Lecture 3 – Frame Theory

David Walnut
Department of Mathematical Sciences
George Mason University
Fairfax, VA USA

Chapman Lectures, Chapman University, Orange, CA
6-10 November 2017

Outline

- Finite Frames
- Abstract Frames
- The Frame operator
- Riesz bases
- Some historical remarks
- The BLT from the perspective of frames
- Applications of frames

Finite Frames and Linear Algebra

- We start from the elementary fact that a collection of d vectors in $\mathbf{C}^d$,

$$X = \{x_1, x_2, \ldots, x_d\}$$

is a *basis* for $\mathbf{C}^d$ if and only if X is *linearly independent*, that is, if

$$c_1 x_1 + c_2 x_2 + \cdots + c_d x_d = 0 \Longrightarrow c_1 = c_2 = \cdots = c_d = 0.$$

- Every $x \in \mathbf{C}^d$ can be written uniquely as

$$x = a_1 x_1 + a_2 x_2 + \cdots + a_d x_d$$

where $a_i = \langle x, \widetilde{x}_i \rangle$ and $\widetilde{X} = \{\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_d\}$ is the *dual basis* of X.

Matrix notation

- Let

$$B = \begin{bmatrix} | & | & & | \\ x_1 & x_2 & \cdots & x_d \\ | & | & & | \end{bmatrix}$$

Since X is linearly independent, B is invertible.

- $B^{-1} = \begin{bmatrix} — & \overline{\widetilde{x}_1} & — \\ — & \overline{\widetilde{x}_2} & — \\ & \vdots & \\ — & \overline{\widetilde{x}_d} & — \end{bmatrix}$ where $\{\widetilde{x}_j\}_{j=1}^{d}$ is the dual basis.

- $x = BB^{-1}x = B \begin{bmatrix} \langle x, \widetilde{x}_1 \rangle \\ \langle x, \widetilde{x}_2 \rangle \\ \vdots \\ \langle x, \widetilde{x}_d \rangle \end{bmatrix} = \sum_{j=1}^{d} \langle x, \widetilde{x}_j \rangle \, x_j.$

Finite Frames

- Let $Y = \{y_1, y_2, \dots, y_n\} \subseteq \mathbf{C}^d$, where $n > d$ be a spanning set for $\mathbf{C}^d$. We say then that Y is a *frame* for $\mathbf{C}^d$.
- Since Y contains a basis for $\mathbf{C}^d$, for all $x \in \mathbf{C}^d$ there exist coefficients α_j such that

$$x = \alpha_1 y_1 + \alpha_2 y_2 + \cdots + \alpha_n y_n.$$

  However these coefficients need not be unique.
- Frames mimic bases but are *redundant* in the sense that
  - Not all of the frame elements need be present in order to represent elements in the vector space, and
  - There are multiple representations of each vector in the space.

Matrix notation

- Let
$$F = \left[\begin{array}{cccc} y_1 & y_2 & \cdots & y_n \end{array} \right]$$
Since Y is a frame, F has full rank, the $d \times d$ matrix FF^* is invertible.
- For all $x \in \mathbf{C}^d$, $x = FF^*(FF^*)^{-1}x$, and note that $F^\dagger = F^*(FF^*)^{-1}$ is the *pseudoinverse* of F. Letting

$$F^\dagger = \left[\begin{array}{c} \overline{\widetilde{y}_1} \\ \overline{\widetilde{y}_2} \\ \vdots \\ \overline{\widetilde{y}_n} \end{array} \right], \quad x = FF^\dagger x = F \left[\begin{array}{c} \langle x, \widetilde{y}_1 \rangle \\ \langle x, \widetilde{y}_2 \rangle \\ \vdots \\ \langle x, \widetilde{y}_n \rangle \end{array} \right] = \sum_{j=1}^{n} \langle x, \widetilde{y}_j \rangle \, y_j.$$

- By properties of the pseudoinverse, $c = F^\dagger x$ is the solution to $Fc = x$ with the smallest norm.
- Hence of all coefficients $\alpha_1, \alpha_2, \ldots, \alpha_n$ that satisfy

$$x = \sum_{j=1}^{n} \alpha_j y_j,$$

$\alpha_j = \langle x, \widetilde{y}_j \rangle$ has minimal norm.

