Generalized sampling and infinite-dimensional compressed sensing

Ben Adcock
Department of Mathematics
Simon Fraser University

Joint work with Anders Hansen (University of Cambridge).

Outline of the talk

Introduction

Compressed sensing for finite-dimensional signals

Can we apply finite-dimensional compressed sensing techniques to infinite-dimensional signals?

Generalized sampling: signal reconstruction in arbitrary bases

Generalized sampling with compressed sensing

Introduction

Compressed sensing for finite-dimensional signals

Can we apply finite-dimensional compressed sensing techniques to infinite-dimensional signals?

Generalized sampling: signal reconstruction in arbitrary bases

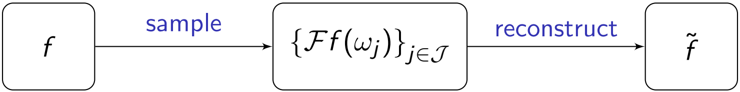
Generalized sampling with compressed sensing

Reconstructions from the Fourier transform

Fundamental problem – recover a image/signal f from samples of its Fourier transform (FT)

$$\mathcal{F}f(\omega) = \int_{\mathbb{R}^d} f(x) \mathrm{e}^{2\mathrm{i}\pi\omega\cdot x} \,\mathrm{d}x.$$

E.g. Magnetic Resonance Imaging (MRI).



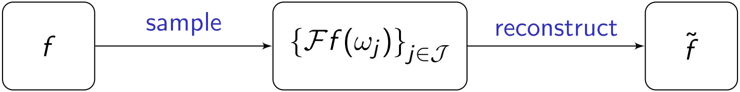
- Note that the sampling scheme is fixed and cannot be altered.
- Taking many samples is expensive/infeasible. Thus, one wants to reconstruct f using as few samples as possible.

Reconstructions from the Fourier transform

Fundamental problem – recover a image/signal f from samples of its Fourier transform (FT)

$$\mathcal{F}f(\omega) = \int_{\mathbb{R}^d} f(x) \mathrm{e}^{2\mathrm{i}\pi\omega\cdot x} \, \mathrm{d}x.$$

E.g. Magnetic Resonance Imaging (MRI).



- Note that the sampling scheme is fixed and cannot be altered.
- Taking many samples is expensive/infeasible. Thus, one wants to reconstruct f using as few samples as possible.

The Shannon Sampling Theorem

Let $f \in \mathrm{L}^2(\mathbb{R})$ with $\mathrm{supp}(f) \subseteq [-T, T]$. Then f is determined uniquely by $\{\mathcal{F}f(j\epsilon)\}_{j\in\mathbb{Z}}$, $\epsilon \leq \frac{1}{2T}$. Specifically,

$$f(x) = \epsilon \sum_{j\in\mathbb{Z}} \mathcal{F}f(j\epsilon)\mathrm{e}^{2\pi \mathrm{i}\epsilon jx}.$$

However, in practice, we cannot access or process all samples of f. Thus, Shannon's Theorem gives rise to the approximation

$$f_N(x) = \epsilon \sum_{j=-N}^{N-1} \mathcal{F}f(j\epsilon)\mathrm{e}^{2\pi \mathrm{i}\epsilon jx}.$$

- In other words, the partial Fourier series of f.

The Shannon Sampling Theorem

Let $f \in \mathrm{L}^2(\mathbb{R})$ with $\mathrm{supp}(f) \subseteq [-T, T]$. Then f is determined uniquely by $\{\mathcal{F}f(j\epsilon)\}_{j\in\mathbb{Z}}$, $\epsilon \leq \frac{1}{2T}$. Specifically,

$$f(x) = \epsilon \sum_{j\in\mathbb{Z}} \mathcal{F}f(j\epsilon)\mathrm{e}^{2\pi\mathrm{i}\epsilon jx}.$$

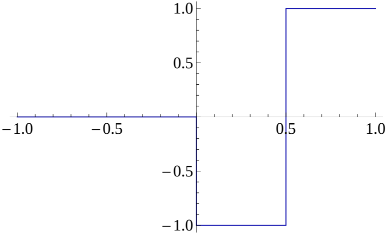
However, in practice, we cannot access or process all samples of f. Thus, Shannon's Theorem gives rise to the approximation

$$f_N(x) = \epsilon \sum_{j=-N}^{N-1} \mathcal{F}f(j\epsilon)\mathrm{e}^{2\pi\mathrm{i}\epsilon jx}.$$

- In other words, the partial Fourier series of f.

Fourier series reconstruction

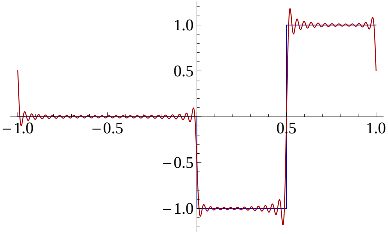
Let $\epsilon = \frac{1}{2}$, $N = 50$:



The function $f(x)$

Fourier series reconstruction

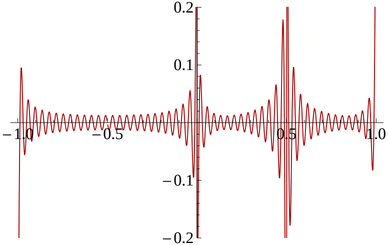
Let $\epsilon = \frac{1}{2}$, $N = 50$:



The functions $f(x)$ and $f_N(x)$

Fourier series reconstruction

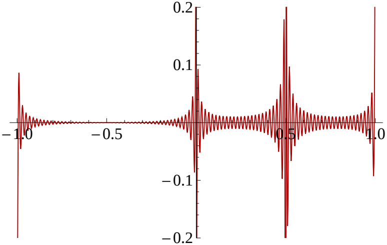
Let $\epsilon = \frac{1}{2}$, $N = 50$:



