2585-21

Joint ICTP-TWAS School on Coherent State Transforms, Time-Frequency and Time-Scale Analysis, Applications

2 - 20 June 2014

Group-theoretical methods for the design and analysis of higher-dimensional wavelet systems III Sparse signals and function spaces

H. Fuhr

RWTH, Aachen, Germany

Group-theoretical methods for the design and analysis of higher-dimensional wavelet systems III
Sparse signals and function spaces

Hartmut Führ
fuehr@matha.rwth-aachen.de

Trieste, June 2014

Lehrstuhl A für Mathematik, RWTH

Agenda for this week

Agenda for this week

I Wavelet transforms associated to groups of affine mappings (Monday)

Agenda for this week

I Wavelet transforms associated to groups of affine mappings (Monday)

II Wavelet inversion, admissibility and the Plancherel formula (Tuesday)

Agenda for this week

I Wavelet transforms associated to groups of affine mappings (Monday)

II Wavelet inversion, admissibility and the Plancherel formula (Tuesday)

III Sparse signals and function spaces (Wednesday)

Agenda for this week

I Wavelet transforms associated to groups of affine mappings (Monday)

II Wavelet inversion, admissibility and the Plancherel formula (Tuesday)

III Sparse signals and function spaces (Wednesday)

IV Wavelet approximation theory over general dilation groups (Thursday, Friday)

Outline

1. Motivation for sparse signals

Outline

1. Motivation for sparse signals

2. ℓ^1-sparse vectors over general ONB's

Outline

1. Motivation for sparse signals

2. ℓ^1-sparse vectors over general ONB's

3. Nonlinear approximation

Outline

1. Motivation for sparse signals
2. ℓ^1-sparse vectors over general ONB's
3. Nonlinear approximation
4. Homogeneous Besov spaces

Outline

1. Motivation for sparse signals
2. ℓ^1-sparse vectors over general ONB's
3. Nonlinear approximation
4. Homogeneous Besov spaces
5. ℓ^1-sparsity w.r.t. wavelet ONB's

Outline

Overview

Why are wavelets useful?

Nice features of wavelet systems:

Why are wavelets useful?

Nice features of wavelet systems:

- Expansions of arbitrary vectors in terms of a “meaningful” system of building blocks.

Why are wavelets useful?

Nice features of wavelet systems:

- Expansions of arbitrary vectors in terms of a “meaningful” system of building blocks.
- Resulting from that: Heuristics for signal processing and analysis

Why are wavelets useful?

Nice features of wavelet systems:

- Expansions of arbitrary vectors in terms of a “meaningful” system of building blocks.
- Resulting from that: Heuristics for signal processing and analysis
- Underlying useful structures: Groups, multiresolution analysis, etc. (Useful both for mathematics and applications)

Why are wavelets useful?

Nice features of wavelet systems:

- Expansions of arbitrary vectors in terms of a “meaningful” system of building blocks.
- Resulting from that: Heuristics for signal processing and analysis
- Underlying useful structures: Groups, multiresolution analysis, etc. (Useful both for mathematics and applications)
- Very important: Wavelet description of realistic signals is efficient,

Why are wavelets useful?

Nice features of wavelet systems:

- Expansions of arbitrary vectors in terms of a "meaningful" system of building blocks.
- Resulting from that: Heuristics for signal processing and analysis
- Underlying useful structures: Groups, multiresolution analysis, etc. (Useful both for mathematics and applications)
- Very important: Wavelet description of realistic signals is efficient, i.e., a realistic signal is well-described by a few building blocks.

Aim of this talk: Make the last statement precise.

Overview

Recall: ONB's and norm preservation

Theorem 1

Let $\mathcal{H}$ be a Hilbert space, and $(x_i)_{i \in I} \subset \mathcal{H}$ an ONS. Then the following are equivalent:

Recall: ONB's and norm preservation

Theorem 1

Let $\mathcal{H}$ be a Hilbert space, and $(x_i)_{i \in I} \subset \mathcal{H}$ an ONS. Then the following are equivalent:

(a) *For all $y \in \mathcal{H}$: $\|y\|^2 = \sum_{i \in I} |\langle y, x_i \rangle|^2$*

Recall: ONB's and norm preservation

Theorem 1

Let $\mathcal{H}$ be a Hilbert space, and $(x_i)_{i \in I} \subset \mathcal{H}$ an ONS. Then the following are equivalent:

(a) *For all $y \in \mathcal{H} : \|y\|^2 = \sum_{i \in I} |\langle y, x_i \rangle|^2$*

