcal{H}\times\ell^2}\right|$$

ensures that the support of the minimizer of the elastic-net functional $\Phi_{\alpha,\beta}$ is contained in I.

The parameter rule in proposition 6.1.1 can be simplified in the following way. By the definition of $x_I^{\alpha,\beta}$ in the assumptions of proposition 6.1.1 we have that $\mathrm{supp}(x_I^{\alpha,\beta}) \subset I$ leading for $i \in I^c$ to

$$\begin{aligned} &\left\langle \begin{pmatrix} K \\ \sqrt{\beta}\,\mathrm{id} \end{pmatrix} x_I^{\alpha,\beta} - \begin{pmatrix} y^\delta \\ 0 \end{pmatrix}, \begin{pmatrix} K \\ \sqrt{\beta}\,\mathrm{id} \end{pmatrix} e_i \right\rangle_{\mathcal{H}\times\ell^2} \\ &= \langle Kx_I^{\alpha,\beta} - y^\delta, Ke_i\rangle_{\mathcal{H}} + \langle \beta x_I^{\alpha,\beta}, e_i\rangle_{\ell^2} \\ &= \langle Kx_I^{\alpha,\beta} - y^\delta, Ke_i\rangle_{\mathcal{H}} \end{aligned}$$

Thus, the above proposition can be formulated as follows, we state the proof for the sake of completeness.

Proposition 6.1.2. *[53, Lemma 2] Let $\mathcal{H}$ be a Hilbert space, $K \in L(\ell^2, \mathcal{H})$ injective, $y \in R(K)$ and $x^\dagger$ the sparse solution of $Kx = y$. Further, let y^δ be a noisy observation of y and $\alpha, \beta \geq 0$. We denote $I = \operatorname{supp}(x^\dagger)$ and $x_I^{\alpha,\beta} = \operatorname{argmin}_{\operatorname{supp}(x)\subset I} \Phi_{\alpha,\beta}(x)$. The implicit parameter rule*

$$\alpha > \sup_{i \in I^c} |\langle K x_I^{\alpha,\beta} - y^\delta, K e_i \rangle_{\mathcal{H}}| \tag{6.1}$$

ensures that the support of the minimizer of the elastic-net functional $\Phi_{\alpha,\beta}$ is contained in I.

Proof. The idea of this proof is to show that for small $h \in \ell^2$, $h \neq 0$ there holds

$$\Phi_{\alpha,\beta}(x_I^{\alpha,\beta} + h) - \Phi_{\alpha,\beta}(x_I^{\alpha,\beta}) > 0, \tag{6.2}$$

which results in $x_I^{\alpha,\beta}$ being a local minimizer of $\Phi_{\alpha,\beta}$. Together with the strict convexity of $\Phi_{\alpha,\beta}$ this guarantees that $x_I^{\alpha,\beta}$ is the global minimizer. The definition of $x_I^{\alpha,\beta}$ then ensures that the support of the minimizer of $\Phi_{\alpha,\beta}$ is contained in I.

Note that the support of $x_I^{\alpha,\beta}$ is a subset of the finite set I and hence $x_I^{\alpha,\beta} \in \ell^1$. Thus, for $h \in \ell^2 \backslash \ell^1$ we have that $x_I^{\alpha,\beta} + h \notin \ell^1$, i.e. $\|x_I^{\alpha,\beta} + h\|_{\ell^1} = \infty$ and since $\alpha > 0$ it follows that $\Phi_{\alpha,\beta}(x_I^{\alpha,\beta} + h) = \infty$. From $\Phi_{\alpha,\beta}(0) = \frac{1}{2}\|y^\delta\|_{\mathcal{H}}^2 < \infty$ it is clear that the minimizing property of $x_I^{\alpha,\beta}$ ensures that $\Phi_{\alpha,\beta}(x_I^{\alpha,\beta}) < \infty$ which leads to $\Phi_{\alpha,\beta}(x_I^{\alpha,\beta} + h) - \Phi_{\alpha,\beta}(x_I^{\alpha,\beta}) = \infty > 0$. Thus, we restrict ourself to $h \in \ell^1$.

Splitting h into the components $h = h_I + h_{I^c}$ with $\operatorname{supp}(h_I) \subset I$ and $\operatorname{supp}(h_{I^c}) \subset I^c$ allows us to reformulate $\Phi_{\alpha,\beta}(x_I^{\alpha,\beta} + h) - \Phi_{\alpha,\beta}(x_I^{\alpha,\beta})$ equivalently as

$$\begin{aligned}
&\Phi_{\alpha,\beta}(x_I^{\alpha,\beta} + h) - \Phi_{\alpha,\beta}(x_I^{\alpha,\beta}) \\
&= \Phi_{\alpha,\beta}(x_I^{\alpha,\beta} + h_I) - \Phi_{\alpha,\beta}(x_I^{\alpha,\beta}) + \tfrac{1}{2}\|K h_{I^c}\|_{\mathcal{H}}^2 + \operatorname{Re}\langle K x_I^{\alpha,\beta} - y^\delta, K h_{I^c} \rangle_{\mathcal{H}} \\
&+ \operatorname{Re}\langle K h_I, K h_{I^c} \rangle_{\mathcal{H}} + \alpha \|h_{I^c}\|_{\ell^1} + \tfrac{\beta}{2}\|h_{I^c}\|_{\ell^2}^2.
\end{aligned}$$

Using $\operatorname{supp}(x_I^{\alpha,\beta} + h_I) \subset I$, the minimizing property of $x_I^{\alpha,\beta}$ implies that

$\Phi_{\alpha,\beta}(x_I^{\alpha,\beta} + h_I) - \Phi_{\alpha,\beta}(x_I^{\alpha,\beta}) > 0$ allowing us to estimate

$$\begin{aligned}
&\Phi_{\alpha,\beta}(x_I^{\alpha,\beta} + h) - \Phi_{\alpha,\beta}(x_I^{\alpha,\beta}) \\
&> \mathrm{Re}\langle Kx_I^{\alpha,\beta} - y^\delta, Kh_{I^c}\rangle_{\mathcal{H}} + \mathrm{Re}\langle Kh_I, Kh_{I^c}\rangle_{\mathcal{H}} + \alpha\|h_{I^c}\|_{\ell^1} \\
&\geq \alpha\|h_{I^c}\|_{\ell^1} - |\langle Kx_I^{\alpha,\beta} - y^\delta, Kh_{I^c}\rangle_{\mathcal{H}}| - |\langle Kh_I, Kh_{I^c}\rangle_{\mathcal{H}}|. \qquad (6.3)
\end{aligned}$$

For further estimation of the first scalar product, we utilize the basis expansion of h to see that $Kh_{I^c} = \sum_{i\in I^c} h_i Ke_i$. Using the linearity of the scalar product followed by the Hölder inequality allows us to estimate

$$\begin{aligned}
|\langle Kx_I^{\alpha,\beta} - y^\delta, Kh_{I^c}\rangle_{\mathcal{H}}| &\leq \sum_{i\in I^c} |h_i||\langle Kx_I^{\alpha,\beta} - y^\delta, Ke_i\rangle_{\mathcal{H}}| \\
&\leq \|h_{I^c}\|_{\ell^1} \sup_{i\in I^c} |\langle Kx_I^{\alpha,\beta} - y^\delta, Ke_i\rangle_{\mathcal{H}}|.
\end{aligned}$$

The second scalar product in (6.3) can be estimated by

$$|\langle Kh_I, Kh_{I^c}\rangle_{\mathcal{H}}| \leq \|K^*Kh_I\|_{\ell^\infty}\|h_{I^c}\|_{\ell^1}.$$

Plugging both estimates into equation (6.3) we continue the estimation by

$$\begin{aligned}
&\Phi_{\alpha,\beta}(x_I^{\alpha,\beta} + h) - \Phi_{\alpha,\beta}(x_I^{\alpha,\beta}) \\
&> \|h_{I^c}\|_{\ell^1}(\alpha - \sup_{i\in I^c} |\langle Kx_I^{\alpha,\beta} - y^\delta, Ke_i\rangle_{\mathcal{H}}| - \|K^*Kh_I\|_{\ell^\infty}).
\end{aligned}$$

Since it suffices to proof that $x_I^{\alpha,\beta}$ is a local minimizer we can restrict h to have $\|h\|_{\ell^1}$ arbitrarily small. For K is bounded this implies that $\|K^*Kh_I\|_{\ell^\infty}$ can be scaled to arbitrary small values. Hence, condition (6.2) is fulfilled if

$$\alpha > \sup_{i\in I^c} |\langle Kx_I^{\alpha,\beta} - y^\delta, Ke_i\rangle_{\mathcal{H}}|.$$

□

This proposition provides a rule on α which is hard to check a priori since it involves the knowledge of $x_I^{\alpha,\beta}$ and I. The dependence of $x_I^{\alpha,\beta}$ on α makes it an implicit rule. Tropp [52, 53] continues with a result on the ℓ^1 functional which reads in our infinite dimensional setting as follows.