The error $f(x) - f_N(x)$

Fourier series reconstruction

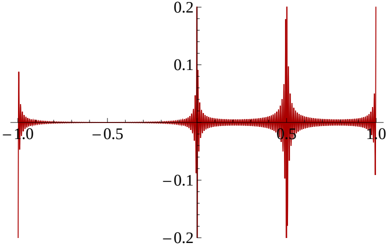
Let $\epsilon = \frac{1}{2}$, $N = 100$:



The error $f(x) - f_N(x)$

Fourier series reconstruction

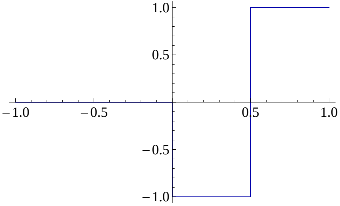
Let $\epsilon = \frac{1}{2}$, $N = 200$:



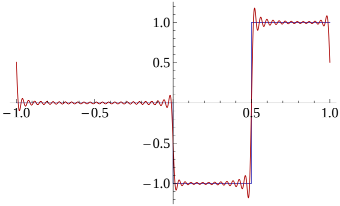
The error $f(x) - f_N(x)$

The coefficients $\mathcal{F}f(j\epsilon)$ decay very slowly as $|j| \to \infty$.

Other ways to reconstruct f



$f(x)$



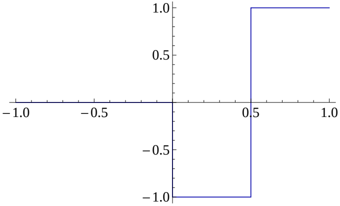
$f_N(x)$

The function f is very poorly represented by f_N. However, f has an extremely good representation in the Haar wavelet basis $\{\phi_j\}_{j\in\mathbb{N}}$.

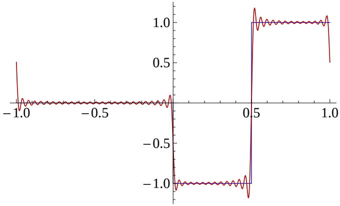
Question:

Given this additional information, is there a better way to recover f than via the Fourier series f_N?

Other ways to reconstruct f



$f(x)$



$f_N(x)$

The function f is very poorly represented by f_N. However, f has an extremely good representation in the Haar wavelet basis $\{\phi_j\}_{j\in\mathbb{N}}$.

Question:

Given this additional information, is there a better way to recover f than via the Fourier series f_N?

Other ways to reconstruct f

More generally, suppose that f is well represented in an arbitrary basis $\{\phi_j\}_{j\in\mathbb{N}}$, i.e.

(i) $f = \sum_{j=1}^{\infty} \alpha_j \phi_j$ with $\alpha_j \to 0$ rapidly,

(ii) f is k-sparse in $\{\phi_j\}_{j\in\mathbb{N}}$, i.e.

$$|\{j : \alpha_j \neq 0\}| = k \ll \infty.$$

Questions:

1. How can we recover f in terms of $\{\phi_j\}_{j\in\mathbb{N}}$ from its samples?

2. If f is actually sparse in $\{\phi_j\}_{j\in\mathbb{N}}$, can we subsample? Specifically, can we recover f using only $\mathcal{O}(k)$ of its samples?

Other ways to reconstruct f

More generally, suppose that f is well represented in an arbitrary basis $\{\phi_j\}_{j\in\mathbb{N}}$, i.e.

(i) $f = \sum_{j=1}^{\infty} \alpha_j \phi_j$ with $\alpha_j \to 0$ rapidly,

(ii) f is k-sparse in $\{\phi_j\}_{j\in\mathbb{N}}$, i.e.

$$|\{j : \alpha_j \neq 0\}| = k \ll \infty.$$

Questions:

1. How can we recover f in terms of $\{\phi_j\}_{j\in\mathbb{N}}$ from its samples?

2. If f is actually sparse in $\{\phi_j\}_{j\in\mathbb{N}}$, can we subsample? Specifically, can we recover f using only $\mathcal{O}(k)$ of its samples?

Introduction

Compressed sensing for finite-dimensional signals

Can we apply finite-dimensional compressed sensing techniques to infinite-dimensional signals?

Generalized sampling: signal reconstruction in arbitrary bases

Generalized sampling with compressed sensing

Sparse finite-dimensional signals

Suppose that we model signals as vectors in $\mathbb{C}^N$.

Let $\{\phi_j\}_{j=1}^N$, $\{\psi_j\}_{j=1}^N$ be two orthonormal bases of $\mathbb{C}^N$, and write

$$f = \sum_{j=1}^{N} \alpha_j \phi_j.$$

- Suppose that f is k-sparse in $\{\phi_j\}_{j=1}^N$, i.e.

$$|\{j : \alpha_j \neq 0\}| = k.$$

Let f be sampled via

$$\hat{f}_j = \langle f, \psi_j \rangle, \quad j = 1, \ldots, N.$$

We want to recover f using only $m = \mathcal{O}(k)$ of its samples $\hat{f}_j$.

Sparse finite-dimensional signals

Suppose that we model signals as vectors in $\mathbb{C}^N$.

Let $\{\phi_j\}_{j=1}^N$, $\{\psi_j\}_{j=1}^N$ be two orthonormal bases of $\mathbb{C}^N$, and write

$$f = \sum_{j=1}^{N} \alpha_j \phi_j.$$

- Suppose that f is k-sparse in $\{\phi_j\}_{j=1}^N$, i.e.

$$|\{j : \alpha_j \neq 0\}| = k.$$

Let f be sampled via

$$\hat{f}_j = \langle f, \psi_j \rangle, \quad j = 1, \ldots, N.$$

We want to recover f using only $m = \mathcal{O}(k)$ of its samples $\hat{f}_j$.

The measurement matrix

Define the change-of-basis matrix

$$U = \begin{pmatrix} u_{11} & \cdots & u_{1N} \\ \vdots & \ddots & \vdots \\ u_{N1} & \cdots & u_{NN} \end{pmatrix} \in \mathbb{C}^{N\times N}, \qquad u_{ij} = \langle \phi_j, \psi_i \rangle.$$

- Note that U is unitary: $U^* U = I$.

