PHASE TRANSITIONS IN GENERALIZED COMPRESSED SENSING

A thesis submitted to the University of Manchester for the degree of Doctor of Philosophy in the Faculty of Science & Engineering

2019

Jake Walvin
School of Mathematics
# Contents

Abstract 5

Declaration 6

Copyright Statement 7

Acknowledgements 9

1 Introduction 10
   1.1 Practical limitations ................................. 14
   1.2 Related work ........................................... 15
   1.3 Notation ................................................. 15
   1.4 Outline .................................................. 15

2 Compressed sensing 17
   2.1 Basics of compressed sensing ............................ 18
      2.1.1 Efficiency .......................................... 21
      2.1.2 Equivalence to $\ell_0$-minimization .................. 22
      2.1.3 Alternative algorithms ............................... 25
      2.1.4 Some experiments .................................. 26
   2.2 Compressibility, stability and robustness .............. 27
   2.3 A general setting ........................................ 29
   2.4 Regularization .......................................... 30
2.4.1 Subdifferentials and descent cones ........................................ 32

3 Integral geometry .............................................................. 36
  3.1 Intrinsic volumes .......................................................... 37
    3.1.1 The statistical dimension ........................................... 40
    3.1.2 The kinematic formulas ............................................. 42
  3.2 Concentration of measure ................................................ 43
    3.2.1 Phase transitions in convex regularization ....................... 45
  3.3 Computing the statistical dimension ................................... 46
    3.3.1 Descent cones of the 1-norm ...................................... 48

4 TV regularization ............................................................... 51
  4.1 Analysis and synthesis sparsity ......................................... 52
  4.2 One dimensional TV norm regularization .............................. 53
    4.2.1 One dimensional finite differences ............................... 55
    4.2.2 Examples ............................................................ 60
    4.2.3 Computing the expected distance ................................ 62
  4.3 2D Total Variation and Graphs ........................................ 66

5 Condition number bounds .................................................... 68
  5.1 Duality ........................................................................ 71
  5.2 Randomized preconditioning .............................................. 76
    5.2.1 The TQC Lemma ..................................................... 77
    5.2.2 On the probabilistic reduction .................................... 79

6 Conclusions ...................................................................... 84
  6.1 TV regularization in two dimensions ................................... 85
  6.2 Conditioning of projected matrices .................................... 85
  6.3 Computational methods and sampling ................................ 85
Compressed Sensing is the problem of recovering sparse solutions to underdetermined systems of linear equations efficiently, when the number of equations is roughly proportional to the sparsity of the solution. This problem has received significant attention in the last 15 years, with applications in many fields of science and technology, including medical imaging, astronomy, machine learning, and signal processing. The most popular approach has been to find a sparse solution $x^0$ of the underdetermined system $Ax = y$ by solving the optimization problem

$$\text{minimize } \|x\|_1 \text{ subject to } Ax = y.$$ 

If $A$ has $m$ rows, the question is how large $m$ has to be so that solving this minimization problem recovers $x^0$ with high probability for random $A$. In groundbreaking work by Donoho, and by Candès and Tao, it was shown that this optimization problem recovers a sparse solution with high probability if $m$ is greater than a constant times $s \log(n/s)$, where $s$ is the sparsity of the solution and $n$ the dimension of the problem. More generally, Amelunxen, Lotz, McCoy and Tropp have shown that

$$\text{minimize } f(x) \text{ subject to } Ax = y$$

recovers a solution $x^0$ of an underdetermined system $Ax = y$ with high probability for random $A$ if and only if $m$ exceeds a critical threshold, the statistical dimension $\delta(f, x^0)$ of $f$ at $x^0$.

The thesis is concerned with the theoretical problem of studying this critical threshold $\delta(f, x^0)$ in the case where $f(x) = g(Dx)$ for some convex function $g$ and matrix $D$. The focus is on cases where computing $\delta(g, x^0)$ is easy, and we aim to estimate $\delta(f, x^0)$ in terms of properties of $g$ and $D$. A practically important example is when $f(x) = \|\Omega x\|_1$, where $\Omega$ is the difference operator $(\Omega x)_i = x_i - x_{i-1}$, which is a one-dimensional case of a Total Variation (TV) norm. The main contributions are:

1. Sharp analytical bounds on the statistical dimension of $\|D \cdot \|_1$ in the case where $D$ is a difference matrix, denoted $\Omega$ (total variation minimization);

2. Bounds on the statistical dimension of $g(D \cdot)$ in terms of that of $g$ using the condition number of $D$;

3. Improvements to condition number bounds by using random projections.

Overall, the results contribute to understanding the role of conditioning of $D$ for the success of solving linear inverse problems by convex regularization.
Declaration

No portion of the work referred to in the thesis has been submitted in support of an application for another degree or qualification of this or any other university or other institute of learning.
Copyright Statement

i. The author of this thesis (including any appendices and/or schedules to this thesis) owns certain copyright or related rights in it (the “Copyright”) and s/he has given The University of Manchester certain rights to use such Copyright, including for administrative purposes.

ii. Copies of this thesis, either in full or in extracts and whether in hard or electronic copy, may be made only in accordance with the Copyright, Designs and Patents Act 1988 (as amended) and regulations issued under it or, where appropriate, in accordance with licensing agreements which the University has from time to time. This page must form part of any such copies made.

iii. The ownership of certain Copyright, patents, designs, trade marks and other intellectual property (the “Intellectual Property”) and any reproductions of copyright works in the thesis, for example graphs and tables (“Reproductions”), which may be described in this thesis, may not be owned by the author and may be owned by third parties. Such Intellectual Property and Reproductions cannot and must not be made available for use without the prior written permission of the owner(s) of the relevant Intellectual Property and/or Reproductions.

iv. Further information on the conditions under which disclosure, publication and commercialisation of this thesis, the Copyright and any Intellectual Property and/or Reproductions described in it may take place is available in the University IP Policy (see http://documents.manchester.ac.uk/DocuInfo.aspx?DocID=487),
in any relevant Thesis restriction declarations deposited in the University Library, The University Library’s regulations (see http://www.manchester.ac.uk/library/aboutus/regulations) and in The University’s Policy on Presentation of Theses.
Acknowledgements

I would like to thank my supervisor, Martin Lotz, for being a great mentor and a very understanding, encouraging and helpful person, both academically and non-academically. I have also had the great fortune of having many advisors at Manchester, both academically and personally. I would also like to thank my family, my mum Michelle, sister Keshia, and my brothers Daniel and Ryan, as well as my great auntie Barbara (deceased), my grandparents Sally, Edwin (deceased), June (deceased) and even my father, John, as well as all other family members who have helped me along the way. I would also like to thank my ex girlfriend Lauren, for helping me become a better person. Finally my friends, many of whom still reside in the Maths department, and in particular office 2.210.
Chapter 1

Introduction

Applications in science and technology depend increasingly on vast amounts of data. In many cases, the data we are considering can be modelled as a superposition, or sum, of datasets. Each dataset will exhibit a particularly simple structure when viewed on its own. It is an important challenge to devise methods that handle large data sets efficiently by exploiting the underlying simple structure. A popular example is image compression. When a picture is taken, it is typically sampled as a grid of points, each one representing a colour or light intensity. The amount of data is proportional to the number of these pixels. The relevant information content, however, is much smaller, the JPEG image compression standard is a good example of this. Expanding the image using the 2D discrete cosine transform reveals that the image is sparse: most relevant information is concentrated around a few coefficients. Other examples include signal denoising and separating mutually coherent structures. One of the big breakthroughs in recent years, generally referred to as compressed or compressive sensing [28, 22, 15], has been the realization that for sparse or compressible data, it is possible to acquire, reconstruct and separate datasets, using a number of measurements that is proportional to the information content of the data rather than the size of the data itself.

A typical situation is when one has an observation $y \approx Ax$, where $x \in \mathbb{R}^n$ represents
data one wants to recover, \( y \in \mathbb{R}^m \), with \( m \leq n \), represents observations, and \( A \in \mathbb{R}^{m \times n} \) is a measurement operator. If \( m < n \) but the solution is known to be sparse or approximately sparse, then one can solve the regularized problem

\[
\text{minimize } \|Ax - y\|^2 + \lambda \|x\|_1,
\]

where \( \|x\|_1 = \sum_{i=1}^{n} |x_i| \) is the 1-norm of \( x \) and \( \lambda \) is a regularization parameter. Problem (1.0.1) is mathematically equivalent to

\[
\text{minimize } \|x\|_1 \quad \text{subject to } \|Ax - y\|^2 \leq \epsilon,
\]

for some \( \epsilon > 0 \), see Proposition 2.4.2. A simplified version is the basis pursuit problem, which will be referred to throughout the thesis

\[
\text{minimize } \|x\|_1 \quad \text{subject to } Ax = y,
\]

where \( \| \cdot \|_1 = \sum_{i=1}^{n} |x_i| \) is the 1-norm. In compressed sensing [22, 15] it was shown that if \( Ax = y \) has a solution \( x^0 \) that is \( s \)-sparse, which means that it has at most \( s \) non-zero entries, then the basic minimization problem (1.0.2) recovers \( x^0 \) with high probability if \( m \), the number of rows of the random matrix \( A \), is larger than \( cs \log(n/s) \), where \( c \) is a constant. If \( m \geq 2s \) and \( A \) has full rank, then \( x^0 \) can also be found directly, but the problem is NP-hard. The importance of the compressed sensing result is that (1.0.2) can be solved efficiently using linear programming.

In many applications one is interested in structures other than sparsity. For example, \( x^0 \) can be a vector with sparse difference, or a matrix of low rank. An image with large blocks of one colour is not sparse, but is sparse when one considers the differences of neighbouring pixels. Instead of the 1-norm, such applications make use of other norms such as the total variation norm [16]. For signals in one dimension, one can
define \( f(x) = \|\Omega x\|_1 \), where

\[
\Omega = \begin{pmatrix}
-1 & 1 & 0 & \cdots & 0 \\
0 & -1 & 1 & \cdots & 0 \\
0 & 0 & -1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & -1
\end{pmatrix}.
\] (1.0.3)

Replacing the 1-norm in (1.0.2) with functions of this form is known as the cosparse, or analysis \( \ell_1 \) setting [42]. More generally, one can consider the problem

\[
\text{minimize } f(x) \text{ subject to } Ax = y,
\] (1.0.4)

where \( f \) is a convex function.

If \( A \) has \( m \) rows, the question is how large \( m \) has to be so that solving this minimization problem recovers \( x^0 \) with high probability for random \( A \). In groundbreaking work by Donoho, and by Candès and Tao, it was shown that this optimization problem recovers a sparse solution with high probability if \( m \) is greater than a constant times \( s \log(n/s) \), where \( s \) is the sparsity of the solution and \( n \) the dimension of the problem.

More generally, Amelunxen, Lotz, McCoy and Tropp have shown in [4] that if \( A \) is a matrix whose entries are standard Gaussian random variables and \( y = Ax^0 \), then (1.0.4) has the unique solution \( x^0 \) with overwhelming probability if the number of rows \( m \) of \( A \) is slightly larger than a critical threshold, the \textit{statistical dimension} \( \delta(f,x^0) \) of \( f \) at \( x^0 \), which we will discuss in detail in Chapter 3, and does not recover \( x^0 \) if \( m \) is slightly smaller. This shows that there is a \textit{phase transition} at a location \( \delta(f,x^0) \). In the case \( f(x) = \|x\|_1 \) this has been observed earlier by Donoho and Tanner [24] and Stojnic [50].

In [4] it was shown that phase transitions exist, but the problem of determining the location \( \delta(f,x^0) \) at which these happen is difficult. For 1-norm minimization, the statistical dimension depends only on the sparsity of \( x^0 \). A method for accurately
Figure 1.1: Phase transitions for three different structures and regularizers. Figure from [4]

determining the statistical dimension in some cases that include the 1-norm and some other examples has been developed by Stojnic [50], see also [4, Recipe 4.1]. However, it does not work as easily for $f(x) = \|\Omega\|_1$ with $\Omega$ the difference matrix above. In fact, while $\delta(\|\cdot\|_1, x^0)$ depends only on the sparsity pattern of $x^0$, the corresponding problem for the TV norm also depends on the location of the support, and the problem of determining this threshold location has been stated as an open problem in [34]. From now on we will refer to the minimization problem (1.0.4) with $f(x) = \|\Omega x\|_1$ as the Total Variation (TV) problem.

This thesis studies different methods to establish the statistical dimension, and the permissible undersampling, in cases where $f(x) = g(Dx)$, with emphasis on the total variation problem. These are

1. An analytical estimate of upper bounds for the statistical dimension of $f(x) = \|Dx\|_1$;

2. General bounds on the statistical dimension of $f(x) = g(Dx)$ in terms of the condition number of random projections of $D$.

The total variation phase transition curve has been elusive. An approximation features in [36], while experimental work has been carried out in [34]. The paper [56] shows that the statistical dimension of the one dimensional TV regularizer coincides asymptotically with the minimax risk of total variation denoising, as studied in [23].
The analysis presented in this thesis gives a reason for why this phase transition is difficult to determine: the large condition number of $D$.

Basic number bounds on the statistical dimension of a composite regularizer $f(x) = g(Dx)$ have the form

$$
\delta(f,x^0) \leq \kappa(D)^2 \cdot \delta(g,Dx^0),
$$

where $\kappa$ is the matrix condition number (actually, the bounds are in terms of Renegar’s condition number with respect to a descent cone, which is stronger). In the case where $D$ is the difference matrix (1.0.3), the condition number is of order $n$ and the bound is of little use. We therefore study an improvement in terms of random projections. If $D$ is invertible, then

$$
\delta(f,x_0) \lesssim \kappa_m^2(D^{-1}) \cdot \delta(g,Dx_0),
$$

where $\kappa_m(A)$ is obtained by averaging the condition number under random orthogonal projections of rows of $A$ onto $m$ dimensions. Even though the bounds are not too sharp (they can’t be), they give non-trivial bounds that illustrate the role of conditioning.

1.1 Practical limitations

The main focus of this thesis is on theoretical questions about properties of convex functions and their geometry. In reality, the measurement matrices $A$ are not random Gaussian matrices, but there are universality results [25, 8, 45] that show that the results extend to other distributions. In practical applications, the matrix $A$ can arise from subsampling directions in computed tomography (CT), and experiments [34] have shown that phase transitions can also be observed in this setting. In general, the statistical dimension can serve as a rule of thumb on how much one can reduce the number of measurements needed to recover a structured signal.
1.2 Related work

The problem of determining the phase transition of total variation minimization has received some attention recently. The works [43, 36] studied the performance of TV minimization and obtained bounds on when this would recover a solution with sparse gradient with high probability, but their bounds are not sharp and do not determine the phase transition of TV minimization. In [56], the authors show that the statistical dimension equals asymptotically the minimax risk of denoising studied by [23], but their work does not give explicit expressions for the statistical dimension of $\|Dx\|_1$. More recently, improved bounds on the statistical dimension of analysis $\ell_1$ problems have been proposed [21, 29], but from this work it becomes clear that the condition of $D$ remains an obstacle to getting sharp bounds.

1.3 Notation

For an integer $n$, we write $[n] = \{1, \ldots, n\}$. For a set $S \subseteq [n]$, we denote by $\overline{S}$ the complement of $S$ in $[n]$. We use the notation $\mathbb{P}\{\cdot\}$ and $\mathbb{E}[\cdot]$ for the probability and expectation of an event. Depending on context $g$ will be a function or a Gaussian vector and denote the normal (Gaussian) distribution by $N(0, 1)$. Depending on context, this will refer to a Gaussian distribution in $\mathbb{R}^n$ or in $\mathbb{R}$. For vectors $x$ and matrices $A$, we usually write $x_S$ and $A_S$ for the subvector and submatrix with rows indexed by a set $S \subseteq [n]$.

1.4 Outline

Chapter 2 contains an introduction to the problem of compressed sensing and its generalizations. Chapter 3 introduces the tools of integral geometry that are needed in this thesis. While the content of Chapters 2 and 3 is known material, the results in
Chapters 4 and 5 are new. Chapter 4 contains a detailed computation of a parameter related to the statistical dimension of TV minimization, which is then applied to obtain improved bounds. Chapter 5 is based on joint work with D. Amelunxen and M. Lotz, and studies condition number bounds on the statistical dimension of $f(Dx)$. Chapter 6 concludes with some discussion and future directions.
Chapter 2

Compressed sensing and low-complexity models

A vector $x \in \mathbb{R}^n$ is called $s$-sparse, if it has at most $s$ non-zero entries. An initial way to conceptualise compressed sensing is to analyse the simplified and idealised setting where $Ax^0 = y$ for an $s$-sparse vector $x^0$, and the goal is to recover $x^0$ in a situation in which the number of rows $m$ is less than the number of columns $n$. We begin by considering the optimization problem as within problem (1.0.2)

$$\text{minimize} \; \|x\|_1 \quad \text{subject to} \quad Ax = y.$$ 

where $\|\cdot\|_1 = \sum_{i=1}^{n} |x_i|$ is the 1-norm. What we hope to accomplish by solving this problem, is to find the sparsest solution. As an optimization problem, we see this as a relaxation to the problem of minimizing the $\ell_0$ “norm” $\|x\|_0$, which is simply counting the number of non-zero entries in the vector. After a general discussion of sparsity, we relax the conditions and define compressibility as an approximate notion of sparsity, and discuss robustness and stability for the case where $Ax = y$ is replaced by $\|Ax - y\| \leq \varepsilon$. After this, we replace the 1-norm in problem (1.0.2) with a general convex function and study conditions under which solving this problem recovers a given solution $x^0$ of $Ax = y$. This leads to a discussion of notions convex
analysis and geometry, such as the subdifferential and descent cones.