Proposition 6.1.3. *[52, Corollary 7] Let $\mathcal{H}$ be a Hilbert space, $K \in L(\ell^2, \mathcal{H})$ injective, $y \in R(K)$ and $x^\dagger$ the sparse solution of $Kx = y$. Further, let y^δ be a noisy observation of y and $\eta = y^\delta - y$. We denote $I = \operatorname{supp}(x^\dagger)$, $P_I^* : \ell^2(I) \to \ell^2$ the canonical embedding and $K_I = KP_I^*$. If*

$$\sup_{i \in I^c} \|K_I^\dagger K e_i\|_{\ell^1(I)} < 1$$

the parameter rule

$$\alpha > \frac{\sup_{i \in I^c} |\langle \eta, P_{R(K_I)^\perp} K e_i \rangle_{\mathcal{H}}|}{1 - \sup_{i \in I^c} \|K_I^\dagger K e_i\|_{\ell^1(I)}}$$

ensures that the support of the minimizer of the ℓ^1 functional Ψ_α is contained in I.

Using the redrafted elastic-net functional given in equation (4.1) and denoting by $P_\perp$ the projection onto $R(\begin{pmatrix} K_I \\ \sqrt{\beta} P_I^* \end{pmatrix})^\perp$ leads to the following version of proposition 6.1.3 for the elastic net.

Proposition 6.1.4. *[52, Corollary 7] Let $\mathcal{H}$ be a Hilbert space, $K \in L(\ell^2, \mathcal{H})$, K injective or $\beta > 0$, $y \in R(K)$ and $x^\dagger$ the sparse solution of $Kx = y$. Further, let y^δ be a noisy observation of y and $\eta = y^\delta - y$. We denote $I = \operatorname{supp}(x^\dagger)$, $P_I^* : \ell^2(I) \to \ell^2$ the canonical embedding and $K_I = KP_I^*$. If*

$$\sup_{i \in I^c} \|(\beta \operatorname{id}_I + K_I^* K_I)^{-1} K_I^* K e_i\|_{\ell^1(I)} < 1$$

the parameter rule

$$\alpha > \frac{\sup_{i \in I^c} |\langle \begin{pmatrix} \eta \\ 0 \end{pmatrix}, P_\perp \begin{pmatrix} K \\ \sqrt{\beta} \operatorname{id} \end{pmatrix} e_i \rangle_{\mathcal{H} \times \ell^2}|}{1 - \sup_{i \in I^c} \|(\beta \operatorname{id}_I + K_I^* K_I)^{-1} K_I^* K e_i\|_{\ell^1(I)}}$$

ensures that the support of the minimizer of the elastic-net functional $\Phi_{\alpha,\beta}$ is contained in I.

The projection $P_\perp$ is hard to evaluate for a general operator K since

$$R(\begin{pmatrix} K_I \\ \sqrt{\beta}P_I^* \end{pmatrix})^\perp = \{x \in \mathcal{H} \times \ell^2 : \langle x, \begin{pmatrix} K_I \\ \sqrt{\beta}P_I^* \end{pmatrix} y\rangle_{\mathcal{H}\times\ell^2} = 0, y \in \ell^2(I)\}$$
$$= \{(x_1, x_2) \in \mathcal{H} \times \ell^2 : \langle x_1, K_I y\rangle_{\mathcal{H}} + \langle x_2, \sqrt{\beta}P_I^* y\rangle_{\ell^2} = 0, y \in \ell^2(I)\}.$$

However, we prove the following, similar result.

Proposition 6.1.5. *Let $\mathcal{H}$ be a Hilbert space, $K \in L(\ell^2, \mathcal{H})$ injective, $y \in R(K)$ and $x^\dagger$ the sparse solution of $Kx = y$. Further, let y^δ be a noisy observation of y and $\eta = y^\delta - y$. We denote $I = \operatorname{supp}(x^\dagger)$, $x_I^{\alpha,\beta} = \operatorname{argmin}_{\operatorname{supp}(x)\subset I} \Phi_{\alpha,\beta}(x)$, $P_I^* : \ell^2(I) \to \ell^2$ the canonical embedding and $K_I = KP_I^*$. If*

$$\sup_{i\in I^c} \|K_I^\dagger K e_i\|_{\ell^1(I)} < 1 \tag{6.4}$$

the implicit parameter rule

$$\alpha > \frac{\sup_{i\in I^c} |\langle \eta, P_{R(K_I)^\perp} K e_i\rangle_{\mathcal{H}}| + \beta\|x_I^{\alpha,\beta}\|_{\ell^\infty} \sup_{i\in I^c} \|K_I^\dagger K e_i\|_{\ell^1(I)}}{1 - \sup_{i\in I^c} \|K_I^\dagger K e_i\|_{\ell^1(I)}}$$

ensures that the support of the minimizer of the elastic-net functional $\Phi_{\alpha,\beta}$ is contained in I.

Proof. We continue the proof of the preceding theorem. By the definition of $x_I^{\alpha,\beta}$ and from the optimality condition, we know that $-K_I^*(K_I P_I x_I^{\alpha,\beta} - y^\delta) \in \alpha \operatorname{Sign}(P_I x_I^{\alpha,\beta}) + \beta P_I x_I^{\alpha,\beta}$. Multiplying both sides by $(K_I^* K_I)^{-1}$ we see its equivalence to $-P_I x_I^{\alpha,\beta} + K_I^\dagger y^\delta \in \alpha(K_I^* K_I)^{-1} \operatorname{Sign}(P_I x_I^{\alpha,\beta}) + \beta(K_I^* K_I)^{-1} P_I x_I^{\alpha,\beta}$ and hence, with a suitable $g \in \operatorname{Sign}(P_I x_I^{\alpha,\beta})$ we have

$$\begin{aligned} Kx_I^{\alpha,\beta} - y^\delta &= K_I(P_I x_I^{\alpha,\beta} - K_I^\dagger y^\delta) + (K_I K_I^\dagger y^\delta - y^\delta) \\ &= -\alpha(K_I^*)^\dagger g - \beta(K_I^*)^\dagger P_I x_I^{\alpha,\beta} + (K_I K_I^\dagger y^\delta - y^\delta). \end{aligned}$$

In this formula, we can use the fact that $K_I K_I^\dagger y^\delta = P_{R(K_I)} y^\delta$ and $y = K_I P_I x^\dagger \in R(K_I)$ to find

$$K_I K_I^\dagger y^\delta - y^\delta = -P_{R(K_I)^\perp} y^\delta = -P_{R(K_I)^\perp}(y^\delta - y) = -P_{R(K_I)^\perp}\eta.$$

Equipped with these equalities, we further estimate condition (6.1) component wise for each $i \in I^c$

$$\begin{aligned}
&|\langle Kx_I^{\alpha,\beta} - y^\delta, Ke_i\rangle_{\mathcal{H}}| \\
&\leq \alpha|\langle g, K_I^\dagger Ke_i\rangle_{\ell^2(I)}| + \beta|\langle P_I x_I^{\alpha,\beta}, K_I^\dagger Ke_i\rangle_{\ell^2(I)}| + |\langle P_{R(K_I)^\perp}\eta, Ke_i\rangle_{\mathcal{H}}| \\
&\leq |\langle P_{R(K_I)^\perp}\eta, Ke_i\rangle_{\mathcal{H}}| + \alpha\|K_I^\dagger Ke_i\|_{\ell^1(I)} + \beta\|x_I^{\alpha,\beta}\|_{\ell^\infty}\|K_I^\dagger Ke_i\|_{\ell^1(I)}
\end{aligned}$$

where we used the Hölder inequality and $\|g\|_{\ell^\infty} \leq 1$.

Now, using assumption (6.4), condition (6.1) is in force, whenever

$$\alpha > \frac{\sup_{i\in I^c}|\langle \eta, P_{R(K_I)^\perp}Ke_i\rangle_{\mathcal{H}}| + \beta\|x_I^{\alpha,\beta}\|_{\ell^\infty}\sup_{i\in I^c}\|K_I^\dagger Ke_i\|_{\ell^1(I)}}{1 - \sup_{i\in I^c}\|K_I^\dagger Ke_i\|_{\ell^1(I)}}.$$

□

We still follow the footsteps of Tropp [52,53] establishing an upper bound for the selection of α which guarantees the equity of the supports of the minimizer of the elastic-net functional $\Phi_{\alpha,\beta}$ and $x^\dagger$. Tropp's results for ℓ^1 minimization transcribed to our setting yields the next results on ℓ^1 minimization.