Standard compressed sensing setup: form an $m \times N$ measurement matrix A by drawing rows of U uniformly at random, i.e.

$$A = P_\Omega U,$$

where P_Ω is a projection matrix and $\Omega \subseteq \{1, \ldots, N\}$, $|\Omega| = m$ is chosen uniformly at random.

The measurement matrix

Define the change-of-basis matrix

$$U = \begin{pmatrix} u_{11} & \cdots & u_{1N} \\ \vdots & \ddots & \vdots \\ u_{N1} & \cdots & u_{NN} \end{pmatrix} \in \mathbb{C}^{N \times N}, \qquad u_{ij} = \langle \phi_j, \psi_i \rangle.$$

- Note that U is unitary: $U^* U = I$.

Standard compressed sensing setup: form an $m \times N$ measurement matrix A by drawing rows of U uniformly at random, i.e.

$$A = P_\Omega U,$$

where P_Ω is a projection matrix and $\Omega \subseteq \{1, \ldots, N\}$, $|\Omega| = m$ is chosen uniformly at random.

Signal reconstruction with compressed sensing

Let $\hat{f} = (\hat{f}_1, \ldots, \hat{f}_N)^\top$. We solve:

$$\min_{\eta \in \mathbb{C}^N} \|\eta\|_{l^1} \quad \text{subject to} \quad P_\Omega U \eta = P_\Omega \hat{f}. \qquad (\star)$$

- Define the incoherence

$$\mu(U) = \max_{i,j} |u_{ij}|.$$

- Then, with high probability, $\alpha = (\alpha_1, \ldots, \alpha_N)^\top$ is the unique minimizer of $(\star)$, provided m is proportional to

$$k \cdot N \cdot \mu(U)^2 \cdot \log N.$$

Signal reconstruction with compressed sensing

Let $\hat{f} = (\hat{f}_1, \ldots, \hat{f}_N)^\top$. We solve:

$$\min_{\eta \in \mathbb{C}^N} \|\eta\|_{l^1} \ \text{ subject to } \ P_\Omega U \eta = P_\Omega \hat{f}. \qquad (\star)$$

- Define the incoherence

$$\mu(U) = \max_{i,j} |u_{ij}|.$$

- Then, with high probability, $\alpha = (\alpha_1, \ldots, \alpha_N)^\top$ is the unique minimizer of $(\star)$, provided m is proportional to

$$k \cdot N \cdot \mu(U)^2 \cdot \log N.$$

Example

Let $u_{ij} = \frac{1}{\sqrt{N}} \mathrm{e}^{2\pi \mathrm{i} ij/N}$

- i.e. U is the DFT matrix.

In this case, $\mu(U) = \frac{1}{\sqrt{N}}$.

- Known as perfect incoherence.
- Hence, we require only $m = \mathcal{O}(k \log N)$ samples to recover a k-sparse $\alpha \in \mathbb{C}^N$.

Introduction

Compressed sensing for finite-dimensional signals

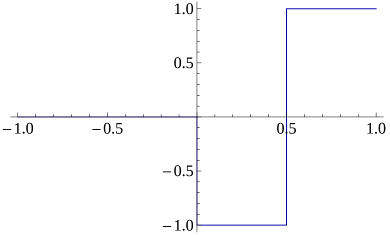
Can we apply finite-dimensional compressed sensing techniques to infinite-dimensional signals?

Generalized sampling: signal reconstruction in arbitrary bases

Generalized sampling with compressed sensing

Example revisited

Recall the example



- f is very sparse in the Haar wavelet basis.

We want to recover f by subsampling from its Fourier coefficients

$$\mathcal{F}f(j\epsilon), \quad j \in \mathbb{Z}.$$

To apply standard CS techniques, we need to discretize.

The standard approach...

Let

$$\hat{f} = (\mathcal{F}f(-N\epsilon), \ldots, \mathcal{F}f((N-1)\epsilon))^\top \in \mathbb{C}^{2N},$$

and suppose that $V_{\mathrm{df}}, V_{\mathrm{dw}} \in \mathbb{C}^{2N \times 2N}$ are the matrices of the DFT and DWT respectively. We set

$$U = V_{\mathrm{df}} V_{\mathrm{dw}}^{-1},$$

and solve

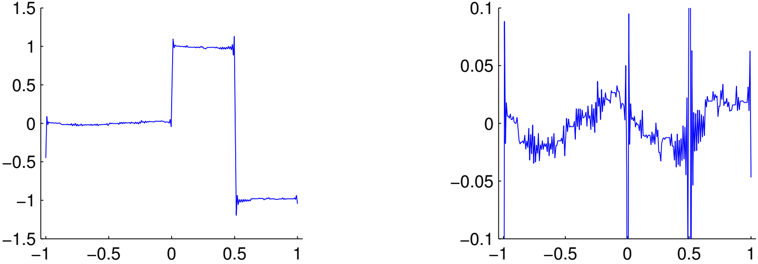
$$\min_{\eta \in \mathbb{C}^{2N}} \|\eta\|_{l^1} \quad \text{subject to} \quad P_\Omega U \eta = P_\Omega \hat{f}, \qquad (\star)$$

where $\Omega \subset \{1, \ldots, 2N\}$, $|\Omega| = m$ is chosen uniformly at random.

- The great advantage of this approach is that $V_{\mathrm{df}}, V_{\mathrm{dw}}$ can be manipulated easily. Hence $(\star)$ can be solved efficiently.

The standard approach...may not work

Let $N = 256$, $m = 130$ and suppose that $f_{N,m}(x) = \sum_{j=1}^{2N} \eta_j \phi_j$:



$f_{N,m}(x)$ $\qquad$ $f(x) - f_{N,m}(x)$

Even though we used nearly half to total number of Fourier coefficients, we still recover f very poorly.