\section*{2.1 Basics of compressed sensing}

In this chapter we discuss compressed sensing in a general linear algebra setting. We start with a simplified and idealised setting for signal recovery. For the duration of this thesis, we will refer to a signal as a real \( n \)-dimensional vector \( x \in \mathbb{R}^n \). Linear measurements or observations will be represented by a linear map \( y = Ax \), with \( A \in \mathbb{R}^{m \times n} \). In other words, when dealing with signals, we observe data \( y \in \mathbb{R}^m \), this is connected to a signal of interest, \( x \in \mathbb{R}^n \) by the system of equations \( y = Ax \) where \( A \in \mathbb{R}^{m \times n} \) is a matrix with \( m \) rows and \( n \) columns. Such a system is solvable, for example using least squares, as long as \( m \geq n \). If \( m < n \) we have an underdetermined system; if a solution exists, infinitely many solutions exist. Without further information we cannot recover \( x \) from the system \( y = Ax \) when \( m < n \). This is where additional structural information, such as sparsity proves useful, as under certain sparsity assumptions we can reconstruct signals even when \( m < n \), throughout this thesis we refer to recovering the sparse solution \( x \) as the minimization problem succeeding.

Many signals occurring in practical applications from signal and image processing are compressible, which means that they can be well approximated by vectors that are sparse in a certain basis or frame, typically a discrete cosine, wavelet, or curvelet basis. A simple example is when \( f \) is the superposition of two sine functions. It is clearly not sparse as it is, but its Fourier transform is 2-sparse as Fourier uses trigonometric functions as its basis. In the figure 2.1 we have an example where the signal is the superposition of a 50Hz sine wave and a 120Hz sine wave.

Some real world examples include JPEG, MPEG and MP3. JPEG2000, for example, uses a wavelet basis, and stores only the largest coefficients in this basis. The image below, for example, shows the Discrete Cosine Transform applied to \( 8 \times 8 \) blocks.
Figure 2.1: Superposition of two sine waves at 50 Hz and 120 Hz

Figure 2.2: A compressed image and the Discrete Cosine Transform applied to $8 \times 8$ blocks.

For the remainder of this section we will not be concerned with specific signal representations and will assume our signals to be strictly sparse vectors. The crucial difficulty in sparse recovery lies in the fact that we do not know the location of the non-zero entries of the vector $x$; if we did, the matrix $A$ would be easily reducible and we could solve a linear system. The difficulty is one of non-linearity: the set of $s$-sparse vectors is a union of hyperplanes. The fact that adding two $s$ sparse vectors gives a $2s$-sparse vector, however, can be used to derive a first recovery result.

2.1.1 Theorem (Uniqueness of Sparse Solution). [28] Let $A$ be an $m \times n$ matrix with $m \leq n$ and assume that all $m \times m$ minors are non-singular. Assuming the linear system $y = Ax$ has an $s$-sparse solution, then this solution will be unique,
provided that \( 2s \leq m \).

**Proof.** Let \( A \in \mathbb{R}^{m \times n} \), assume there exist two \( s \)-sparse solutions \( x^0, x^1 \) such that \( y = Ax^0, y = Ax^1 \) and set \( z = x^0 - x^1 \). Since both \( x^0 \) and \( x^1 \) are \( s \)-sparse we have that \( z = x^0 - x^1 \) is \( 2s \) - sparse. With this, we have that

\[
Az = A(x^0 - x^1) = y - y = 0. \tag{2.1.1}
\]

Let \( I \) be the set defined by \( I = \{ i : z_i \neq 0 \} \) and define \( A_I \) as the matrix consisting of columns indexed by \( I \), that is, we have \( A_I z_I = 0 \) where \( A_I \) is an \( m \times 2s \) matrix. Since \( A_I z_I = 0 \), it follows that \( z_I \in \ker(A) \), but since \( 2s \leq m \) and \( A \) has rank \( 2s \), it follows that \( z_I = 0 \) and therefore \( z = 0 \) and \( x^0 = x^1 \). \hfill \square

2.1.2 Remark. **The condition on the invertibility of all minors can be relaxed.** The \textit{spark} \( \text{sp}(A) \) of the matrix \( A \) is defined as the minimal size of a linearly dependent set of columns of \( A \). From this definition we also have that every \( s \)-sparse vector \( x \in \mathbb{R}^n \) is the unique sparse solution of \( Az = Ax \) if and only if \( \text{sp}(A) > 2s \), see [28] for a discussion.

Finding the support of the signal (the non-zero values) may require a search among all \( \binom{n}{s} m \times s \) sub-matrices. In fact, the task of finding the support is NP-hard [28, 2.3]. From this it follows that in general, there is no hope for finding a sparse solution efficiently. However, one may ask whether for some matrices, if there could be methods that work considerably better than trial-and-error.

This leads to a few important questions:

1. How do we find suitable matrices \( A \) to allow for reconstruction?
2. How do we reconstruct \( x \) from \( y = Ax \)?
3. How can we create \textit{efficient} reconstruction algorithms?
2.1. BASICS OF COMPRESSED SENSING

Starting with (1), we have a simple example as to why compressive sensing is not fitted for arbitrary matrices \( A \in \mathbb{R}^{m \times n} \).

Let \( A \) be constructed from rows of the identity matrix, then \( y = Ax \) simply takes some entries of \( x \), so would contain mostly zero entries for sparse \( x \), and so no information of the non-zero entries caught in \( y \) would be obtained, so reconstruction is not possible for such a matrix. So compressed sensing is concerned with both the design of the matrix \( A \), and the recovery algorithm. We also want \( A \) to work for all signals \( x \) simultaneously.

In terms of algorithms, we wish to have fast reconstruction algorithms. The exact problem that we want to solve is:

\[
\minimize \|x\|_0 \quad \text{subject to} \quad y = Ax. \tag{2.1.2}
\]

The approach that is the basis of compressed sensing is replacing the \( \|\cdot\|_0 \) function by the 1-norm \( \|x\|_1 \), an approach known as Basis Pursuit [18], as within equation (1.0.2).

At first sight, there is no reason why basis pursuit as in equation (1.0.2) should have anything to do with finding the sparsest vector, and if it does, why it should be more efficient than minimizing \( \|x\|_0 \). Both issues will be addressed in the following.

We begin by addressing the efficiency issue by noting that basis pursuit can be reformulated as a linear programming problem, which can be solved in polynomial time.

---

2.1.1 Efficiency

The problem (1.0.2) is solvable in polynomial time by transforming it into a linear programming problem. To see how this works, introduce new variables \( u_1, u_2, \ldots, u_n \).
and change the problem to

\[
\begin{aligned}
\text{minimize} & \sum_{i=1}^{n} u_i \\
\text{subject to} & -u_i \leq x_i \leq u_i \\
& u_i \geq 0, \quad 1 \leq i \leq n \\
y &= Ax
\end{aligned}
\] (2.1.3)

We can see that this problem is equivalent to Basis Pursuit by noting that at an optimal solution, \( x_i = u_i \) or \( x_i = -u_i \); if not, we can make \( u_i \) smaller until \( u_i = x_i \) if \( x_i \geq 0 \) or \( u_i = -x_i \) if \( x_i < 0 \). From this it follows that at an optimal solution we have

\[
|x_1| + |x_2| + \ldots + |x_n| = u_1 + u_2 + \ldots + u_n,
\] (2.1.4)

and the solution to (1.0.2) is obtained by taking the \( x \) part of the solution of (2.1.3).

### 2.1.2 Equivalence to \( \ell_0 \)-minimization

We next discuss when (1.0.2) solves (2.1.2).

**2.1.3 Definition (BP-exactness).** [39] A matrix \( A \) is called BP-exact for sparsity \( s \) if for all \( y \in \mathbb{R}^m \) such that \( y = Ax \) has a unique sparse solution \( \tilde{x} \) with \( |\text{supp}(\tilde{x})| \leq s \), the basis pursuit problem (1.0.2) has \( \tilde{x} \) as the unique minimum.

Clearly, not every matrix \( A \) is BP-exact, since (1.0.2) is solvable in polynomial time and (2.1.2) is NP-hard. There is a simple way to characterise matrices that are BP-exact.

**2.1.4 Lemma (null space Property).** [28] A matrix \( A \) is BP-exact for sparsity \( s \), if and only if for every \( s \)-sparse vector \( x \) and every vector \( z \in \ker A \) with \( z \neq 0 \),

\[
\|x + z\|_1 > \|x\|_1
\]

holds.

*Proof.* Assume that \( A \) is BP-exact for sparsity \( s \). Let \( x^0 \) be \( s \)-sparse and assume that \( \|x^0 + z\|_1 \leq \|x^0\|_1 \) for some \( z \in \ker A, z \neq 0 \). Then \( A(x^0 + z) = Ax \) and...
∥x^0 + z∥_1 ≤ ∥x^0∥_1, so that ∥x^0∥_1 is not the unique optimal solution to the basis pursuit problem y = Ax, in contradiction to the BP-exactness of A.

Assume now that for every s-sparse x^0 and z ∈ ker A, ∥x^0 + z∥_1 > ∥x^0∥_1 holds. Let x^1 be a solution of the basis pursuit problem y = Ax with y given by y = A x^0. Then ∥x^1∥_1 = ∥x^0 + (x^1 - x^0)∥_1 ≤ ∥x^0∥_1 (since x^1 is a minimizer of this norm) and A(x^1 - x^0) = Ax^1 - Ax^0 = 0, since both x^0 and x^1 are solutions. It follows that x^1 - x^0 ∈ ker A, so that by assumption, x^1 = x^0 has to hold.

2.1.5 Remark. The above lemma can be formulated as the condition that the null space of A, ker A, does not intersect the cone of descent directions of the 1-norm in any point other than the origin. The above property holds more generally for minimizing a convex function with linear constraints [4]. Figure 2.3 illustrates the null space property: in this example, the point on y = Ax with the smallest 1 norm is a 1-sparse vector x^0. Note that it would be highly unlikely for the point with the smallest 2-norm on y = Ax to be 1-sparse.

![Figure 2.3: The null space Property](image)

Producing BP-exact matrices A with small m is still very much an open problem. The best constructions require m to be of order s^2 [12]. Recall that exact sparse reconstruction is possible if m ≥ 2s.

However, it has been found that the use of random matrices can yield good results. An important result is that, with a random matrix A ∈ R^m×n, with entries drawn
from the normal distribution, \( s \)-sparse vectors \( x \) can be reconstructed from \( y = Ax \) with very high probability if we have \( m \geq Cs \log(n/s) \) with \( C > 0 \) a real constant [22, 15].

The following version of this result is from [39] and [7].

2.1.6 Theorem. There exist constants \( C, c_1 \in \mathbb{R} \) such that \( C, c_1 > 0 \) and such that if \( m, n, s \in \mathbb{Z} \) and \( 1 \leq s \leq n/C \) and \( m \geq Cs \log(n/s) \) and if \( A \in \mathbb{R}^{m \times n} \) is a random normal distribution matrix, that is, entries are drawn independently from the standard normal distribution \( N(0,1) \), then we have,

\[
P\{ A \text{ is BP-exact for sparsity } s \} \geq 1 - e^{-c_1m}.
\]

(2.1.5)

This bound tells us that the amount of data needed to recover \( s \)-sparse vectors is proportional to a linear scale of \( s \), and the signal length \( n \) only affects the scale logarithmically. In simple terms, if \( s \) is small compared to \( n \) then the number \( m \) of measurements can also be chosen small in comparison to \( n \), meaning it is plausible that we can acquire exact solutions of underdetermined systems of linear equations, solving this issue in compressive sensing.

The above result is derived by connecting BP-exactness of a matrix \( A \) to another property of \( A \) using the notion of restricted \( \varepsilon \)-almost isometry.

2.1.7 Definition. [39] A matrix \( A \) has the property of \( t \)-restricted \( \varepsilon \)-almost isometry if the corresponding linear map satisfies the condition of \( \varepsilon \)-almost isometry for every \( t \)-sparse \( x \); that is, if

\[
(1 - \varepsilon)\|x\|_2 \leq \|Ax\|_2 \leq (1 + \varepsilon)\|x\|_2,
\]

(2.1.6)

for all \( x \in \mathbb{R}^n \) with \( |\text{supp}(x)| \leq t \).

Using ideas related to the Johnson-Lindenstrauss Theorem [38], one can show that random matrices satisfy a restricted isometry property with high probability. The result on BP-exactness then follows from the following lemma.
2.1. Lemma. There is a constant $\varepsilon > 0$ such that if a matrix $A$ has the property of $3s$-restricted $\varepsilon$-almost isometry, then it is BP-exact for sparsity $s$.

The proof can be found in [39].

2.1.3 Alternative algorithms

The basis pursuit approach is very efficient in sparse recovery, however it is not the only attempt to solve the sparse recovery problem efficiently. Some alternative classes of algorithms are described in [51] or in [28]. Examples of such methods are

- **Orthogonal Matching Pursuit (OMP)** Orthogonal matching pursuit is a method which builds up the support set of the reconstructed sparse vector iteratively by adding an index to the current support at each iteration, the index is chosen to minimize the residual at each iteration.

- **Iterative Hard Thresholding (IHT)** Starting with the zero vector, say $x^0 = 0$, the iterative hard thresholding method iteratively computes

$$x^{n+1} = H_s(x^n + A^*(y - Ax^n)).$$

Where $H_s$ is the hard thresholding operator which keeps the $s$ largest entries of a vector in absolute value and sets the other entries to 0. IHT is a variant of gradient descent with an additional coordinate projection step at every iteration.

Under suitable assumptions these methods can also recover sparse vectors. These algorithms are typically faster than Basis Pursuit, but the recovery guarantees are less clear.
2.1.4 Some experiments

Using Matlab package SGPL [53] we can perform basis pursuit to recover sparse vectors. We generate random \( m \times n \) matrices \( A \) and sparse vectors \( x^0 \) then solve \( y = Ax \) with \( y = Ax^0 \). With fixed \( n \) and fixed sparsity \( s \) the code runs \( m \) increasing from 1 to \( n \) and recreates \( x^0 \) using 1-norm minimization, the vectors that are within a given tolerance (in this case \( 10^{-4} \)) to the original sparse vectors \( x^0 \) are recorded as a success, the following graphs for different fixed \( n \) and \( s \) show the proportion of successes (\( y \)-axis) against the values of \( m \) which run from 1 to \( n \) (\( x \)-axis).

In Figure 2.4 we have \( n = 128 \), on the left \( s = 10 \) and on the right \( s = 20 \). It is clear that for larger values of \( s \), the reconstruction is less likely for lower values of \( m \). For \( s = 10 \) we have probability close to 0 until \( m \approx 30 \) and almost guaranteed reconstruction when \( m \geq 60 \). For \( s = 20 \) we have probability close to 0 until \( m \approx 45 \) and almost guaranteed reconstruction for \( m \geq 85 \). In Figure 2.5 we have \( n = 256 \), on the left \( s = 20 \) and on the right \( s = 40 \). It is clear that for larger values of \( s \), the reconstruction is less likely for lower values of \( m \). For \( s = 20 \) we have probability close to 0 until \( m \approx 70 \) and almost guaranteed reconstruction when \( m \geq 110 \). For \( s = 40 \) we have probability close to 0 until \( m \approx 115 \) and almost guaranteed reconstruction for \( m \geq 155 \). The observed phase transition behaviour is explained in Chapter 3.

![Graphs](image)

Figure 2.4: For \( n = 128 \): left: \( s = 10 \), right: \( s = 20 \), \( m \) = number of measurements, \( p \) = probability
2.2 Compressibility, stability and robustness

In many practical examples we do not have exact sparsity. While we may not often have sparse signals, we often may have compressible or weakly sparse signals. If \( x \in \mathbb{R}^n \) is not sparse, we would like to recover a solution that is close to being sparse, where the error is controlled by its distance to \( s \)-sparse vectors. For \( p > 0 \), the \( \ell_p \)-error of the best \( s \)-term approximation to \( v \in \mathbb{R}^n \) is defined by

\[
\sigma_s(v)_p = \inf_{\|z\|_0 \leq s} \|v - z\|_p.
\] (2.2.1)

2.2.1 Definition (Stable Null Space Property). \([28]\) A matrix \( A \in \mathbb{R}^{m \times n} \) is said to satisfy the stable null space property with constant \( 0 < p < 1 \) relative to a set \( S \subset [n] \) if

\[
\|v_S\|_1 < p\|v_S\|_1, \quad \forall v \in \ker A \setminus \{0\}.
\] (2.2.2)

The matrix \( A \) is said to satisfy the stable null space property of order \( s \) with constant \( p \in (0, 1) \), if it satisfies the stable null space property with constant \( p \in (0, 1) \) relative to any set \( S \subset [n] \) that satisfies \( |S| \leq s \).

2.2.2 Theorem. \([28]\) Suppose a matrix \( A \in \mathbb{R}^{m \times n} \) satisfies the stable null space property of order \( s \) with constant \( p \in (0, 1) \), then for any \( x \in \mathbb{R}^n \), a solution \( \hat{x} \) of

Figure 2.5: For \( n = 256 \): left: \( s = 20 \), right: \( s = 40 \), \( m = \) number of measurements, \( p = \) probability
the $\ell_1$ minimization problem \eqref{eq:1.0.2} with $y = Ax$ approximates the vector $x$ with $\ell_1$-error

$$\|x - \hat{x}\|_1 \leq \frac{2(1 + p)}{(1 - p)} \sigma_s(x). \quad (2.2.3)$$

**2.2.3 Theorem.** [28] The matrix $A \in \mathbb{R}^{m \times n}$ satisfies the stable null space property with constant $p \in (0, 1)$ relative to $S$ if and only if, for all vectors $x^0, x^1 \in \mathbb{R}^n$ with $Ax^0 = Ax^1$, we have

$$\|x^1 - x^0\|_1 \leq \frac{1 + p}{1 - p} \left(\|x^1\|_1 - \|x^0\|_1 + 2\|(x^0)_{\bar{S}}\|_1\right). \quad (2.2.4)$$

As well as issues in regards to stability, we can also discuss the notion of robustness in regards to sparse approximation schemes. If we replace the Basis Pursuit problem as within equation \eqref{eq:1.0.4} with the convex optimization problem,

$$\min_{x \in \mathbb{R}^n} \|x\|_1 \quad \text{subject to} \quad \|y - Ax\| \leq \eta, \quad (2.2.5)$$

then the robustness of the basis pursuit algorithm is guaranteed by the following property.