- The collection of vectors $\{\widetilde{y}_j\}_{j=1}^{n}$ is called the *dual frame* of Y.

Abstract frames

Definition

A *frame* in a separable Hilbert space H is a sequence of vectors $\{x_k\}_{k \in K}$ with the property that there exist constants A, $B > 0$, called the *frame bounds* such that for all x in the Hilbert space

$$A \|x\|^2 \leq \sum_{k \in K} |\langle x, x_k \rangle|^2 \leq B \|x\|^2 .$$

A frame is *tight* if $A = B$ and is *uniform* if $\|x_j\| = \|x_k\|$ for all j and k. A tight frame with frame bounds $A = B = 1$ is a *Parseval frame*.

- Note that in the case of finite frames, this inequality is $A \|x\|^2_{\mathbf{C}^d} \leq \|F^* x\|_{\mathbf{C}^n} \leq B \|x\|^2_{\mathbf{C}^d}$.
- This inequality is always satisfied if F has full rank.

The frame operator

Definition

Given a frame $\{x_k\}$, we define the *analysis operator* by

$$T\colon H \to \ell^2(K); \quad x \mapsto \{\langle x, x_k \rangle\}.$$

Its adjoint is *synthesis operator*

$$T^*\colon \ell^2(K) \to H; \quad \{c_k\} \mapsto \sum c_k \, x_k.$$

The *frame operator* for $\{x_k\}_{k \in K}$ is

$$S = T^* T\colon H \to H; \quad x \mapsto \sum \langle x, x_k \rangle \, x_k.$$

- The analysis operator corresponds to the mapping $F^*: \mathbf{C}^d \to \mathbf{C}^n$, $x \mapsto F^*x$, and the synthesis operator to $F: \mathbf{C}^n \to \mathbf{C}^d$, $y \mapsto Fy$.
- The frame inequality can be written as

$$A\langle x, x\rangle \leq \langle Sx, x\rangle \leq B\langle x, x\rangle$$

or as

$$A\,I \leq S \leq BI$$

so that S is a self-adjoint positive operator on H.

The frame operator

- After some manipulation, we arrive at the operator inequality

$$\frac{A-B}{A+B} I \leq I - \frac{2}{B+A} S \leq \frac{B-A}{B+A} I.$$

- This implies that S is an isomorphism of H and that

$$\left\| I - \frac{2}{B+A} S \right\| \leq \frac{B-A}{B+A} < 1.$$

- If the frame is tight ($A = B$) then S is a multiple of the identity.
- Numerical inversion of S converges very rapidly if $A \approx B$, so that good frame bounds are important.

The dual frame

- We have the following representations of $x \in H$.

$$x = S\,S^{-1}x = \sum \langle S^{-1}x, x_k\rangle\, x_k = \sum \langle x, S^{-1}x_k\rangle\, x_k,$$

$$x = S^{-1}Sx = \sum \langle x, x_k\rangle\, S^{-1}x_k.$$

- The collection $\{S^{-1}x_k\}$ is called the *dual frame* of $\{x_k\}$.
- In general frames are redundant, that is, there exist sequences $c = \{c_k\} \in \ell^2 \setminus \{0\}$ such that

$$T^*c = \sum c_k\, x_k = 0.$$

- For all sequences $\{c_k\}$ such that $x = \sum c_k x_k$,

$$\sum |c_n|^2 = \sum |\langle x, S^{-1}x_n\rangle|^2 + \sum |\langle x, S^{-1}x_n\rangle - c_n|^2.$$

- This implies that of all such c_k, $c_k = \langle x, S^{-1}x_k\rangle$ has the smallest norm.

