Hauptseminar Uncertainty Relation

Prof. M. Wolf
Dr. David Reeb

„Uncertainty Relations in Signal Recovery II"

Dennis Elbrächter

13th June 2013

1. Introduction

We will show that a function cannot be highly concentrated on the time as well as on the frequency domain (using the L_1 norm). This will enable us to see how signals can be perfectly reconstructed from a noise-distorted measurement. In the second part we will see some of the results from the recent field of compressed sensing, and look at the basic underlying ideas.

2. Notation

2.1 Fourier transform

Let $f: \mathbb{R} \longrightarrow \mathbb{C}$, we define its Fourier transform as

$\hat{f}(w) \coloneqq \int_{\mathbb{R}} e^{-2\pi i w t} f(t)\, dt$

2.2 L_1-Norm and supremum norm

Let $f \in L_1$

$\|f\|_1 := \int |f(t)|\, dt$ is the L_1-Norm of f

$\|f\|_\infty := \text{esssup}_{t \epsilon \mathbb{R}} |f(t)|$ is the supremum norm of f

2.3 ε-concentration (w.r.t the L_1-norm)

Let $f \in L_1$

f is ε-concentrated to T, if $\|f - P_T f\|_1 \leq \varepsilon \|f\|_1$, where P_T

is the time-limiting operator for L_1

2.4 Bandlimited functions

$B_1(W) :=\{ f \in L_1 \mid \text{supp}(\hat{f}) = W\}$ is the set of functions bandlimited to a Borel-set $W \subseteq \mathbb{R}$.

3. L_1 Uncertainty Principle

Theorem Let $f \in L_1$ be ε-concentrated to T and bandlimited to W, then $|W||T| \geq 1 - \varepsilon$

Proof Let $f \in B_1(W)$. We write f as reverse Fourier-transform of its Fourier-transform:

$$f(t) = \int_W e^{2\pi iwt} \hat{f}(w)\, dw = \int_W e^{2\pi iwt} \int_{\mathbb{R}} e^{-2\pi iws} f(s)\, ds\ dw$$

We use Fubini's theorem to swap the order of integration, and pull f(s) out of the inner integral since it does not depend on w:

$$= \int_W \int_{\mathbb{R}} e^{2\pi iw(t-s)} f(s)\, ds\ dw = \int_{\mathbb{R}} f(s) \int_W e^{2\pi iw(t-s)}\ dw\, ds$$

We look at the absolute of f(t) and use the fact that $\left|e^{2\pi ix}\right| = 1$ for every $x \in \mathbb{R}$:

$$|f(t)| \leq \int_{\mathbb{R}} |f(s)| \int_W 1\ dw\, ds = |W| \int_{\mathbb{R}} |f(s)|\ ds$$

Therefore: $\|f\|_\infty \leq |W|\, \|f\|_1$

Together with

$$\|P_T f\|_1 = \int_T |f(s)|\, ds \leq \int_T \sup_{s\in\mathbb{R}} |f(s)|\ ds = \|f\|_\infty |T|$$

We get: $$\frac{\|P_T f\|_1}{\|f\|_1} \leq |T||W| \qquad (1)$$

Using the assumption that f is ε-concentrated gives us:

$$\|f\|_1 = \|f - P_T f + P_T f\|_1 \leq \|f - P_T f\|_1 + \|P_T f\|_1$$

$$\Rightarrow \|P_T f\|_1 \geq \|f\|_1 - \|f - P_T f\|_1 \geq \|f\|_1 (1 - \varepsilon)$$

$$\Rightarrow \frac{\|P_T f\|_1}{\|f\|_1} \geq 1 - \varepsilon \qquad (2)$$

The claim follows by combining (1) and (2). ■

4. Logan's Phenomenon

We have shown that a function cannot be both highly concentrated on the time as well as on the frequency domain. We now apply it to the following kind of signal reconstruction problem.

Consider a signal s, which we know to be bandlimited to W. We can measure it perfectly except on a set of time T where we have some noise n. Although we know nothing about T except that is small, and n can be arbitrarily large (as long as it is finite), we get an impressive result:

Theorem Let $r = s + n$, where s ∈ $B_1(W)$, $n = P_T n$ and $|W||T| < \frac{1}{2}$. Using L1-minimisation we can recover s perfectly:

$$\tilde{s} = argmin_{g \in B1(W)} \|r - g\|_1 \;\Rightarrow\; s = \tilde{s}$$

L1-minimisation is a convex optimization problem, meaning there is a unique solution, which can be computed effectively. It will work no matter how strong the noise is as long as it sparse (meaning supp(n) is small), even if we do not know where the noise is. To prove this is true we'll first need another statement:

Lemma Let $|W||T| < \frac{1}{2}$. If n vanishes outside of T, then its best bandlimited L1-norm approximation $g \in$ $B_1(W)$ is zero.

