Uncertainty and Information in Time-Frequency Analysis

Suren Jayasuriya
University of Pittsburgh

REU/MCTP/UBM Summer Research Conference,
Texas A & M University, July 28, 2011

Preliminaries

$L^2(\mathbb{R})$ is the space of all functions from $\mathbb{R} \to \mathbb{C}$ such that their L^2 norm: $||f||_2 = (\int_{\mathbb{R}} |f(t)|^2 dt)^{1/2} < \infty$ is finite.

Definition

The Fourier Transform $\mathcal{F} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ of a function f is defined as

$$\mathcal{F}(f)(\omega) = \hat{f}(\omega) = \int_{\mathbb{R}} f(t) e^{-2\pi i \omega t} dt.$$

Plancherel's Theorem: $||f||_2 = ||\hat{f}||_2$

Preliminaries continued

Discrete Setting:

Definition

Let $x \in \mathbb{R}^N$, i.e. $x = (x_t)_1^N = (x_1, x_2, \ldots, x_N)$. The Discrete Fourier Transform (DFT) of x is given by:

$$\mathcal{F}x(\omega) = \hat{x}(\omega) = \frac{1}{\sqrt{N}} \sum_{t=1}^{N} x_t \cdot e^{-2\pi i \omega t/N}, \omega = 1, 2, \ldots, N.$$

Plancherel's Theorem: $\sum_{t=1}^{N} |x_t|^2 = \sum_{\omega=1}^{N} |\hat{x}_\omega|^2$.

Classical Uncertainty Principle

Let $\Delta_f t = \left(\int_{\mathbb{R}} (t - t_0)^2 |f(t)|^2 dt\right)^{1/2}$ where $t_0 \in \mathbb{R}$.

Let $\Delta_f \omega = \left(\int_{\mathbb{R}} (\omega - \omega_0)^2 |\hat{f}(\omega)|^2 d\omega\right)^{1/2}$ where $\omega_0 \in \mathbb{R}$.

Theorem (Heisenberg's Inequality)

If $f \in L^2(\mathbb{R})$ with $||f||_2 = 1$, then

$$\Delta_f t \cdot \Delta_f \omega \geq \frac{1}{4\pi}$$

Quantum Mechanics: $\Delta_f t$ = position "uncertainty",
$\Delta_f \omega$ = momentum "uncertainty"

Uncertainty Principle of Donoho and Stark

Definition

A function f is ϵ-concentrated on a set T if

$$||f - \chi_T f||_2 \leq \epsilon$$

where χ_T is the characteristic function of the set T.

Theorem (Donoho/Stark 1989)

Suppose f is ϵ_T-concentrated on T, and its Fourier transform $\hat{f}$ is ϵ_W - concentrated on a set W with $||f||_2 = 1$. Then

$$m(T) \cdot m(W) \geq (1 - \epsilon_T - \epsilon_W)^2.$$

Information Theory

Introduced by C.E. Shannon's 1948 paper: "A Mathematical Theory of Communication"

Sentence 1: “The sun will set in the west tomorrow"

Sentence 2: “There will be a solar eclipse tomorrow"

Which has more information?

Caveat- Received message: “WZHSLNRU?@TG"

Information Theory

Introduced by C.E. Shannon's 1948 paper: "A Mathematical Theory of Communication"

Sentence 1: "The sun will set in the west tomorrow"

Sentence 2: "There will be a solar eclipse tomorrow"

Which has more information?

Caveat- Received message: "WZHSLNRU?@TG"

Three Intuitive Postulates for Information:

1. If E, F are events such that $P(E) \leq P(F)$, then $I(E) \geq I(F)$.
2. If E, F are independent events, $I(E \cap F) = I(E) + I(F)$.
3. For all events E, $I(E) \geq 0$.

(Shannon 1948) The only function that satisfies 1,2,3 is of the form:

$$I(E) = -K log_a(P(E))$$

where a, K are positive constants.

Information as a Random Variable

Consider a discrete random variable $X : \mathcal{S} \to \{x_1, \ldots, x_n\} \subset \mathbb{R}$ with associated probability distribution specified by $p_i = P(X = x_i)$.

Example: Let $\mathcal{S} = \{\text{heads}, \text{tails}\}$. Then $X : \mathcal{S} \to \{0, 1\}$ is a random variable where $X(\text{heads}) = 1$ and $X(\text{tails}) = 0$ with associated probabilities $p_0 = p_1 = \frac{1}{2}$.

Warning: $X = 1$ is commonly written instead of $X(\cdot) = 1$.

Definition

The information of a random variable X is given by $I(X) : \{x_1, \ldots, x_n\} \to \mathbb{R}$ by $I(X) = -\log_2(P(X))$. The units of information with respect to $\log_2$ are called bits.