Riesz bases

- A frame that ceases to be a frame upon the removal of one element is called an *exact frame*. An exact frame satisfies the following.
  - The analysis operator $T\colon H \to \ell^2$, $x \mapsto \{\langle x, x_k\rangle\}$ is an isomorphism.
  - The synthesis operator $T^*\colon \ell^2 \to H$ is injective, i.e., the frame is linearly independent.
  - The dual frame $\{S^{-1}x_k\}$ is biorthogonal to $\{x_k\}$.
  - For every sequence $c_k \in \ell^2$,

$$B^{-1}\sum |c_k|^2 \le \left\| \sum c_k\, x_k \right\|^2 \le A^{-1}\sum |c_k|^2.$$

- We say in this case that $\{x_k\}$ is a *Riesz basis* for H.

Some historical remarks

- The notion of a frame was first introduced in 1952 by Duffin and Schaeffer in the context of nonharmonic Fourier series.
- The paper referenced results of Paley and Wiener (1930) on basicity properties of sets

$$\mathcal{E}(\Lambda) = \{e^{2\pi i\lambda x} \colon \lambda \in \Lambda\}$$

in $L^2(-1/2, 1/2)$ where $\Lambda \subseteq \mathbb{R}$ is a perturbation of $\mathbb{Z}$.

Theorem (Duffin-Schaeffer)

The collection $\mathcal{E}(\Lambda)$ is a frame for $L^2(-\gamma, \gamma)$ for all $0 < \gamma < 1/2$ if the set Λ has uniform density 1*, that is, if there exist constants δ, $L > 0$ such that for all $n \in \mathbb{Z}$, $|\lambda_n - n| \leq L$ and for all $n \neq m$, $|\lambda_n - \lambda_m| \geq \delta$ (Λ is* uniformly discrete*).*

Beurling density

- Uniform density is a special case of *Beurling density* (Landau, 1967).
- Given $\Lambda \subseteq \mathbb{R}$, uniformly discrete, define for $r > 0$

$$n^+(r) = \sup_{x\in\mathbb{R}} |\Lambda \cap [x, x+r]| \text{ and } n^-(r) = \inf_{x\in\mathbb{R}} |\Lambda \cap [x, x+r]|.$$

- Let $D^+(\Lambda) = \lim_{r\to\infty} \frac{n^+(r)}{r}$ and $D^-(\Lambda) = \lim_{r\to\infty} \frac{n^-(r)}{r}$ denote the upper and lower Beurling densities of Λ.
- If $D^+(\Lambda) = D^-(\Lambda)$ then the common value $D(\Lambda)$ is the Beurling density of Λ.

- H. J. Landau (1967) showed, among other things, that, given a uniformly discrete subset Λ of $\mathbb{R}$, $\mathcal{E}(\Lambda)$ is a frame for $L^2(-\gamma,\gamma)$, then $D^-(\Lambda) \geq 2\gamma$, and that if $\mathcal{E}(\Lambda)$ is a Riesz basis for $L^2(-\gamma,\gamma)$, then $D^+(\Lambda) \leq 2\gamma$.
- As a consequence, if Λ has uniform density 1, frames $\mathcal{E}(\Lambda)$ must necessarily be overcomplete for $L^2(-\gamma,\gamma)$ whenever $0 < \gamma < 1/2$.
- Hence the Duffin and Schaeffer result describes necessarily redundant systems.

- Most of the work on frames in the 60s and 70s focused on properties of non-redundant systems of exponentials.
- Daubechies, Grossman, and Meyer (1986) connected explicitly the notion of a frame with the expansion of a function in terms of so-called *coherent states*.
- By this was meant the image of a single function by a fixed collection of transformations based on the Weil-Heisenberg group (Gabor expansions) and the affine group (wavelets).
- Having a frame as opposed to a Riesz basis is *necessary* in order to have stable expansions in terms of atoms with desirable properties.