Proof Assume $g \neq 0$. Using the inequality (1) from the proof of the first theorem, we get:

$$\frac{\|P_T g\|_1}{\|g\|_1} \leq |T||W| < \frac{1}{2}$$

It follows, that less than half of the volume of g is on T:

$$\|P_T g\|_1 < \frac{1}{2}\|g\|_1 \Rightarrow \|P_T g\|_1 < \|P_U g\|_1$$

where U is the complement of T. This enables us to estimate how well g approximates:

$$\|n - g\|_1 = \|P_T(n - g)\|_1 + \|P_U g\|_1$$

$$\geq \|P_T n\|_1 - \|P_T g\|_1 + \|P_U g\|_1$$

$$> \|P_T n\|_1 = \|n\|_1 = \|n - 0\|_1$$

This means that any possible bandlimited approximation would be worse than 0. ■

The idea behind this Lemma is that if $|W||T| < \frac{1}{2}$ is true, any function that is bandlimited to W has more than half of its volume outside of T. That means if we want to use it to approximate a function that vanishes outside of T we will always lose more on the complement of T (where n is zero) than we win on T.

Proof of Theorem Let $r = s + n$, $s \in B_1(W)$, $n = P_T n$ and $|W||T| < \frac{1}{2}$. We now consider:

$\|r - g\|_1 = \|s + n - g\|_1 = \|n + (s - g)\|_1$, where $g \in B_1(W)$

We can apply the Lemma and see that g minimizes the term for $(s - g) = 0$. Therefore solving $\tilde{s} = argmin_{g \in B_1(W)} \|r - g\|_1$ will give us $\tilde{s} = s$, reconstructing the original signal perfectly. ■

5. Compressed Sensing

Another interesting problem arises if we want to reconstruct a signal from incomplete frequency measurements. In our model, the signal we try to measure will be represented by a complex function $f: \mathbb{Z}/N\mathbb{Z} \longrightarrow \mathbb{C}$ on the cyclic group of N elements (or in other words a vector of N complex numbers). Its discrete Fourier transform is $f: \mathbb{Z}/N\mathbb{Z} \longrightarrow \mathbb{C}$ with $\hat{f}(\xi) = \frac{1}{\sqrt{N}} \sum_{x \epsilon \mathbb{Z}/N\mathbb{Z}} f(x) e^{-2\pi i x \xi / N}$. We assume that we can measure some of the Fourier coefficients $c_\xi = \hat{f}(\xi)$ for $\xi \in \Lambda \subseteq \mathbb{Z}/N\mathbb{Z}$ and want to know how many we need to reconstruct f.

We call f S-sparse if f has at most S positions, where it is non-zero. Using a discrete analog of Logan's phenomenon, one can see that it is possible to recover f perfectly if $S(N - |\Lambda|) < \frac{1}{2} N$ $\left(\Rightarrow |\Lambda| > N - \frac{N}{2S}\right)$, which would mean that we still have to measure most of the frequencies.

Unfortunately the discrete uncertainty principle

$$|supp(f)||supp(\hat{f})| \geq N, if\ f\ is\ not\ identically\ zero$$

on which Logan's phenomenon is based, is sharp.

This can be seen by considering the example of the Dirac-comb, where we assume N to be square and f to be the indicator function of $\{0, \sqrt{N}, 2\sqrt{N}, \ldots, N - \sqrt{N}\}$, the multiples of $\sqrt{N}$. From Fourier analysis we know that the Dirac-comb is its own Fourier transform, resulting in $|supp(f)||supp(\hat{f})| = \sqrt{N} * \sqrt{N} = N$.

Looking at this example we notice that the Dirac-comb has a very regular structure. We would expect a more irregular function with the same sparsity in the time domain to be less sparse in the frequency domain (one could imagine it as needing more different frequencies to cancel each other out if one has peaks of different height and with different distances from each other). This leads to the conjecture that $|supp(f)||supp(\hat{f})| \geq N'$ with $N' \gg N$ should be true for almost all $f: \mathbb{Z}/N\mathbb{Z} \longrightarrow \mathbb{C}$. (Except a relatively small amount of regularly structured functions).

In fact Candes, Romberg and Tao have shown in their paper that assuming a random set $\Lambda \subseteq \mathbb{Z}/N\mathbb{Z}$ and a S-sparse function $f: \mathbb{Z}/N\mathbb{Z} \longrightarrow \mathbb{C}$, $|\Lambda|$ only needs to be a few times larger than S to recover f perfectly with high probability. Numerical experiments with N in the range of a few hundred to a few thousand showed that in practice we can expect a successful recovery more than 50% of the time for $|\Lambda| > 4S$, and more than 90% of the time for $|\Lambda| > 8S$.

References

[1] David L. Donoho, Philip B.Stark,
"Uncertainty Relations and Signal Recovery"

[2] Emmanuel Candes, Justin Romberg, Terence Tao,
"Robust Uncertainty Principles: Exact Signal Reconstruction from Highly Incomplete Frequency Information"

[3] Terence Tao
"http://terrytao.wordpress.com/2007/04/15/ostrowski-lecture-the-uniform-uncertainty-principle-and-compressed-sensing/"