(b) *For all $y \in \mathcal{H} : y = \sum_{i \in I} \langle y, x_i \rangle x_i$*

Recall: ONB's and norm preservation

Theorem 1

Let $\mathcal{H}$ be a Hilbert space, and $(x_i)_{i \in I} \subset \mathcal{H}$ an ONS. Then the following are equivalent:

(a) *For all $y \in \mathcal{H}$: $\|y\|^2 = \sum_{i \in I} |\langle y, x_i \rangle|^2$*

(b) *For all $y \in \mathcal{H}$: $y = \sum_{i \in I} \langle y, x_i \rangle x_i$*

(c) *The coefficient operator*

$$\mathcal{H} \ni y \mapsto (\langle y, x_i \rangle)_{i \in I} \in \ell^2(I)$$

is unitary.

Recall: ONB's and norm preservation

Theorem 1

Let $\mathcal{H}$ be a Hilbert space, and $(x_i)_{i \in I} \subset \mathcal{H}$ an ONS. Then the following are equivalent:

(a) *For all $y \in \mathcal{H}$: $\|y\|^2 = \sum_{i \in I} |\langle y, x_i \rangle|^2$*

(b) *For all $y \in \mathcal{H}$: $y = \sum_{i \in I} \langle y, x_i \rangle x_i$*

(c) *The coefficient operator*

$$\mathcal{H} \ni y \mapsto (\langle y, x_i \rangle)_{i \in I} \in \ell^2(I)$$

is unitary.

(d) *For all $y \in \mathcal{H}$: If $\langle y, x_i \rangle = 0$, for all $i \in I$, then $y = 0$.*

If the equivalent conditions hold, then the system is an orthonormal basis (ONB).

ℓ^1-sparse vectors

Note

Given any infinite set I, one has $\ell^1(I) \subset \ell^2(I)$ properly, i.e., the condition

$$\sum_{i \in I} |c_i| < \infty \Rightarrow \sum_{i \in I} |c_i|^2 < \infty \quad .$$

ℓ^1-sparse vectors

Note

Given any infinite set I, one has $\ell^1(I) \subset \ell^2(I)$ properly, i.e., the condition

$$\sum_{i \in I} |c_i| < \infty \Rightarrow \sum_{i \in I} |c_i|^2 < \infty \quad .$$

Counterexample for converse: $I = \mathbb{N}$, $c_n = 1/n$.

ℓ^1-sparse vectors

Note

Given any infinite set I, one has $\ell^1(I) \subset \ell^2(I)$ properly, i.e., the condition

$$\sum_{i \in I} |c_i| < \infty \Rightarrow \sum_{i \in I} |c_i|^2 < \infty \quad .$$

Counterexample for converse: $I = \mathbb{N}$, $c_n = 1/n$.

Definition 2

Let $X = (x_i)_{i \in I} \subset \mathbb{N}$ denote an ONB. We call $y \in \ell^2(I)$ an ℓ^1-sparse vector(w.r.t. X) if the coefficient sequence $(\langle y, x_i \rangle)_{i \in I}$ is in ℓ^1.

ℓ^1-sparse vectors

Note

Given any infinite set I, one has $\ell^1(I) \subset \ell^2(I)$ properly, i.e., the condition

$$\sum_{i \in I} |c_i| < \infty \Rightarrow \sum_{i \in I} |c_i|^2 < \infty \quad .$$

Counterexample for converse: $I = \mathbb{N}$, $c_n = 1/n$.

Definition 2

Let $X = (x_i)_{i \in I} \subset \mathbb{N}$ denote an ONB. We call $y \in \ell^2(I)$ an ℓ^1-sparse vector(w.r.t. X) if the coefficient sequence $(\langle y, x_i \rangle)_{i \in I}$ is in ℓ^1. I.e., the ℓ^1-sparse vectors are precisely those that have an expansion

$$y = \sum_{i \in I} c_i x_i \quad , \quad \sum_{i \in I} |c_i| < \infty \ .$$

Other notions of sparsity

Other notions of sparsity

Remark

There are many useful ways of measuring sparsity:

Other notions of sparsity

Remark

There are many useful ways of measuring sparsity:

- Replace ℓ^1 by ℓ^p, for $0 < p < 1$. Quasi-norm on coefficients, more and more restrictive as $p \to 0$.

Other notions of sparsity

Remark

There are many useful ways of measuring sparsity:

- Replace ℓ^1 by ℓ^p, for $0 < p < 1$. Quasi-norm on coefficients, more and more restrictive as $p \to 0$.
- ℓ^0-"norm", i.e, number of nonzero coefficients. See compressive sensing!