**2.2.4 Definition.** [28] The matrix $A \in \mathbb{R}^{m \times n}$ is said to satisfy the robust null space property, with respect to $\|\cdot\|$, with constants $0 < p < 1$ and $\tau > 0$ relative to a set $S \subset [n]$ if

$$\|v_S\|_1 < p\|v_{\bar{S}}\|_1 + \tau \|Av\|, \quad \forall v \in \mathbb{R}^n. \quad (2.2.6)$$

The matrix $A$ is said to satisfy the robust null space property with constants $0 < p < 1$ and $\tau > 0$ if it satisfies the robust null space property with constants $p, \tau$ relative to any set $S \subset [n]$ with $|S| \leq s$.

With this we have the theorem.

**2.2.5 Theorem.** [28] Let a matrix $A \in \mathbb{R}^{m \times n}$ satisfy the robust null space property of order $s$, with constants $0 < p < 1$ and $\tau > 0$. Then, an optimal solution to the convex minimization problem \eqref{eq:2.2.5}, with $y = Ax + e$ and $\|e\|_1 \leq \eta$ approximates the vector $x$ with $\ell_1$-error

$$\|x - \hat{x}\|_1 \leq \frac{2(1 + p)}{(1 - p)} \sigma_s(x) + \frac{4\tau}{(1 - p)} \eta. \quad (2.2.7)$$
A different, but equivalent and common practical formulation of the robust $\ell_1$ minimization property is given by

$$\min \|y - Ax\|^2 + \lambda \|x\|_1,$$

for some regularization parameter $\lambda$. This formulation is known in statistics as the LASSO [31].

### 2.3 A general setting

Many of the considerations so far extend beyond the 1-norm to general convex functions. This will be useful later on. Often, the signal of interest can be written as a non-linear combination of a few atoms from a set $\mathcal{A}$ [17]. An example would be a low-rank matrix, that can be represented as superposition of rank one matrices. In such cases it is often possible to recover such a signal from linear observations by minimizing the gauge function,

$$\minimize \|x\|_{\mathcal{A}} \quad \text{subject to} \quad y = Ax,$$

where $\|\cdot\|_{\mathcal{A}}$ is defined as the smallest $\lambda$ such that $x \in \lambda \text{conv}\{\mathcal{A}\}$, the convex hull of the set of atoms. If the set of atoms consists of $\pm e_i$, with $e_i$ the standard basis vectors, then this is just the 1-norm, while if $\mathcal{A} = \{e_i e^T : i = 1, \ldots, n\}$ which consists of rank 1 matrices, then we get the so-called nuclear norm, which is the sum of the singular values of a matrix. It is important to note that the above norm is a convex function, and it is therefore worth looking more closely at general convex optimization problems. Rather than looking at only atomic norms, we will consider general convex functions in the following. We will assume all convex functions $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ to be proper, that is, assume that they take on a finite value at some point.
2.4 Regularization

In this section we discuss the problems that underlie co-sparse analysis and our analysis in general. First we introduce different formulations of the convex regularization problem, then we introduce the problem of convex denoising. In this section we make use of the first-order optimality conditions that were previously discussed, and general properties of cones and their polars, as these help us study descent cones of regularizers.

2.4.1 Definition (Regularization). Let \( f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\} \) be a proper convex function. The exact form of the general constrained convex regularization problem is

\[
\text{minimize } f(x) \quad \text{subject to } y = Ax, \tag{2.4.1}
\]

while the noisy form is

\[
\text{minimize } f(x) \quad \text{subject to } \|y - Ax\|_2 \leq \epsilon \tag{2.4.2}
\]

Where \( \epsilon \) is some error. If we interchange the role of the function \( f \) and the residual, we get the generalised LASSO

\[
\text{minimize } \|y - Ax\| \quad \text{subject to } f(x) \leq \tau. \tag{2.4.3}
\]

Finally, we have the Lagrangian form,

\[
\text{minimize } \|y - Ax\|_2^2 + \lambda f(x). \tag{2.4.4}
\]

The problems (2.4.2), (2.4.3), and (2.4.4) are equivalent. The practical problem is to effectively find the parameters involved.

2.4.2 Proposition. [28] Let \( f : \mathbb{R}^n \to \mathbb{R} \) be a strictly convex function, with \( f(0) = 0 \). Then the three problems (2.4.2), (2.4.3) and (2.4.4) are equivalent:

1. If \( \hat{x} \) is the unique solution of (2.4.2) with \( \epsilon \geq 0 \), then there exists \( \tau \geq 0 \) such that \( \hat{x} \) is the unique solution of (2.4.3).
2. If \( \hat{x} \) is the solution of (2.4.3) with \( \tau > 0 \), then there exists \( \lambda \geq 0 \) such that \( \hat{x} \) is the solution of (2.4.4).

3. If \( \hat{x} \) is the solution of (2.4.4) with \( \lambda > 0 \), then there exists \( \epsilon \geq 0 \) such that \( \hat{x} \) is the solution of (2.4.2).

Proof. 1. Let \( \hat{x} \) be the unique minimizer of (2.4.2) with \( \epsilon \geq 0 \). Set \( \tau := f(\hat{x}) \), then any \( x \neq \hat{x} \) with \( f(x) \leq \tau \) must satisfy \( \|y - Ax\|_2 > \epsilon \geq \|y - A\hat{x}\|_2 \), which shows that \( \hat{x} \) is the unique minimizer of (2.4.3).

2. Let \( \hat{x} \) be the minimizer of (2.4.3) with \( \tau > 0 \), and let

\[
\mathcal{L}(x, \lambda) = \frac{1}{2}\|y - Ax\|^2_2 + \lambda(f(x) - \tau)
\]

(2.4.5)

be the Langrangian of the equivalent problem of minimizing the squared residual. Since we have that \( \tau > 0 \) and \( f \) is continuous and strictly convex, the set \( \{x : f(x) < \tau\} \) is non-empty. This means that strong duality holds, that is

\[
\mathcal{L}(\hat{x}, \lambda^*) = \max_{\lambda \geq 0} \min_x \mathcal{L}(x, \lambda) = \min_x \max_{\lambda \geq 0} \mathcal{L}(x, \lambda).
\]

(2.4.6)

This means that for all \( x \), we have \( \mathcal{L}(\hat{x}, \lambda^*) \leq \mathcal{L}(x, \lambda^*) \), so we have that \( \hat{x} \) is also the minimizer of (2.4.4) with \( \lambda = \lambda^* \).

3. Let \( \hat{x} \) be the minimizer of (2.4.4) with \( \lambda > 0 \). Then we set \( \epsilon := \|y - A\hat{x}\|_2 \). Any \( x \neq \hat{x} \) that satisfies the constraint \( \|y - Ax\|_2 \leq \epsilon \) also has to satisfy \( f(x) > f(\hat{x}) \), as otherwise \( \hat{x} \) could not be a minimizer of (2.4.4). Therefore \( \hat{x} \) is also a minimizer of (2.4.2).

Note that there is symmetry between (2.4.2) and (2.4.3). It is possible to generalise this further, by replacing \( \|y - Ax\| \) by a convex function \( g(x) \), the equivalence of the two statements follows.
2.4.1 Subdifferentials and descent cones

We can consider convex cones through use of the subdifferential [6], first we define a convex cone and its polar. A convex cone \( C \subset \mathbb{R}^n \) is a convex set such that \( C = \tau C, \forall \tau > 0 \). We denote the polar of a cone \( C \) as \( C^\circ \), which is defined as the set of outward normals of \( C \), that is

\[
C^\circ := \{ u \in \mathbb{R}^n : \langle u, x \rangle \leq 0, \forall x \in C \}. \tag{2.4.7}
\]

The polar cone is always closed and convex.

2.4.3 Definition (Subdifferential, Subgradients). [6] Let \( f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\} \) be a proper convex function and \( x \) be a vector in \( \mathbb{R}^n \), we define the subdifferential of \( f \) at the vector \( x^0 \) as follows

\[
\partial f(x^0) := \{ z \in \mathbb{R}^n : f(x^0) + \langle z, x^0 \rangle \leq f(x^0 + z), \forall x \in \mathbb{R}^n \}, \tag{2.4.8}
\]

The elements of \( \partial f(x^0) \) are the subgradients of \( f \) at a point \( x^0 \).

2.4.4 Proposition (First Order Optimality Condition). [6, 5] Let \( f : \mathbb{R}^n \to \mathbb{R} \) be a convex function, and recall (2.4.1), the first order optimality condition is that \( \hat{x} \) is a unique solution if and only if

\[
\exists u \neq 0 : A^T u \in \partial f(\hat{x}). \tag{2.4.9}
\]

The concept of the subdifferential is mostly useful for non-smooth functions.

If we have a function \( f : \mathbb{R}^n \to \mathbb{R} \) which is differentiable at \( \hat{x} \), then the subdifferential simply contains the gradient of \( f \) at \( \hat{x} \), and the vector \( u \) in (2.4.9) consists of the lagrange multipliers [5].

Let \( f : \mathbb{R}^n \to \mathbb{R} \) be a convex function, we can rewrite the optimality condition in terms of the descent cone \( D(f, x^0) \) of \( f \) at \( x^0 \), defined as the convex cone of all
directions in which $f$ decreases from $x^0$,

$$
\mathcal{D}(f, x^0) = \text{cone}\{z : f(x^0 + z) \leq f(x^0)\}. 
$$

(2.4.10)

One can show [32] that if the subdifferential is compact, non-empty and does not contain the origin, then

$$
\mathcal{D}(f, x^0)^\circ = \text{cone}(\partial f(x^0)).
$$

Moreover, an application of a conic form of Farkas’ Lemma [3, Lemma 2.4] states that if we have a linear subspace $L$ and a polyhedral cone $C$, then $L$ intersects $C$ non-trivially if and only if $C^\circ$ intersects the orthogonal complement $L^\perp$ only at the origin. With the descent cone defined above we can rewrite the condition in (2.4.9) as

$$
\ker A \cap \mathcal{D}(f, \hat{x}) = \{0\}.
$$

(2.4.11)

That is, the kernel of $A$ does not intersect the descent cone nontrivially. In the case of the noisy problem (2.4.2), if we have a point $x^0$ that satisfies the constraints, we can introduce the smallest cone-restricted singular value

$$
\sigma_{f, x^0}(A) := \sigma_{\mathcal{D}(f, x^0)}(A) := \min_{x \in \mathcal{D}(f, x^0)} \frac{\|Ax\|_2}{\|x\|_2}.
$$

(2.4.12)

If $\hat{x}$ is the solution of (2.4.2) then we see that the error satisfies

$$
\|x^0 - \hat{x}\| \leq 2\epsilon \cdot \sigma_{f, x^0}(A).
$$

(2.4.13)

We can express subdifferentials with use of the dual norm of a norm $f$. We denote the dual norm of a norm $f$ as $f^\circ$, it is defined for any vector $x \in \mathbb{R}^n$ by

$$
\begin{align*}
    f^\circ(z) &:= \max_{x \in \mathbb{R}^n} (z^T x) \text{ such that } f(x) \leq 1. 
    
\end{align*}
$$

(2.4.14)

The dual norm of $f^\circ$ is $f$ itself, so the formula above holds if we swap the roles of $f$ and $f^\circ$. If $f$ is a norm with a dual norm $f^\circ$, then the subdifferential of $f$ at a point
\( x^0 \) can be written as
\[
\partial f(x^0) = \begin{cases} 
\{ z \in \mathbb{R}^n : f^o(z) = 1, \langle z, x^0 \rangle = f(x^0) \} & \text{if } x^0 \neq 0 \\
\{ z \in \mathbb{R}^n : f^o(z) = 1 \} & \text{if } x^0 \leq 0.
\end{cases}
\]

For the \( \ell_1 \)-norm at an \( s \)-sparse vector \( \hat{x} \), we have
\[
\partial \| \hat{x} \|_1 = \{ z \in \mathbb{R}^n : \| z \|_\infty = 1, \langle z, \hat{x} \rangle = \| \hat{x} \|_1 \} \tag{2.4.15}
\]
We require \( \langle z, \hat{x} \rangle = \| \hat{x} \|_1 \) which is equivalent to \( z_i = \text{sgn}(\hat{x}_i) \) for all entries \( \hat{x}_i \neq 0 \) of the vector \( \hat{x} \). The equation (2.4.15) gives us that the subdifferential of the 1-norm at an \( s \)-sparse vector is a \((n - s - 1)\)-dimensional face of a unit hypercube.

An important class of regularizers are of the form \( f(x) := g(Ax) + h(Bx) \), where \( A \) and \( B \) are linear maps, the subdifferential is easy to compute and is given by
\[
\partial f(x) = A^* \partial g(Ax) + B^* \partial h(Bx) \tag{2.4.16}
\]
As we are often interested in the descent cone rather than the subdifferential, the following identity can be useful.

**2.4.5 Lemma.** [5] Let \( C \subset \mathbb{R}^n \) be a closed convex cone. If \( A \in \mathbb{R}^{m \times n} \), then
\[
A^{-1}(C^o) = (A^* C)^o, \tag{2.4.17}
\]
where \( A^{-1}(C^o) = \{ x \in \mathbb{R}^n \text{ such that } Ax \in C^o \} \) denotes the inverse image of \( C^o \) under \( A \). In particular, for the descent cone of a regularizer of the form \( g(x) = f(Ax) \) we have
\[
\mathcal{D}(g, x^0) = A^{-1} \mathcal{D}(f, Ax^0). \tag{2.4.18}
\]

**Proof.** Let \( x \in A^{-1}(C^o) \) and \( y \in C \). Then \( \langle x, A^* y \rangle = \langle Ax, y \rangle \leq 0 \), as \( Ax \in C^o \). Therefore \( A^{-1}(C^o) \subset (A^* C)^o \).

On the other hand, let \( x \in (A^* C)^o \) and \( y \in C \), then \( \langle Ax, y \rangle = \langle x, A^* y \rangle \leq 0 \), so that \( Ax \in C^o \), therefore \((A^* C)^o \subset A^{-1}(C^o)\). \( \square \)
An important example is that of the $\ell_1$-norm.

**2.4.6 Example (1-norm).** Let $x \in \mathbb{R}^n$, the 1-norm of the vector can be written as

$$\|x\|_1 = \sum_{i=1}^n |\Pi_i(x)|,$$  \hspace{1cm} (2.4.19)

where $\Pi_i(x)$ is the projection on the $i$-th component of $x$.

We can compute the subdifferential and we get

$$\partial \|x\|_1 = \sum_{i=1}^n \Pi_i^T \partial |x_i|$$  \hspace{1cm} (2.4.20)

and the subdifferential of the absolute value is

$$\partial |x| = \begin{cases} \frac{x}{|x|}, & x \neq 0 \\ [-1, 1], & x = 0. \end{cases}$$

This again leads to the description of the subdifferential of the $\ell_1$-norm as a $(n-s-1)$-dimensional face of a unit hypercube, just as in (2.4.15).
Chapter 3

Integral geometry

As seen the previous chapter, solving
\[
\text{minimize } f(x) \text{ subject to } Ax = y
\]
will recover a solution \( x^0 \) of \( Ax = y \) if and only if \( \ker A \cap D(f, x^0) = \{0\} \), where \( D(f, x^0) \) is the cone of descent directions of \( f \) at \( x^0 \). Verifying this condition is not easy, but if \( A \) is random we can say something about the probability of success of finding the sparse solution. The tools used for this come from the field of Integral Geometry or Geometric Probability [49].

Integral Geometry deals with the properties of randomly transformed geometric objects. When dealing with random subspaces and convex cones, the theory is called Conic Integral Geometry. The theory answers questions such as the probability that a randomly oriented convex cone intersects another one:

\[
\mathbb{P}\{C \cap QK \neq \{0\}\}, \quad (3.0.1)
\]

where \( Q \) is a random orthogonal matrix from the Haar probability measure on the orthogonal group and \( C, K \) are convex cones. The probability of intersection can be expressed in terms of invariants of the cones \( C \) and \( K \), the intrinsic volumes.

Consider as an example the problem in two dimensions. Convex cones \( C \) and \( K \) that
are not linear spaces are, in this setting, equivalent to arcs on the circle. It is then easy to verify that
\[ P\{C \cap QK \neq \{0\}\} = v_2(C) + v_2(K), \]
where \( v_2 \) is the proportion of the unit circle covered by the corresponding arc.

In higher dimensions the formulation for intrinsic volumes is much more difficult, since convex cones themselves become much more complicated. We will see that the intersection probability (3.0.1) undergoes a phase transition: it becomes almost 1 if the sum of certain parameters, the statistical dimensions \( \delta(C) \) and \( \delta(K) \), is greater than the ambient dimension, and almost 0 if not. The theory of these phase transitions was developed in [4] and we will recall it here. We will then see how this theory applies in the setting of null spaces of random linear maps and descent cones, which leads to the observed phase transition phenomena in compressed sensing.

### 3.1 Intrinsic volumes

We first define the intrinsic volumes of polyhedral cones. A cone is polyhedral if it can be written as the intersection of a finite number of closed halfspaces. Polyhedral cones are closed and convex. A detailed study of intrinsic volumes of polyhedral cones and their properties, with applications, can be found in [3].

#### 3.1.1 Definition (Intrinsic Volumes: Polyhedral Case)

Let \( C \subset \mathbb{R}^n \) be a polyhedral cone. For each \( k \in \mathbb{N}_0 \), the \( k \)-th conic intrinsic volume \( v_k(C) \) is given by
\[
v_k(C) := P\{\Pi_C(g) \text{ lies in the relative interior of a } k\text{-dimensional face of } C\},
\]
where \( \Pi_C(g) = \arg\min_{x \in C} \|g - x\| \) is the metric projection of \( g \) onto \( C \) and \( g \in \mathbb{R}^n \) is a standard normal random vector.

