Nonuniform sampling of multivariate functions using derivatives

Ben Adcock
Department of Mathematics
Purdue University

Joint work with Milana Gataric and Anders Hansen (Cambridge)

Outline

Outline

Motivation

Recently-developed sensors can record functions values and spatial gradient information.

Question: What can we expect to gain from this additional data?

Two potential answers:

(i) Efficient acquisition. Recording gradient information means sensors can be placed further apart than in the case where only function values are measured.

(ii) Improved reconstruction quality. In sparsity-regularized reconstructions, one can exploit joint sparsity of functions and their derivatives to attain better accuracy.

This talk: Mathematical understanding of (i) using sampling theory.

- Necessary first step towards (ii).

Motivation

Recently-developed sensors can record functions values and spatial gradient information.

Question: What can we expect to gain from this additional data?

Two potential answers:

(i) Efficient acquisition. Recording gradient information means sensors can be placed further apart than in the case where only function values are measured.

(ii) Improved reconstruction quality. In sparsity-regularized reconstructions, one can exploit joint sparsity of functions and their derivatives to attain better accuracy.

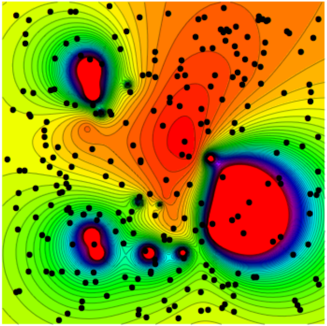
This talk: Mathematical understanding of (i) using sampling theory.

- Necessary first step towards (ii).

Formulation

Let

- $f : \mathbb{R}^d \to \mathbb{R}$ be the function to recover,
- $X = \{x_n : n \in I\}$ be the set of sensor locations.



Question: Under what conditions on f and on X is it possible to stably recover f from the measurements

$$\{f(x_n) : n \in I\} \cup \{\nabla f(x_n) : n \in I\}.$$

Or more generally, from the first k derivatives

$$\{D^\alpha f(x_n) : n \in I, |\alpha|_1 \leq k\}.$$

Outline

Classical sampling theory

Let $\Omega \subseteq \mathbb{R}^d$ be compact. The Paley–Wiener space

$$\mathrm{B}(\Omega) = \left\{ f \in \mathrm{L}^2(\mathbb{R}^d) : \mathrm{supp}(\hat{f}) \subseteq \Omega \right\}.$$

is the space of functions which are bandlimited to Ω.

Shannon Sampling Theorem: Let $\Omega = (-\omega, \omega)^d$. Any function $f \in \mathrm{B}((-\omega, \omega)^d)$ is uniquely defined by the samples

$$f(x_n), \quad x_n = \frac{n\pi}{\omega}, \qquad n \in \mathbb{Z}^d.$$

Moreover,

$$f(x) = \sum_{n \in \mathbb{Z}^d} f\left(\frac{n\pi}{\omega}\right) \mathrm{sinc}\left(\omega x - n\pi\right).$$

In particular, Parseval's identity holds

$$\sum_{n \in \mathbb{Z}^d} \left| f\left(\frac{n\pi}{\omega}\right)\right|^2 = (2\omega)^d \|f\|_{L^2}^2.$$

- We refer to the constant $\frac{\pi}{2\omega}$ as the Nyquist rate.

Nonuniform sampling theory

Shannon's theorem requires uniformly-spaced samples. This is rarely the case in practice. Moreover, Ω must be a hypercube.

A collection of nonuniformly-spaced sample points

$$X = \{x_n : n \in I\} \subseteq \mathbb{R}^d,$$

is called a stable set of sampling for $\mathrm{B}(\Omega)$ if

$$A\|f\|_{L^2}^2 \leq \sum_{n\in I} |f(x_n)|^2 \leq B\|f\|_{L^2}^2, \quad \forall f \in \mathrm{B}(\Omega).$$

The ratio B/A is a measure of stability.

Known results:

- $d = 1$, $\Omega = (-\omega, \omega)$. Almost complete characterization in terms of Beurling density (Jaffard 1991, Seip 1995).
- $d \geq 2$, Ω compact, convex and symmetric. Sufficient sharp condition in terms of polar set of Ω (Beurling 1960s, Benedetto & Wu 2000).

Nonuniform sampling theory

Limitations:

- Requires separation of points: $|x_n - x_m| \geq \eta$. $B/A \to \infty$ as $\eta \to 0$.
- No explicit estimates for A and B.