Theorem 6.1.6. *[52, Theorem 8] Let $\mathcal{H}$ be a Hilbert space, $K \in L(\ell^2, \mathcal{H})$, K injective, $y \in R(K)$ and $x^\dagger$ the sparse solution of $Kx = y$. Further, let y^δ be a noisy observation of y and $\eta = y^\delta - y$. We denote $I = \operatorname{supp}(x^\dagger)$, $P_I^* : \ell^2(I) \to \ell^2$ the canonical embedding and $K_I = KP_I^*$. If*

$$\sup_{i\in I^c}\|K_I^\dagger Ke_i\|_{\ell^1(I)} < 1$$

the parameter rule

$$\frac{\sup_{i\in I^c}|\langle \eta, P_{R(K_I)^\perp}Ke_i\rangle_{\mathcal{H}}|}{1 - \sup_{i\in I^c}\|K_I^\dagger Ke_i\|_{\ell^1(I)}} < \alpha < \frac{\min_{i\in I}|x^\dagger|_i}{\|(K_I^*K_I)^{-1}\|_{\ell^1(I),\ell^1(I)}} - \sup_{i\in I}|\langle \eta, Ke_i\rangle_{\mathcal{H}}|$$

ensures that the support of the minimizer of the ℓ^1 functional Ψ_α is equal to I.

This result can be applied to the redrafted elastic-net functional $\Phi_{\alpha,\beta}$, but, as mentioned for proposition 6.1.4, $P_\perp$ is hard to evaluate and we want to continue using our lower bound and the corresponding condition given in proposition 6.1.5. Hence, again we state a modification of Tropp's results.

Theorem 6.1.7. *Let $\mathcal{H}$ be a Hilbert space, $K \in L(\ell^2, \mathcal{H})$ injective, $y \in R(K)$ and $x^\dagger$ the sparse solution of $Kx = y$. Further, let y^δ be a noisy observation of y and $\eta = y^\delta - y$. We denote $I = \operatorname{supp}(x^\dagger)$, $P_I^* : \ell^2(I) \to \ell^2$ the canonical embedding and $K_I = KP_I^*$. If*

$$\sup_{i \in I^c} \|K_I^\dagger K e_i\|_{\ell^1(I)} < 1 \tag{6.5}$$

we find that the parameter rule

$$\frac{\sup_{i\in I^c} |\langle \eta, P_{R(K_I)^\perp} K e_i\rangle_{\mathcal{H}}| + \beta \frac{1}{2\alpha}\|y^\delta\|_{\mathcal{H}}^2 \sup_{i\in I^c} \|K_I^\dagger K e_i\|_{\ell^1(I)}}{1 - \sup_{i\in I^c} \|K_I^\dagger K e_i\|_{\ell^1(I)}} < \alpha$$

$$\frac{\min_{i\in I} |x^\dagger|_i}{\|(K_I^* K_I)^{-1}\|_{\ell^1(I),\ell^1(I)}} - \beta \frac{1}{2\alpha}\|y^\delta\|_{\mathcal{H}}^2 - \sup_{i\in I} |\langle \eta, K e_i\rangle_{\mathcal{H}}| > \alpha$$

ensures that the support of the minimizer of the elastic-net functional $\Phi_{\alpha,\beta}$ is equal to I.

Proof. We remark that for every $\alpha > 0$ and $\beta \geq 0$ the minimizer $x_I^{\alpha,\beta} = \operatorname{argmin}_{\operatorname{supp}(x)\subset I} \Phi_{\alpha,\beta}(x)$ fulfills

$$\alpha \|x_I^{\alpha,\beta}\|_{\ell^\infty} \leq \alpha \|x_I^{\alpha,\beta}\|_{\ell^1} \leq \Phi_{\alpha,\beta}(x_I^{\alpha,\beta}) \leq \Phi_{\alpha,\beta}(0) = \tfrac{1}{2}\|y^\delta\|_{\mathcal{H}}^2 \tag{6.6}$$

since $x_I^{\alpha,\beta}$ is the minimizer among all elements with support in I, especially 0 has such a support. The same inequality holds for the minimizer $x^{\alpha,\beta}$ of $\Phi_{\alpha,\beta}$.

Now, the lower bound on α ensures together with the previous corollary that the support of the minimizer of $\Phi_{\alpha,\beta}$ is a subset of $I = \operatorname{supp}(x^\dagger)$. Further, the optimality condition (4.2) for the minimizer of $\Phi_{\alpha,\beta}$, denoted by $x^{\alpha,\beta}$, guarantees that

$$\begin{aligned} K_I^* K_I (P_I x^{\alpha,\beta} - P_I x^\dagger) &= K_I^*(K_I P_I x^{\alpha,\beta} - y) \\ &= K_I^*(K x^{\alpha,\beta} - y^\delta) + K_I^* \eta \\ &\in -\alpha \operatorname{Sign}(P_I x^{\alpha,\beta}) - \beta P_I x^{\alpha,\beta} + K_I^* \eta \\ \implies P_I(x^{\alpha,\beta} - x^\dagger) &\in (K_I^* K_I)^{-1}(-\alpha \operatorname{Sign}(P_I x^{\alpha,\beta}) - \beta P_I x^{\alpha,\beta} + K_I^* \eta). \end{aligned}$$

Hence, there exists $u \in \operatorname{Sign}(P_I x^{\alpha,\beta})$ such that for every $j \in I$

$$\begin{aligned} |x^{\alpha,\beta} - x^\dagger|_j &= |\langle P_I(x^{\alpha,\beta} - x^\dagger), P_I e_j\rangle_{\ell^2(I)}| \\ &= |\langle (K_I^* K_I)^{-1}(\alpha u + \beta P_I x^{\alpha,\beta} - K_I^* \eta), P_I e_j\rangle_{\ell^2(I)}|. \end{aligned}$$

Utilizing the symmetry property of $(K_I^* K_I)^{-1}$ and the Hölder inequality we conclude

$$\begin{aligned}|x^{\alpha,\beta} - x^\dagger|_j &= |\langle \alpha u + \beta P_I x^{\alpha,\beta} - K_I^* \eta, (K_I^* K_I)^{-1} P_I e_j \rangle_{\ell^2(I)}| \\ &\leq \|\alpha u + \beta P_I x^{\alpha,\beta} - K_I^* \eta\|_{\ell^\infty(I)} \|(K_I^* K_I)^{-1} P_I e_j\|_{\ell^1(I)} \\ &\leq (\alpha + \beta \|P_I x^{\alpha,\beta}\|_{\ell^\infty(I)} + \|K_I^* \eta\|_{\ell^\infty(I)}) \|(K_I^* K_I)^{-1} P_I e_j\|_{\ell^1(I)}.\end{aligned}$$

Since $\operatorname{supp}(x^{\alpha,\beta}) \subset I$ and by equation (6.6) we have that $\|P_I x^{\alpha,\beta}\|_{\ell^\infty(I)} = \|x^{\alpha,\beta}\|_{\ell^\infty} \leq \frac{1}{2\alpha} \|y^\delta\|_{\mathcal{H}}^2$. Further, we can reformulate

$$\|K_I^* \eta\|_{\ell^\infty(I)} = \sup_{i \in I} |\langle \eta, K e_i \rangle_{\mathcal{H}}|$$

and

$$\|(K_I^* K_I)^{-1} P_I e_j\|_{\ell^1(I)} \leq \sup_{j \in I} \|(K_I^* K_I)^{-1} P_I e_j\|_{\ell^1(I)} = \|(K_I^* K_I)^{-1}\|_{\ell^1(I),\ell^1(I)}.$$

The right hand side is finite since $K_I^* K_I$ maps between finite dimensional spaces and hence is a matrix. Plugging this into the above calculation and applying the parameter rule ensures that for every $j \in I$

$$\begin{aligned}|x^{\alpha,\beta} - x^\dagger|_j &\leq (\alpha + \beta \frac{1}{2\alpha} \|y^\delta\|_{\mathcal{H}}^2 + \sup_{i \in I} |\langle \eta, K e_i \rangle_{\mathcal{H}}|) \|(K_I^* K_I)^{-1}\|_{\ell^1(I),\ell^1(I)} \\ &< \min_{i \in I} |x^\dagger|\end{aligned}$$

which proofs the claim. □

We leave the footsteps of Tropp addressing the still problematic estimation of $\|(K_I^* K_I)^{-1}\|_{\ell^1(I),\ell^1(I)}$ and $\|K_I^\dagger K e_i\|_{\ell^1(I)}$ in the parameter choice rule of Theorem 6.1.7. We can estimate

$$\sup_{i \in I^c} \|K_I^\dagger K e_i\|_{\ell^1(I)} \leq \|(K_I^* K_I)^{-1}\|_{\ell^1(I),\ell^1(I)} \cdot \sup_{i \in I^c} \|K_I^* K e_i\|_{\ell^1(I)}.$$