Explanation: the DFT destroys sparsity

Note that

$$V_{\mathrm{df}}^{-1} : \{\text{Fourier coeffics.}\} \to \{\text{grid values}\}.$$

However, this mapping is only approximate. In particular, if

$$\chi(f) = V_{\mathrm{df}}^{-1} \hat{f},$$

then $\chi(f)$ is the vector of grid values of the approximation

$$f_N(x) = \epsilon \sum_{j=-N}^{N-1} \mathcal{F}f(j\epsilon)\mathrm{e}^{2\pi \mathrm{i} \epsilon j x}.$$

Explanation: the DFT destroys sparsity

Recall that we solve

$$\min_{\eta \in \mathbb{C}^{2N}} \|\eta\|_{l^1} \quad \text{subject to} \quad P_\Omega U \eta = P_\Omega \hat{f}. \qquad (\star)$$

If $\xi = V_{\mathrm{dw}} \chi(f)$, then the RHS is given by

$$P_\Omega V_{\mathrm{df}} V_{\mathrm{dw}}^{-1} \xi.$$

Thus, for $(\star)$ to have a sparse solution, we require $\xi = V_{\mathrm{dw}} \chi(f)$ to be a sparse vector.

- This cannot happen: ξ is the vector of Haar wavelet coefficients of the truncated Fourier series f_N.
- f_N is smooth, and therefore cannot be sparse in Haar wavelets.

Another approach

One could consider replacing the DFT by the continuous Fourier transform, i.e. we let

$$U = \begin{pmatrix} u_{-N,1} & \cdots & u_{-N,2N} \\ \vdots & \ddots & \vdots \\ u_{N-1,1} & \cdots & u_{N-1,2N} \end{pmatrix}, \qquad u_{i,j} = \mathcal{F}(\phi_j)(i\epsilon),$$

and solve

$$\min_{\eta \in \mathbb{C}^{2N}} \|\eta\|_{l^1} \text{ subject to } P_\Omega U\eta = P_\Omega \hat{f}, \qquad (\star)$$

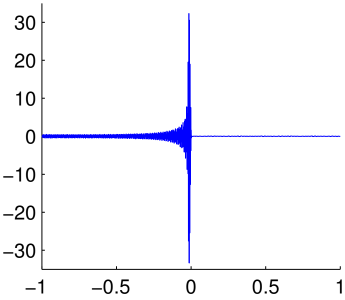
once more.

Another approach

Example: consider

$$f = \sum_{j=1}^{2N} \alpha_j \phi_j,$$

where $2N = 768$ and $k = 5$. Set $m = 760$:



The error $f(x) - f_{N,m}(x)$

Explanation: loss of unitary structure

The systems $\{\phi_j\}_{j\in\mathbb{N}}$ and $\{\sqrt{\epsilon}\exp(2\pi\mathrm{i}\epsilon j\cdot)\}_{j\in\mathbb{Z}}$ are orthonormal bases. Therefore the infinite change-of-basis matrix

$$U = \begin{pmatrix} \vdots & \vdots & \iddots \\ u_{-1,1} & u_{-1,2} & \cdots \\ u_{0,1} & u_{0,2} & \cdots \\ u_{1,1} & u_{1,2} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}, \qquad u_{i,j} = \sqrt{\epsilon}\int_0^1 \phi_j(x)\mathrm{e}^{-2\pi\mathrm{i}\epsilon ix}\,\mathrm{d}x,$$

is a unitary operator $l^2(\mathbb{N}) \to l^2(\mathbb{N})$.

However, this is not the case for the finite matrix formed from the first N basis functions $\{\phi_j\}_{j=1}^{2N}$ and $\{\sqrt{\epsilon}\exp(2\pi\mathrm{i}\epsilon j\cdot)\}_{j=-N}^{N-1}$.

Summary

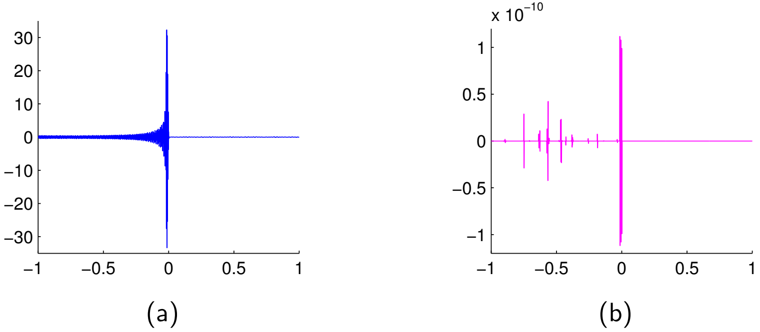
The two most obvious discretizations failed because they destroyed the key structure of the problem, i.e. sparsity and unitarity.

This suggests an alternative approach. We should:

Formulate the problem in infinite dimensions, and then seek to discretize in a manner that preserves as much of the key infinite-dimensional structure as possible.

Numerical example

Recall the example $f = \sum_{j=1}^{2N} \alpha_j \phi_j$, where $2N = 768$ and $k = 5$:



(a) reconstruction with $m = 760$ via the previous finite-dimensional approach (b) reconstruction using the new approach using $m = 50$ samples.

Introduction

Compressed sensing for finite-dimensional signals

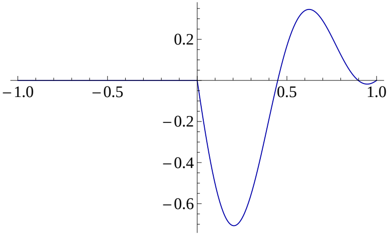
Can we apply finite-dimensional compressed sensing techniques to infinite-dimensional signals?

Generalized sampling: signal reconstruction in arbitrary bases

Generalized sampling with compressed sensing

Example problem

Ignore sparsity for the moment and consider an arbitrary signal:



If $\{\phi_j\}_{j\in\mathbb{N}}$ are, for example, Haar wavelets, we may write

$$f = \sum_{j=1}^{\infty} \alpha_j \phi_j.$$

Key point: recovering $\{\alpha_j\}$ is an infinite-dimensional problem in a Hilbert space, as opposed to $\mathbb{C}^N$.

