Lecture 3: Uncertainty Principles

Note Title 1/18/2014

☆ Heisenberg's Inequality (L^2)

Impossible for a signal to be both time (or space) limited and band limited.

Band-Limited

Let $f \in L^2$, and $\hat{f}(\xi) = 0$ for $|\xi| > \frac{\Omega}{2}$

$$\Rightarrow \quad f(x) = \int_{-\Omega/2}^{\Omega/2} \hat{f}(\xi) e^{2\pi i \xi x} d\xi$$

Now consider the following fcn of $z \in \mathbb{C}$:

$$F(z) = \int_{-\Omega/2}^{\Omega/2} e^{2\pi i \xi z} \hat{f}(\xi) d\xi$$

This makes sense for any $z \in \mathbb{C}$, and becomes an analytic fcn, in fact an entire fcn of $z \in \mathbb{C}$.

$\Rightarrow$ analytic on the whole $\mathbb{C}$.

Therefore, f being a restriction of F on the real axis, f cannot vanish except at some isolated points unless $f \equiv 0$.

space-lim.

By the same token, if $f(x) = 0$ for $|x| > \frac{A}{2}$, then $\hat{f}(\xi)$ cannot vanish unless $\hat{f} \equiv 0$.

$\Rightarrow$ f cannot be localized in both domains!

Heisenberg's inequality (a.k.a. uncertainty principle) gives us more precise statement.

Def. Dispersion (or spread) of f about $x=a$

$$\Delta_a f := \int (x-a)^2 |f(x)|^2 dx \Big/ \int |f(x)|^2 dx.$$

If we define $p(x) := |f(x)|^2 / \int |f(x)|^2 dx$, then $\Delta_a f$ is the 2nd moment of p around $x=a$.
$p(x)$ can be viewed as a pdf $\leftarrow \int p(x)\,dx = 1$.

Thm (Heisenberg's Inequality)

$$\forall f \in L^2, \quad (\Delta_{x_0} f)(\Delta_{\xi_0} \hat{f}) \geq \frac{1}{16\pi^2}$$

for all $x_0, \xi_0 \in \mathbb{R}$.

(Pf) For technical convenience, in addition to $f \in L^2$, assume : $\begin{cases} f \in C(\mathbb{R}) \cap PS(\mathbb{R}) \\ xf \in L^2, \ f' \in L^2 \end{cases}$ piecewise smooth fcns

The reasoning behind these assumptions :

$$xf \notin L^2 \iff \int x^2 f^2(x)\,dx = \infty \iff \Delta_{x_0} f = \infty$$

$$f' \notin L^2 \iff \int |i2\pi\xi \hat{f}(\xi)|^2 dx = \infty \iff \Delta_{\xi_0} \hat{f} = \infty$$

So, we are removing these cases from our consideration.

Let's consider first $x_0 = \xi_0 = 0$ case and consider an integral :

$$\int_a^b x \overline{f(x)} f'(x)\,dx = x|f(x)|^2 \Big|_a^b - \int_a^b \left(|f(x)|^2 + x\overline{f'(x)} f(x)\right) dx$$

Since $xf \in L^2$, $\hookrightarrow$ this goes to 0 as $\begin{cases} b \to +\infty \\ a \to -\infty \end{cases}$.

$$\underbrace{\int_{-\infty}^{\infty} |f(x)|^2 dx}_{= \|f\|_2^2} = -2\,\mathrm{Re} \int_{-\infty}^{\infty} x \overline{f(x)} f'(x)\,dx.$$

By the Cauchy-Schwarz ineq., we have

$$\|f\|_2^4 = 4\left(\mathrm{Re}\int_{-\infty}^{\infty} \overline{x f(x)} f'(x)\,dx\right)^2$$

$$\leq 4\int_{-\infty}^{\infty} x^2|f(x)|^2\,dx \cdot \int_{-\infty}^{\infty} |f'(x)|^2\,dx$$

Now the Plancherel Thm tells us $\|f\|_2^2 = \|\hat{f}\|_2^2$

also, $\int |f'(x)|^2 dx = \int |\widehat{f'}(\xi)|^2 d\xi$

$$= 4\pi^2 \int \xi^2 |\hat{f}(\xi)|^2 dx$$

Hence,

$$\|f\|_2^4 = \|f\|_2^2 \cdot \|\hat{f}\|_2^2 \leq 16\pi^2 \underbrace{\int x^2|f(x)|^2 dx}_{\Delta_0 f\cdot\|f\|_2^2} \cdot \underbrace{\int \xi^2|\hat{f}(\xi)|^2 d\xi}_{\Delta_0\hat{f}\cdot\|\hat{f}\|_2^2}$$

$$\iff \Delta_0 f \cdot \Delta_0 \hat{f} \geq \frac{1}{16\pi^2}$$

General case is the same, i.e., apply the above process to $F(x) = e^{-2\pi i \xi_0 x} f(x+x_0)$.