If we can achieve the Neumann condition, i.e. that

$$\|K_I^* K_I - \mathrm{id}_I\|_{\ell^1(I),\ell^1(I)} < 1, \tag{6.7}$$

we may apply the Neumann series estimate to find that

$$\|(K_I^* K_I)^{-1}\|_{\ell^1(I),\ell^1(I)} \leq \frac{1}{1 - \|K_I^* K_I - \mathrm{id}_I\|_{\ell^1(I),\ell^1(I)}}.$$

Combining these results we observe the estimate

$$\sup_{i\in I^c} \|K_I^\dagger K e_i\|_{\ell^1(I)} \leq \frac{\sup_{i\in I^c} \|K_I^* K e_i\|_{\ell^1(I)}}{1 - \|K_I^* K_I - \mathrm{id}_I\|_{\ell^1(I),\ell^1(I)}}. \tag{6.8}$$

To use this inequality we take a closer look on the Neumann condition (6.7).

Lemma 6.1.8. *In the situation of the preceeding theorem for properly scaled operators, i.e.* $\|Ke_i\|_{\mathcal{H}}^2 \leq 1$ *for all* $i \in I$*, define*

$$on := \min_{i\in I} \|Ke_i\|_{\mathcal{H}}^2, \quad off_I := \max_{i\in I} \sum_{j\in I, j\neq i} |\langle Ke_i, Ke_j\rangle_{\mathcal{H}}|,$$

$$off_{I^c} := \sup_{i\in I^c} \sum_{j\in I} |\langle Ke_j, Ke_i\rangle_{\mathcal{H}}|.$$

If $off_I - on < 0$ *the Neumann series estimate can be applied leading to*

$$\sup_{i\in I^c} \|K_I^\dagger K e_i\|_{\ell^1(I)} \leq \frac{off_{I^c}}{on - off_I}.$$

Proof. Motivated by the structure of K^*K, we start by splitting $K_I^* K_I$ into its diagonal M and off diagonal M_{off} finding that

$$\begin{aligned}\|K_I^* K_I - \mathrm{id}_I\|_{\ell^1(I),\ell^1(I)} &\leq \|M - \mathrm{id}_I\|_{\ell^1(I),\ell^1(I)} + \|M_{off}\|_{\ell^1(I),\ell^1(I)} \\ &= \max_{i\in I} |1 - \|Ke_i\|_{\mathcal{H}}^2| + \max_{i\in I} \sum_{j\in I, j\neq i} |\langle Ke_i, Ke_j\rangle_{\mathcal{H}}|.\end{aligned}$$

The assumption $\|Ke_i\|_{\mathcal{H}}^2 \leq 1$ for all $i \in I$ leads to the fact that the Neumann condition (6.7) is fulfilled whenever

$$1 - \min_{i\in I} \|Ke_i\|_{\mathcal{H}}^2 + \max_{i\in I} \sum_{j\in I, j\neq i} |\langle Ke_i, Ke_j\rangle_{\mathcal{H}}| < 1. \tag{6.9}$$

Since $off_I - on < 0$ the Neumann condition (6.7) is applicable and it holds

$$\|K_I^* K_I - \mathrm{id}_I\|_{\ell^1(I),\ell^1(I)} \leq 1 - on + off_I < 1.$$

Collecting the information of inequality 6.8 with the above estimate and the equation

$$\sup_{i\in I^c} \|K_I^* K e_i\|_{\ell^1(I)} = \sup_{i\in I^c} \sum_{j\in I} |K_I^* K e_i|_j = \sup_{i\in I^c} \sum_{j\in I} |\langle Ke_j, Ke_i\rangle_{\mathcal{H}}| = off_{I^c}$$

proves the lemma. □

Using this lemma in the parameter rules of theorem 6.1.7 we find our final result.

Lemma 6.1.9. *In the situation of the preceeding theorem 6.1.7 let K be properly scaled, i.e. $\|Ke_i\|_{\mathcal{H}}^2 \leq 1$ for all $i \in I$, and define*

$$on := \min_{i\in I} \|Ke_i\|_{\mathcal{H}}^2, \quad off_I := \max_{i\in I} \sum_{j\in I, j\neq i} |\langle Ke_i, Ke_j\rangle_{\mathcal{H}}|,$$

$$off_{I^c} := \sup_{i\in I^c} \sum_{j\in I} |\langle Ke_j, Ke_i\rangle_{\mathcal{H}}|.$$

Further, denote $\delta = \|\eta\|_{\mathcal{H}}$. If $\frac{off_{I^c}}{on-off_I} < 1$ the parameter rule

$$\frac{\delta(on - off_I) + \beta\frac{1}{2\alpha}\|y^\delta\|_{\mathcal{H}}^2 off_{I^c}}{on - off_I - off_{I^c}} < \alpha,$$

$$\min_{i\in I} |x^\dagger|_i (on - off_I) - \beta\frac{\|y^\delta\|_{\mathcal{H}}^2}{2\alpha} - \delta > \alpha$$

ensures that the support of the minimizer of the elastic-net functional $\Phi_{\alpha,\beta}$ is equal to I.

Proof. First, using the implication

$$\frac{off_{I^c}}{on - off_I} < 1 \iff off_{I^c} + off_I - on < 0 \implies off_I - on < 0.$$

assures the applicability of lemma 6.1.8 leading to

$$\sup_{i\in I^c} \|K_I^\dagger Ke_i\|_{\ell^1(I)} \leq \frac{off_{I^c}}{on - off_I} < 1.$$

Hence, the known parameter rule of theorem 6.1.7 can be used where we estimate

- $\sup_{i\in I^c} \|K_I^\dagger Ke_i\|_{\ell^1(I)}$ by $\frac{off_{I^c}}{on-off_I}$,
- $\|(K_I^* K_I)^{-1}\|_{\ell^1(I),\ell^1(I)}$ by $\frac{1}{on-off_I}$.

We have seen that all these replacements increase the values of the corresponding term. By Cauchy-Schwarz and the properties of projections we also know that

$$\langle \eta, P_{R(K_I)^\perp} Ke_i\rangle_{\mathcal{H}} \leq \|\eta\|_{\mathcal{H}} \|P_{R(K_I)^\perp} Ke_i\|_{\mathcal{H}} \leq \delta.$$

These replacements lead to the stated parameter rules. □

6.2 Parameter-choice recipe

In this section we use the situation of lemma 6.1.9. We first consider the parameter rule for $\beta = 0$ leading to the definition

$$\alpha_{\min} := \frac{\delta(on - off_I)}{on - off_I - off_{I^c}}, \quad \alpha_{\max} := \min_{i \in I} |x^\dagger|_i (on - off_I) - \delta.$$

In this notation the support of the minimizer of $\Phi_{\alpha,\beta}$ is equal to I if $\beta = 0$ and $\alpha_{\min} < \alpha < \alpha_{\max}$. Selecting $\beta > 0$ downsizes the interval, i.e. $\alpha_{\min} < \alpha < \alpha_{\max}$ is true for every choice of α and β fulfilling the parameter rules. Starting with the left hand side of the parameter rule of lemma 6.1.9, we find that

$$\beta < (\alpha - \alpha_{\min}) \frac{2\alpha(on - off_I - off_{I^c})}{\|y^\delta\|_{\mathcal{H}}^2 off_{I^c}} =: u_1(\alpha). \tag{6.10}$$

This constitutes a quadratic upper bound $u_1(\alpha)$ on β with zeros at 0 and $\alpha_{\min}$ and hence is monotonically increasing on $[\alpha_{\min}, \alpha_{\max}]$. Analogously, the right hand side of the parameter rule of lemma 6.1.9 leads to

$$\beta < (\alpha_{\max} - \alpha) \frac{2\alpha}{\|y^\delta\|_{\mathcal{H}}^2} =: u_2(\alpha). \tag{6.11}$$

This constitutes another quadratic upper bound $u_2(\alpha)$ on β with zeros at 0 and $\alpha_{\max}$. We aim to select β as big as possible to obtain maximal regularization for algorithms as the RFSS. A closer look on the two inequalities (6.10) and (6.11) yields that an upper bound $\beta_{\max}$ for β and a corresponding optimal value for α, which we call α_{opt}, can be calculated either as the intersection point of the two quadratic functions $u_1(\alpha)$ and $u_2(\alpha)$ on the interval $[\alpha_{\min}, \alpha_{\max}]$ or as the maximal point of the quadratic function $u_2(\alpha)$. Both situations are drafted in figure 6.2.1 and figure 6.2.2

A short calculation shows that the intersection point is attained at

$$\alpha = \frac{\alpha_{\min}(on - off_I - off_{I^c}) + \alpha_{\max} off_{I^c}}{on - off_I}.$$

The maximal point of $u_2(\alpha)$ is attained at

$$\alpha = \frac{\alpha_{\max}}{2}.$$

Hence, we have an explicit formula for α_{opt} and $\beta_{\max}$ which we summarize in the next lemma.