Hilbert space formulation

Suppose that H is a separable Hilbert space with norm $\|\cdot\|$ and inner product $\langle \cdot, \cdot \rangle$.

- Let $\{\psi_j\}_{j\in\mathbb{N}}$ be an o.n. basis of H (the sampling basis).
- Let $\{\phi_j\}_{j\in\mathbb{N}}$ be an o.n. basis of H (the reconstruction basis).

Let $f = \sum_{j=1}^{\infty} \alpha_j \phi_j \in \mathrm{H}$ and

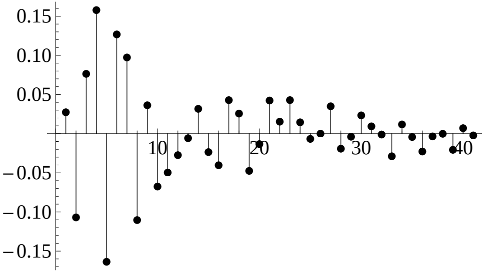
$$\hat{f}_j = \langle f, \psi_j \rangle, \quad j \in \mathbb{N}, \qquad \text{(samples of } f\text{)}.$$

Problem: reconstruct f in terms of $\{\phi_j\}_{j\in\mathbb{N}}$ from its first N samples $\hat{f}_1, \ldots, \hat{f}_N$.

Key consideration I: good recovery of coefficients

Suppose that we have a mapping

$$\mathcal{L} : \{\hat{f}_1, \ldots, \hat{f}_N\} \longmapsto \{\tilde{\alpha}_1, \ldots, \tilde{\alpha}_M\}.$$

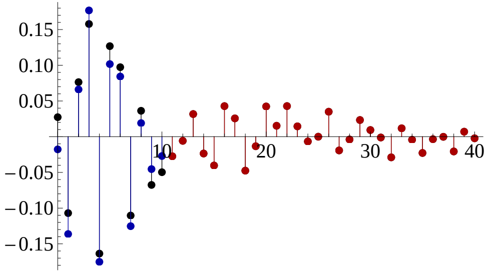


The coefficients $\alpha_1, \alpha_2, \ldots$.

Key consideration I: good recovery of coefficients

Suppose that we have a mapping

$$\mathcal{L} : \{\hat{f}_1, \ldots, \hat{f}_N\} \longmapsto \{\tilde{\alpha}_1, \ldots, \tilde{\alpha}_M\}.$$



The coefficients $\alpha_1, \alpha_2, \ldots$ and approximate coefficients $\tilde{\alpha}_1, \ldots, \tilde{\alpha}_M$.

Key consideration I: good recovery of coefficients

Suppose that we have a mapping

$$\mathcal{L} : \{\hat{f}_1, \ldots, \hat{f}_N\} \longmapsto \{\tilde{\alpha}_1, \ldots, \tilde{\alpha}_M\}.$$

Error equation:

$$f - \sum_{j=1}^{M} \tilde{\alpha}_j \phi_j = \sum_{j=1}^{M} (\alpha_j - \tilde{\alpha}_j)\phi_j + \sum_{j=M+1}^{\infty} \alpha_j \phi_j$$

total error — regularization error — truncation error

Its important that

regularization error $\approx$ truncation error, (quasi-optimality).

Key consideration II: numerical stability

The mapping $\mathcal{L} : \{\hat{f}_1, \dots, \hat{f}_N\} \longmapsto \{\tilde{\alpha}_1, \dots, \tilde{\alpha}_M\}$ should be numerically stable, i.e. the condition number

$$\|\mathcal{L}\|\|\mathcal{L}^{-1}\| \ll \infty,$$

to avoid large errors due to, for example,

- round-off,
- noise,
- jitter.

Infinite-dimensional formulation

Note that

$$\begin{pmatrix} u_{11} & u_{12} & \cdots \\ u_{21} & u_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} \hat{f}_1 \\ \hat{f}_2 \\ \vdots \end{pmatrix}, \qquad u_{ij} = \langle \phi_j, \psi_i \rangle.$$

i.e. there is an invertible mapping $U : l^2(\mathbb{N}) \to l^2(\mathbb{N})$ with

$$U^{-1}\{\hat{f}_j\}_{j\in\mathbb{N}} = \{\alpha_j\}_{j\in\mathbb{N}}$$

We cannot compute U^{-1}. However,

- The operator $U = \{u_{ij}\}_{i,j\in\mathbb{N}} : l^2(\mathbb{N}) \to l^2(\mathbb{N})$ is unitary.
- Thus, we now seek a discretization that preserves this structure.

Infinite-dimensional formulation

Note that

$$\begin{pmatrix} u_{11} & u_{12} & \cdots \\ u_{21} & u_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} \hat{f}_1 \\ \hat{f}_2 \\ \vdots \end{pmatrix}, \qquad u_{ij} = \langle \phi_j, \psi_i \rangle.$$

i.e. there is an invertible mapping $U : l^2(\mathbb{N}) \to l^2(\mathbb{N})$ with

$$U^{-1}\{\hat{f}_j\}_{j\in\mathbb{N}} = \{\alpha_j\}_{j\in\mathbb{N}}$$

We cannot compute U^{-1}. However,

- The operator $U = \{u_{ij}\}_{i,j\in\mathbb{N}} : l^2(\mathbb{N}) \to l^2(\mathbb{N})$ is unitary.
- Thus, we now seek a discretization that preserves this structure.

Sections of U

Let $P_N : l^2(\mathbb{N}) \to l^2(\mathbb{N})$ be the orthogonal projection.

$$P_N\{\beta_1, \beta_2, \ldots\} = \{\beta_1, \ldots, \beta_N, 0, 0, \ldots\}.$$

The finite section

$$P_N U P_N = \begin{pmatrix} u_{11} & \cdots & u_{1N} \\ \vdots & \ddots & \vdots \\ u_{N1} & \cdots & u_{NN} \end{pmatrix} \in \mathbb{C}^{N\times N},$$

will not be unitary in general.