For a polyhedral cone, the sequence of intrinsic volumes forms a probability distribution on \( \{0, 1, 2, \ldots, n\} \). Important examples are linear subspaces and orthants.
3.1.2 Example (Linear Subspaces). Let \( L_j \) be a \( j \)-dimensional subspace in \( \mathbb{R}^n \). Then \( L_j \) is a polyhedral cone with precisely one face, so the map \( \Pi_{L_j} \) projects every point onto this \( j \)-dimensional face. So we have,

\[
v_k(L_j) = \begin{cases} 
  1, & \text{if } k = j, \\
  0, & \text{if } k \neq j \text{ for } k = 0, 1, 2, \ldots, n.
\end{cases}
\]

3.1.3 Example (The Non-negative Orthant). The non-negative orthant \( \mathbb{R}^n_+ \) is a polyhedral cone. The projection \( \Pi_{\mathbb{R}^n_+}(g) \) lies in the relative interior of a \( k \)-dimensional face of the orthant if and only if exactly \( k \) coordinates of \( g \) are positive. Each coordinate of \( g \) is positive with the probability one half, and negative with probability one half, and the coordinates are independent. So the intrinsic volumes of the orthant are given by

\[
v_k(\mathbb{R}^n_+) = 2^{-n} \binom{n}{k} \text{ for } k = 0, 1, 2, \ldots, n.
\]

(3.1.2)

The intrinsic volumes coincide with the probability density of a Binomial \((n, 1/2)\) random variable.

We only sketch the general definition of intrinsic volumes by means of polyhedral approximation, without going into the details. We can equip the family of closed convex cones in \( \mathbb{R}^n \) with the conic Hausdorff metric to form a compact metric space. The polyhedral cones are dense in this metric space, and the conic intrinsic volumes are continuous with respect to the metric, so we may define the intrinsic volumes of a general closed convex cone by approximation. The conic Hausdorff metric is obtained by identifying each closed convex cone \( C \subset \mathbb{R}^n \) with the spherical convex set \( C \cap S^{n-1} \). Then we use the known construction of the Hausdorff metric on the sphere.

3.1.4 Definition (Intrinsic Volumes: General Case). Let \( C \in \mathbb{R}^n \) be a closed convex cone and let \( \{C_i : i = 1, 2, 3, \ldots, n\} \subset \mathbb{R}^n \) be a sequence of polyhedral cones that converges to \( C \) in the conic Hausdorff metric. For each \( k \in \mathbb{N}_0 \), the \( k \)th conic intrinsic
volume \( v_k(C) \) is given by the limit

\[
v_k(C) := \lim_{i \to +\infty} v_k(C_i),
\]

where the limit does not depend on the approximating sequence.

The conic intrinsic volumes of closed convex cones satisfy a number of important relationships which are discussed here.

1. **Distribution.** The intrinsic volumes describe a probability distribution on \( \{0, 1, 2, \ldots, n\} \), that is

\[
\sum_{k=0}^{n} v_k(C) = 1 \text{ and } v_k(C) \geq 0 \text{ for } k = 0, 1, 2, \ldots, n.
\]

2. **Orthogonal invariance.** For an orthogonal transformation \( Q \in O(n) \),

\[
v_k(QC) = v_k(C);
\]

3. **Polarity.** The intrinsic volumes reverse under polarity, that is

\[
v_k(C) = v_{n-k}(C^o) \text{ for } k = 0, 1, 2, \ldots, n.
\]

4. **Gauss-Bonnet Formula.** When \( C \) is not a subspace we have,

\[
\sum_{k=0, k \text{ even}}^{n} v_k(C) = \sum_{k=0, k \text{ odd}}^{n} v_k(C) = \frac{1}{2}.
\]

5. **Direct Products.** For a product of convex cones \( C \times K \subset \mathbb{R}^n \times \mathbb{R}^{n'} \), we have

\[
v_k(C \times K) = \sum_{i+j=k} v_i(C) \cdot v_j(K) \text{ for } k = 0, 1, 2, \ldots, n + n'.
\]

The intrinsic volumes are the essential ingredients in the (generalized) Steiner formula, which in its original formulation gives an expression for the measure of the neighborhood of a convex cone:

\[
P\{\|\Pi_C(g)\| \geq r\} = \sum_{i=0}^{n} v_i(C) \cdot P\{\chi_i \geq r\},
\]
Figure 3.1: Intrinsic volumes of the cone \( C = \{ x : x_1 \leq \cdots \leq x_n \} \).

where \( \chi_0 = 0 \) and \( \chi_1, \ldots, \chi_n \) are independent chi-distributed random variables with \( \chi_i \) having \( i \) degrees of freedom.

A powerful generalization of the Steiner formula (3.1.8) was derived in [40], which we state for completeness: if \( f : \mathbb{R}^2_+ \to \mathbb{R} \) is a Borel function and \( C \subseteq \mathbb{R}^n \) a closed convex cone, then

\[
\mathbb{E} \left[ f(\|\Pi_C(g)\|_1, \|\Pi_{C^\perp}(g)\|_1) \right] = \sum_{i=0}^n v_i(C) \mathbb{E} \left[ f(\chi_i, \chi'_n) \right],
\]

(3.1.9)

where \( \chi_0 = \chi'_0 = 0 \) and \( \chi_1, \ldots, \chi_n, \chi'_1, \ldots, \chi'_n \) are independent chi-distributed random variables with \( \chi_i \) and \( \chi'_i \) having \( i \) degrees of freedom. What this statement essentially says is that the intrinsic volumes can be characterised as the weights of a mixture of chi distributed random variables.

3.1.1 The statistical dimension

It is sometimes convenient to work with reparametrizations of the intrinsic volumes, the tail and half-tail functionals

\[
t_k(C) = \sum_{i \geq 0} v_{k+i}(C), \quad h_k(C) = 2 \sum_{i \geq 0 \text{ even}} v_{k+i}(C),
\]
Intrinsic Volumes}

which are defined for $0 \leq k \leq n$. Adding (or subtracting) the Gauss-Bonnet relation (3.1.6) to the identity $\sum_{i \geq 0} v_i(C) = 1$, we see that $h_0(C) = h_1(C) = 1$ if $C$ is not a linear subspace, so that the sequences $2v_0(C), 2v_2(C), \ldots$ and $2v_1(C), 2v_3(C), \ldots$ are also probability distributions. We also have the interleaving property

$$t_{i+1}(C) \leq h_i(C) \leq t_i(C).$$

The intrinsic volumes can be recovered from the half-tail functionals as

$$v_i(C) = \begin{cases} \frac{1}{2}(h_i(C) - h_{i+2}(C)) & \text{for } 0 \leq i \leq n - 2, \\ \frac{1}{2}h_i(C) & \text{else.} \end{cases} \quad (3.1.10)$$

An important summary parameter is the \textit{statistical dimension} of a cone $C$, defined as the expected value of the intrinsic volumes considered as probability distribution:

$$\delta(C) = \sum_{k=0}^{n} k v_k(C) = \frac{1}{2} h_1(C) + \sum_{i \geq 2} h_i(C).$$

The statistical dimension coincides with the expected squared norm of the projection of a Gaussian vector on the cone, $\delta(C) = \mathbb{E} \left[ \|C(g)\|^2 \right]$. The statistical dimension is closely related to the Gaussian width $w(C) = \mathbb{E} \left[ \max_{x \in C \cap B_n} \langle x, g \rangle \right]$. It differs from the squared Gaussian width by at most 1,

$$w^2(C) \leq \delta(C) \leq w^2(C) + 1,$$

see [4, Proposition 10.2].

The important thing about the statistical dimension is that it reduces to the usual dimension for linear subspaces, and also extends various properties of the dimension to closed convex cones $C \subseteq \mathbb{R}^n$:

(a) \textbf{Orthogonal invariance.} For an orthogonal transformation $Q \in O(n)$,

$$\delta(QC) = \delta(C);$$
(b) **Complementarity.**

\[ \delta(C) + \delta(C^c) = n; \quad (3.1.11) \]

This generalizes the relation \( \dim L + \dim L^\perp = n \) for a linear subspace \( L \subseteq \mathbb{R}^n \).

(c) **Additivity.**

\[ \delta(C \times K) = \delta(C) + \delta(K). \]

(d) **Monotonicity.**

\[ \delta(C) \leq \delta(K) \text{ if } C \subseteq K. \]

The analogy with linear subspaces will be taken further when discussing concentration of intrinsic volumes, see Section 3.2.

### 3.1.2 The kinematic formulas

The intrinsic volumes allow to study the properties of random intersections of cones through the kinematic formula. A proof of these formulas for polyhedral cones is given in [3, Section 5]. When we say that \( Q \) is drawn uniformly at random from the orthogonal group \( O(n) \), we mean that it is drawn from the Haar probability measure \( \nu \) on \( O(n) \), which is the unique regular Borel measure on \( O(n) \) that is left and right invariant (\( \nu(QA) = \nu(AQ) = \nu(A) \) for \( Q \in O(n) \) and a Borel measurable \( A \subseteq O(n) \)) and satisfies \( \nu(O(n)) = 1 \). Moreover, for measurable \( f: O(n) \to \mathbb{R}^+ \), we write

\[ \mathbb{E}_{Q \in O(n)}[f(Q)] := \int_{Q \in O(n)} f(Q) \nu(dQ) \]

for the integral with respect to the Haar probability measure, and we will occasionally omit the subscript \( Q \in O(n) \), or just write \( Q \) in the subscript, when there is no ambiguity.

#### 3.1.5 Theorem (Kinematic Formula). Let \( C, K \subseteq \mathbb{R}^n \) be polyhedral cones. Then, for \( Q \in O(n) \) uniformly at random, and \( k > 0 \),

\[ \mathbb{E}[v_k(C \cap QK)] = v_{k+n}(C \times K), \quad \mathbb{E}[v_0(C \cap QK)] = t_0(C \times K) = 1. \quad (3.1.12) \]
If \( K = L \) is a linear subspace of dimension \( n - m \), then

\[
\mathbb{E}[v_k(C \cap QL)] = v_{k+m}(C), \quad \mathbb{E}[v_0(C \cap QL)] = \sum_{j=0}^{m} v_j(C).
\] (3.1.13)

Combining Theorem 3.1.5 with the Gauss-Bonnet relation (3.1.6) yields the so-called Crofton formulas, which we formulate in the following corollary.

**3.1.6 Corollary.** Let \( C, K \subseteq \mathbb{R}^n \) be polyhedral cones such that not both of \( C \) and \( K \) are linear subspaces, and let \( L \subseteq \mathbb{R}^n \) be a linear subspace of dimension \( n - m \). Then, for \( Q \in O(n) \) uniformly at random,

\[
\mathbb{P}\{C \cap QK \neq 0\} = h_{n+1}(C \times K), \quad \mathbb{P}\{C \cap QL \neq 0\} = h_{m+1}(C).
\]

Applying the polarity relation \((C \cap K)^\circ = C^\circ + K^\circ\) (see [3, Proposition 2.5]) to the kinematic formulas, we obtain a polar version of the kinematic formula, for \( k > 0 \),

\[
\mathbb{E}[v_{n-k}(C + QK)] = v_{n-k}(C \times K), \quad \mathbb{E}[v_n(C + QK)] = t_n(C \times K). \quad (3.1.14)
\]

A consequence of this polar form is a projection formula for intrinsic volumes, due to Glasauer [30]. Let \( Q \in O(n) \) uniform at random and \( P \in \mathbb{R}^{n \times n} \) a fixed orthogonal projection onto a linear subspace \( L \) of dimension \( m \). Then for \( 0 < k \leq m \),

\[
\mathbb{E}[v_{m-k}(PQC)] = v_{m-k}(C), \quad \mathbb{E}[v_m(PQC)] = t_m(C). \quad (3.1.15)
\]

As we will see in Section 5.2.1, this result holds for any full rank \( T \in \mathbb{R}^{m \times n} \), instead of just for projections \( P \).

### 3.2 Concentration of measure

It was shown in [4] and [40] that the intrinsic volumes concentrate sharply around their mean, the statistical dimension. For a closed convex cone \( C \), let \( X_C \) be a random variable satisfying

\[
\mathbb{P}\{X_C = k\} = v_k(C).
\]
CHAPTER 3. INTEGRAL GEOMETRY

The following is from [40].

3.2.1 Theorem. Let $\lambda \geq 0$. Then
\[
\mathbb{P}\{|X_C - \delta(C)| \geq \lambda\} \leq 2 \exp\left(\frac{-\lambda^2/4}{\min\{\delta(C), \delta(C^o)\}} + \lambda/3\right).
\]
This suggests that $X_C$ is similar to a Gaussian variable with mean $\delta(C)$ and variance $2 \min\{\delta(C), \delta(C^o)\}$. That is, the only intrinsic volumes of a convex cone that are large are the $v_k$ where we have $k$ in the range $\delta(C) \pm\text{const} \times \sqrt{\min\{\delta(C), \delta(C^o)\}}$ so a cone with statistical dimension $\delta(C)$ essentially behaves like a linear subspace of dimension close to $\delta(C)$ [4]. The tail is reminiscent of the Bernstein Inequality for the concentration of sums of independent Bernoulli random variables. Bernstein’s Inequality is similar to Hoeffding’s Inequality. More details on these inequalities and their relations can be found in [11]. The concentration result 3.2.1 can be combined with the kinematic formula into the approximate kinematic formula, which underlies the phase transition results from [4].

3.2.2 Theorem. Let $\eta \in (0, 1)$ and let $C$ and $K$ be convex cones in $\mathbb{R}^n$. Then for $Q \in \mathcal{Q}(n)$ uniformly at random,
\[
\delta(C) + \delta(K) \leq n - a_\eta \sqrt{\delta(K)} \implies \mathbb{P}\{C \cap QK \neq \{0\}\} \leq \eta;
\]
\[
\delta(C) + \delta(K) \geq n + a_\eta \sqrt{\delta(K)} \implies \mathbb{P}\{C \cap QK \neq \{0\}\} \geq 1 - \eta,
\]
with $a_\eta := \sqrt{8 \log(4/\eta)}$. If $K = L$ is a linear subspace of dimension $n - m$, then
\[
\delta(C) \leq m - a_\eta \sqrt{\delta(K)} \implies \mathbb{P}\{C \cap QL \neq \{0\}\} = h_{m+1}(C) \leq \eta;
\]
\[
\delta(C) \geq m + a_\eta \sqrt{\delta(K)} \implies \mathbb{P}\{C \cap QL \neq \{0\}\} = h_{m+1}(C) \geq 1 - \eta.
\]
Applying the above to the statistical dimension, we get the following expression.

3.2.3 Corollary. Let $\eta \in (0, 1)$ and assume that $m \geq \delta(C) + a_\eta \sqrt{m}$, with $a_\eta = 2\sqrt{\log(2/\eta)}$. Let $Q$ be a random orthogonal matrix and $P$ the projection on the first $m$ coordinates, then
\[
\delta(C) - (n - m)\eta \leq \mathbb{E}_Q[\delta(PQC)] \leq \delta(C),
\]
3.2. Phase transitions in convex regularization

We now return to the problem

\[
\begin{align*}
\text{minimize } & f(x) \quad \text{subject to } \ Ax = y.
\end{align*}
\]

(3.2.1)

index $\delta(D(f, x^0))$ - Statistical dimension of $f$ at $x^0$. To make the notation easier, we define $\delta(f, x^0) = \delta(D(f, x^0))$ and call this the statistical dimension of $f$ at $x^0$. Recall that the above problem has $x^0$ as unique solution if and only if $\ker A \cap D(f, x^0) = \{0\}$. If $A$ is an $m \times n$ matrix with random Gaussian entries, then it can be shown that $\ker A$ has the same distribution as $QL$, where $Q$ is a random orthogonal matrix and $L$ is a linear subspace of dimension $n - m$, the dimension of the null space of $A$ if $A$ has full rank $m$. The statistical dimension of $L$ is $n - m$, the dimension of $L$, and we have

\[\delta(f, x^0) + n - m \geq n \iff \delta(f, x^0) \geq m.\]

The Approximate Kinematic Formula, Theorem 3.2.2, can be restated in this context as follows.

3.2.4 Theorem (Phase Transitions in Linear Inverse Problems). [4] Let $\eta \in (0, 1)$ and let $x^0 \in \mathbb{R}^n$ be a fixed vector and $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ a proper convex function. Let $A \in \mathbb{R}^{m \times n}$ have independent standard normal entries, that is, we have random measurement matrices, and $y = Ax^0$, then

\[
m \leq \delta(f, x^0) - a_\eta \sqrt{n - m} \quad \Rightarrow \quad (3.2.1) \text{ recovers sparse solution with probability } \leq \eta
\]

\[
m \geq \delta(f, x^0) + a_\eta \sqrt{n - m} \quad \Rightarrow \quad (3.2.1) \text{ recovers sparse solution with probability } \geq 1 - \eta,
\]

where $a_\eta := \sqrt{8 \log(4/\eta)}$.

A phase transition is a sharp change in the character of a computational problem as its parameters vary [4]. Phase transitions have been observed in experimental work, such as [34], where it is used to quantify the amount of admissible undersampling.
in X-ray tomography. The above result shows the existence of phase transitions and locates them at the statistical dimension, but it does not tell us how to find these locations.

3.3 Computing the statistical dimension

We have seen that the statistical dimension of linear subspaces is just the dimension, and for self-dual cones (cones that look the same as their polars, which includes the orthant and the cone of positive semidefinite matrices), the identity $\delta(C) + \delta(C^\circ) = n$ implies that $\delta(C) = n/2$. Unfortunately, computing the statistical dimension of regularizers $f$ is in general not easy. In this section we present a recipe, due to Stojnic [50] and later refined in [4], that allows to accurately estimate the statistical dimension under certain conditions.

We can compute the statistical dimension of a descent cone with use of the subdifferential. Recall from (2.4.10) that the polar of the descent cone $D(f, x^0)$ is the cone spanned by the subdifferential, $\text{cone}(\partial f(x^0))$. By the decomposition $\|g\|^2 = \|\Pi_C(g)\|^2 + \|\Pi_{C^\circ}(g)\|^2$, we also see that the projected length onto a cone is equal to the distance to the polar cone, $\|\Pi_C(g)\| = \text{dist}(g, C^\circ)$, where $\text{dist}(g, C^\circ) = \min_{x \in C^\circ} \|g - x\|$. We can therefore write the statistical dimension of a cone as

$$\delta(C) = \mathbb{E} [\text{dist}(g, C^\circ)^2]. \quad (3.3.1)$$

Since $\text{dist}(g, \text{cone}(\partial f(x^0))) = \inf_{\tau > 0} \text{dist}(g, \tau \partial f(x^0))$, we have

$$\delta(f, x^0) = \mathbb{E} [\inf_{\tau > 0} \text{dist}(g, \tau \partial f(x^0))^2].$$

The following result from [4] shows that we can get at least a bound for the statistical dimension if we exchange the expectation and the infimum.