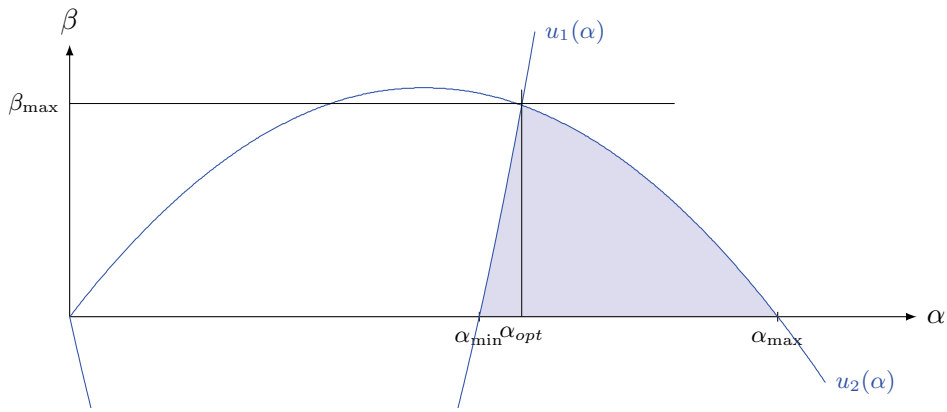


Figure 6.2.1: For all pairs (α, β) inside the blue region we obtain that the support of the minimizer of $\Phi_{\alpha,\beta}$ is equal to I. The maximal value for β can be chosen at the intersection of $u_1(\alpha)$ and $u_2(\alpha)$.

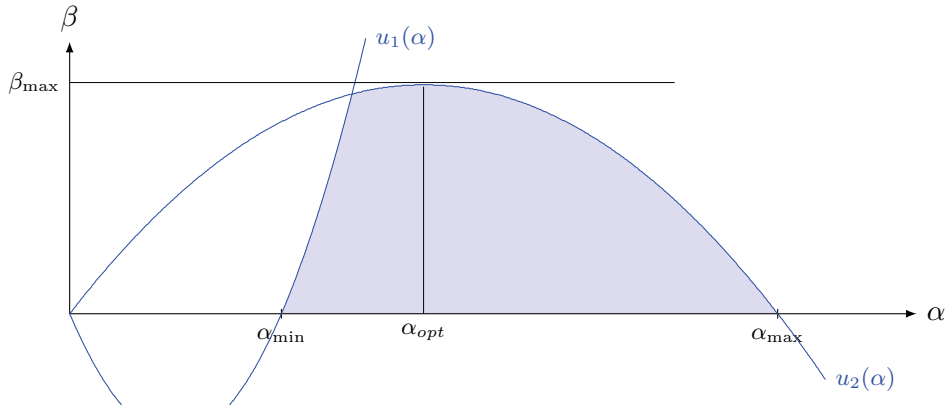


Figure 6.2.2: For all pairs (α, β) inside the blue region we obtain that the support of the minimizer of $\Phi_{\alpha,\beta}$ is equal to I. The maximal value for β can be chosen at the maximal point $u_2(\alpha)$.

Lemma 6.2.1. *In the situation of theorem 6.1.7 let K be properly scaled, i.e. $\|Ke_i\|_{\mathcal{H}}^2 \leq 1$ for all $i \in I$, and define*

$$\begin{aligned}
on &:= \min_{i\in I} \|Ke_i\|_{\mathcal{H}}^2, \\
off_I &:= \max_{i\in I} \sum_{j\in I, j\neq i} |\langle Ke_i, Ke_j\rangle_{\mathcal{H}}|, \\
off_{I^c} &:= \sup_{i\in I^c} \sum_{j\in I} |\langle Ke_j, Ke_i\rangle_{\mathcal{H}}|, \\
\alpha_{\min} &:= \frac{\delta(on - off_I)}{on - off_I - off_{I^c}}, \\
\alpha_{\max} &:= \min_{i\in I} |x^\dagger|_i (on - off_I) - \delta.
\end{aligned}$$

Further, denote $\delta = \|\eta\|_{\mathcal{H}}$. If $\frac{off_{I^c}}{on - off_I} < 1$ and $\alpha_{\min} < \alpha_{\max}$ the parameter rule

$$\begin{aligned}
\alpha &= \max\Big(\frac{\alpha_{\min}(on - off_I - off_{I^c}) + \alpha_{\max} off_{I^c}}{on - off_I}, \frac{\alpha_{\max}}{2}\Big), \\
\beta &< \frac{2\alpha(\alpha_{\max} - \alpha)}{\|y^\delta\|_{\mathcal{H}}^2}.
\end{aligned}$$

ensures that the support of the minimizer of the elastic-net functional $\Phi_{\alpha,\beta}$ is equal to I.

6.3 Numerical example

For demonstrating the applicability of the parameter rules of lemma 6.2.1, we recycle the mass spectrometry example of section 5.2. We have a vector $x^\dagger \in \ell^2$ supported on $I = \{63, 67, 373, 567, 1013\}$ with corresponding values $5, 4, 2, 9, 3$. Further, we consider the linear operator $K : \ell^2 \to L^2$ defined via

$$Ke_i(x) = \frac{1}{\sqrt{\sigma\sqrt{\pi}}} \mathrm{e}^{-\frac{(x-i)^2}{2\sigma^2}} \in L^2.$$

This operator corresponds to a convolution operator with a Gaussian kernel. The noisy measurement y^δ is generated by applying K and adding Gaussian noise η with a noise level of 5%

$$y^\delta = Kx^\dagger + \eta.$$

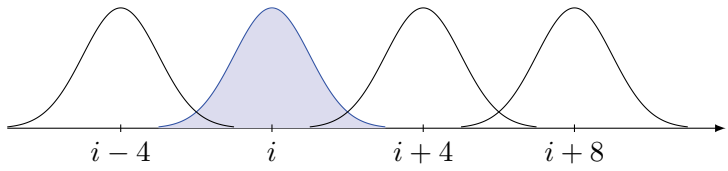


Figure 6.3.1: Correlating Ke_i (blue) with peaks Ke_j for $i, j \in I$, $i \neq j$ and the prior information that the minimal peak distance is 4.

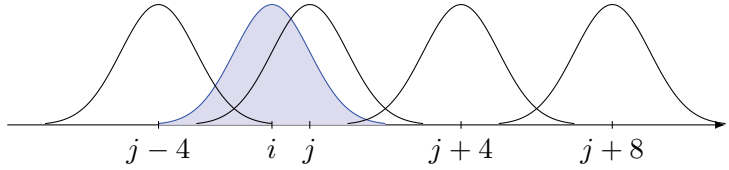


Figure 6.3.2: Correlating Ke_i (blue) with peaks Ke_j for $i \in I^c$, $j \in I$ and the prior information that the minimal peak distance is 4.

