Frame Theory

Yue Zhang

This review is based on materials in 2015 IMA summer school: modern harmonic analysis and its applications .

September 10, 2015

Outline

Frame Definition

Definition: A sequence $\{x_n\}_{n \in N}$ in a Hilbert space H is a frame for H if there exist constants A, $B > 0$ such that the following *pseudo-Plancherel formula* holds for $\forall x \in H$:

$$A\|x\|^2 \leq \sum_{n=1}^{\infty} |\langle x, x_n \rangle|^2 \leq B\|x\|^2.$$

- *Plancherel Equality* is $\|x\|^2 = \sum_{n=1}^{\infty} |\langle x, x_n \rangle|^2$;
- A and B are called frame bounds. If A = B, the frame is called a *tight* frame, if $A = B = 1$, it's *Parseval frame*.
- A frame is *exact* if it fails to be a frame whenever any single element is deleted from the sequence.
- ONB $\Leftrightarrow$ exact Parseval frame.

Example

- **Mercedes frame:** Let $H = \mathbb{R}^2$, and set $x_1 = (0,1)$, $x_2 = (-\frac{\sqrt{3}}{2}, -\frac{1}{2})$, $x_3 = (\frac{\sqrt{3}}{2}, -\frac{1}{2})$, then $\sum_{n=1}^{3} |\langle x, x_n \rangle|^2 = \frac{3}{2}\|x\|^2$ for all $x \in \mathbb{R}^2$.
  If we set $c = (2/3)^{1/2}$, then $\{cx_1, cx_2, cx_3\}$ is a tight frame. It's not orthogonal, not a basis (not unique representation).
- **Trigonometric system:** Let $H = L^2[0,1]$, $\langle f, g \rangle = \int_0^1 f(t)\overline{g(t)}dt$. Then $\{e_n\}_{n \in Z} = \{e^{2\pi int}\}_{n \in Z}$ is an ONB for H.
- What's more, it can be shown that, given $b < 1$, $\{e_{bn}\}_{n \in Z} = \{e^{2\pi ibnt}\}_{n \in Z}$ is a tight frame for $L^2[0,1]$, with $A = B = 1/b$.

Properties of FT

- If $f \in C_c^2(\mathbb{R})$, then $\hat{f} \in L^2(\mathbb{R})$ and $\|\hat{f}\|_2 = \|f\|_2$. Hence $\mathcal{F} : f \to \hat{f}$ is an isometric map from a dense subset of $L^2(\mathbb{R})$ into $L^2(\mathbb{R})$.
- Recall that $\{e_{bn}\}_{n \in Z}$ is a tight frame for $L^2[1/2, 1/2]$, note here we restrict $\{e_{bn}\}_{n \in Z}$ to be itself within this interval and zero outside. Then $\{s_{bn}\} = \{\mathcal{F}^{-1}(e_{bn}\chi)\}$ is a tight frame for PaleyWiener space:

$$PW = \{f \in L^2(\mathbb{R}) : supp(\hat{f}) \subseteq [1/2, 1/2]\}$$

Hence for all $f \in PW$, $f = b \sum_{n \in Z} \langle f, s_{bn} \rangle s_{bn}$ and then we work out the calculation to get the Shannon Sampling theorem!

Classical Sampling Theorem

Shannon Sampling Theorem: Fix $f \in PW$, i.e., $f \in L^2(\mathbb{R})$ is bandlimited to [1/2, 1/2]. Then for $0 < b \leq 1$,

$$f(x) = b \sum_{n \in Z} \langle f, s_{bn} \rangle s_{bn}(x) = b \sum_{n \in Z} f(bn) \frac{\sin \pi(x - bn)}{\pi(x - bn)}$$

Outline

FUNTFs

- Recall that a frame $\Phi = \{\varphi_k\}_{k=1}^{M} \subseteq \mathbb{R}^N$ is a tight frame for $\mathbb{R}^N$ if $A = B$.
- If Φ is a tight frame of unit-norm vectors, that is, $\|\varphi_k\| = 1$, we say that Φ is a finite unit-norm tight frame (FUNTF). In this case, we can reconstruct x by

$$x = \frac{N}{M} \sum_{k=1}^{M} \langle x, \varphi_k \rangle \varphi_k.$$

- Why are we interested in FUNTF?
- It turns out that it's closely related to the condition number of the frame.

Frame Potential

Theorem (Benedetto and Fickus, 2003)

For each $\Phi = \{\varphi_k\}_{k=1}^{M} \subseteq \mathbb{R}^N$, such that $\|\varphi_k\| = 1$ for each k, we have

$$FP(\Phi) = \sum_{i=1}^{M} \sum_{k=1}^{k} |\langle \varphi_j, \varphi_k \rangle|^2 \geq \frac{M}{N} \max(M, N)$$

Furthermore,
• If $M \leq N$, the minimum of FP is M and is achieved by orthonormal systems for $\mathbb{R}^N$ with M elements.
• If $M \geq N$, the minimum of FP is $\frac{M^2}{N}$ and is achieved by FUNTFs.
$FP(\Phi)$ is the frame potential.