3.3.1 Proposition (Statistical Dimension of a Descent Cone). Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be a proper convex function and let $x \in \mathbb{R}^n$. We assume the subdifferential
3.3. COMPUTING THE STATISTICAL DIMENSION

\( \partial f(x) \) is non-empty and compact and that it doesn’t contain the origin, then we define

\[ J(\tau) := J(\tau; \partial f(x)) := \mathbb{E}[\text{dist}^2 (g, \tau \cdot \partial f(x))] \]
for \( \tau \geq 0 \) \hspace{1cm} (3.3.2)

where \( g \) is a standard normal Gaussian vector. Then

\[ \delta(D(f, x)) \leq \inf_{\tau \geq 0} J(\tau). \] \hspace{1cm} (3.3.3)

Furthermore, the function \( J \) is strictly convex, continuous at \( \tau = 0 \), differentiable for \( \tau \geq 0 \), and has a unique minimum.

We now give a complete formulation of how to calculate the statistical dimension of a descent cone, see also [4, Recipe 4.1]. Assume that \( f : \mathbb{R}^n \to \mathbb{R} \) is a proper convex function and \( x \in \mathbb{R}^n \). With some assumptions that the subdifferential \( \partial f(x) \) is nonempty, compact and does not contain the origin, then

1. Identify the subdifferential \( S = \partial f(x) \).

2. For each \( \tau \geq 0 \), compute \( J(\tau) = \mathbb{E}[\text{dist}^2 (g, \tau S)] \).

3. Find the unique solution to the first derivative stationary equation \( J'(\tau) = 0 \).

   There may be no solution.

4. If the stationary equation has a solution \( \tau_0 \), then \( \delta(D(f, x)) \leq J(\tau_0) \).

5. Otherwise the bound is simply \( \delta(D(f, x)) \leq J(0) = n \).

This method suggests a way to study the statistical dimension of a descent cone, you minimize the function \( J \) by setting the derivative to 0. We also have the error estimates for this method, also derived in [4].

3.3.2 Theorem (Error Bound for Descent Cone Method). Let \( f : \mathbb{R}^n \to \mathbb{R} \) be a norm on \( \mathbb{R}^n \) and let \( x \in \mathbb{R}^n \) be a non-zero point then we have the following inequalities

\[ 0 \leq \inf_{\tau \geq 0} J(\tau; \partial f(x)) - \delta(f, x) \leq \frac{2 \sup \{ \|s\|_2 : s \in \partial f(x) \}}{f(x/\|x\|_2)} \] \hspace{1cm} (3.3.4)
3.3.3 Remark. The bound (3.3.4) gives an indication of when the upper bound on the statistical dimension is sharp, namely when the subdifferential is not too large.

3.3.1 Descent cones of the $1$-norm

The following result summarises the method to obtain the statistical dimension of the descent cone of the $1$-norm at a sparse vector, combining this with the previous theorems and we find the location of the phase transition for $1$ regularized inverse problems when we have a large dimension.

3.3.4 Proposition. [4] Let $x \in \mathbb{R}^n$ be a vector with $s$ non-zero entries, then the normalized statistical dimension of the descent cone of the $1$-norm at $x$ satisfies the bounds

$$\psi\left(\frac{s}{n}\right) - \frac{2}{\sqrt{sn}} \leq \frac{\delta(\|\cdot\|_1, x)}{n} \leq \psi\left(\frac{s}{n}\right).$$

The function $\psi : [0, 1] \to [0, 1]$ is defined as

$$\psi(\rho) := \inf_{\tau \geq 0} \left( \rho(1 + \tau^2) + (1 - \rho) \int_{\tau}^{\infty} (u - \tau)^2 \cdot \varphi(u) du \right).$$

The integral kernel $\varphi(u) := \sqrt{2/\pi}e^{-u^2/2}$ is a probability density with support $[0, +\infty)$. We obtain the infimum by finding the unique positive $\tau$ which solves the stationary equation

$$\int_{\tau}^{\infty} \left( \frac{u}{\tau} - 1 \right) \cdot \varphi(u) du = \frac{\rho}{1 - \rho}. \quad (3.3.7)$$

We include a proof for the upper bound, as it is a template for the calculations in the next chapter, for the lower bound see [4].

Proof. We may assume that an $s$-sparse vector $x$ has the form $x = (x_1, x_2, \ldots, x_s, 0, \ldots, 0)$, where $x_i \geq 0$. This is due to the fact that the $1$-norm is invariant under signed permutations. Recall the subdifferential bound

$$\delta(f, x) \leq \inf_{\tau \geq 0} J(\tau). \quad (3.3.8)$$
3.3. COMPUTING THE STATISTICAL DIMENSION

Let \( f(x) = \|x\|_1 \), using equation (3.3.2) we see that

\[
\delta(\|\cdot\|_1, x) \leq \inf_{\tau \geq 0} \mathbb{E}[\text{dist}^2(g, \tau \cdot \partial \|x\|_1)]
\] (3.3.9)

With the ordering assumption, the subdifferential of the 1-norm at \( x \) has the following structure

\[
\partial \|x\|_1 = \{ z \in \mathbb{R}^n : \|z\|_\infty = 1, \langle z, x \rangle = \|x\|_1 \} = \begin{cases} 
\{ z_i = 1, \ i = 1, \ldots, s, \\
\{ z_i \in [-1, 1], \ i = s + 1, \ldots, n. 
\end{cases}
\]

We compute the distance from a random Gaussian vector \( g \) to the dilated subdifferential as follows

\[
\text{dist}^2(g, \tau \cdot \partial \|x\|_1) = \sum_{i=1}^{s} (g_i - \tau)^2 + \sum_{i=s+1}^{n} \text{Pos}^2(\|g_i\| - \tau),
\] (3.3.10)

where \( \text{Pos}(a) := \max(a, 0) \). Taking expectation we get

\[
\mathbb{E}[\text{dist}^2(g, \tau \cdot \partial \|x\|_1)] = s(1 + \tau^2) + (n - s) \sqrt{\frac{2}{\pi}} \int_{\tau}^{\infty} (u - \tau)^2 e^{-\frac{u^2}{2}} du.
\] (3.3.11)

Using equation (3.3.9) and normalising we obtain

\[
\frac{\delta(\|\cdot\|_1, x)}{n} \leq \inf \left\{ (s/n)(1 + \tau^2) + (1 - s/n) \sqrt{\frac{2}{\pi}} \int_{\tau}^{\infty} (u - \tau)^2 e^{-\frac{u^2}{2}} du \right\}.
\] (3.3.12)

Which is the upper bound as desired.

We conclude by noting that while the statistical dimension of the 1-norm descent cone does not have a closed form expression, it can be computed numerically for any parameters \( n, s \). Figure 3.2 shows a plot that shows the success probability of recovering the sparse solution of (1.0.2) for combinations of relative sparsity and relative undersampling, and shows the upper bound computed for the relative statistical dimension of this norm.
Figure 3.2: Phase transitions for 1-norm minimization
Chapter 4

Phase transitions for TV regularization

In this chapter we consider a problem of the form

\[
\text{minimize } \|\Omega x\|_1 \text{ subject to } Ax = y, \tag{4.0.1}
\]

where \( \Omega \in \mathbb{R}^{p \times n} \). The goal is to apply the method from Section 3.3 to the composite regularizer \( \|\Omega \cdot \|_1 \) in the special case when \( \Omega \) maps a vector \( x \) to the vector of its differences. Such a regularizer is designed to promote sparsity in the discrete gradient of a signal. A two-dimensional version of this regularizer is popular in imaging and knows as discrete total variation (TV) regularization [16], and it is common to also refer to the one-dimensional case by this name. While this chapter focusses on one dimensional signals, at the end of the chapter we outline how the methods can be generalized. We begin by looking at the analysis-\( \ell_1 \), or cosparse signal recovery model in general.
4.1 Analysis and synthesis sparsity

In the classical compressed sensing setting, we assume that the signal of interest is sparse in some basis or frame,

\[ x = Bz, \quad z \in \mathbb{R}^q, \]

with \( z \) \( s \)-sparse. For the analysis sparsity model, we assume the signal \( x \) we are trying to recover is sparse after a transformation, that is, \( \Omega x \) is \( s \)-sparse. We call \( \Omega \) an analysis operator, as opposed to the synthesis operator \( B \).

Although the models of sparse recovery (synthesis sparsity) and cosparsity (analysis sparsity) are superficially similar, there are some key differences: in cosparsity the emphasis is on the zero entries of \( \Omega x \) (hence the name), rather than the non-zero entries which is the focus of synthesis sparse recovery.

The recovery of cosparsive vectors using (4.0.1) has been studied extensively, see [35] but also [14, 27, 41, 42, 52].

A typical example is discrete total variation. Here, the analysis operator is a one or two dimensional difference operator. In this section we will cover the one dimensional case. Let \( \Omega = \Omega_{\text{diff}} \in \mathbb{R}^{(n-1) \times n} \) be the following matrix

\[
\Omega = \begin{pmatrix}
-1 & 1 & 0 & \ldots & 0 & 0 \\
0 & -1 & 1 & \ldots & 0 & 0 \\
0 & 0 & -1 & \ldots & 0 & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\
0 & 0 & 0 & \ldots & -1 & 1
\end{pmatrix}
\]

Problem (4.0.1) with the operator (4.1.2) is useful for piecewise constant signals that only make few changes: that is, the vector of differences of adjacent entries is sparse. Before discussing the one-dimensional TV regularization, we briefly mention some results in the more general setting of arbitrary \( \Omega \).
Another important case covered in [35] is where rows $\omega_j$ of $\Omega \in \mathbb{R}^{p \times n}$ form a frame, this is also covered in the literature including [27], that is, there exist constants $0 < a \leq b < \infty$ such that
\[
a \|x\|_2^2 \leq \|\Omega x\|_2^2 = \sum_{j=1}^{p} |\langle \omega_j, x \rangle|^2 \leq b \|x\|_2^2.
\] (4.1.3)

In our finite dimensional case, the constants $a$ and $b$ always exist if the $\omega_j$ span $\mathbb{R}^n$, we often refer to $\Omega$ as a frame, and if $a = b$ then we call $\Omega$ a tight frame.

Frames are more general than bases and allow for stable expansions. They are useful when bases with certain properties do not exist, see for example [10]. Similar to synthesis sparsity, if $p \geq n$ and the columns of $\Omega$ are linearly independent, then a signal $x$ can be uniquely determined by its frame coefficients, we can reconstruct $x$ from $\Omega x$ by using the canonical dual frame of $\Omega$, the elements are given by the columns of the Moore-Penrose inverse, that is
\[
\Omega^\dagger = (\Omega^T \Omega)^{-1} \Omega^T
\] (4.1.4)

and for any $x$ we have $x = \Omega^\dagger (\Omega x)$. The lower bound of the canonical dual frame is $b^{-1}$ and the upper bound is $a^{-1}$.

One should notice that frame bounds are equivalent to condition number bounds on the operator $\Omega$, a topic discussed in more detail in Chapter 5. As we will also see in Chapter 5, the frame bounds (or condition number) of the particular matrix (4.1.2) are not very tight: they differ by a factor of the order of $n$.

Much of the theory of cosparse recovery has been developed along the lines of the standard theory, including stability and robustness, incorporating the condition of the analysis operator implicitly into the analysis.

### 4.2 One dimensional TV norm regularization

Let $f(x) = \|\Omega x\|_1$, with $\Omega$ as defined in (4.1.2). From Chapter 3 we know that in order to find the phase transition curves of (4.0.1), we need to compute the statistical
We also know from (2.4.16) that the subdifferential of a function $f = \|\Omega x\|_1$ is given by

$$\partial f(x^0) = \Omega^T \partial \|\Omega x^0\|_1.$$ 

Using this, and assuming that $y^0 = \Omega x^0$ is $s$-sparse, we have the upper bound

$$\delta(f, x^0) \leq \inf_{\tau > 0} E[\text{dist}^2(g, \tau \Omega^T \partial \|y^0\|_1)].$$ (4.2.1)

How good is this upper bound? Recall that the subdifferential of the 1-norm at a point $x^0$ is given by

$$\partial \|x^0\|_1 = \{z \in \mathbb{R}^n : \|z\|_\infty = 1, \langle z, x^0 \rangle = \|x^0\|_1 \} \subset \{z : \|z\|_\infty \leq 1 \}.$$ 

Moreover, by (3.3.4) we have the bound on the error

$$\frac{2 \max\{\|s\|_2 : s \in \Omega^T \partial \|y^0\|_1\}}{\|y^0\|_1/\|x^0\|_2} \leq \frac{2}{\|y^0\|_1/\|x^0\|_2} \max_{\|x\|_\infty \leq 1} \|\Omega^T x\|.$$ 

Using the norm inequality $\|x\|_2 \leq \sqrt{n} \|x\|_\infty$, we get the bound

$$\max_{\|x\|_\infty \leq 1} \|\Omega^T x\|_2 \leq \sqrt{n} \max_{\|x\|_2 \leq 1} \|\Omega^T x\|_2 = \sqrt{n} \sigma_1(\Omega),$$

where $\sigma_1(\Omega)$ is the largest singular value of $\Omega$. On the other hand, by the norm inequality $\|x\|_2 \leq \|x\|_1$, we have that

$$\frac{\|y^0\|_1}{\|x^0\|_2} = \frac{\|\Omega x^0\|_1}{\|x^0\|_2} \geq \frac{\|\Omega x^0\|_2}{\|x^0\|_2} \geq \sigma_n(\Omega),$$

where $\sigma_n(\Omega)$ is the smallest singular value. The condition number of $\Omega$, $\kappa(\Omega)$, is defined as the largest divided by the smallest singular value. We therefore have the following bound on the error of the approximation,

$$\frac{1}{n} \left| \delta(f, x^0) - \inf_{\tau > 0} E[\text{dist}^2(g, \tau \Omega^T \partial \|y^0\|_1)] \right| \leq 2 \frac{\kappa(\Omega)}{\sqrt{n}}.$$
From this we see that we can guarantee good bounds if the condition number of $\Omega$ is small. The bound can actually be improved when considering that we only need to maximize and minimize over certain subspaces in the definition of the singular values. Unfortunately, in the case of TV minimization it is known that the condition number is $\kappa(\Omega) \geq 2(n+1)/\pi$, which makes the bound useless. Nevertheless, having in mind that the upper bound may be pessimistic, we will attempt to use it to get an upper bound for the statistical dimension of TV minimization.

### 4.2.1 One dimensional finite differences

Let $x \in \mathbb{R}^n$ and let $\Omega \in \mathbb{R}^{(n-1) \times n}$ be the finite difference matrix as in (4.1.2)

So we have

$$\Omega x = (x_2 - x_1, x_3 - x_2, \ldots, x_n - x_{n-1})^T \quad (4.2.2)$$

Let $I \subset [n-1]$ be the support of $y$ with $|I| = s$. For $y_i \neq 0$ define the sign $\text{sgn}(y_i)$.

Assume $y_i \neq 0$ for $i \in I$ and $y_j = 0$ otherwise, and for simplicity also assume that $1, n-1 \notin I$ (that is, the signal starts out flat and ends flat). An example of a typical signal can be seen in figure (4.1).

![Figure 4.1: A signal of length 15 whose gradient is 7-sparse.](image)

If we define the regularizer $f(x) = \|\Omega x\|_1$, then for a fixed $x$, by (2.4.16) and (2.4.15), the subdifferential of $f$ at $x$ is given by

$$\partial f(x) = \{\Omega^T z : z_i \in [-1, 1] \text{ for } i \notin I, z_j = \text{sgn}(y_j) \text{ for } j \in I\}. \quad (4.2.3)$$
We want to compute the expression

$$\inf_{\tau > 0} E[\text{dist}^2(g, \tau \Omega^T \partial \|y\|_1)], \quad (4.2.4)$$

where $g$ is a random standard Gaussian vector, as this provides an upper bound to the statistical dimension. With the case of the 1-norm it was easy to compute the Gaussian distance function, however with the case of finite differences this is much more difficult. The reason is that since the subdifferential of the 1-norm is spanned by a faces of a hypercube, we can split the squared distance of a Gaussian vector to this subdifferential into a sum of the squared distances of one-dimensional Gaussians, and from this get an explicit expression in terms of simple Gaussian integrals. As we will see shortly, the seemingly harmless fact that we have a linear image $\Omega^T$ of this subdifferential makes such a characterization more difficult, if not impossible. We now makes this more precise.

The transpose $\Omega^T$ is given by

$$\Omega^T = \begin{pmatrix}
-1 & 0 & 0 & \ldots & 0 \\
1 & -1 & 0 & \ldots & 0 \\
0 & 1 & -1 & \ldots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \ldots & -1 \\
0 & 0 & 0 & \ldots & 1 \\
\end{pmatrix} \quad (4.2.5)$$

so that for any $z \in \mathbb{R}^{n-1}$ we have

$$\Omega^T z = (-z_1, z_1 - z_2, \ldots, z_{n-2} - z_{n-1}, z_{n-1})^T. \quad (4.2.6)$$

Using the characterisation of the 1-norm as described in equation (4.2.3), we know
that

\[\omega \in \partial f(x) \implies \omega_i \in \Delta_i, \quad \Delta_i = \begin{cases} [-1, 1], & i = 1, n \\ [-2, 2], & i = 1, i \notin I, \\ \text{sgn}(y_{i-1}) - [-1, 1], & i - 1 \in I, i \notin I, \\ [-1, 1] - \text{sgn}(y_i), & i - 1 \notin I, i \in I, \\ \{\text{sgn}(y_{i-1}) - \text{sgn}(y_i)\}, & \{i - 1, i\} \subset I. \end{cases}\]

The problem with this subdifferential, as opposed to the one of the 1-norm, is that the above does not completely characterise the subdifferential; in fact, we have \(s + 1\) dependencies which we will now describe.