Then $\Delta_0 F = \Delta_{x_0} f$, $\Delta_0 \hat{F} = \Delta_{\xi_0}\hat{f}$

$$\Rightarrow \Delta_{x_0} f \cdot \Delta_{\xi_0}\hat{f} \geq \frac{1}{16\pi^2} \quad ///$$

<u>Exercise</u> : Show that the equality in Heisenberg's inequality holds iff f is a Gaussian fcn.

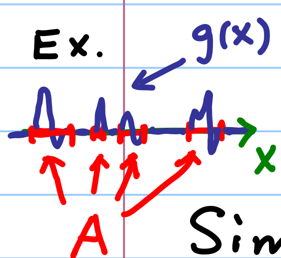
★ More General Uncertainty Principles of Donoho-Stark (1989)

Rather than dealing with dispersions of f & $\hat{f}$, let's consider practical concentration of f & $\hat{f}$ on measurable sets (much more general than intervals).

Def. $f \in L^2$ is said to be ε-concentrated on a measurable set A if $\exists\, g \in L^2$ s.t. $\operatorname{supp} g = A$ and $\|f - g\|_2 \le \varepsilon$.

$$\operatorname{supp} g = \{x \in \mathbb{R} \mid g(x) \neq 0\}$$

Ex.



Similarly we can define the ε-concentration of $\hat{f}$.

Thm (Donoho-Stark, 1989)

Let A & Ω be measurable sets and suppose $\exists$ the Fourier transform pair $(f, \hat{f})$ with $\|f\|_2 = \|\hat{f}\|_2 = 1$ s.t. f is ε_A-concentrated on A and $\hat{f}$ is ε_Ω-concentrated on Ω.

Then,
$$|A| \cdot |\Omega| \ge \left(1 - (\varepsilon_A + \varepsilon_\Omega)\right)^2$$

In short, f & $\hat{f}$ cannot both be highly concentrated no matter what sets of concentration A & Ω we choose.

(Pf) See Donoho-Stark (1989)

This thm has a deep & surprising counterpart in the discrete setting.

Let $\mathbf{f} = (f_0, \cdots, f_{N-1})^T$ be a vector in $\mathbb{C}^N$ and $\hat{\mathbf{f}} = (\hat{f}_0, \cdots, \hat{f}_{N-1})^T \in \mathbb{C}^N$ be its discrete Fourier Transf.:
(DFT)

$$\hat{f}_k := \frac{1}{\sqrt{N}} \sum_{\ell=0}^{N-1} f_\ell \, e^{-2\pi i k \ell / N}, \quad k = 0, 1, \cdots, N-1.$$

(We'll cover DFT much more in our later lectures! Here, we just want to show an interesting consequence of the discrete version of the uncertainty principle.)

Let $\|\mathbf{f}\|_0 := \#\{\ell \in [0, N-1] \mid f_\ell \neq 0\}$.
This is the so-called ℓ^0 quasi-norm.
A typical measure of sparsity of $\mathbf{f}$.
(more precisely, $\|\mathbf{f}\|_0$ small $\Longleftrightarrow$ $\mathbf{f}$: sparse)

It's not a norm since $\|\cdot\|_0$ does not satisfy the homogeneity : $\|a\mathbf{f}\| = |a| \cdot \|\mathbf{f}\|, \ \forall a \in \mathbb{C}$.

<u>Thm (Donoho–Stark, 1989)</u>

$$\|\mathbf{f}\|_0 \cdot \|\hat{\mathbf{f}}\|_0 \geq N.$$

→ Pf : See their article.

<u>Cor (Donoho–Stark, 1989)</u>

$$\|\mathbf{f}\|_0 + \|\hat{\mathbf{f}}\|_0 \geq 2\sqrt{N}.$$

(Pf) Easy application of AM ≥ GM. ///

Here is a more general discrete version corresponding to the continuous version.

Thm (Donoho - Stark, 1989)

Let $(f, \hat{f})$ be a DFT pair of unit norm with

f : ε_A-concentrated on the index set A and
$\hat{f}$: ε_Ω-concentrated on the index set Ω.
Then, $\|f\|_0 \cdot \|\hat{f}\|_0 \geq N(1-(\varepsilon_A+\varepsilon_\Omega))^2$.

☆ Recovery of a sparse wide-band signal from narrow-band measurements

As an application of the above generalized uncertainty principles, consider the following signal processing problem, often appears in practice (e.g., astronomical imaging, spectroscopy, geophysics, ...).