The case $d = 1$. Gröchenig (1992): define the density

$$\delta = \sup_{x \in \mathbb{R}} \inf_{n \in I} |x - x_n|.$$

If $\delta < \frac{\pi}{2\omega}$ then, for all $f \in \mathrm{B}(-\omega, \omega)$,

$$\left(1 - \frac{2\omega\delta}{\pi}\right)^2 \|f\|_{L^2}^2 \leq \sum_{n \in I} \mu_n |f(x_n)|^2 \leq \left(1 + \frac{2\omega\delta}{\pi}\right)^2 \|f\|_{L^2}^2,$$

where the weights $\mu_n = \frac{1}{2}(x_{n+1} - x_{n-1})$ compensate for local clustering.

$\Rightarrow$ It suffices to take nonuniform samples just above the Nyquist rate.

Nonuniform sampling theory

Limitations:

- Requires separation of points: $|x_n - x_m| \geq \eta$. $B/A \to \infty$ as $\eta \to 0$.
- No explicit estimates for A and B.

The case $d = 1$. Gröchenig (1992): define the density

$$\delta = \sup_{x \in \mathbb{R}} \inf_{n \in I} |x - x_n|.$$

If $\delta < \frac{\pi}{2\omega}$ then, for all $f \in \mathrm{B}(-\omega, \omega)$,

$$\left(1 - \frac{2\omega\delta}{\pi}\right)^2 \|f\|_{L^2}^2 \leq \sum_{n \in I} \mu_n |f(x_n)|^2 \leq \left(1 + \frac{2\omega\delta}{\pi}\right)^2 \|f\|_{L^2}^2,$$

where the weights $\mu_n = \frac{1}{2}(x_{n+1} - x_{n-1})$ compensate for local clustering.

$\Rightarrow$ It suffices to take nonuniform samples just above the Nyquist rate.

Nonuniform sampling theory

The case $d \geq 2$. Let $X = \{x_n : n \in I\} \subseteq \mathbb{R}^d$. Let

- $\delta = \sup_{x \in \mathbb{R}^d} \inf_{n \in I} |x - x_n|$
- $\mu_n = \mathrm{Vol}(V_n)$, where $\{V_n\}_{n \in I}$ are the Voronoi cells of X.

Gröchenig (1992,2001): If $\Omega \subseteq (-\omega, \omega)^d$ and

$$\delta < \frac{\log(2)}{\omega d},$$

then X is a weighted stable set of sampling with

$$B/A \leq (2\exp(-\omega\delta d) - 1)^{-2}.$$

Limitation: Not sharp – deteriorates linearly with d. Conversely, Beurling's sharp condition is dimension independent.

Nonuniform sampling theory

The case $d \geq 2$. Let $X = \{x_n : n \in I\} \subseteq \mathbb{R}^d$. Let

- $\delta = \sup_{x \in \mathbb{R}^d} \inf_{n \in I} |x - x_n|$
- $\mu_n = \mathrm{Vol}(V_n)$, where $\{V_n\}_{n \in I}$ are the Voronoi cells of X.

Gröchenig (1992,2001): If $\Omega \subseteq (-\omega, \omega)^d$ and

$$\delta < \frac{\log(2)}{\omega d},$$

then X is a weighted stable set of sampling with

$$B/A \leq (2\exp(-\omega\delta d) - 1)^{-2}.$$

Limitation: Not sharp – deteriorates linearly with d. Conversely, Beurling's sharp condition is dimension independent.

An improvement of Gröchenig's result

Theorem (BA, Gataric, Hansen (2014))

Suppose that $\Omega \subseteq \mathcal{B}(a,\omega)$ for $a \in \mathbb{R}^d$ and $\omega > 0$ and that

$$\delta < \frac{\log(2)}{\omega} \approx \frac{0.6931}{\omega}.$$

Then X is a weighted stable set of sampling with

$$B/A \leq \left(2\exp(-\omega\delta) - 1\right)^{-2}.$$

- If $\Omega = (-\omega,\omega)^d$ then $\delta < \frac{\log(2)}{\omega\sqrt{d}}$. Factor of $\sqrt{d}$ improvement over Gröchenig's bound.
- If $\Omega = \mathcal{B}(a,\omega)$, then the estimate is sharp with respect to d. However, it is strictly less than the sharp condition $\delta \approx 1.5708/\omega$ of Beurling (albeit with explicit bounds).

Outline

Sampling theory with derivative measurements

We now consider the data

$$\{D^\alpha f(x_n) : n \in I, |\alpha|_1 \le k\},$$

where $\alpha = (\alpha_1, \ldots, \alpha_d)$ is a multi-index and $|\alpha|_1 = \alpha_1 + \ldots + \alpha_d$.

Objective: Show that increasing k allows for a larger maximum density δ.

Uniform sampling with derivatives

Classical problem: Shannon (1950s), Jagerman & Fogel (1956), Linden & Abramson (1960), Papoulis (1977), Rawn (1989),....