We start the calculation with the diagonal part using standard theory on the Gaussian kernel as can be found in standard statistics lectures

$$on(\sigma) = \min_{i \in I} \|Ke_i\|_{L^2}^2 = \frac{1}{\sigma\sqrt{\pi}} \int_{\mathbb{R}} \mathrm{e}^{-\frac{x^2}{\sigma^2}} \, \mathrm{dx} = 1.$$

We continue with the off-diagonal part, i.e. $j \neq i$,

$$\langle Ke_i, Ke_j \rangle_{L^2} = \int_{\mathbb{R}} Ke_i(x) Ke_j(x) \, \mathrm{dx} = e^{-\frac{(i-j)^2}{4\sigma^2}}.$$

This inner product depends only on the distance $d = j - i$ and, as drafted in figure 6.3.1, exploiting the prior knowledge that the minimal distance of the peaks is 4 leads to

$$\mathit{off}_I(\sigma) = \max_{i \in I} \sum_{j \in I, j \neq i} \langle Ke_i, Ke_j \rangle_{L^2} \leq 2 \sum_{d=1}^{\infty} e^{-\frac{(4d)^2}{4\sigma^2}}.$$

Analogously, as drafted in figure 6.3.2, we have

$$\mathit{off}_{I^c}(\sigma) = \sup_{i \in I^c} \sum_{j \in I} \langle Ke_i, Ke_j \rangle_{L^2} \leq \max_{k=1,2,3} \sum_{d=0}^{\infty} e^{-\frac{(k+4d)^2}{4\sigma^2}} + \sum_{d=0}^{\infty} e^{-\frac{((4-k)+4d)^2}{4\sigma^2}}.$$

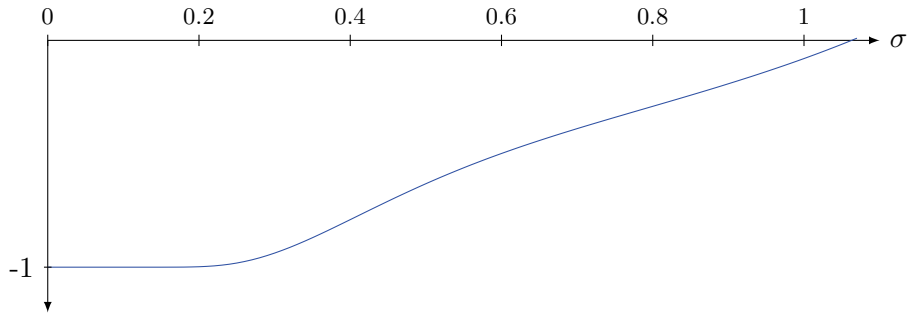


Figure 6.3.3: The condition of lemma 6.1.9 is fulfilled for all σ where the graph is negative.

Next, we check the assumptions of lemma 6.1.9 noticing that our operator is properly scaled since $on(\sigma) = 1$. Considering the condition

$$\frac{off_{I^c}(\sigma)}{on(\sigma) - off_I(\sigma)} - 1 < 0,$$

we can see in figure 6.3.3 for which values of σ the lemma can be applied. Before applying the parameter rule of lemma 6.1.9, we check if the interval for the parameter α is not empty, i.e. in the notation of lemma 6.1.9 if $\alpha_{\min} < \alpha_{\max}$ or more precisely

$$\frac{\delta(on(\sigma) - off_I(\sigma))}{on(\sigma) - off_I(\sigma) - off_{I^c}(\sigma)} < 2(on(\sigma) - off_I(\sigma)) - \delta \tag{6.12}$$

where we applied the prior knowledge that $\min_{i\in I} |x^\dagger|_i \geq 2$. Condition (6.12) is visualized in figure 6.3.4 as a function $\delta(\sigma)$.

Choosing $\sigma = 0.6$ we have $\|y^\delta\|_{\ell^2} = 11.6453$, $\delta = 0.5809$, $off_I(\sigma) = 0.0000299$, $off_{I^c}(\sigma) = 0.5013$ and $on_\sigma = 1$. Figure 6.3.5 illustrates the region of pairs (α, β) where exact recovery is granted. We obtain

$$\alpha_{\min} = 1.1649, \quad \alpha_{\max} = 1.4190.$$

Now, the parameter rule leads to the parameter choice of

$$\alpha = 1.2923, \quad \beta = 0.0024.$$

This choice in fact delivers the correct support (cf. figure 6.3.6).

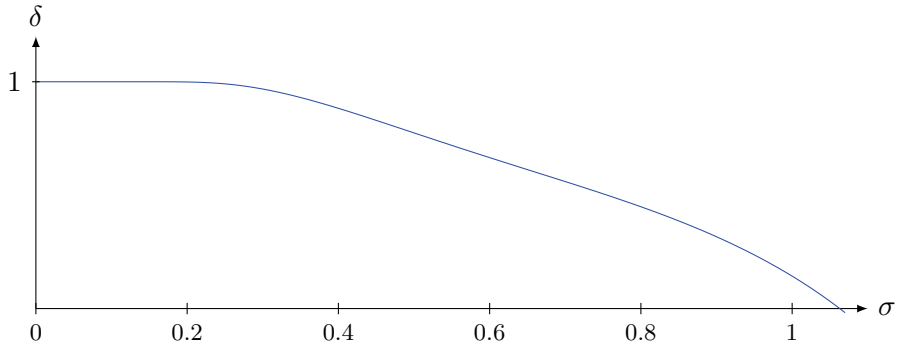


Figure 6.3.4: There exists a parameter α for exact recovery if the pair (σ, δ) is below the graph.

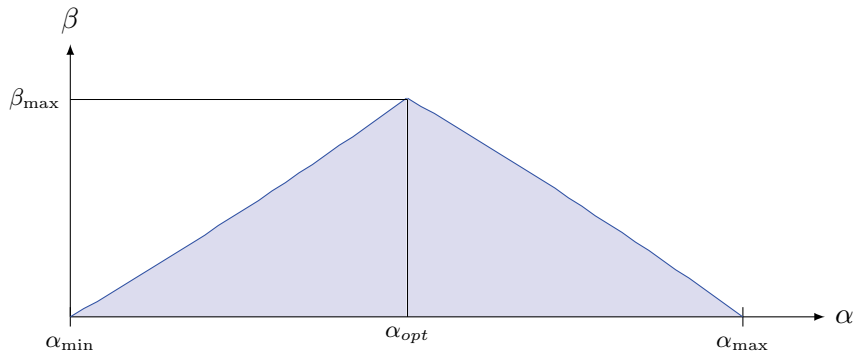


Figure 6.3.5: For all pairs (α, β) inside the blue region we obtain that the support of the minimizer of $\Phi_{\alpha,\beta}$ is equal to I. The maximal value for β can be chosen at the intersection of $u_1(\alpha)$ and $u_2(\alpha)$.

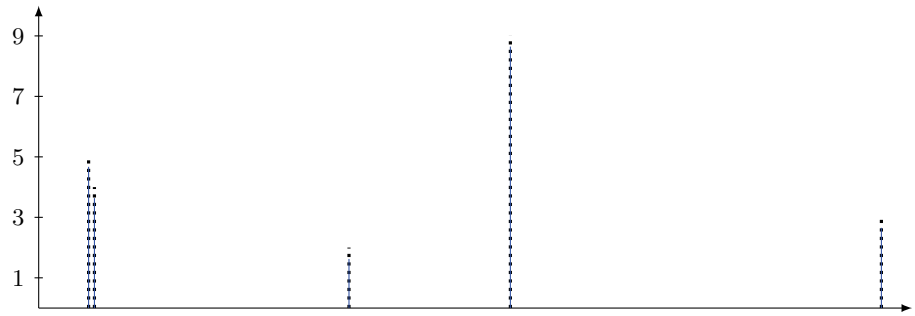


Figure 6.3.6: Source signal $x^\dagger$ (black) and the recovered signal $x^{\alpha,\beta}$ (blue), the supports coincide.

7 Conclusion & future work

We started this work with a motivation of sparsity in various applications as in inverse problems, compressed sensing and image processing. As a concrete framework, we considered a Hilbert space $\mathcal{H}$ and a bounded linear operator $K \in L(\ell^2, \mathcal{H})$. Having $y \in R(K)$ we assumed that $Kx = y$ has a sparse solution $x^\dagger$ and we denoted $I = \operatorname{supp}(x^\dagger)$. y^δ is a noisy observation of y with $\eta = y^\delta - y$ and $\delta = \|\eta\|_{\mathcal{H}}$. Some limitations of this framework were discussed. After some facts and basic principles we made a trip to the existing theory of ℓ^1 minimization, i.e. the minimization of the functional

$$\Psi_\alpha = \tfrac{1}{2}\|Kx - y^\delta\|_{\mathcal{H}}^2 + \alpha\|x\|_{\ell^1}$$

and corresponding algorithms. So far, this work constitutes an introduction to the state of the art of sparse approximation. Observing numerical troubles using ℓ^1-minimization algorithms, we have motivated the use of the elastic-net functional

$$\Phi_{\alpha,\beta} = \tfrac{1}{2}\|Kx - y^\delta\|_{\mathcal{H}}^2 + \alpha\|x\|_{\ell^1} + \tfrac{1}{2}\beta\|x\|_{\ell^2}^2,$$

showing that it stabilizes ℓ^1-minimization algorithms. We have shown how a part of the theory for ℓ^1 minimization carries over to the elastic net by reformulating the elastic-net functional as an ℓ^1 functional via

$$\Phi_{\alpha,\beta} = \frac{1}{2}\left\| \begin{pmatrix} K \\ \sqrt{\beta}\,\mathrm{id} \end{pmatrix} x - \begin{pmatrix} y^\delta \\ 0 \end{pmatrix} \right\|_{\mathcal{H}\times\ell^2}^2 + \alpha\|x\|_{\ell^1}.$$