Entropy as a Measure of Uncertainty

Examples of Calculating Information of Events:

I(“coin landing heads")
$= -\log_2(1/2) = 1$ bit

I(“card being Ace of Spades")
= - $\log_2(\frac{1}{52}) \approx 5.70$bits

Definition (Shannon 1948)

The entropy of a random variable X is the expected value of $I(X)$ given by

$$H(X) = \mathbb{E}(I(X)) = -\sum_{j=1}^{n} p_j \log_2(p_j).$$

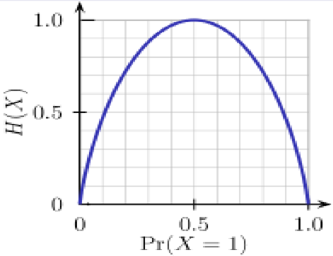


Figure: Entropy of a “Weighted" Coin Flip

Hirschman Uncertainty Principle

Let x_t, $\hat{x}_\omega \in \mathbb{R}^N$ such that $||x|| = 1$.

Let X, Y be random variables who map into $\{1, 2, \ldots, N\}$ with associated probability distributions given by $P(X = i) = |x_i|^2$ and $P(Y = i) = |\hat{x}_i|^2$.

Theorem (Hirschman's Uncertainty Principle (Dembo et al. 1991))

Let x_t and $\hat{x}_\omega$ be a Fourier transform pair such that $||x|| = 1$. Then defining random variables X, Y as given above, we have

$$H(X) + H(Y) \geq \log_2(N).$$

New Approach: Approximate Concentration of Entropy

Let x_t and $\hat{x}_\omega$ be a Fourier transform pair in $\mathbb{R}^N$ such that $||x|| = 1$ and X and Y be defined as before.

Let $T \subseteq \{1, \ldots, N\}$. Define $H(X|_T) = -\sum_{j \in T} p_j \log_2(p_j)$.

Definition

X is ϵ-concentrated in entropy to a set $T \subseteq \{1, 2, \ldots, N\}$ if

$$H(X) - H(X|_T) = -\sum_{j \notin T} p_j \log_2(p_j) \leq \epsilon.$$

Question: Are there lower bounds for $H(X|_T), H(Y|_W)$ that can be established?

Numerical Simulations

$H(X) + H(Y) =$ Sum of Entropies,
$H(X|_T) + H(Y|_W) =$ Sum of Approximate Entropies

Figure: $\epsilon_T = \epsilon_W = 1/10$

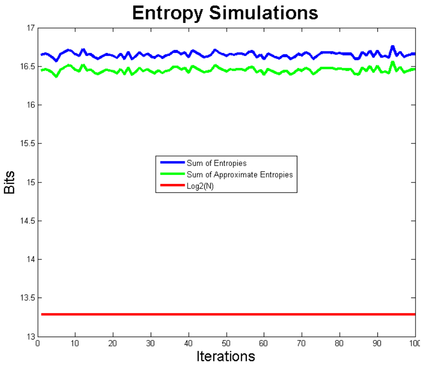
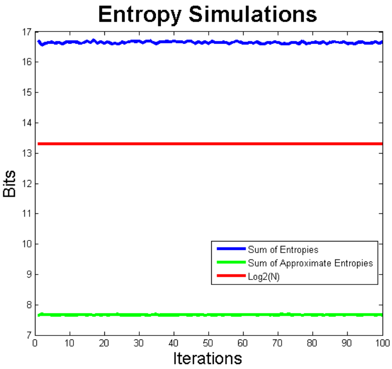


Figure: $\epsilon_T = \epsilon_W = 5$



An Uncertainty Result for Approximate Concentration of Entropy

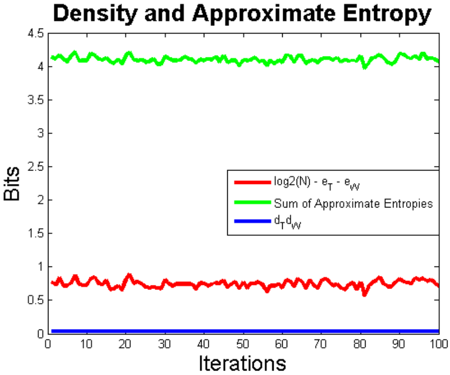
Theorem

Let x_t and $\hat{x}_\omega$ be a Fourier transform pair in $\mathbb{R}^N$ such that $||x|| = 1$ and two random variables X, Y who share the same range, where $P(X = i) = |x_i|^2$ and $P(Y = i) = |\hat{x}_i|^2$. Suppose X is ϵ_T-concentrated in entropy to a set T, and Y is ϵ_W-concentrated in entropy to a set W. Then we have

$$\log_2(N) - \epsilon_T - \epsilon_W \leq H(X|_T) + H(Y|_W).$$

Density of the sets T and W

We define the density of T to be $d_T = \frac{N_T}{N}$ where N_T is the number of non-zero entries in T. Similarly, we define the density of W, $d_W = \frac{N_W}{N}$.