However, the uneven section

$$P_N U P_M = \begin{pmatrix} u_{11} & \cdots & u_{1M} \\ \vdots & \ddots & \vdots \\ u_{N1} & \cdots & u_{NM} \end{pmatrix} \in \mathbb{C}^{N\times M},$$

will be unitary in the limit $N \to \infty$ for fixed M.

Generalized sampling

Let $\tilde{\alpha} = \{\tilde{\alpha}_1, \ldots, \tilde{\alpha}_M, 0, 0, \ldots\}$ be defined by

$$P_M U^* P_N U P_M \tilde{\alpha} = P_M U^* P_N \hat{f}.$$

Theorem (BA, Hansen)

For each $M \in \mathbb{N}$, there exists a $N_0 \in \mathbb{N}$ such that $\tilde{\alpha}$ exists uniquely for all $N \geq N_0$ and satisfies

$$\|\alpha - \tilde{\alpha}\|_{l^2} \leq \frac{1}{\sqrt{1 - C_{N,M}}} \|\alpha - P_M \alpha\|_{l^2},$$

where $C_{N,M} = \|P_M - P_M U^* P_N U P_M\|_{l^2 \to l^2}$.

The quantity $C_{N,M}$

The quantity

$$C_{N,M} = \|P_M - P_M U^* P_N U P_M\|_{l^2 \to l^2}$$

measures how close to unitary $P_N U P_M$ is. Thus, aside from the error $\|\alpha - \tilde{\alpha}\|$, $C_{N,M}$ also determines stability.

Key point:

- $C_{N,M} \to 0$ as $N \to \infty$ for fixed M.
- Thus, we can ensure accuracy of the reconstructed coefficients and stability by varying N suitably.
- Note that we typically have $M < N$, i.e. we reconstruct fewer coefficients $\tilde{\alpha}_1, \ldots, \tilde{\alpha}_M$ than the number of samples.

The stable sampling rate

Define the stable sampling rate:

$$\Theta(M;\theta) = \min\left\{N \in \mathbb{N} : C_{N,M} < \theta\right\}, \quad \theta \in (0,1).$$

For given M, setting $N \geq \Theta(M;\theta)$ ensures

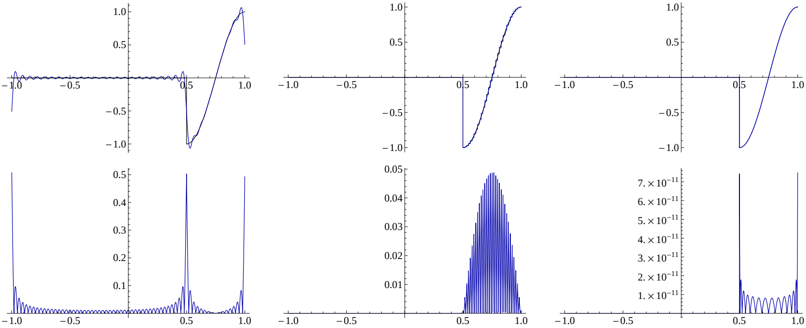
1. Existence and uniqueness of $\tilde{\alpha}$.
2. Numerical stability: $\|\mathcal{L}\|\|\mathcal{L}^{-1}\| \leq \frac{1}{\sqrt{1-\theta}}$.
3. Quasi-optimality: $\|\alpha - \tilde{\alpha}\|_{l^2} \leq \frac{1}{\sqrt{1-\theta}}\|\alpha - P_M\alpha\|_{l^2}$.

Examples:

- Fourier sampling, Haar wavelets: $\Theta(M;\theta) = \mathcal{O}(M)$.
- Fourier sampling, piecewise polynomials: $\Theta(M;\theta) = \mathcal{O}\left(M^2\right)$.

Example

Consider $f(x) = \cos 2\pi x \mathbb{I}_{[\frac{1}{2},1]}(x)$ with $N = 128$ Fourier coefficients:



Fourier

Haar ($M = 64$)
($C_{N,M} = 0.586$)

Polynomial ($M = 13$)
($C_{N,M} = 0.536$)

Introduction

Compressed sensing for finite-dimensional signals

Can we apply finite-dimensional compressed sensing techniques to infinite-dimensional signals?

Generalized sampling: signal reconstruction in arbitrary bases

Generalized sampling with compressed sensing

Exploiting sparsity

Now suppose that $f \in \mathrm{H}$ is k-sparse in $\{\phi_j\}_{j \in \mathbb{N}}$, i.e.

$$f = \sum_{j=1}^{M} \alpha_j \phi_j, \qquad |\{j : \alpha_j \neq 0\}| = k \ll M.$$

- M is the largest index of a nonzero coefficient α_j.

We can certainly use generalized sampling to recover f exactly. However, we require at least $N = M$ samples.

Questions:

1. How can we take advantage of the sparsity?
2. How many samples do we require, and how does this depend on k and M?

Exploiting sparsity

Now suppose that $f \in \mathrm{H}$ is k-sparse in $\{\phi_j\}_{j \in \mathbb{N}}$, i.e.

$$f = \sum_{j=1}^{M} \alpha_j \phi_j, \qquad |\{j : \alpha_j \neq 0\}| = k \ll M.$$

- M is the largest index of a nonzero coefficient α_j.

We can certainly use generalized sampling to recover f exactly. However, we require at least $N = M$ samples.

Questions:

1. How can we take advantage of the sparsity?
2. How many samples do we require, and how does this depend on k and M?

Notation

Define

- k – the sparsity,
- M – range of nonzero coefficients α_j of f, i.e.

$$\{j : \alpha_j \neq 0\} \subseteq \{1, \ldots, M\}.$$

- m – number of samples used,
- N – range of indices from which m samples are drawn,
- Ω – subset of $\{1, \ldots, N\}$ of cardinality $|\Omega| = m$ chosen uniformly at random.