Moreover, we have derived properties not obvious from this reformulation such as a version of the optimality condition, later exploited for the RSSN algorithm, and stability results for the minimizers. We have derived two active-set algorithms for the minimization of the elastic-net functional, the RSSN and RFSS, and have shown several active-set choices for the RSSN. The choice of the active-set is essentially influencing the convergence of the RSSN algorithm. The stabilizing properties of the elastic-net algorithms have been verified on numerical examples. Along with the verification we have compared different ℓ^1 algorithms in terms of their stability. Finally, parameter choice rules have been established which guarantee exact recovery, i.e. the support of the obtained elastic-net minimizer coincides with the support of $x^\dagger$. The parameter-choice rules culminate in a recipe containing an explicit formula for the choice of α and β, motivated by maximizing the stabilizing effect of the elastic net. The applicability of these rules have been shown for an example from mass spectrometry or rather for several applications involving convolution with a Gaussian kernel.

Last but not least, we give a short outlook on interesting open questions related to elastic-net minimization. An interesting tool for parameter selection would be to know under which conditions the supports of the minimizers of $\Phi_{\alpha,\beta}$ are monotonically increasing for decreasing α and fixed β. In that case the sparsity curve, i.e. plotting the sparsity of the minimizers in dependence of α, can be used to iteratively adapt α to find a value leading to the desired sparsity. Using the least squares solution and the fact that $\alpha = \max_{i\in\mathbb{N}} |K^* y^\delta|_i$ is the value where the minimizer becomes 0 results in two easy to calculate points on the sparsity curve. These can be used for interpolation to find the desired value for α. Iterating this procedure as using bisection for zero detection did work well in first experiments. Though, not every sparsity level is attained and hence a suitable α does not need to exist. The continuity results on the minimizers of $\Phi_{\alpha,\beta}$ in dependence of α promises a fast convergence of the algorithms during the iterations.

In this paper we have examined the abilities to use the elastic net for stabilizing ℓ^1-minimization algorithms. We aimed on still obtaining results close to those of ℓ^1 minimization. Hence, similarly to the exact recovery conditions it would be exciting to have conditions under which there exist $\beta > 0$ such that the support of the elastic-net minimizer coincides with the support of the ℓ^1 minimizer. Especially, an upper bound for such β values would be interesting.

We already mentioned the possibility to replace α, β by vectors $(\alpha_i)_{i\in\mathbb{N}}$, $(\beta_i)_{i\in\mathbb{N}}$ penalizing each component individually. It might be possible to improve known parameter rules by taking this modifications into account, i.e. in mass spectrometry one may have the a priory knowledge that different parts of the spectra have different noise levels. Hence, for different parts of the spectra different α values might be desirable.

On the algorithmic side, e.g. the IST has a guaranteed convergence speed for $\alpha = 0$, $\beta > 0$. An interesting thing is to check if this convergence speed can be guaranteed also for $\alpha > 0$, $\beta > 0$ which is not the case for $\alpha > 0$, $\beta = 0$. This might be used in combination with a path following strategy, i.e. successively reducing β, to obtain ℓ^1 minimizers with a guaranteed convergence speed, but benefiting from the regularizing properties of the elastic net. For the RSSN, it would be interesting to know how β influences the domain where the RSSN converges. First experiments suggest that the impact is positive and it might be possible also by strategies of varying β to transfer the RSSN into a globally convergent algorithm for ℓ^1 minimization, at least under some conditions on the operator K. The hope is that the super-linear convergence sustains.

In chapter 6, we presented exact-recovery conditions which constitute a worst case analysis. In compressed sensing, there are first efforts for probabilistic results (phase transition [5]). Those are still limited to greedy methods, i.e. matching pursuit, and a development of similar methods for ℓ^1 minimization or the elastic net are desirable. The most obvious problem is the lack of parameter-selection methods.

In chapter 5, we used the following condition for convergence of an elastic-net algorithm: For a given tolerance $\varepsilon > 0$ and $\alpha, \beta > 0$ let the output x of an elastic-net algorithm fulfill

$$\begin{aligned} |K^*(Kx - y) + \beta x + \alpha \operatorname{sign}(x)|_i &< \varepsilon, & x_i &\neq 0, \\ |K^*(Kx - y) + \beta x|_i - \alpha &< \varepsilon, & x_i &= 0, \end{aligned}$$

for every component x_i of x. It is an interesting question to study the following situation. Given $\varepsilon_0 > 0$ is there a maximal value ε_1 such that every output of an elastic-net minimization algorithm fulfilling the above conditions for $\varepsilon 1$ does also fulfill these conditions for $\beta = 0$ and tolerance ε_0, i.e. the output is also an ℓ^1 minimizer with tolerance ε_0.

Notation index

δ	Noise level $\|\eta\|_{\mathcal{H}}$
η	Noise $y^\delta - y$
$\mathcal{H}$	Hilbert space
K_I	KP_I^*
K	Bounded linear operator $\ell^2 \to \mathcal{H}$
P_I	Projection $\ell^2 \to \ell^2(I)$ for $I \subset \mathbb{N}$
S	Soft shrinkage function $S_\xi(x) = \max\{0, \|x\| - \xi\} \cdot \operatorname{sign}(x)$
sign	Sign function
Sign	Set valued sign function
$x^\dagger$	Sparse solution of $Kx = y$
y	Noise free data
y^δ	Noisy observation/measurement of y

Bibliography

[1] Theodore Alexandrov, Oliver Keszoecze, Dirk A. Lorenz, Stefan Schiffler, and Klaus Steinhorst. An active set approach to the elastic-net and its applications in mass spectrometry. In *Proceedings of the Second Workshop "'Signal Processing with Adaptive Sparse Structured Representations"'*, 2009.

[2] P. Babu, K. Pelckmans, P. Stoica, and J. Li. Linear systems, sparse solutions, and sudoku. *IEEE Signal Processing*, 17:40–42, 2010.

[3] Amir Beck and Marc Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM J. Imaging Sciences*, 2(1):183–202, 2009.

[4] E. van den Berg, M. P. Friedlander, G. Hennenfent, F. Herrmann, R. Saab, and Ö. Yılmaz. Sparco: A testing framework for sparse reconstruction. Technical Report TR-2007-20, Dept. Computer Science, University of British Columbia, Vancouver, October 2007.

[5] Jeffrey D. Blanchard, Coralia Cartis, Jared Tanner, and Andrew Thompson. Phase transitions for greedy sparse approximation algorithms. Technical Report ERGO 09-010, School of Mathematics, University of Edinburgh, 2009.

[6] Kristian Bredies, Theodore Alexandrov, Jens Decker, Dirk Lorenz, and Herbert Thiele. Sparse deconvolution for peak picking and ion charge estimation in mass spectrometry. In *Proceedings of the 15th European Conference on Mathematics for Industry*, 2008.

[7] Kristian Bredies and Dirk A. Lorenz. Iterated hard shrinkage for minimization problems with sparsity constraints. *SIAM Journal on Scientific Computing*, 30(2):657–683, 2008.

[8] Alfred M. Bruckstein, David L. Donoho, and Michael Elad. From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM Review*, 51(1):34–81, 2009.

[9] Emmanuel J. Candés and Michael B. Wakin. An introduction to compressive sampling. *IEEE Signal Processing Magazine*, 25(2):21–30, 2008.

[10] R.M. Caprioli, T.B. Farmer, and J. Gile. Molecular imaging of biological samples : Localization of peptides and proteins using MALDI-TOF MS. *Analytical Chemistry*, 69(23):4751–4760, 1997.

[11] Scott Shaobing Chen, David L. Donoho, and Michael A. Saunders. Atomic decomposition by basis pursuit. *SIAM Journal on Scientific Computing*, 20(1):33–61, 1998.

[12] Xiaojun Chen, Zuhair Nashed, and Liqun Qi. Smoothing methods and semismooth methods for nondifferentiable operator equations. *SIAM Journal on Numerical Analysis*, 38(4):1200–1216, 2000.

[13] Ingrid Daubechies, Michel Defrise, and Christine De Mol. An iterative thresholding algorithm for linear inverse problems with a sparsity constraint. *Communications in Pure and Applied Mathematics*, 57(11):1413–1457, 2004.