The BLT and frames

Theorem

Given $\Omega, \Omega' > 0$ *and any function* $g(t) \in L^2(\mathbb{R})$, *define*

$$g_{n/\Omega, m\Omega'}(t) = g(t - n/\Omega)\, e^{2\pi i m \Omega' t}.$$

- *If* $\Omega' > \Omega$ *then* $\{g_{n/\Omega, m\Omega'}\}_{n,m \in \mathbb{Z}}$ *is incomplete in* $L^2(\mathbb{R})$ *and hence not a frame.*
- *If* $\Omega' < \Omega$ *and* $\{g_{n/\Omega, m\Omega'}\}_{n,m \in \mathbb{Z}}$ *is a frame, then it is not a Riesz basis, i.e., the frame is overcomplete.*
- *If* $\Omega' = \Omega$ *and* $\{g_{n/\Omega, m\Omega'}\}_{n,m \in \mathbb{Z}}$ *is a frame, then it is a Riesz basis, i.e., it is exact and*

$$\int_{-\infty}^{\infty} |t\, g(t)|^2\, dt \cdot \int_{-\infty}^{\infty} |\omega\, \hat{g}(\omega)|^2\, d\omega = \infty.$$

Frames in Communication Theory

- Consider the following model for transmission of a signal over a channel. Suppose that the signal of interest is the vector $x \in \mathbf{C}^d$, and that we have fixed a $d \times n$ frame matrix F.
- We store the vector by forming its frame coefficients $y = F^* x \in \mathbf{C}^n$, then transmit y over the channel.
- The received signal $\widehat{y}$ will be corrupted by quantization error and by noise, that is, $\widehat{y} = y + \epsilon$ where ϵ is a random vector in $\mathbb{R}^n$. $\epsilon = 0$ means perfect reconstruction is possible.
- The extent to which the original signal can be reconstructed from the noisy coefficients $\widehat{y}$ is a measure of the *robustness to noise* of the coding scheme.
- This question was investigated by Goyal, Kovačević and Kelner, 2001 for the above coding scheme.

Noise Reduction

Theorem

If the transmission error ϵ is modelled as zero-mean uncorrelated noise, the mean square error of the reconstructed signal is minimized if and only if the frame is uniform and tight.

- The range of the mapping $F^* : \mathbb{R}^d \to \mathbb{R}^n$ is only a d-dimensional subspace of $\mathbb{R}^n$, and the transmitted vector y sits in this subspace, but the distorted vector $\widehat{y}$ is unlikely to be.

- The reconstruction scheme

$$\widehat{x} = (FF^*)^{-1} F\widehat{y}$$

  solves the least-squares problem $\min_x \|F^*x - \widehat{y}\|_2$ with the minimum-norm solution $\widehat{x}$.
- In fact, this scheme projects the noise vector ϵ onto the range of F^* then reconstructs $\widehat{x}$ from those frame coefficients. Hence the noise power is automatically reduced.

Erasure errors

- Suppose in addition that the channel distorts the transmitted vector $\widehat{y}$ by erasing components at random.
- Robustness to this sort of distortion means maximizing the number of components that can be erased while still allowing reconstruction of the signal as accurately as possible from the remaining coefficients.

Definition

A frame $\mathcal{F} = \{x_k\}_{k=1}^n$ in $\mathbf{C}^d$ is *maximally robust to erasures* if the removal of any $l \leq n - d$ vectors from $\mathcal{F}$ leaves a frame. The *Spark* of an $d \times n$ matrix M is the size of the smallest linearly dependent subset of columns of M. Hence a frame with frame matrix F is maximally robust to erasures if $Spark(F^*) = d + 1$.

Compressive sensing.