Consider $d = 1$ and let $\Omega = (-\omega, \omega)$. If

$$x_n = \frac{(k+1)n\pi}{\omega}, \quad n \in \mathbb{Z},$$

then $\{x_n : n \in \mathbb{Z}\}$ is a stable set of sampling for $\mathrm{B}(\Omega)$, i.e.

$$A\|f\|_{L^2}^2 \leq \sum_{n\in\mathbb{Z}} \sum_{l=0}^{k} \left| f^{(l)} \left(\frac{(k+1)n\pi}{\omega} \right) \right|^2 \leq B\|f\|_{L^2}^2.$$

Moreover, there exist functions $h_0(x), \ldots, h_k(x)$ such that

$$f(x) = \sum_{n\in\mathbb{Z}} \sum_{l=0}^{k} f^{(l)} \left(\frac{(k+1)n\pi}{\omega} \right) h_l\left(\omega x - (k+1)n\pi\right), \quad f \in \mathrm{B}(\Omega).$$

Multivariate nonuniform sampling theorem with derivatives

Setup:

- Let $X = \{x_n : n \in I\} \subseteq \mathbb{R}^d$, and define the weights

$$\mu_{n,\alpha} = \frac{1}{\alpha!} \int_{V_n} (x - x_n)^{2\alpha} \, \mathrm{d}x, \quad n \in I, \alpha \in \mathbb{N}_0^d,$$

  where $\{V_n : n \in I\}$ are the Voronoi cells of X.

- Define the function

$$h_k(z) = \exp(z) \left(\exp(z) - \sum_{r=0}^{k} z^r / r! \right), \quad z \in (0, \infty).$$

  This function is increasing on $(0, \infty)$. Write $H_k(w)$ for its inverse.

Multivariate nonuniform sampling theorem with derivatives

Theorem (BA, Gataric, Hansen (2014))

Suppose that $\Omega \subseteq \mathcal{B}(a, \omega)$. *If*

$$\delta < \frac{H_k(1)}{\omega}$$

then X *is a weighted stable set of sampling for derivatives, i.e.*

$$A\|f\|_{L^2}^2 \leq \sum_{n \in I} \sum_{|\alpha|_1 \leq k} \mu_{n,\alpha} |D^\alpha f(x_n)|^2 \leq B\|f\|_{L^2}^2, \quad \forall f \in \mathrm{B}(\Omega),$$

where

$$B/A \leq \frac{\exp((\omega\delta + 1)^2 + d - 1)}{(1 - h_k(\omega\delta)^2)}.$$

Discussion

The key part is the estimate

$$\delta < \frac{H_k(1)}{\omega}. \qquad (\star)$$

Remarks:

- As in the $k = 0$ case, if $\Omega = \mathcal{B}(a, \omega)$ then $(\star)$ is independent of d.
- If $\Omega = (-\omega, \omega)^d$ is a hybercube, then $(\star)$ reads

$$\delta < \frac{H_k(1)}{\sqrt{d}\omega},$$

i.e. it decays like $1/\sqrt{d}$ for large d.

Discussion

k	1	2	3	4	5	6	7
$H_k(1)$	0.8141	1.1268	1.4304	1.7290	2.0416	2.3170	2.6080
NYQ*	3.1416	4.7124	6.2832	7.8540	9.4248	10.9956	12.5664

*Conjectured: currently no existing analogue of Beurling's theorem for nonuniform sampling with derivatives

Large k asymptotics:

Proposition (BA, Gataric, Hansen (2014))

If $W(\cdot)$ is the Lambert-W function, then

$$H_k(1) \sim W(1/\mathrm{e})k \approx 0.2785k, \qquad k \to \infty.$$

Recall this holds for general domains Ω and arbitrary nonuniform samples, and gives explicit bounds. However, it is substantially less than the conjectured Nyquist rate, which is $\sim 1.5708k$.

The univariate case

Wirtinger–Sobolev inequality: Let $f \in \mathrm{H}^k(a,b)$ with $f^{(r)}(a) = 0$, $r = 0, \dots, k-1$. Then there is a constant $c_k > 0$ such that

$$\|f\|_{L^2(a,b)} \le (c_k)^k (b-a)^k \|f^{(k)}\|_{L^2(a,b)}$$

Theorem

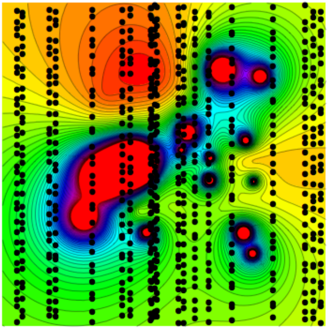
Suppose that $\delta < \frac{1}{c_{k+1}\omega}$. Then X is a weighted stable set of sampling for derivatives for $\mathrm{B}(-\omega,\omega)$ with $B/A \le \frac{\mathrm{e}^{(\delta\omega)^2+1}(1+2\delta\omega/\pi)^2}{(1-(c_{k+1}\delta\omega)^{k+1})^2}$.