[14] Edmond de Hoffmann and Vincent Stroobant. *Mass Spectrometry.* Wiley & Sons, third edition, 2007.

[15] C. De Mol, E. De Vito, and L. Rosasco. Elastic-net regularization in learning theory. *Journal of Complexity*, 25(2):201–230, 2009.

[16] Loïc Denis, Dirk A. Lorenz, and Dennis Trede. Greedy solution of ill-posed problems: Error bounds and exact inversion. *Inverse Problems*, 25(11):115017 (24pp), 2009.

[17] David Donoho. Compressed sensing. *IEEE Transactions on Information Theory*, 52(4):1289–1306, 2006.

[18] David L. Donoho and Michael Elad. Optimally-sparse representation in general (non-orthogonal) dictionaries via ℓ^1 minimization. *Proceedings of the National Academy of Sciences*, 100:2197–2202, 2003.

[19] David L. Donoho and X. Huo. Uncertainty principles and ideal atomic decomposition. *IEEE Trans. Inform. Theory*, 47:2845–2862, 1999.

[20] David L. Donoho and Yaakov Tsaig. Fast solution of l1-norm minimization problems when the solution may be sparse. *IEEE trans. on inf. theory*, 54(11):4789–4812, 2008.

[21] Bradley Efron, Trevor Hastie, Iain Johnstone, and Robert Tibshirani. Least angle regression. *The Annals of Statistics*, 32(2):407–451, 2004.

[22] Heinz W. Engl, Martin Hanke, and Andreas Neubauer. *Regularization of Inverse Problems*, volume 375 of *Mathematics and its Applications*. Kluwer Academic Publishers Group, Dordrecht, 2000.

[23] Mário A. T. Figueiredo, Robert D. Nowak, and Stephen J. Wright. Gradient projection for sparse reconstruction: Applications to compressed sensing and other inverse problems. *IEEE Journal of Selected Topics in Signal Processing*, 1(4):586–597, 2007.

[24] Jean-Jacques Fuchs. Recovery of exact sparse representations in the presence of bounded noise. *IEEE Transactions on Information Theory*, 51(10):3601–3608, 2005.

[25] Markus Grasmair. Well-posedness and convergence rates for sparse regularization with sublinear l^q penalty term. *Inverse Problems and Imaging*, 3(3):383–387, 2009.

[26] Markus Grasmair, Markus Haltmeier, and Otmar Scherzer. Sparse regularization with ℓ^q penalty term. *Inverse Problems*, 24(5):055020 (13pp), 2008.

[27] Rémi Gribonval and Morten Nielsen. Beyond sparsity: Recovering structured representations by ℓ^1 minimization and greedy algorithms. *Advances in Computational Mathematics*, 28(1):23–41, 2008.

[28] Roland Griesse and Dirk A. Lorenz. A semismooth Newton method for Tikhonov functionals with sparsity constraints. *Inverse Problems*, 24(3):035007 (19pp), 2008.

[29] Elaine T. Hale, Wotao Yin, and Yin Zhang. Fixed-point continuation for l1-minimization: Methodology and convergence. *SIAM J. Optim.*, 19(3):1107–1130, 2008.

[30] Per Christian Hansen. Regularization Tools Version 4.0 for Matlab 7.3. *Numerical Algorithms*, 46:189–194, 2007.

[31] Michael Hintermüller, Kazufumi Ito, and Karl Kunisch. The primal-dual active set strategy as a semismooth Newton method. *SIAM Journal on Optimization*, 13(3):865–888, 2003.

[32] Michael Hintermüller and Karl Kunisch. Path-following methods for a class of constrained minimization problems in function space. *SIAM J. Optim.*, 17(1):159–187, 2006.

[33] Bangti Jin, Dirk A. Lorenz, and Stefan Schiffler. Elastic-net regularization: Error estimates and active set methods. *Inverse Problems*, 25(11):115022 (26pp), 2009.

[34] Seung-Jean Kim, K. Koh, M. Lustig, Stephen Boyd, and Dimitry Gorinevsky. An interior-point method for large-scale l1-regularized least squares. *IEEE J. Sel. Top. Signal Process.*, 1(4):606–17, 2007.

[35] Esther Klann, Michael Kuhn, Dirk A. Lorenz, Peter Maass, and Herbert Thiele. Shrinkage versus deconvolution. *Inverse Problems*, 23(5):2231–2248, 2007.

[36] Jagadeesha Kumar and M.S. Shunmugam. Fitting of robust reference surface based on least absolute deviations. *Precision engineering*, 31:102–113, 2007.

[37] Honglak Lee, Alexis Battle, Rajat Raina, and Andrew Y. Ng. Efficient sparse coding algorithms. In B. Schölkopf, J. Platt, and T. Hoffman,

editors, *Advances in Neural Information Processing Systems 19*, pages 801–808. MIT Press, Cambridge, MA, 2007.

[38] Dirk A. Lorenz. Convergence rates and source conditions for Tikhonov regularization with sparsity constraints. *Journal of Inverse and Ill-Posed Problems*, 16(5):463–478, 2008.

[39] Dirk A. Lorenz, Stefan Schiffler, and Dennis Trede. Beyond convergence rates: Exact inversion of ill-posed problems with sparsity constraints. Preprint, 2009.

[40] Ignace Loris. On the performance of algorithms for the minimization of ℓ^1-penalized functionals. *Inverse Problems*, 25(3):035008 (16pp), 2009.

[41] Stéphane Mallat, Geoff Davis, and Zhifeng Zhang. Adaptive time-frequency decompositions. *SPIE Journal of Optical Engineering*, 33(7):2183–2191, July 1994.

[42] L.A. McDonnell and R.M.A. Heeren. Imaging mass spectrometry. *Mass Spectrometry Reviews*, 26(4):606–643, 2007.

[43] B.K. Natarajan. Sparse approximate solutions to linear systems. *SIAM J. Comput.*, 24(2):227–234, 1995.

[44] Albrecht Pietsch. *History of Banach Spaces and Linear Operators*. Birkhäuser, 2007.

[45] Andreas Rieder. *Keine Probleme mit inversen Problemen*. Vieweg+Teubner, 2003.

[46] R. Tyrell Rockafellar and Roger J-B. Wets. *Variational Analysis*. Springer, 1998.

[47] Justin Romberg. Imaging via compressive sampling. *IEEE Signal Processing Magazine*, 25(2):14–20, 2008.

[48] Volker Roth. The generalized lasso. *IEEE Trans On Neural Networks*, 15(1):16–28, 2004.

[49] Stefan Schiffler. Exact inversion with the elastic net. *preprint*, 2010.

[50] Ralph Edwin Showalter. *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*, volume 49 of *Mathematical Surveys and Monographs*. American Mathematical Society, 1997.

[51] Joel A. Tropp. Greed is good: Algorithmic results for sparse approximation. *IEEE Transactions on Information Theory*, 50(10):2231–2242, 2004.

[52] Joel A. Tropp. Just relax: Convex programming methods for identifying sparse signals in noise. *IEEE Transactions on Information Theory*, 52(3):1030–1051, 2006.

[53] Joel A. Tropp. Corrigendum in "Just relax: Convex programming methods for identifying sparse signals in noise". *IEEE Transactions on Information Theory*, 55(2):917–918, 2009.

[54] Joel A. Tropp and Stephen J. Wright. Computational methods for sparse solution of linear inverse problems. Technical report, California Institute of Technology, 2009.

[55] Yaakov Tsaig and David L. Donoho. Breakdown of equivalence between the minimal ℓ^1-norm solution and the sparsest solution. *Signal Processing*, 86:533–548, 2006.

[56] S. Ulbrich. Semismooth Newton methods for operator equations in function spaces. *SIAM Journal on Optimization*, 13(3):805–842, 2002.

[57] J. Throck Watson and O. David Sparkman. *Introduction to Mass Spectrometry*. Wiley & Sons, fourth edition, 2007.

[58] Dirk Werner. *Funktionalanalysis*. Springer, 2005.

[59] S. J. Wright, R. D. Nowak, and M. A. T. Figueiredo. Sparse reconstruction by separable aproximation. *IEEE Transactions on Signal Processing*, 57(7):2479–2493, 2009.

[60] Clemens A. Zarzer. On Tikhonov regularization with non-convex sparsity constraints. *Inverse Problems*, 25(2):025006 (13pp), 2009.

[61] Peng Zhao and Bin Yu. Stagewise lasso. *Journal of Machine learning Res.*, 8:2701–2726, 2007.

[62] Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B*, 67(2):301–320, 2005.