- Problem: Recover an unknown vector $x \in \mathbb{R}^n$ from $d < n$ linear measurements under the assumption that x is *sparse*, i.e., for some $s \in \mathbf{N}$, x has no more than s nonzero elements (that is, that x is s-sparse).
- If we define a $d \times n$ matrix F to be the *measurement matrix*, then the problem becomes to recover x from $y = Fx$ under the assumption that x is s-sparse.
- Without the assumption of sparsity, the problem is clearly underdetermined and hence not solvable.

- A necessary and sufficient condition on F guaranteeing that the problem has a solution is given in the following theorem.

Theorem

The collection of s-sparse vectors in $\mathbb{R}^n$ is uniquely determined by the measurements Fx in the sense that for all s-sparse vectors x_1 and x_2, $Fx_1 = Fx_2$ implies $x_1 = x_2$ if and only if $Spark(F) > 2s$.

- If we think of the measurement matrix F as a frame matrix, that is, as a matrix whose columns form a frame for $\mathbb{R}^d$, then it is clear that in order for the problem to be meaningful, it is required that the frame be redundant.

Phaseless recovery

- Problem: Given a frame $\{f_i\}_{i=1}^n$ for $\mathbb{C}^d$, recover a vector $x \in \mathbb{R}^d$ from the magnitudes of its frame coefficients, i.e.,

$$\{|\langle x, f_i \rangle|\}_{i=1}^n.$$

- The initial breakthrough in this work is due to Balan, Casazza and Edidin, 2006.

Definition

A frame $\mathcal{F} = \{f_i\}_{i=1}^n$ for $\mathbb{C}^d$ is called *phase retrievable* if the mapping

$$\alpha \colon \mathbb{C}^d \to \mathbb{R}^n; \quad x \mapsto \{|\langle x, f_i \rangle|\}_{i=1}^n$$

is injective up to a constant phase factor.

The Real Case

Theorem

Assume that $\mathcal{F} = \{f_i\}_{i=1}^n$ is a frame for $\mathbb{R}^d$. Then

- *If $\mathcal{F}$ is phase retrievable for $\mathbb{R}^d$ then $n \geq 2d - 1$.*
- *If $n = 2d - 1$ then $\mathcal{F}$ is phase retrievable if and only if the frame matrix F corresponding to $\mathcal{F}$ has full Spark.*

Complex case

Theorem (Various authors)

Assume that $\mathcal{F} = \{f_i\}_{i=1}^n$ is a frame for $\mathbb{C}^d$. Then

- *(2013) If $\mathcal{F}$ is phase retrievable for $\mathbb{C}^d$ then*

$$n \geq 4d - 2 - 2b(n) + \begin{cases} 2 & \text{if } n \text{ is odd and } b = 3 \mod 4 \\ 1 & \text{if } n \text{ is odd and } b = 2 \mod 4 \\ 0 & \text{otherwise} \end{cases}$$

  where $b(n)$ denotes the number of 1s in the binary expansion of $n - 1$.

- *(2015) For any positive integer d, a phase retrievable frame $\mathcal{F}$ for $\mathbb{C}^d$ can be constructed that contains $n = 4d - 4$ vectors.*
- *(2015) If $n \geq 4d - 4$ then for generic frames, $\mathcal{F}$ is phase retrievable, and if $d = 2^k + 1$ and $n < 4d - 4$ then no frame $\mathcal{F}$ for $\mathbb{C}^d$ is phase retrievable.*

Feichtinger conjecture

Theorem (Marcus, Spielman and Srivastava, 2015)

Any uniform frame can be partitioned into a finite union of Riesz sequences.

- This theorem was conjectured by Feichtinger in the early 1990s.
- It was shown to be equivalent to several long-standing conjectures in operator theory, graph theory, mathematical physics, and signal processing, namely The Kadisson-Singer Conjecture (1959), the Paving Conjecture (1979), and the Bourgain-Tzafriri Conjecture (1991).
- The above named settled the Kadisson-Singer Conjecture in 2015.