Linear Algebra and its Applications 504 (2016) 406–432

Contents lists available at ScienceDirect

Linear Algebra and its Applications

www.elsevier.com/locate/laa

A geometrical stability condition for compressed sensing

Axel Flinth

Institut für Mathematik, Technische Universität Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany

A R T I C L E   I N F O

Article history:
Received 28 October 2015
Accepted 11 April 2016
Available online 19 April 2016
Submitted by H. Rauhut

MSC:
primary 52A20, 90C25
secondary 94A12

Keywords:
Compressed sensing
Convex geometry
Grassmannian condition number
Sparse recovery

A B S T R A C T

During the last decade, the paradigm of compressed sensing has gained significant importance in the signal processing community. While the original idea was to utilize sparsity assumptions to design powerful recovery algorithms of vectors $x \in \mathbb{R}^d$, the concept has been extended to cover many other types of problems. A noteable example is low-rank matrix recovery. Many methods used for recovery rely on solving convex programs.

A particularly nice trait of compressed sensing is its geometrical intuition. In recent papers, a classical optimality condition has been used together with tools from convex geometry and probability theory to prove beautiful results concerning the recovery of signals from Gaussian measurements. In this paper, we aim to formulate a geometrical condition for stability and robustness, i.e. for the recovery of approximately structured signals from noisy measurements.

We will investigate the connection between the new condition with the notion of *restricted singular values*, classical stability and robustness conditions in compressed sensing, and also to important geometrical concepts from complexity theory. We will also prove the maybe somewhat surprising fact that for many convex programs, exact recovery of a signal x_0 imme-

E-mail address: flinth@math.tu-berlin.de.

http://dx.doi.org/10.1016/j.laa.2016.04.017

diately implies some stability and robustness when recovering signals close to x_0.

1. Introduction

Suppose that we are given linear measurements $b \in \mathbb{R}^m$ of a signal $x_0 \in \mathbb{R}^d$, i.e. $b = Ax_0$ for some matrix $A \in \mathbb{R}^{m,d}$, and are asked to recover the signal from them. If $d > m$, this will not be trivial, since the map $x_0 \mapsto b$ in that case won't be injective. If, however, one assumes that x_0 in some sense is *sparse*, e.g., that many of x_0's entries vanish, we can still recover the signal, e.g. with the help of ℓ_1-minimization [8]:

$$\min \|x\|_1 \text{ subject to } Ax = b \qquad (\mathcal{P}_1)$$

This is the philosophy of *compressed sensing*, an area of mathematics which has achieved major attention over the last decade. It has become a standard technique to choose A at random, and then to ask the question how large the number of measurements m has to be in order for $(\mathcal{P}_1)$ to be successful with high probability. A popular assumption is that A has the Gaussian distribution, i.e., that the entries are i.i.d. standard normally distributed.

A widely used criterion to ensure that $(\mathcal{P}_1)$ is successful is the *RIP-property*. Put a bit informally, a matrix A is said to possess the RIP-property if its RIP-constants;

$$\delta_k = \min\left\{\delta > 0 : \forall x \text{ k-sparse } : (1-\delta)\|x\|_2^2 \leq \|Ax\|_2^2 \leq (1+\delta)\|x\|_2^2\right\},$$

are small.

The idea of using convex programs like $(\mathcal{P}_1)$ to recover structured signals has come to be used in a much wider sense than the one above. Some examples of structure assumptions that have been considered in the literature are dictionary sparsity [10], block sparsity [20], sparsity with prior information [13,16,17], saturated vectors (i.e. vectors with $|x(i)| = \|x\|_\infty$ for many i) [14] and low-rank assumptions for matrix completion [7]. Although these problems may seem very different at first sight, they can all be solved with the help of a convex program of the form

$$\min f(x) \text{ subject to } Ax = b, \qquad (\mathcal{P}_f)$$

where f is some convex function defined on an appropriate space. In all of the mentioned examples above, f is chosen to be a norm, but this is not per se necessary.

The connection between the different convex program approaches was thoroughly investigated in [9], in which the very general case of f being an *atomic norm* was in-