Generalized sampling with compressed sensing

We solve

$$\inf_{\beta\in\mathbb{C}^M} \|\beta\|_1 \text{ subject to } P_\Omega U P_M \beta = P_\Omega \hat{f}.$$

Note that $P_\Omega U P_M = P_\Omega P_N U P_M$ is the result of randomly selecting rows of the uneven section $P_N U P_M$.

- e.g. with $k = 3$, $M = 6$, $m = 4$ and $N = 8$:

$$\begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} & u_{15} & u_{16} \\ u_{21} & u_{22} & u_{23} & u_{24} & u_{25} & u_{26} \\ u_{31} & u_{32} & u_{33} & u_{34} & u_{35} & u_{36} \\ u_{41} & u_{42} & u_{43} & u_{44} & u_{45} & u_{46} \\ u_{51} & u_{52} & u_{53} & u_{54} & u_{55} & u_{56} \\ u_{61} & u_{62} & u_{63} & u_{64} & u_{65} & u_{66} \\ u_{71} & u_{72} & u_{73} & u_{74} & u_{75} & u_{76} \\ u_{81} & u_{82} & u_{83} & u_{84} & u_{85} & u_{86} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ 0 \\ 0 \\ \alpha_4 \\ \alpha_5 \\ 0 \end{pmatrix} = \begin{pmatrix} \hat{f}_1 \\ \hat{f}_2 \\ \hat{f}_3 \\ \hat{f}_4 \\ \hat{f}_5 \\ \hat{f}_6 \\ \hat{f}_7 \\ \hat{f}_8 \end{pmatrix}.$$

Generalized sampling with compressed sensing

We solve

$$\inf_{\beta \in \mathbb{C}^M} \|\beta\|_1 \text{ subject to } P_\Omega U P_M \beta = P_\Omega \hat{f}.$$

Note that $P_\Omega U P_M = P_\Omega P_N U P_M$ is the result of randomly selecting rows of the uneven section $P_N U P_M$.

- e.g. with $k = 3$, $M = 6$, $m = 4$ and $N = 8$:

$$\begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} & u_{15} & u_{16} \\ u_{21} & u_{22} & u_{23} & u_{24} & u_{25} & u_{26} \\ u_{31} & u_{32} & u_{33} & u_{34} & u_{35} & u_{36} \\ u_{41} & u_{42} & u_{43} & u_{44} & u_{45} & u_{46} \\ u_{51} & u_{52} & u_{53} & u_{54} & u_{55} & u_{56} \\ u_{61} & u_{62} & u_{63} & u_{64} & u_{65} & u_{66} \\ u_{71} & u_{72} & u_{73} & u_{74} & u_{75} & u_{76} \\ u_{81} & u_{82} & u_{83} & u_{84} & u_{85} & u_{86} \end{pmatrix} \begin{pmatrix} \alpha_1 \\ 0 \\ 0 \\ \alpha_4 \\ \alpha_5 \\ 0 \end{pmatrix} = \begin{pmatrix} \hat{f}_1 \\ \hat{f}_2 \\ \hat{f}_3 \\ \hat{f}_4 \\ \hat{f}_5 \\ \hat{f}_6 \\ \hat{f}_7 \\ \hat{f}_8 \end{pmatrix}.$$

Infinite-dimensional incoherence

Define the incoherence

$$\mu(U) = \max_{i,j \in \mathbb{N}} |u_{ij}|.$$

Note that μ depends only on the infinite matrix U, and not N, M.

We shall also write

$$\|U\|_{\mathrm{mr}} = \sqrt{\sup_{i \in \mathbb{N}} \sum_{j \in \mathbb{N}} |u_{ij}|^2}.$$

Balancing property

The stable sampling rate controls how to take the uneven section of U in GS. For GS with CS, we define the balancing property:

Definition

Let $M \in \mathbb{N}$, $\epsilon > 0$ and $k \in \mathbb{N}$ be given. Then $N, m \in \mathbb{N}$ satisfy the weak balancing property with respect to M, ϵ and k if

$$\|P_M U^* P_N U P_M - P_M\| \leq \left[4\sqrt{\log_2\left(4N\sqrt{k}/m\right)}\right]^{-1},$$

and

$$\max_{\substack{\Gamma \subseteq \{1,\ldots,M\} \\ |\Gamma| = k}} \|P_M P_\Gamma^\perp U^* P_N U P_\Gamma\|_{\mathrm{mr}} \leq \frac{1}{8\sqrt{k}}.$$

Main result

Recall that we solve

$$\inf_{\beta \in \mathbb{C}^M} \|\beta\|_1 \text{ subject to } P_\Omega U P_M \beta = P_\Omega \hat{f}. \qquad (\star)$$

Theorem (BA, Hansen)

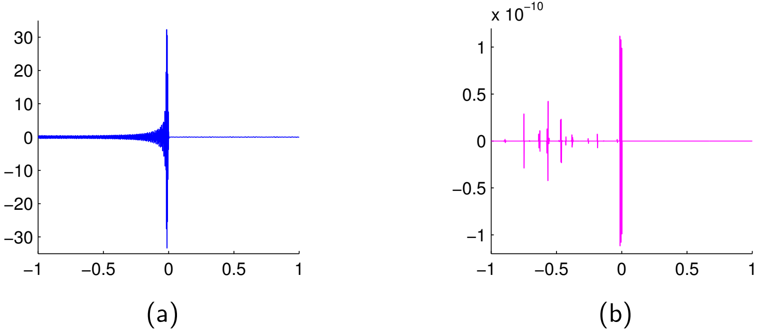
Let $\epsilon > 0$, and suppose that $N, m \in \mathbb{N}$ satisfy the weak balancing property with respect to M, ϵ and k. Suppose also that

$$m \geq C \cdot k \cdot N \cdot \mu(U)^2 \cdot (1 - \log \epsilon) \cdot \log\left(\frac{MN\sqrt{k}}{m}\right),$$

for some universal constant C. Then, with probability $p > 1 - \epsilon$, there exists a unique solution to $(\star)$ and this solution coincides with $\alpha = \{\alpha_1, \alpha_2, \ldots\}$. If $m = N$ then this holds with $p = 1$.