Results and Conjectures

Let X and Y be defined for the unit-normalized Fourier transform pair $x, \hat{x}$ as given before.

Theorem

Let X be ϵ_T-concentrated in entropy to a set T, Y be ϵ_W-concentrated in entropy to a set W. Then for $N \geq 2^{1+\epsilon_T+\epsilon_W}$,

$$d_T d_W \leq \log_2(N) - \epsilon_T - \epsilon_W \leq H(X|_T) + H(Y|_W).$$

We also know that $1 - \epsilon_T - \epsilon_W \leq \log_2(N) - \epsilon_T - \epsilon_W$. The following conjecture is suggested by numerical simulations:

Conjecture

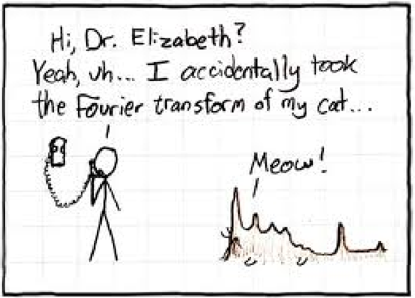
$$1 - \epsilon_T - \epsilon_W \leq d_T d_W.$$

References

1. A. Dembo, T.M. Cover, J.A. Thomas, *Information Theoretic Inequalities*, IEEE Transactions On Information Theory, Vol. 37, No.6, 1991.
2. D. L. Donoho and P. B. Stark. "Uncertainty Principles and Signal Recovery". SIAM Journal of Applied Mathematics 49 (1989), 906-931.
3. Gröchenig, Karlheinz."Foundations of Time-Frequency Analysis". Series: Applied and Numerical Harmonic Analysis, Birkhaüser, 2001

Thanks

Thanks to Dr. David Larson, Dr. Lewis Bowen, Dr. Yunus Zeytuncu, and Stephen Rowe for their advice and guidance as well as the Math REU program at Texas A & M University for this opportunity



This work is funded by NSF grant 0850470, "REU Site: Undergraduate Research in Mathematical Sciences and its Applications."

Appendix

Theorem (Heisenberg's Inequality)

If $f \in L^2(\mathbb{R})$ with $||f||_2 = 1$, then

$$\left(\int_{\mathbb{R}} (t - t_0)^2 |f(t)|^2 dt\right)^{1/2} \cdot \left(\int_{\mathbb{R}} (\omega - \omega_0)^2 |\hat{f}(\omega)|^2 d\omega\right)^{1/2} \geq \frac{1}{4\pi}$$

Lemma

Let A, B be self-adjoint operators on a Hilbert space $\mathcal{H}$. We define the commutator of A, B to be $[A, B] := AB - BA$. Then we have that

$$||(A - a)f|| \cdot ||(B - b)f|| \geq \frac{1}{2} |\langle [A, B]f, f \rangle|$$

for $a, b \in \mathbb{R}$ and f in the domain of $AB \cap BA$.

Proof of Lemma

Proof.

$$
\begin{aligned}
|\langle [A,B]f, f\rangle| &= |\langle ((A-a)(B-b)-(B-b)(A-a))\, f, f\rangle| \\
&= |\langle (B-b)f, (A-a)f\rangle - \langle (A-a)f, (B-b)f\rangle| \\
&\leq |\langle (B-b)f, (A-a)f\rangle| + |\langle (A-a)f, (B-b)f\rangle| \\
&\leq 2||(B-b)f|| \cdot ||(A-a)f||
\end{aligned}
$$

from which the lemma follows. □

Proof of Heisenberg's Inequality

With this lemma, we may continue with the proof of the theorem. Let the operators $A, B \in B(L^2(\mathbb{R}))$ by

$$Af = tf(t), B = \frac{1}{2\pi i} f'(t).$$

A, B are self-adjoint operators. By the lemma, we have then that

$$||(A-a)f|| \cdot ||(B-b)f|| \geq \frac{1}{2}|\langle [A,B]f, f\rangle|.$$

Observe that

$$\frac{1}{2}|\langle [A,B]f, f\rangle| = \frac{1}{2}|\int_{\mathbb{R}} \frac{1}{2\pi i}|f(t)|^2 dt| = 1/4\pi.$$

Then,

$||(B-b)f|| = ||\mathcal{F}(B-b)(f)|| = \left(\int_{\mathbb{R}}(\omega-\omega_0)^2|\hat{f}(\omega)|^2 d\omega\right)^{1/2}$ and $||(A-a)f|| = \left(\int_{\mathbb{R}}(t-t_0)^2|f(t)|^2 dt\right)^{1/2}$. QED.