Let \(I = \{i_1, \ldots, i_s\}\) with \(i_k \leq i_{k+1}\) for \(1 \leq k \leq s - 1\). For \(1 \leq k \leq s - 1\), define the index sets \(I_k = \{i_k + 1, \ldots, i_{k+1}\}\) and \(I_0 = \{1, \ldots, i_1\}, I_s = \{i_s + 1, \ldots, n\}\).

Denote by \(\omega_I\) the projection of a vector \(\omega\) onto the coordinates indexed by \(I\). The relevant vectors in the subgradient can be characterised using the following \(s + 1\) affine conditions

\[S_0 := \left\{ x \in \mathbb{R}^{i_1} : x_1 \in [-1, 1], x_j \in [-2, 2] \text{ for } 1 < j < i_1, x_{i_1} \in [-1, 1] - \text{sgn}(y_{i_1}), \right\}, \]

\[S_k := \left\{ x \in \mathbb{R}^{i_{k+1} - i_k} : x_{i_k + 1} \in \text{sgn}(y_{i_k}) - [-1, 1], x_j \in [-2, 2] \text{ for } i_k + 1 < j < i_{k+1}, \right\} \]

\[x_{i_k + 1} \in [-1, 1] - \text{sgn}(y_{i_{k+1}}), \quad \sum_{j=i_k+1}^{i_{k+1}} x_i = \text{sgn}(y_{i_k}) - \text{sgn}(y_{i_{k+1}}) \text{ for } 1 \leq k \leq s - 1 \} \].
$S_s := \left\{ x \in \mathbb{R}^{s-i_s} : x_{i_s} + 1 \in [-1, 1] - \text{sgn}(y_s), x_j \in [-2, 2] \text{ for } i_s + 1 < j < s, x_s \in [-1, 1], \sum_{i=i_s+1}^{n} x_i = \text{sgn}(y_{i_s}) \right\}.$

The subdifferential can therefore be characterised concisely as

$$\partial f(x) = \{ \omega \in \mathbb{R}^n : \omega_{I_k} \in S_k, 0 \leq k \leq s \}. \quad (4.2.8)$$

The expected distance to the dilated subdifferential is given as the sum

$$E[ \text{dist}^2(g, \tau \Omega^T \partial\|y\|_1)] = \sum_{i=0}^{s} E[ \text{dist}^2(g_{I_k}, \tau S_i)]. \quad (4.2.9)$$

With the description from (4.2.7), we have two types of terms in the above sum, corresponding to singletons and intervals. We look at all the possible summands.

1. If $i_{k+1} = i_k + 1$ for some $1 \leq k \leq s - 1$, then

$$E[ \text{dist}^2(g_{I_k}, \tau S_k)] = E[(g_{I_k} - \tau c)^2] = \tau^2 c^2 + 1,$$

where $c = \text{sgn}(y_{i_{k+1}}) - \text{sgn}(y_{i_k})$. If $\text{sgn}(y_{i_{k+1}}) \text{sgn}(y_{i_k}) < 0$, that is, there is a sign change, then $c^2 = 4$, otherwise $c = 0$.

2. The sets $S_0$ and $S_s$ are just sections of a rectangle, and from the definitions we have that for $k \in \{0, s\}, E[ \text{dist}^2(g_{I_k}, \tau S_k)]$ does not depend on the signs of $y_1$ and $y_s$. We therefore assume

$$S_0 = \left\{ x \in [-1, 1] \times [-2, 2]^{i_1} \times [0, 2] : \sum_{i=1}^{i_1} x_i = 1 \right\},$$

$$S_s = \left\{ x \in [0, 2] \times [-2, 2]^{s-i_s} \times [-1, 1] : \sum_{i=i_s+1}^{n} x_i = 1 \right\}.$$

3. If $i_k + 1 < i_{k+1}$ and $\text{sgn}(y_{i_{k+1}}) \cdot \text{sgn}(y_{i_k}) > 0$, then

$$S_k = S^+(i_{k+1} - i_k) := \left\{ x \in [0, 2] \times [-2, 2]^{i_{k+1}} \times [-2, 0] : \sum_{j=i_{k+1}}^{i_{k+1}} x_j = 0 \right\}.$$
4.2. ONE DIMENSIONAL TV NORM REGULARIZATION

4. If $i_k + 1 < i_{k+1}$ and $\text{sgn}(y_{i_{k+1}}) \cdot \text{sgn}(y_{i_k}) < 0$, then

$$S_k = S^-(i_{k+1} - i_k) := \{ x \in [0, 2] \times [-2, 2]^{i_{k+1} - i_k - 2} \times [0, 2] : \sum_{j=i_k+1}^{i_{k+1}} x_j = 2 \}.$$ 

Note that the cases $S^-$ and $S^+$ are related. Specifically, we have $S^-(m) = S^+(m) + (0, \ldots, 0, 2)^T$. Using the symmetry of the Gaussian distribution, we get

$$
\mathbb{E}[\text{dist}^2(g, \tau S^-)(m))] = \mathbb{E}[\min_{y \in S^+(m)} \sum_{i=1}^{m} (g_i - \tau y_i)^2 + 4\tau - 4\tau(g_m - \tau y_m)] \\
= \mathbb{E}[\min_{y \in S^+(m)} \sum_{i=1}^{m} (g_i - \tau y_i)^2 + 4\tau^2(1 + y_m)] \\
= \mathbb{E}[\text{dist}^2(g, \tau S^+)] + 4\tau^2.
$$

Apart from the boundary intervals, we therefore only have one type of set to consider: the intersection if the linear hyperplane orthogonal to $(1, \ldots, 1)^T$ with a rectangle.

Note that Case 1 above, where $i_{k+1} = i_k + 1$, corresponds to the expected squared distance to the set $S^+(1)$ or $S^-(1)$.

If we denote by $C$ the number of sign changes in the part of the signal indexed by the support. For $\sigma \in \{+, 0\}$, define

$$J^\sigma(k, \tau) = \mathbb{E}[\text{dist}^2(g, \tau S^\sigma(m))]$$

The function we wish to minimize over $\tau$ is then

$$
\mathbb{E}[\text{dist}^2(g, \tau \Omega^T \partial ||y||_1)] = \tau^2 4C + \sum_{k=1}^{s-1} J^+(i_{k+1} - i_k, \tau) + J^0(i_1, \tau) + J^0(s - i_s, \tau). \quad (4.2.10)
$$

For monotone signals, that is, signals that only change in one direction, the term $\tau^2 4C$ disappears. Before proceeding with the computation of (4.2.10), we present a few small examples to show how the sets $S^+$ look like. These examples illustrate why getting closed form expressions for the expected distances to these sets is not possible.
4.2.2 Examples

To get a better idea of the problem, we look at what these sets look like.

4.2.1 Example. Let $i_{k+1} = i_k + 3$ and $\text{sgn}(y_{i_k}) = 1, \text{sgn}(y_{i_{k+1}}) = 1$. Then the set with these conditions is isometric to

$$S = \{x \in [0, 2] \times [-2, 2] \times [-2, 0], x_1 + x_2 + x_3 = 0\}. \quad (4.2.11)$$

Geometrically this is simply the intersection of a cuboid cut from a cube with a linear hyperplane.

![Figure 4.2: A set $S_k$: a parallelogram with side length $2\sqrt{2}$ and angles $\pi/3$ and $2\pi/3$.](image)

The set $S$ is a parallelogram, denote this $P$, with side lengths $2\sqrt{2}$, angles $\pi/3$ and $2\pi/3$. Using the symmetry of the Gaussian distribution, we can assume that $P$ is embedded in the $x - y$ plane, hence we write

$$E[\text{dist}^2(g, S)] = E[(g_0)^2] + E[\min_{(x,y) \in P} (g_1 - x)^2 + (g_2 - y)^2] = 1 + E[\min_{(x,y) \in P} (g_1 - x)^2 + (g_2 - y)^2]. \quad (4.2.12)$$

Here the typical problem with these sets is clear, they are not separable, that is we can not reduce the expectation to a sum of univariable expectations. Figure (4.3) shows different regions around the set $S$.

If we know which region the random vector $g$ is in, then we can easily determine the distance to $S$, as it is simply the distance to the corresponding face of $S$, in higher dimensions, however, this can be more cumbersome.
4.2. ONE DIMENSIONAL TV NORM REGULARIZATION

One way to deal with this is to first realize that we are dealing with a scaled version \( \tau S \), then realize that the probability that a Gaussian vector has very large magnitude is small. So with high probability, we can restrict to vectors within the circle shown in figure (4.4). Inside this circle, the distance to \( S \) is the distance to the cone generated by \( S \).

4.2.2 Example. We can also illustrate the resulting set in one dimension higher, we get a three-dimensional parallelepiped, it is possible that this is a general form for these sets, that they can be assembled recursively from lower dimensional ones, this may be stated as a conjecture for future work.
4.2.3 Example. Let $i_{k+1} = i_k + 3$ and $\text{sgn}(y_k) = 1$, $\text{sgn}(y_{k+1}) = -1$. Then the set with these conditions is isometric to

$$S = \{x \in [0, 2] \times [-2, 2] \times [-2, 0], x_1 + x_2 + x_3 = 2\}. \quad (4.2.13)$$

Geometrically, this is the intersection of a rectangle cut out from a box with a linear hyperplane, which in this case is a triangle with side length $2\sqrt{2}$.

![Figure 4.5: A set $S_k$: an equilateral triangle with side length $2\sqrt{2}$.](image)

4.2.3 Computing the expected distance

As computing the expression $J^+(m, \tau)$ explicitly is not feasible, we determine these numerically. The problem is to minimize the distance of a vector $g$ to a set of the form

$$S = B \cap L, \quad (4.2.14)$$

where $B$ and $L$ are as follows

$$B = [a_1, b_1] \times \cdots \times [a_m, b_m], \quad L = \{x \in \mathbb{R}^m : \sum_{i=1}^{k} x_i = 0\}. \quad (4.2.15)$$

The goal is then compute the expected value of the squared distance of $g$ to $S$ when $g$ is Gaussian. The optimization problem is

$$\text{minimize} \|g - x\|^2 \quad \text{subject to} \quad x \in S.$$
If $D$ is a matrix whose columns span $L$, for example, $D$ could be the transpose of the difference matrix (4.1.2) in $\mathbb{R}^{(m-1)\times m}$, then the problem becomes

$$\text{minimize} \|g - Dy\|^2 \quad \text{subject to} \quad Dy \in B. \quad (4.2.16)$$

In this way, the constraint of lying in $L$ is already incorporated into the problem and we only have linear inequalities as constraints. A problem like (4.2.16) is a typical quadratic programming problem, and there are many ways of solving it. One approach that appears to work well with our example is the Frank-Wolfe algorithm [33, 19], also known as conditional gradient. This algorithm aims to solve a problem of the form

$$\text{minimize} f(x) \quad \text{subject to} \quad x \in S$$

with $S$ convex, and (one version of it) works as follows:

1. Start with an initial guess $x^0$;

2. For each $k \geq 0$, solve the linear programming problem

$$\text{minimize} \langle \nabla f(x^k), s \rangle \quad \text{subject to} \quad s \in S$$

and set $d^k = s - x^k$;

3. If $-\langle \nabla f(x^k), d^k \rangle < \delta$ for some predefined tolerance, return $x^k$;

4. Compute a step length $\gamma^k$ and set

$$x^{k+1} = x^k + \gamma^k d^k$$

The algorithm solves our quadratic optimization problem by a sequence of linear optimization problems. In our case, we compute the step length explicitly using line search as with gradient descent, but in general other rules are typically used. For a detailed convergence analysis of this algorithm, see [33, 19].
4.2.4 Remark. A potentially alternative approach get the squared distance to $S$ is to use alternating projections. This approach would start with an initial value, and the project it alternatingly to $L$ and $B$. This approach is fast, as each of the projections can be computed quickly. For example, the projection onto the box $B$ can be computed as

$$\Pi_B(x) = \min(\max(a, x), b),$$

where $a, b$ denote the vectors of left and right boundaries. This algorithm converges to a point in $B \cap L$ [20] but not necessarily to the closest point. It converges to a closest point when both $B$ and $L$ are linear subspaces.

As we have seen, the bound on the statistical dimension depends on the number of sign changes in the signal. We now have to figure out what values of $C$ to expect for a typical signal. For this, we assume the the support $I$ of the difference vector and the signs of the entries are chosen uniformly at random, and look at the expected value of $C$. Note that we looks at sign changes within the support.

4.2.5 Lemma. The expected number of sign changes satisfies

$$E[C] = (s - 1)/2.$$  

Proof. The support of a difference vector $y \in \mathbb{R}^{n-1}$ is a subset $I \subset \{2, \ldots, n-2\}$, that is, we choose $s$ out of $n - 3$ vectors for the location. For each choice of $I$, we choose a random sign vector of length $s$. Because each support $I$ is equally likely and the distribution of the sign changes within each $I$ is the same, we can restrict to a fixed $I$. The probability of having exactly $k$ sign changes is

$$\mathbb{P}\{C = k\} = \binom{s-1}{k} \frac{1}{2^{s-1}}.$$  

This is a binomial distribution, and the expected value is $E[C] = (s - 1)/2$.  

Figure 4.6 shows the phase transition curve for one dimensional TV regularization with $n = 100$ when not only the measurement matrix but also the signal is chosen at
random. To generate the curve, we minimized the expression (4.2.10) over $\tau$ using the precomputed values. The empirical phase transition curve was computed by solving the minimization problem (4.0.1) repeatedly using random measurements and random signals.

![Figure 4.6: Phase transition curve for 1-D TV regularization.](image)

In the phase transition experiment in figure 4.6 we used random signals in addition to a random measurement matrix, but the phase transition curve changes if you consider monotone signals (for example, where all the changes are up-changes) or if every change in the signal goes in the opposite direction as the previous change.
(maximally alternating).

4.3 2D Total Variation and Graphs

The approach for the 1-D TV regularization can be generalized. Let $G = (V, E)$ be a directed graph. That is, $V = \{1, \ldots, n\}$ is a set of vertices and $E \subset V \times V$ is a subset of directed edges. We can define a function

$$\Omega: \mathbb{R}^n \to \mathbb{R}^{|E|}, \quad x \mapsto \Omega x,$$

where $(\Omega x)_e = x_j - x_i$ if $e = (i, j) \in E$ is an edge. Define then

$$f(x) = \|\Omega x\|_1.$$  

The 1-D TV regularizer is the special case when the graph $G$ is a linear graph. For 2-D TV regularization we consider a two-dimensional lattice, where the vertices correspond to coordinates $x_{ij}$, $1 \leq i, j \leq n$, and we have edges $((i + 1, j), (i, j))$ and $((i, j), (i, j + 1))$. This corresponds to differences

$$x_{i+1,j} - x_{ij}, \quad x_{i,j+1} - x_{ij}.$$  

Note that the resulting regularizer is just a special case of what is generally known under TV regularization [16].

Assume that $y^0 = \Omega x^0$ is $s$-sparse with support $I \subset E$. The support can now be seen as characterizing active edges in the graph. Using the characterisation of the 1-norm as described in equation (4.2.3), we can get a characterisation of the subdifferential just as with 1-D TV. One added complication is that now each vertex can be adjacent to many other vertices rather than just one. For example, consider the 2D total variation case and assume we have pixels $x_{i,j}$, $x_{i,j-1}$ and $x_{i-1,j}$ that all have the same value. It follows that the corresponding edges $((i, j-1), (i, j))$ and $((i-1, j), (i, j))$ are not in the support. The $(i,j)$-th coordinate of an element in the subdifferential will
then come from the set $[-2, 2] + [-2, 2]$. The dependencies between the coordinates of different parts of the subdifferential correspond to paths between “monochromatic” paths between vertices that are part of active edges. These are paths where all but the first and last edge are inactive. From this characterization, we get a procedure that associates to any graph an upper bound on the statistical dimension of the associated subdifferential, even though the computations will get considerably more involved.
Chapter 5

Condition number bounds

So far we have looked at estimating the statistical dimension of a composite regularizer $f(\Omega x)$ directly. In this chapter we study bounds on the statistical dimension in terms of the conditioning of $\Omega$. These bounds involve Renegar’s condition number of a linear map with respect to a convex cone. General information on condition numbers can be found in [13]. The results in this chapter are from [5] and are based on previous work by Amelunxen and Lotz [1].

The condition number of a numerical computation problem measures the sensitivity of the output to small perturbations in the input. The classical condition number of a matrix $A \in \mathbb{R}^{m \times n}$, which measures the sensitivity of matrix inversion or linear system solving to perturbations, is equal to the ratio of the operator norm and the smallest singular value. With the notation $\|A\| := \max_{x \in S^{n-1}} \|Ax\|$, $\sigma(A) := \min_{x \in S^{n-1}} \|Ax\|$, where $S^{n-1}$ is the unit sphere in $n$ dimensions, this condition number is

$$
\kappa(A) = \min \left\{ \frac{\|A\|}{\sigma(A)}, \frac{\|A^T\|}{\sigma(A^T)} \right\}.
$$

Renegar’s condition number is the generalization when replacing the source and target vector spaces $\mathbb{R}^n$ and $\mathbb{R}^m$ with convex cones. Let $C \subseteq \mathbb{R}^n$, $K \subseteq \mathbb{R}^m$ be closed convex cones, and let $A \in \mathbb{R}^{m \times n}$. Define restricted versions of the norm and the
singular value:
\[
\|A\|_{C \to K} := \max_{x \in C \cap S^{n-1}} \|\Pi_K(Ax)\|, \quad \sigma_{C \to K}(A) := \min_{x \in C \cap S^{n-1}} \|\Pi_K(Ax)\|,
\]
where \(\Pi_K\) denotes the orthogonal projection onto the cone \(K\). In what follows we will only concerned with the case \(K = \mathbb{R}^m\), but the symmetric formulation does have conceptual advantages.