- Based on earlier work of Gröchenig ($k = 0$) and Razafinjatovo ($k = 1$).

k	1	2	3	4	5	6	7
$H_k(1)$	0.8141	1.1268	1.4304	1.7290	2.0416	2.3170	2.6080
$1/c_{k+1}$	1.8751	2.2248	2.5903	2.9621	3.3367	3.7125	4.0888
NYQ	3.1416	4.7124	6.2832	7.8540	9.4248	10.9956	12.5664

For large k, $1/c_{k+1} \sim \mathrm{e}^{-1}k \approx 0.3679k$, as opposed to $0.2785k$.

Spatial-temporal sampling



In some applications, we consider $f = f(x, t)$, where $x \in \mathbb{R}^d$ and $t \in [0, \infty)$. Acquisition occurs sparsely in space and densely in time. Measurements may not be taken at the same times at different sensors, and only spatial derivatives are acquired.

Spatial-temporal sampling

Define the set

$$Z = \left\{ (x_n, t_{m,n}) \in \mathbb{R}^d \times [0, \infty) : n \in I, m \in J \right\},$$

and consider the spatial derivative measurements:

$$\{ D_x^\alpha f(x_n, t_{m,n}) : n \in I, m \in J, |\alpha|_1 \leq k \}.$$

Theorem

Let δ_x and δ_t be the spatial and temporal sampling densities. Suppose that $f \in \mathrm{B}(\Omega)$, where $\Omega \subseteq \mathcal{B}(a, \omega) \times (-\nu, \nu)$. If

$$\delta_t < \frac{\pi}{2\nu}, \qquad \delta_x < \frac{C(k)}{\omega}, \qquad C(k) = \begin{cases} 1/c_{k+1} & d = 1 \\ H_k(1) & d \geq 2 \end{cases}$$

Then Z is a stable weighted set of sampling for derivatives.

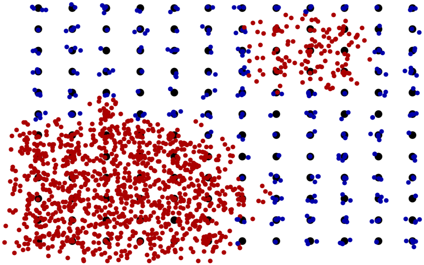
Perturbation theory and larger sampling gaps

Perturbation theory: Kadec-1/4 theorem (1960s). Also Balan (1997) and Christensen (1999).

Suppose $X = \{x_n : n \in I\}$ is a (weighted) stable set of sampling for derivatives, e.g. uniform samples. Consider the perturbed sample points:

$$\{\tilde{x}_{n,m} : m \in J_n,\ n \in I\},$$

with sufficiently small perturbation $\epsilon = \sup_{n \in I} \sup_{m \in J_n} |x_n - \tilde{x}_{n,m}|$.



Black dots are Nyquist rate samples. Red dots are high density samples. Blue dots are low density samples.

Note: different numbers of sensors allowed in different locations.

Perturbation theorem for derivatives

Theorem (BA, Gataric, Hansen (2014))

Suppose that $X = \{x_n : n \in I\}$ is a (weighted) stable set of sampling for derivatives for $\mathrm{B}(\Omega)$ with bounds A and B, where $\Omega \subseteq \mathcal{B}(a, \omega)$. Let

$$\tilde{X} = \{\tilde{x}_{n,m} : m \in J_n,\ n \in I\}, \qquad \sup_n |J_n| < \infty$$

and suppose that

$$\epsilon = \sup_{n \in I} \sup_{m \in J_n} |x_n - \tilde{x}_{n,m}| < \frac{\log(1 + \sqrt{A/B})}{\omega},$$

Then $\tilde{X}$ is a stable weighted set of sampling for derivatives.

$\Rightarrow$ Nyquist-sized gaps $\approx k\pi/\omega$ between samples are permitted, as long as the perturbation ϵ is small.

Outline

Introduction

Nonuniform sampling theory

Sampling theory with derivatives

Conclusions and Outlook

Conclusions and Outlook

Conclusion: Nonuniform sampling with k derivatives allows for a larger maximal sampling density δ, scaling linearly with k. Optimal constants remain elusive, but perturbation theory permit larger sampling gaps.

Future work:

- Improved density conditions, e.g. using random samples (Bass & Gröchenig 2004)
- Other function spaces, e.g. shift-invariant spaces
- Reconstructions from derivative samples
  - Infinite-dimensional compressed sensing (BA & Hansen, 2011)
  - Joint ℓ^1 regularization – exploiting common sparsity of f and $D^\alpha f$
  - Current work proves the existence of stably invertible sampling operators, a necessary prerequisite for CS