Suppose the discrete measurement r is a noisy, band-limited version of the ideal signal s, i.e.,

$$(*) \qquad r = \underbrace{P_\Omega}_{\text{BL-operator}} s + \underbrace{n}_{\text{noise}}$$

$$(P_\Omega s)_\ell := \frac{1}{\sqrt{N}} \sum_{k\in\Omega} \underbrace{\hat{s}_k}_{\text{True DFT coeff's of } s} e^{2\pi i k\ell/N}$$

By taking the DFT on both sides of (*)

$$\Rightarrow \quad \hat{r}_k = \begin{cases} \hat{s}_k + \hat{n}_k & \text{if } k \in \Omega \\ 0 & \text{otherwise.} \end{cases}$$

where we also assumed $P_\Omega \mathbf{n} = \mathbf{n}$ (i.e., $\mathbf{n}$ is also band-limited).

Can we recover $\mathbf{s}$ from $\mathbf{r}$?
Enter the uncertainty principle!

<u>Thm (Donoho-Stark, 1989)</u>

Suppose $\mathbf{n} \equiv 0$ (no noise), and we know $\|\mathbf{s}\|_0 \leq N_0 < N$.

If $$N_0 \cdot \underbrace{(N - |\Omega|)}_{\text{\# of missing freq. components.}} < N/2 \qquad (*)$$

then $\mathbf{s}$ can be uniquely reconstructed from $\mathbf{r}$.

$\Rightarrow$ If $\mathbf{s}$ is sparse, $\exists$ a chance of recovery!

(Pf) <u>Uniqueness</u>: Suppose $\mathbf{s}'$ also generates $\mathbf{r}$,
i.e., $P_\Omega \mathbf{s}' = \mathbf{r} = P_\Omega \mathbf{s}$.
Define $\mathbf{h} := \mathbf{s}' - \mathbf{s}$ so that $P_\Omega \mathbf{h} = 0$.

$$\left.\begin{array}{l} \text{Now } \|\mathbf{s}'\|_0 \leq N_0 \quad \Rightarrow \quad \|\mathbf{h}\|_0 \leq 2N_0 \\ \text{Because } P_\Omega \mathbf{h} = 0, \quad \|\hat{\mathbf{h}}\|_0 \leq N - |\Omega| \end{array}\right\} (**)$$

$\Rightarrow \mathbf{h} \equiv 0$ otherwise $\mathbf{h}$ violates the uncertainty principle $\|\mathbf{h}\|_0 \cdot \|\hat{\mathbf{h}}\|_0 \geq N$ because (*) & (**) lead to $\|\mathbf{h}\|_0 \cdot \|\hat{\mathbf{h}}\|_0 < N$.

How about the reconstruction algorithm?

Let $\tilde{s} := \arg\min_{s' \in \mathcal{S}_0} \| r - P_\Omega s' \|$

where $\mathcal{S}_0 := \{ f \in \mathbb{C}^N \mid \|f\|_0 \le N_0 < N \}$

From the uniqueness, we know $\tilde{s} = s$.

But how can we find such $\tilde{s}$?

$\Rightarrow$ a combinatorial algorithm.

Let $\Pi :=$ the $\binom{N}{N_0}$ subsets $\{\tau\}$ of indices

$\{0, 1, \cdots, N-1\}$ with $|\tau| = N_0$.

For a given $\tau \in \Pi$, let

This is a least squares problem! $\rightarrow$ $\tilde{s}_\tau := \arg\min_{s'} \{ \| r - P_\Omega s' \| \mid \operatorname{supp} s' = \tau \}$

$\Rightarrow \exists\, \tau_0 \in \Pi$ s.t. $\tilde{s}_{\tau_0} = \tilde{s}$, i.e.,

$$\tilde{s} = \arg\min_{\tilde{s}_\tau, \tau \in \Pi} \| r - P_\Omega \tilde{s}_\tau \|. \qquad ///$$

$\Rightarrow$ Impractical for large N.

$\Rightarrow$ $\exists$ a much better approach using L^1 (or ℓ^1) compressed sensing. Read Donoho-Stark as its beginning!

How about the noisy case?

<u>Thm (Donoho-Stark, 1989)</u>

Suppose $\|s\|_0 \le N_0 < N$ with $2N_0(N - |\Omega|) < N$. Assume $\|n\| \le \varepsilon$. If $\tilde{s}$ satisfies $\|\tilde{s}\|_0 \le N_0$ and $\| r - P_\Omega \tilde{s} \| \le \varepsilon$, <u>then</u>

$$\| s - \tilde{s} \| \le 2\varepsilon \Big/ \sqrt{1 - 2N_0(1 - |\Omega|/N)} \qquad //$$