A generalized form of Renegar’s condition number is defined as
\[
\mathcal{R}_C(A) := \min \left\{ \frac{\|A\|}{\sigma_C \to K(A)}, \frac{\|A^T\|}{\sigma_K \to C(-A^T)} \right\}.
\]
Renegar’s condition number appears in error bounds to solutions of (2.4.2): if \(x^0\) is a feasible point, meaning it satisfies the constraint \(\|y - Ax^0\|_2 \leq \epsilon\) and if \(\hat{x}\) is a solution of (2.4.2), then \(\|\hat{x} - x^0\| \leq 2\epsilon / \mathcal{R}_D(f,x^0)(A) / \|A\| [5]\). Renegar’s condition number was originally introduced to study the complexity of linear programming [46], see [54] for an analysis of the running time of an interior-point method for the convex feasibility problem in terms of this condition number, and [13] for a discussion and references.

In [47], Renegar’s condition number is used to study restart schemes for algorithms such as NESTA [9] in the context of compressed sensing.

The first result linking the statistical dimension with Renegar’s condition is Theorem 5.0.1, due to Amelunxen and Lotz [1].

5.0.1 **Theorem.** Let \(C \subseteq \mathbb{R}^n\) be a closed convex cone, and \(\delta(C)\) the statistical dimension of \(C\). Then for \(A \in \mathbb{R}^{p \times n}\),
\[
\delta(AC) \leq \mathcal{R}_C(A)^2 \cdot \delta(C),
\]
where \(\mathcal{R}_C(A)\) is Renegar’s condition number associated to the matrix \(A\) and the cone \(C\).

For \(p = n\), and if \(A\) has full rank, then we also have a lower bound in terms of the standard condition number, that is
\[
\frac{\delta(C)}{\kappa(A)^2} \leq \delta(AC) \leq \kappa(A)^2 \cdot \delta(C),
\]
where $\kappa(A)$ denotes the matrix condition number of $A$.

Unfortunately, when using Renegar’s condition number we can only obtain an upper bound in (5.0.3), so we do not have an analogous result for the lower bound as we do for the standard condition number as within (5.0.4).

5.0.2 Lemma. Let $C \subseteq \mathbb{R}^n$ be a closed convex cone with nonempty interior, and let $A \in \mathbb{R}^{p \times n}$, when $p \geq n$ we have

$$\mathcal{R}_C(A) \leq \kappa(A).$$ \hspace{1cm} (5.0.5)

Proof. The proof is given in [1].

5.0.3 Example. This example illustrates the role of the cone and why considering Renegar’s condition number can be beneficial. Consider the finite difference matrix

$$\Omega = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ 0 & 0 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \end{pmatrix}.$$  

The condition number is known to be of order $\Omega(n)$, making condition bounds using the normal matrix condition number useless. Using Renegar’s condition number with respect to a cone, on the other hand, can improve the situation dramatically. Consider, for example, the cone

$$C = \{ x \in \mathbb{R}^n : x_1 \geq 0, \ x_i x_{i+1} \leq 0 \text{ for } 1 \leq i < n \}.$$  

This cone is the orthant consisting of vectors with alternating signs. The cone-restricted singular value of $\Omega$ is given by

$$\sigma_C(\Omega)^2 = \min_{x \in C \cap S^{n-1}} \|\Omega x\|^2$$

$$= \min_{x \in C \cap S^{n-1}} \sum_{i=1}^{n-1} (x_{i+1} - x_i)^2 + x_n^2 = \min_{x \in C \cap S^{n-1}} 2 - x_1^2 - \sum_{i=1}^{n-1} 2x_i x_{i+1} \geq 1.$$
Using the same expression for $\|\Omega x\|^2$, we see that the square of the operator norm is bounded by 4, so that the square of Renegar’s condition number with respect to this cone is bounded by 2. If, on the other hand, $C$ is the non-negative orthant, then Renegar’s condition number coincides with the normal matrix condition number. Intuitively, Renegar’s condition number gives an improvement if the cone $C$ captures a portion of the ellipsoid defined by $\Omega \Omega^T$ that is not too eccentric. Other examples when Renegar’s condition number gives significant improvements is for small cones (such as the cone of increasing sequences) or cones contained in linear subspaces of small dimension (such as subdifferential cones of the 1 or $\infty$ norms).

The proof of Theorem 5.0.1 is given in [5, Section 3].

Theorem 5.0.1 translates into a bound for the statistical dimension of convex regularizers by observing that if $f(x) = g(\Omega x)$ with invertible $\Omega$, then (see Chapter 2) the descent cone of $f$ at $x^0$ is given by $D(f, x^0) = \Omega^{-1} D(g, \Omega x^0)$.

5.0.4 Corollary. Let $f(x) = g(\Omega x)$, where $g$ is a proper convex function. Then for invertible $\Omega \in \mathbb{R}^{n \times n}$,

$$\delta(f, x^0) \leq \mathcal{R}_{D(g, \Omega x^0)}(\Omega^{-1}) \cdot \delta(g, \Omega x^0).$$

In particular,

$$\frac{\delta(g, \Omega x^0)}{\kappa(\Omega)^2} \leq \delta(f, x^0) \leq \kappa(\Omega)^2 \cdot \delta(g, \Omega x^0).$$

In this chapter we study two ways in which this bound can be improved: duality and random projections.

5.1 Duality

Using polarity (3.1.11), we get the following version of the bound that ensures that the right-hand side is always bounded by $n$.  

5.1.1 Proposition. Let \( C \subseteq \mathbb{R}^n \) be a closed convex cone, and \( \delta(C) \) the statistical dimension of \( C \). Let \( A \in \mathbb{R}^{n \times n} \) be non-singular. Then

\[
\delta(AC) \leq \kappa(A)^{-2} \cdot \delta(C) + (1 - \kappa(A)^{-2}) \cdot n.
\]

In particular, if \( f(x) = g(\Omega x) \), where \( g \) is a proper convex function and \( \Omega \in \mathbb{R}^{n \times n} \) is non-singular, then

\[
\delta(f, x^0) \leq \kappa(\Omega)^{-2} \cdot \delta(g, \Omega x^0) + (1 - \kappa(\Omega)^{-2}) \cdot n. \tag{5.1.1}
\]

Proof. We have

\[
\delta(AC) \overset{(1)}{=} n - \delta(A^{-T}C^o) \\
\overset{(2)}{\leq} n - \kappa(A)^{-2} \cdot \delta(C^o) \\
\overset{(3)}{=} n - \kappa(A)^{-2} (n - \delta(C)) \\
= \kappa(A)^{-2} \cdot \delta(C) + (1 - \kappa(A)^{-2}) \cdot n,
\]

where for (1) we used (3.1.11) and Lemma 2.4.5, for (2) we used Theorem 5.0.1, and for (3) we used (3.1.11) again. \( \square \)

We interpret the upper bounds in Proposition 5.1.1 as interpolating between the statistical dimension of \( C \) and the ambient dimension \( n \). Proposition 5.1.1 has implications on the recovery thresholds for 1-analysis minimization. The following is a consequence of the special structure of the descent cone of the 1-norm. In the statement, we use the notation \( A_J \) for the submatrix of a matrix \( A \) with columns indexed by \( J \subset [n] = \{1, \ldots, n\} \), and denote by \( \overline{J} = [n] \setminus J \) the complement of \( J \). Note that this notation is local to this chapter.

5.1.2 Proposition. Let \( \Omega \in \mathbb{R}^{p \times n} \) be of full rank with \( p \geq n \) and \( A \in \mathbb{R}^{m \times n} \) with \( m \leq n \). Consider the problem

\[
\text{minimize } \|\Omega x\|_1 \quad \text{subject to } Ax = b. \tag{5.1.2}
\]
Let $x^0$ be such that $Ax^0 = b$, and such that $y^0 = \Omega x^0$ is $s$-sparse with support $J \subset [p]$. Let $M \in \mathbb{R}^{n \times (p-s+1)}$ be a matrix whose first $p - s$ columns consist of the columns of $\Omega^T$ that are indexed by $J$, and the last column is $m_{p-s+1} = \frac{1}{\sqrt{s}} \sum_{j \in J} \text{sgn}(y^0_j) d_j$, where the vectors $d_j$ denote the columns of $\Omega^T$. Then

$$
\delta(f, x^0) \leq \kappa(M)^{-2} \cdot \delta(||x||_1, \Omega x^0) + \left(1 - (p/n)\kappa(M)^{-2}\right) \cdot n \quad (5.1.3)
$$

In particular, given $\eta \in (0, 1)$, Problem (5.1.2) with Gaussian measurement matrix succeeds at recovering $x^0$ with probability $1 - \eta$ if

$$
m \geq \kappa(M)^{-2} \cdot \delta(||x||_1, \Omega x^0) + \left(1 - (p/n)\kappa(M)^{-2}\right) \cdot n + a_n \sqrt{n},
$$

where $a_n$ as defined in Theorem 3.2.2.

Proof. Set $f(x) = ||\Omega x||_1$ with $\Omega \in \mathbb{R}^{p \times n}$ and $p \geq n$. Let $x^0$ be given such that $y^0 = \Omega x^0$ is $s$-sparse with support $J$. Assuming that $\Omega$ has rank $n$ and $x^0 \neq 0$, $y^0$ has at most $n - 1$ zero entries, and the support therefore satisfies $s \geq p - n + 1$. As shown in Chapter 2, the descent cone $\mathcal{D}(f, x^0)$ is polar to the subdifferential cone $\text{cone}(\partial f(x^0))$. Moreover, the statistical dimension satisfies $\delta(C) + \delta(C^\circ) = n$, so that

$$
\delta(\mathcal{D}(f, x^0)) = n - \delta(\text{cone}(\partial f(x^0))) = n - \delta(\Omega^T \text{cone}(\partial ||y^0||_1)).
$$

Recall the subdifferential of the 1-norm (2.4.15) and denote by $C := \text{cone}(\partial ||y^0||_1)$ the cone generated by this subdifferential.

![Figure 5.1: Cone spanned by $(p - s)$-face of $d$-dimensional hypercube](image)

It follows that the cone generated by this subdifferential is contained in a subspace $L$ of dimension $\dim L = p - s + 1 \leq n$. An orthonormal basis of this subspace is given...
by the rows of a matrix $B = [b_1, \ldots, b_{p-s+1}]$, where for $1 \leq j \leq p-s$, the $b_j$ are the unit vectors $e_j$ for $j \in J$ and $b_{p-s+1} = \frac{1}{\sqrt{s}} \sum_{j \in J} \text{sgn}(y^0_j) e_j$. A moment’s thought shows that $C = B\hat{C}$, where $\hat{C} \subset \mathbb{R}^{p-s+1}$ is the cone in $\mathbb{R}^{p-s+1}$ spanned by vectors of the form $\pm e_i + \sqrt{s}e_{p-s+1}$ for $1 \leq j \leq n - p$ (see Figure 5.1). By the orthogonal invariance and the embedding invariance of the statistical dimension (see Properties (a) and (c) in Section 3.1.1), we get $\delta(C) = \delta(\hat{C})$. With this setup, we have

$$\Omega^T C = \Omega^T B\hat{C} = M\hat{C}. $$

Applying the bounds from Theorem 5.0.1 we thus get

$$\delta(D(f, x^0)) = n - \delta(\Omega^T C)$$

$$= n - \delta(M\hat{C})$$

$$\leq n - \kappa^{-2}(M)\delta(\hat{C})$$

$$= n - \kappa^{-2}(M)\delta(C)$$

$$= \kappa(M)^{-2} \cdot \delta(\|\cdot\|_1, \Omega x^0) + (1 - (p/n)\kappa(M)^{-2}) \cdot n,$$

as was to be shown, the last equality comes from the combining of Theorem 5.0.1 and equation (5.1.3).

5.1.3 Example. An illustrative example is the finite difference matrix $\Omega$ of example 5.0.3. The regularizer $f(x) = \|\Omega x\|_1$ is a one-dimensional version of a total variation regularizer, and is used to promote gradient sparsity. The standard method from Chapter 4 for computing the statistical dimension of the descent cone of $f$ works but is cumbersome. The standard condition number bound Theorem 5.0.1 is also not applicable, as it is known that the condition number satisfies $\kappa(\Omega) \geq \frac{2(n+1)}{\pi}$. Figure 5.2 plots the upper bound of Proposition 5.1.2 for signals with random support location and sparsity ranging from 1 to 200, and compares it to the actual statistical dimension computed by Monte Carlo simulation. As can be seen in this example, the upper bound is not very useful because of the large condition numbers involved.

5.1.4 Remark. It is natural to ask for which dictionaries $\Omega$ Proposition 5.1.2 gives good bounds. This clearly also depends on the support of the signal one wishes to
Figure 5.2: The statistical dimension of $\| \Omega \|_1$ for different sparsity levels and the upper bound (5.1.3). The signal has length 200. Note the different scaling on the right panel, which shows that the upper bound is, while being smaller than 1, not very tight. Note the different scale in the two graphs: the upper bound is very close to a horizontal line if superimposed in the first figure.

recover. A closer look at the matrix $C$ in the case of the finite difference matrix and for monotonely increasing signals shows that $C$ is (up to rows of zeros) itself a finite difference matrix of order $n - s + 1$, and the quality of the bounds increases with the size of the support. Another natural example is when $\Omega \in \mathbb{R}^{p \times n}$ is a Gaussian random matrix (that is, a matrix whose entries are independent standard normal distributed random variables). In this case, the invariance properties of Gaussians imply that the matrix $C$ is again a Gaussian matrix in $\mathbb{R}^{n \times p-s+1}$. For such matrices, the condition number is known to be of order $(\sqrt{n} + \sqrt{p-s+1})/((\sqrt{n} - \sqrt{p-s+1})$ with high probability, see for example [55, Theorem 5.32]. In this example we see again that if the support is large, $s \approx p$, then the condition number is close to 1 and the bound becomes useful.
5.2 Randomized preconditioning

While Corollary 5.1.1 ensures that the upper bound does not become completely trivial, when Ω is ill-conditioned it still does not give satisfactory results. The main contribution of this chapter is an improvement of the condition bounds using randomization: using methods from conic integral geometry, we derive a “preconditioned” version of Theorem 5.0.1. The idea is based on the philosophy that a randomly oriented convex cone $C$ ought to behave roughly like a linear subspace of dimension $\delta(C)$. In that sense, the statistical dimension of a cone $C$ should be approximately invariant under projecting $C$ to a subspace of dimension close to $\delta(C)$.

In fact, in Section 5.2.1 we will see that for $n \geq m \geq \delta(C)$, we have

$$E_Q[\delta(PQC)] \approx \delta(C),$$

where $P$ is the projection on the the first $m$ coordinates and where the expectation is with respect to a random orthogonal matrix $Q$, distributed according to the normalized Haar measure on the orthogonal group. From this it follows that the condition bounds should ideally depend not on the conditioning of $\Omega$ itself, but on a generic projection of $\Omega$ to linear subspaces of dimension of order $\delta(C)$. For $m \leq n$ define

$$\mathcal{R}_m^2(A) := E_Q[\kappa(PQA)^2].$$

For simplicity, in the following we consider the case $p = n$ (square matrices).

5.2.1 Theorem. Let $C \subseteq \mathbb{R}^n$ be a closed convex cone and $A \in \mathbb{R}^{n \times n}$. Let $\eta \in (0,1)$ and assume that $m \geq \delta(C) + 2\sqrt{\log(2/\eta)m}$. Then

$$\delta(AC) \leq E_Q[\mathcal{R}_C(PQA)^2] \cdot \delta(C) + (n - m)\eta.$$

For the matrix condition number,

$$\delta(AC) \leq \mathcal{R}_m^2(A) \cdot \delta(C) + (n - m)\eta. \quad (5.2.1)$$
As a consequence of Theorem 5.2.1 we get the following preconditioned version of the previous bounds.

**5.2.2 Corollary.** If \( f(x) = g(\Omega x) \), where \( g \) is a proper convex function and \( \Omega \in \mathbb{R}^{n \times n} \) is non-singular. Let \( \eta \in (0,1) \) and assume that \( m \geq \delta(g, \Omega x^0) + 2\sqrt{\log(2/\eta)m} \). Then

\[
\delta(f, x^0) \leq \mathbb{E}_Q \left[ R_{D(g, Dx^0)}(PD^{-1})^2 \right] \cdot \delta(g, \Omega x^0) + (n - m)\eta \tag{5.2.2}
\]

and

\[
\delta(f, x^0) \leq \mathbb{E}_m(\Omega^{-1}) \cdot \delta(g, \Omega x^0) + (n - m)\eta.
\]

**5.2.3 Example.** Using the preconditioned bounds we can determine the optimal \( m \) that minimizes the right-hand side of (5.2.2). For example, if \( C = D(g, \Omega x^0) \) is a cone in \( \mathbb{R}^{400} \) with \( \delta(C) = 20 \), then the following plot shows the best upper bound in (5.2.1) for various values of \( m \geq \delta(C) \).

Figure 5.3: The condition number of \( PD \) reduces when projecting to a lower dimension \( m \). However, the error bound in (5.2.1) limits how close to \( \delta(C) \) the target dimension \( m \) can become.