Thus FUNTFs can be considered 'optimally conditioned' frames. In particular the condition number of the frame operator is 1.

Scalable Frame and Optimization

Definition: A frame $\Phi = \{\varphi_k\}_{k=1}^M \subseteq \mathbb{R}^N$ is scalable, if $\exists \{c_k\}_{k=1}^M \subset \mathbb{R}^N$ such that $\{c_k \varphi_k\}_{k=1}^M$ is a tight frame for $\mathbb{R}^N$.

...

Conclusion: a frame is scalable iff there exists an $M \times M$ diagonal matrix $C = diag(c_k)$ s.t. $\Phi X^2 \Phi^T = \alpha I$, for some constant α. Define a measure of scalability,

$$D_\Phi := \min_{C \geq 0 \text{ diagonal}} \|\Phi C \Phi^T - I\|_F.$$

Φ is scalable iff $D_\Phi = 0$.

Where do we need this? It's said that it can be used in improving NDE scheme, as people may want frames with low condition numbers.

Outline

Frames and Time-Frequency Analysis

Preconditioning of Finite Frames

Sparse Fourier Transform

Introduction to Phase Retrieval Problem

Basics

- What do we mean when we say sparse Fourier transform?
- Given any vector $x \in \mathbb{C}^N$, is there a sampling set $S \subset [N]$ with size $|S| = m < n$ and an algorithm $\mathcal{A}$ such that the algorithm on input x(S) can find a(n approximation) k term Fourier representation of x?
- We can tell immediately that sampling is critical here.
- Two approaches:
  - Mathematical Perspective: R.I.P. to L_1 optimization.
  - Computer Science Perspective: Fast Fourier Sampling ('Divide-and-Conquer').

SFT via L_1 optimization

Conclusion: If $\mathcal{F}_S$ (=DFT matrix restricted on rows in sample set S) satisfy R.I.P, then for all x there are algorithms that return $\hat{x}$ with

$$\|x - \hat{x}\|_2 \leq \frac{C}{\sqrt{k}} \|x - x_k\|_1$$

Where x_k = best $k-$ term Fourier representation, $|S| = \Theta(k \log^c(n/k))$, C is a constant.

- L_1 minimization via convex optimization approach.
- CoSaMP (compressive sampling matching pursuit 2008). (Greedy)
- IHT (iterative hard thresholding, 2008), $y = \Phi x + e$, x $K-$ sparse, $x_{n+1} = H_k(x_n + \Phi^*(y - \Phi x_n))$
- OMP (by Anna Gilbert et al. 2007) (Greedy)

Cons: Inefficient! (Really?)

SFT via Fast Fourier Sampling

This idea (maybe) comes from the classical 'divide-and-conquer' binary search algorithm in CS. Suppose the length of signal is n, it can be factored into a_1, a_2, ..., a_m, all are primes, Then the number samples we need (to identify one frequency) are just $\sum_{i=1}^{m} a_i$.
Briefly speaking, all sparse FFTs (repeatedly) perform some version of the three following steps:

- identification of frequencies whose Fourier coefficients are large in magnitude (typically a randomized sub-routine),
- accurate estimation of the Fourier coefficients of the frequencies identified in the first step,
- subtraction of the contribution of the partial Fourier representation computed by the first two steps from the entries of f before any subsequent repetitions.

Outline

Phase Retrieval Problem Formulate

Suppose we have signal $x_0 \in \mathbb{C}^n$, and quadratic measure $\mathbb{A}(x_0) = \{|\langle a_k, x_0 \rangle|^2 : k = 1, ..., m\}$. Phase retrieval problem is a feasibility problem:

$$\begin{aligned} &\text{find } x \\ &\text{obeying } \mathbb{A}(x) = \mathbb{A}(x_0) := b \end{aligned}$$

Indeed, let $X = xx^*$

$$|\langle a_k, x \rangle|^2 = Tr(x^* a_k a_k^* x) = Tr(a_k a_k^* x x^*) := Tr(A_k x)$$

Then we rewrite the optimization problem:

$$\begin{aligned} &\text{find } X \\ &\text{subject to } \mathcal{A}(x) = b, X \succeq 0, rank(X) = 1 \end{aligned}$$

Which is equivalent to

$$\begin{aligned} &\text{minimize } rank(X) \\ &\text{subject to } \mathcal{A}(x) = b, X \succeq 0 \end{aligned}$$