Numerical example I

Recall the example $f = \sum_{j=1}^{2M} \alpha_j \phi_j$, where $2M = 768$ and $k = 5$:



(a) reconstruction with $m = 760$ via the finite-dimensional approach (b) reconstruction using the new approach using $m = 50$ samples.

This demonstrates the importance of the balancing property:

- (a): $N = 2M$, i.e. a finite section.
- (b): $N = 1351 > 2M$, i.e. an uneven section.

Numerical example II

Most signals f are not exactly sparse, but compressible, i.e.

$$f = g + h,$$

where g is sparse and h is small in norm.

- When subsampling, we aim to recover f up to some error determined by $\|h\|$.

Example: let $|\{j : \alpha_j \neq 0\}| = 25$ and

$$f(x) = \sum_{j=1}^{200} \alpha_j \phi_j(x) + \chi_{[\frac{1}{2},\frac{9}{16}]}(x) \cos 2\pi x, \quad x \in [-1, 1].$$

N	(a)	(b)	(c) (avg. 20 trials)
601	1.43e0	4.74e-5	4.73e-5 ($m = 230$)
1201	8.5e-1	2.36e-5	2.38e-5 ($m = 460$)

Error for (a) the partial Fourier series, (b) generalized sampling and (c) generalized sampling with compressed sensing.

Numerical example II

Most signals f are not exactly sparse, but compressible, i.e.

$$f = g + h,$$

where g is sparse and h is small in norm.

- When subsampling, we aim to recover f up to some error determined by $\|h\|$.

Example: let $|\{j : \alpha_j \neq 0\}| = 25$ and

$$f(x) = \sum_{j=1}^{200} \alpha_j \phi_j(x) + \chi_{[\frac{1}{2},\frac{9}{16}]}(x) \cos 2\pi x, \quad x \in [-1, 1].$$

N	(a)	(b)	(c) (avg. 20 trials)
601	1.43e0	4.74e-5	4.73e-5 ($m = 230$)
1201	8.5e-1	2.36e-5	2.38e-5 ($m = 460$)

Error for (a) the partial Fourier series, (b) generalized sampling and (c) generalized sampling with compressed sensing.

Conclusions

1. Many important signals are best modelled as elements of infinite-dimensional function spaces.

2. Standard CS tools are based on finite-dimensional vector spaces, and therefore are not always well suited for infinite-dimensional problems.

3. Generalized sampling allows one to work directly with the infinite-dimensional model, and can be combined with standard l^1 minimization to achieve subsampling.

4. Key idea is to preserve the underlying infinite-dimensional structure: unitarity and sparsity.

Current work – compressed sensing

Recall that we require

$$m \geq C \cdot k \cdot N \cdot \mu(U)^2 \cdot (1 - \log \epsilon) \cdot \log \left(\frac{MN\sqrt{k}}{m} \right).$$

Thus, we can only subsample if $N\mu(U)^2$ is small.

Question: what about when $N\mu(U)^2$ is large?

Answer: In general, nothing can be done. However, in many cases, one has asymptotic incoherence

$$\mu(UP_n^{\perp}) \to 0, \quad n \to \infty.$$

The idea now is to use semi-random sampling: fully sample in regions in coherent regions of U, and subsample elsewhere, i.e.

$$\Omega = \Omega_1 \cup \Omega_2,$$

where $\Omega_1 = \{1, \ldots, N_1\}$ and $\Omega_2 \subseteq \{N_1 + 1, \ldots, N_2\}$, $|\Omega_2| = m$.

Current work – compressed sensing

Recall that we require

$$m \geq C \cdot k \cdot N \cdot \mu(U)^2 \cdot (1 - \log \epsilon) \cdot \log \left(\frac{MN\sqrt{k}}{m} \right).$$

Thus, we can only subsample if $N\mu(U)^2$ is small.

Question: what about when $N\mu(U)^2$ is large?

Answer: In general, nothing can be done. However, in many cases, one has asymptotic incoherence

$$\mu(UP_n^{\perp}) \to 0, \quad n \to \infty.$$

The idea now is to use semi-random sampling: fully sample in regions in coherent regions of U, and subsample elsewhere, i.e.

$$\Omega = \Omega_1 \cup \Omega_2,$$

where $\Omega_1 = \{1, \ldots, N_1\}$ and $\Omega_2 \subseteq \{N_1 + 1, \ldots, N_2\}$, $|\Omega_2| = m$.

Current work – generalized sampling

Nonuniform sampling

- Uniform samples of $\mathcal{F}f$ are often not encountered in applications, e.g. MRI – spiral sampling.
- Nonuniform samples may well be highly clustered in some regions and sparse in others.

Inverse and ill-posed problems

- Suppose that $\mathcal{A}f = g$, where $\mathcal{A}$ is typically ill-posed. We want to recover f in arbitrary bases from samples of g.

References

Generalized sampling

- B. Adcock & A. C. Hansen, *A generalised sampling theorem for stable reconstructions in Hilbert spaces*. J. Fourier Anal. Appl. (to appear), 2011.
- B. Adcock & A. C. Hansen, *Stable reconstructions in Hilbert spaces and the resolution of the Gibbs phenomenon*. Appl. Comput. Harmon. Anal. (to appear), 2011.
- B. Adcock & A. C. Hansen, *Sharp bounds, optimality and a geometric interpretation for generalised sampling in Hilbert spaces*. Preprint, 2011.
- B. Adcock, A. C. Hansen, E. Herrholz & G. Teschke, *Generalized sampling: extensions to frames and ill-posed problems*. In preparation, 2011.

Generalized sampling with compressed sensing

- B. Adcock & A. C. Hansen, *Generalized sampling and infinite-dimensional compressed sensing*. Preprint, 2011.
- B. Adcock, A. C. Hansen, E. Herrholz & G. Teschke, *Generalized sampling, infinite-dimensional compressed sensing, and semi-random sampling for asymptotically incoherent disctionaries*. In preparation, 2011.