**5.2.1 The TQC Lemma**

The following generalization of the projection formulas (3.1.15), first observed by Mike McCoy and Joel Tropp, may at first sight look surprising. We include a
5.2.4 Lemma. Let $T \in \mathbb{R}^{m \times n}$ be of full rank $m \leq n$. Then for $0 \leq k < m$,

$$
\mathbb{E}[v_k(TQC)] = v_k(C), \quad \mathbb{E}[v_m(TQC)] = t_m(C)
$$

(5.2.3)

Proof. In view of (3.1.10), it suffices to show (5.2.3) for the half-tail functionals $h_j$ instead of the intrinsic volumes $v_j$. Let $L \subset \mathbb{R}^n$ be a linear subspace of dimension $\dim L = k \leq m$. From Proposition [3, Proposition 2.2] it follows that

$$
QC \cap T^{-1}L \neq \{0\} \iff TQC \cap L \neq \{0\} \quad \text{or} \quad \ker T \cap QC \neq \{0\},
$$

where in this case, as before, $T^{-1}L$ denotes the pre-image of $L$ under $T$. Denoting by $P$ the orthogonal projection onto the complement $(\ker T)^\perp$, we thus get

$$
PQC \cap (T^{-1}L \cap (\ker T)^\perp) \neq \{0\} \iff TQC \cap L \neq \{0\},
$$

and taking probabilities,

$$
P\{PQC \cap (T^{-1}L \cap (\ker T)^\perp) \neq \{0\}\} = P\{TQC \cap L \neq \{0\}\}. \quad (5.2.4)
$$

To compute the probability on the left, let $Q_0$ be a random orthogonal transformation of the space $(\ker T)^\perp$. Restricting to $(\ker T)^\perp$ as ambient space,

$$
P_Q\{PQC \cap (T^{-1}L \cap (\ker T)^\perp) \neq \{0\}\} = P_Q\{PQC \cap Q_0(T^{-1}L \cap (\ker T)^\perp) \neq \{0\}\}
$$

$$
= \mathbb{E}_{Q_0} P_Q\{PQC \cap Q_0(T^{-1}L \cap (\ker T)^\perp) \neq \{0\}\}
$$

$$
\overset{(1)}{=} \mathbb{E}_{Q} \mathbb{E}_{Q_0}\{PQC \cap Q_0(T^{-1}L \cap (\ker T)^\perp) \neq \{0\}\}
$$

$$
\overset{(2)}{=} \mathbb{E}_{Q}[h_{m-k+1}(PQC)]
$$

where for (1) we summoned Fubini on the representation of the probability as expectation of an indicator variable and for (2) the Crofton formula of Corollary 3.1.6 with $(\ker T)^\perp$ as ambient space, and assuming that $L$ is in general position relative to $\ker T$. A similar argument on the right-hand side of (5.2.4) shows that

$$
P_Q\{TQC \cap L \neq \{0\}\} = \mathbb{E}_{Q}[h_{m-k+1}(TQC)].
5.2. RANDOMIZED PRECONDITIONING

In summary, we have shown that $E_Q[h_{m-k+1}(TQC)] = E_Q[h_{m-k+1}(PQC)]$ for $0 \leq k \leq m$, and hence also $E_Q[v_i(TQC)] = E_Q[v_i(PQC)]$ for $0 \leq i \leq m$. The claim now follows by applying the projection formula (3.1.15).

As with the case where $T$ is a projection, applying the above to the statistical dimension, we get the following expression from Corollary 3.2.3.

5.2.5 Corollary. Let $\eta \in (0,1)$ and assume that $m \geq \delta(C) + \eta \sqrt{m}$, with $\eta = 2\sqrt{\log(2/\eta)}$. Then under the conditions of Lemma 5.2.4, we have

$$\delta(C) - (n - m)\eta \leq E_Q[\delta(TQC)] \leq \delta(C).$$

Proof of Theorem 5.2.1. The upper bound follows from

$$\delta(AC) \leq E_Q[\delta(PQAC)] + (n - m)\eta \leq E_Q \left[ R_C(PQA)^2 \right] \delta(C) + (n - m)\eta,$$

where we used Theorem 5.0.1 for the second inequality. For the upper bound in terms for the matrix condition number we refer to [5, Theorem 5.2].

5.2.2 On the probabilistic reduction

The condition bounds in Theorem 5.2.1 naturally lead to the question of how to compute or bound the condition number of a random projection of a matrix,

$$\kappa(P_mQA) \text{ or } R_C(P_mQA)$$

where $Q \in O(n)$ is a random orthogonal matrix. If $m = \lfloor \rho n \rfloor$ with $\rho \in (0,1)$, then in many cases the condition number $\kappa(P_mQA)$ remains bounded with high probability as $n \to \infty$. Below we sketch how such condition numbers can be bounded.

In what follows, let $A \in \mathbb{R}^{n \times n}$ be fixed and non-singular, and we write $Q_m = P_mQ$ for a random matrix with orthogonal rows, uniformly distributed on the Stiefel manifold. We first reduce to the case of Gaussian matrices, for which tools are readily available. If $G \sim N(0, 1)$ is an $m \times n$ random matrix with Gaussian entries,
then \( Q_m = (GG^T)^{-1/2}G \) is uniformly distributed on the Stiefel manifold, so that 
\( R_C(Q_mA) \) has the same distribution as 
\( R_C((GG^T)^{-1/2}GA) \). We can bound (with probability one)

\[
R_C ((GG^T)^{-1/2}GA) \leq \kappa ((GG^T)^{-1/2}) R_C (GA) = \kappa(G) R_C (GA),
\]

transforming the problem into one in which the orthogonal matrix is replaced with a Gaussian one. There are different ways to estimate such condition numbers, the approach taken here is based on Gordon’s inequality. We restrict the analysis to the classical matrix condition number, a more refined analysis using Renegar’s condition number is likely to incorporate the Gaussian width of the cone. Moreover, using the invariance of the condition number under transposition, we consider \( \kappa(AG) \) with a \( n \times m \) matrix \( G \), \( m \leq n \). An alternative, suggested by Armin Eftekhari, would be to appeal to the Hanson-Wright inequality [48, 26], or more directly, the Bernstein inequality.

5.2.6 Proposition. Let \( A \in \mathbb{R}^{n \times n} \) and \( G \in \mathbb{R}^{n \times m} \), with \( m \leq n \). Then

\[
E[\kappa(AG)] \leq \frac{\|A\|_F + \sqrt{m}\|A\|_2}{\|A\|_F - \sqrt{m}\|A\|_2}.
\]

(5.2.5)

Using a standard procedure one can show that the singular value and the norm will stay close to their expected values with high probability.

The proof uses Gordon’s Lemma and following [28, Section 9.2] and is detailed in [5].

5.2.7 Theorem. (Gordon) Let \( X_{i,j} \) and \( Y_{i,j} \) be finite families of Gaussian random variables with mean zero. If

\[
E|X_{i,j} - X_{k,\ell}|^2 \leq E|Y_{i,j} - Y_{k,\ell}|^2 \text{ for } i \neq k, j, \ell
\]

\[
E|X_{i,j} - X_{i,\ell}|^2 \geq E|Y_{i,j} - Y_{i,\ell}|^2 \text{ for } i, j, \ell,
\]

then

\[
E \min_{i \in [n]} \max_{j \in [m]} X_{i,j} \geq E \min_{i \in [n]} \max_{j \in [m]} Y_{i,j}.
\]
5.2. RANDOMIZED PRECONDITIONING

Proof. We will derive the inequalities

\[ \|A\|_F - \sqrt{m} \|A\|_2 \leq \mathbb{E}[\sigma(AG)] \leq \mathbb{E}[\|AG\|_2] \leq \|A\|_F + \sqrt{m} \|A\|_2. \]

where \(\sigma\) denotes the smallest singular value. We will restrict to showing the lower bound, the upper bound follows similarly by using Slepian’s inequality. Without lack of generality assume \(A = \Sigma\) is diagonal, with entries \(\sigma_1 \geq \cdots \geq \sigma_n\) on the diagonal, and assume \(\sigma_1 = 1\). Define the Gaussian processes

\[ X_{x,y} = \langle Gx, \Sigma y \rangle, \quad Y_{x,y} = \langle g, x \rangle + \langle h, \Sigma y \rangle, \]

indexed by \(x \in S^{m-1}, y \in S^{n-1}\), with \(g \in \mathbb{R}^m\) and \(h \in \mathbb{R}^n\) Gaussian vectors with i.i.d. standard Gaussian coefficients, and so that \(G, g\) and \(h\) are independent. We get

\[ \mathbb{E}[(X_{x,y} - X_{x',y'})^2] = \|\Sigma y\|^2 + \|\Sigma y'\|^2 - 2\langle x, x'\rangle \langle \Sigma y, \Sigma y'\rangle, \]

\[ \mathbb{E}[(Y_{x,y} - Y_{x',y'})^2] = \|\Sigma y\|^2 + \|\Sigma y'\|^2 + 2 - 2\langle x, x'\rangle - 2\langle \Sigma y, \Sigma y'\rangle, \]

so that

\[ \mathbb{E}[(Y_{x,y} - Y_{x',y'})^2] - \mathbb{E}[(X_{x,y} - X_{x',y'})^2] = 2(1 - \langle x, x'\rangle)(1 - \langle \Sigma y, \Sigma y'\rangle) \geq 0. \]

This expression is 0 if \(x = x'\), and non-negative otherwise, since by assumption \(\Sigma\) has largest entry equal to 1. We can therefore apply Gordon’s Theorem 5.2.7 to infer an inequality

\[ \mathbb{E}[\sigma(\Sigma G)] = \mathbb{E}[ \min_{x \in S^{m-1}} \max_{y \in S^{n-1}} \langle Gx, \Sigma y \rangle ] = \mathbb{E}[ \min_{x \in S^{m-1}} \max_{y \in S^{n-1}} X_{x,y} ] \]

\[ \geq \mathbb{E}[ \min_{x \in S^{m-1}} \max_{y \in S^{n-1}} Y_{x,y} ] = \|\Sigma\|_F - \sqrt{m}. \]

In general, if \(\sigma_1 \neq 1\), we replace \(\Sigma\) by \(\Sigma / \|\Sigma\|_2 = \Sigma / \|A\|_2\), and obtain the desired bound.

It would be interesting to characterize those matrices \(A\) for which \(\kappa(P_mQA) \approx 1\) using a kind of restricted isometry property, as for example in [44]. We leave a detailed discussion of the probability distribution of \(\kappa(P_mQA)\) and its ramifications for another occasion, and instead consider a special case.
5.2.8 Example. Consider again the matrix $\Omega$ from Example 5.0.3. For $\rho \in \{0.2, 0.4, 0.6, 0.8\}$ and $n$ ranging from 1 to 400, $m = \lfloor \rho n \rfloor$, we plot the average condition number $\kappa(\Omega G)$, where $G \in \mathbb{R}^{n \times m}$ is a Gaussian random matrix. As $n$ increases, this condition number appears to converge to a constant value.

Figure 5.4: Condition number $\kappa(P_m \Omega G)$ for the matrix $\Omega$ from Example 5.0.3. $P_m$ is the projection to the first $m = \lfloor \rho n \rfloor$ coordinates. The oscillations are an artefact of the simulation for which we currently do not have an explanation.

As we saw in Example 5.0.3, the operator norm is bounded by $\|\sigma\|_\infty \leq 2$. The Frobenius norm, on the other hand, is easily seen to be $\|\Omega\|_F = \|\sigma\|_2 = \sqrt{2n - 1}$. Setting $m = \rho n$, the condition number thus concentrates on a value bounded by

$$\frac{\sqrt{2n - 1} + 2\sqrt{m}}{\sqrt{2n - 1} - 2\sqrt{m}} \approx 1 + \frac{\sqrt{2}\rho}{1 - \sqrt{2}\rho},$$

which is sensible if $\rho < 1/2$.

A note on distributions

The results presented are based in integral geometry, and as such depend crucially on $Q$ being uniformly distributed in the orthogonal group with the Haar measure. By known universality results [45], the results are likely to carry over to other distributions. In the context of this thesis, however, we are not interested in actually
preconditioning the matrices involved. The randomization here is only a tool to improve bounds based on the condition number, and the question of whether this is a “realistic” distribution is of no concern.
Chapter 6

Conclusions

We started with the theoretical problem of recovering structured solutions from underdetermined systems of equations using convex optimization, where the underlying structure of interest is sparsity, and later where the underlying structures are not necessarily sparsity, and where the system of equations is given by \( f(x) = g(Dx) \), for which we require convex regularization.

The main results in this thesis contribute to understanding the role of conditioning of \( D \) for the success of solving such problems, primarily the following:

1. Sharp analytical bounds on the statistical dimension of \( \|Dx\|_1 \) in the case where \( D \) is a difference matrix, denoted \( \Omega \) (total variation minimization in 1-D);

2. Bounds on the statistical dimension of \( g(Dx) \) in terms of that of \( g(x) \) using the condition number of \( D \).
6.1 TV regularization in two dimensions

Further consideration is for the TV-norm regularization case, where we could further explore the conjecture that a general form of the sets of the subdifferential of the TV-norm in higher dimensions can be assembled recursively from lower dimensional sets, and form a parallelepiped.

6.2 Conditioning of projected matrices

Based on the work in Chapter 5 it would be interesting to study the conditioning of randomly projected fixed matrices $A$ more systematically. Interesting is the case of Renegar’s condition number,

$$R_C(PQA),$$

where $Q$ is a random orthogonal matrix and $P$ a projection to the first $m$ rows. What is the limiting behaviour of the distribution of this quantity? The bounds provided in Chapter 5 are just a beginning.

6.3 Computational methods and sampling

In this thesis, attempts were made to compute the statistical dimension $\delta(f, x^0)$ of a convex regularizer. Using the definition $\delta(C) = \mathbb{E}[\|\Pi_C(g)\|^2]$, it is tempting to compute the statistical dimension by generating random vectors $g$, projecting them onto the cone $C$, and determining the length. The average over many such trials will be an estimate of the statistical dimension. Alternatively, in the case of polyhedral cones, one can try to just determine the dimension of the face on which $g$ is projected and average over this dimension. If a cone is given as the conic hull of finitely many vectors $v_1, \ldots, v_p$, collected as columns of a matrix $M$, then a projection can be
computed by non-negative least squares:

$$\minimize_{x \geq 0} \| Mx - g \|_2.$$  

The standard method, the algorithm by Lawson and Hanson [37], looks similar to orthogonal matching pursuit with additional provisions for ensuring non-negativity. One direction is to work out these relations, and to study efficient implementations that would allow to find the dimension of a projection without having to compute the full projection.

As we saw at the beginning of Chapter 3, computing the intrinsic volumes is the same as estimating the coefficients of a mixture of \(\chi\)-distributed random variables. This approach to computing the intrinsic volumes has been implemented in the R package \texttt{conivol} by Dennis Amelunxen [2]. Potential applications include isotonic regression, where the results of [4] have been used and the intrinsic volumes of certain polyhedral cones play a role.
Index

A - Measurement operator, Gaussian random matrix, 11

$A^T$ - Transpose of the matrix $A$, 32

$A^*$ - Conjugate transpose of a matrix $A$, 25

$A_S$ - Submatrix with rows indexed by a set $S \subset [n]$, 15

$B$ - Synthesis operator, 52

$C$ - Convex cone, 32

$C^\circ$ - Polar of cone $C$, 32

$C_i$ - Sequence of polyhedral cones, 39

$D$ - Matrix regularizer, 13, 84

$E$ - Set of edges, 66

$G = (V,E)$ - Directed graph, 66

$H_s$ - Thresholding operator, 25

$I$ - Index set, 57

$J$, 47, 59, 62

$K$ - Convex cone, 36

$L$ - Linear subspace, 36

$L_j^\perp$ - Orthogonal complement of $L$, 33

$L_j$ - $j$ dimensional subspace of $L$, 38

$M$ - Matrix, 86

$N(0,1)$ - Standard Gaussian distribution, 15

$P(\cdot)$ - Projection on the first $m$ coordinates, 85

$Q$ - Random orthogonal matrix, 85

$S \subset [n]$ - Index set $S$, 15

$S_0$, 58

$S_k$, 58

$S_s$, 58

$V$ - Set of vertices, 66

$[n] = \{1, \ldots, n\}$, 15

$E[\cdot]$ - Expectation of an event, 15

$\Omega$ - Finite difference matrix, 12, 51, 52

$\Omega^\dagger$ - Moore-Penrose inverse of $\Omega$, 53

$\Pi_i(x)$ - Projection on the $i$-th component of $x$, 35

$\mathbb{P}$ - Probability of an event, 15

$\mathbb{R}_n^+$ - Non-negative orthant, 38

$\mathcal{R}$ - Renegar’s condition number, 85

$\chi_i$ - Independent chi-distributed random variables, 40

$\delta(C)$ - Statistical dimension of a cone $C$, 41, 44

$\delta(f,x^0)$ - Statistical dimension of $f$ at $x^0$, 12, 45, 54

$\gamma^k$ - Step length, 63
\(\kappa(D)\) - Condition number of matrix \(D\), 14
\(\lambda\) - Regularization parameter, 11, 29
\(\mathcal{D}(f, x)\) - Descent cone of a convex function \(f\) at \(x\), 33
\(\mathcal{L}(x, \lambda)\) - Lagrangian problem, 31
\(P\) - Orthogonal projection, 43
\(\|\cdot\|_A\) - Smallest \(\lambda\) such that \(x \in \lambda \text{conv}\{A\}\), 29
\(\omega_j\) - Rows of \(\Omega\), 52
\(\partial f(x)\) - Subdifferential of \(f\) at \(x\), 32
\(\sigma_s(v)_p\) - \(f_p\)-error of the best \(s\)-term approximation to \(v\), 27
\(\sigma_{f,x^0}(A)\) - Smallest cone restricted singular value, 33
\(Q \in O(n)\) - Orthogonal transform, 41
\(f\) - Convex function, 12, 84
\(f^\circ\) - Dual norm, 33
\(g\) - Convex function, 5, 12, 31, 84
\(g\) - Random standard Gaussian vector, 37, 46, 47, 56, 58, 86
\(h_k\) - half-tail functional, 41
\(s\)-sparse - At most \(s\) non-zero entries, 11
\(sp(A)\) - Spark of a matrix \(A\), the minimal size of a linearly dependent set of columns of \(A\), 20
\(t_k\) - tail functional, 41
\(v\) - Kernel vector, 27
\(v_2(C)\) - Proportion of the unit circle covered by the corresponding arc of \(C\), 37
\(v_k(C)\) - \(k\)-th intrinsic volume, 37
\(x\) - Data vector, 11
\(x^0\) - Vector, 11
\(x_S\) - Subvector indexed by a set \(S \subset [n]\), 15
\(y\) - Observation vector, 11
\(\text{sgn}(x_i)\) - Sign of \(x_i\), 34
\(\text{supp}(x)\) - Support set of a vector \(x\), 22
TV norm - Total variation norm, 13
Bibliography