Other notions of sparsity

Remark

There are many useful ways of measuring sparsity:

- Replace ℓ^1 by ℓ^p, for $0 < p < 1$. Quasi-norm on coefficients, more and more restrictive as $p \to 0$.
- ℓ^0-"norm", i.e, number of nonzero coefficients. See compressive sensing!
- Weighted versions, i.e. multiply coefficients by a suitable unbounded sequence of numbers before taking the norm.

We concentrate on ℓ^1 in the following for simplicity.

Overview

Nonlinear n-term approximation

Nonlinear *n*-term approximation

Aim

Mathematical description of signals that are well approximated by a few building blocks.

Nonlinear n-term approximation

Aim

Mathematical description of signals that are well approximated by a few building blocks.

Definition 3

Let $X = (x_i)_{i \in I}$ denote a system of vectors in a Hilbert space $\mathcal{H}$. Given $y \in \mathcal{H}$ and $n \in \mathbb{N}$, we let

$$\sigma_n(y)_X = \inf \left\{ \left\| y - \sum_{i \in J} c_i x_i \right\| : J \subset I, c_i \in \mathbb{C}, |J| \leq n \right\}$$

denote the (nonlinear) n-term approximation error of y with respect to the system X.

Nonlinear n-term approximation

Aim

Mathematical description of signals that are well approximated by a few building blocks.

Definition 3

Let $X = (x_i)_{i \in I}$ denote a system of vectors in a Hilbert space $\mathcal{H}$. Given $y \in \mathcal{H}$ and $n \in \mathbb{N}$, we let

$$\sigma_n(y)_X = \inf \left\{ \left\| y - \sum_{i \in J} c_i x_i \right\| : J \subset I, c_i \in \mathbb{C}, |J| \le n \right\}$$

denote the (nonlinear) n-term approximation error of y with respect to the system X.

Note: The optimal set J will depend on y (hence nonlinear approximation).

Nonlinear approximation over ONB's

Nonlinear approximation over ONB's

Theorem 4

Let $X = (x_i)_{i \in I}$ *denote an ONB of the Hilbert space* $\mathcal{H}$*, and let* $y \in \mathcal{H}$*.*

Nonlinear approximation over ONB's

Theorem 4

Let $X = (x_i)_{i \in I}$ denote an ONB of the Hilbert space $\mathcal{H}$, and let $y \in \mathcal{H}$. Given $n \in \mathbb{N}$, let J_n denote the set of indices of the n largest expansion coefficients $\langle y, x_i \rangle$.

Nonlinear approximation over ONB's

Theorem 4

Let $X = (x_i)_{i \in I}$ denote an ONB of the Hilbert space $\mathcal{H}$, and let $y \in \mathcal{H}$. Given $n \in \mathbb{N}$, let J_n denote the set of indices of the n largest expansion coefficients $\langle y, x_i \rangle$. Then

$$y_n = \sum_{i \in J_n} \langle y, x_i \rangle x_i$$

is the optimal n-term approximation of y, and we have

$$\sigma_n(y)_X^2 = \sum_{j \notin J_n} |\langle y, x_j \rangle|^2 \quad .$$

Nonlinear approximation over ONB's

Theorem 4

Let $X = (x_i)_{i \in I}$ denote an ONB of the Hilbert space $\mathcal{H}$, and let $y \in \mathcal{H}$. Given $n \in \mathbb{N}$, let J_n denote the set of indices of the n largest expansion coefficients $\langle y, x_i \rangle$. Then

$$y_n = \sum_{i \in J_n} \langle y, x_i \rangle x_i$$

is the optimal n-term approximation of y, and we have

$$\sigma_n(y)_X^2 = \sum_{j \notin J_n} |\langle y, x_j \rangle|^2 \quad .$$

Alternatively, let $(c_n)_{n \in \mathbb{N}}$ denote the sequence of expansion coefficients of y, sorted by decreasing modulus. Then

$$\sigma_n(y)_X^2 = \sum_{k > n} |c_k|^2 \quad .$$

ℓ^1-Sparsity and nonlinear approximation rate

Theorem 5

Let $X = (x_i)_{i \in I} \subset \mathcal{H}$ denote an ONB, and $y \in \mathcal{H}$. Then the following statements are equivalent:

ℓ^1-Sparsity and nonlinear approximation rate

Theorem 5

Let $X = (x_i)_{i \in I} \subset \mathcal{H}$ denote an ONB, and $y \in \mathcal{H}$. Then the following statements are equivalent:

(a) *y is ℓ^1-sparse with respect to X.*

ℓ^1-Sparsity and nonlinear approximation rate

Theorem 5

Let $X = (x_i)_{i \in I} \subset \mathcal{H}$ denote an ONB, and $y \in \mathcal{H}$. Then the following statements are equivalent:

(a) *y is ℓ^1-sparse with respect to X.*

(b) *The n-term approximation errors fulfill*

$$\sum_{n \in \mathbb{N}} n^{-1/2} \sigma_n(y)_X < \infty \quad .$$

ℓ^1-Sparsity and nonlinear approximation rate

Theorem 5

Let $X = (x_i)_{i \in I} \subset \mathcal{H}$ denote an ONB, and $y \in \mathcal{H}$. Then the following statements are equivalent:

(a) *y is ℓ^1-sparse with respect to X.*

(b) *The n-term approximation errors fulfill*

$$\sum_{n \in \mathbb{N}} n^{-1/2} \sigma_n(y)_X < \infty \quad .$$

ℓ^1-Sparsity and nonlinear approximation rate

Theorem 5

Let $X = (x_i)_{i \in I} \subset \mathcal{H}$ denote an ONB, and $y \in \mathcal{H}$. Then the following statements are equivalent:

(a) *y is ℓ^1-sparse with respect to X.*

(b) *The n-term approximation errors fulfill*

$$\sum_{n \in \mathbb{N}} n^{-1/2} \sigma_n(y)_X < \infty \quad .$$

- Informal (slightly incorrect) interpretation:

ℓ^1-Sparsity and nonlinear approximation rate

Theorem 5

Let $X = (x_i)_{i \in I} \subset \mathcal{H}$ denote an ONB, and $y \in \mathcal{H}$. Then the following statements are equivalent:

(a) *y is ℓ^1-sparse with respect to X.*

(b) *The n-term approximation errors fulfill*

$$\sum_{n \in \mathbb{N}} n^{-1/2} \sigma_n(y)_X < \infty \quad .$$

- Informal (slightly incorrect) interpretation:
  y is ℓ^1-sparse iff $\sigma_n(y)_X \to 0$ faster than $n^{-1/2}$.

ℓ^1-Sparsity and nonlinear approximation rate

Theorem 5

Let $X = (x_i)_{i \in I} \subset \mathcal{H}$ denote an ONB, and $y \in \mathcal{H}$. Then the following statements are equivalent:

(a) *y is ℓ^1-sparse with respect to X.*

(b) *The n-term approximation errors fulfill*

$$\sum_{n \in \mathbb{N}} n^{-1/2} \sigma_n(y)_X < \infty \quad .$$

- Informal (slightly incorrect) interpretation:
  y is ℓ^1-sparse iff $\sigma_n(y)_X \to 0$ faster than $n^{-1/2}$.
- Can be extended to other sparsity measures.

ℓ^1-Sparsity and nonlinear approximation rate

Theorem 5

Let $X = (x_i)_{i \in I} \subset \mathcal{H}$ denote an ONB, and $y \in \mathcal{H}$. Then the following statements are equivalent:

(a) *y is ℓ^1-sparse with respect to X.*

(b) *The n-term approximation errors fulfill*

$$\sum_{n \in \mathbb{N}} n^{-1/2} \sigma_n(y)_X < \infty \quad .$$

- Informal (slightly incorrect) interpretation:
  y is ℓ^1-sparse iff $\sigma_n(y)_X \to 0$ faster than $n^{-1/2}$.
- Can be extended to other sparsity measures.
  E.g., $p < 1$ will lead to faster decay.

For later use: Frame version

Theorem 6

Let $X = (x_i)_{i \in I} \subset \mathcal{H}$ denote a frame, and $y \in \mathcal{H}$. Assume that

$$y = \sum_{i \in I} c_i x_i \quad , \quad \sum_{i \in I} |c_i| < \infty \quad .$$

Then the n-term approximation errors fulfill

$$\sum_{n \in \mathbb{N}} n^{-1/2} \sigma_n(y)_X < \infty \quad .$$

For later use: Frame version

Theorem 6

Let $X = (x_i)_{i \in I} \subset \mathcal{H}$ denote a frame, and $y \in \mathcal{H}$. Assume that

$$y = \sum_{i \in I} c_i x_i \quad , \quad \sum_{i \in I} |c_i| < \infty \quad .$$

Then the n-term approximation errors fulfill

$$\sum_{n \in \mathbb{N}} n^{-1/2} \sigma_n(y)_X < \infty \quad .$$

Note

For later use: Frame version

Theorem 6

Let $X = (x_i)_{i \in I} \subset \mathcal{H}$ denote a frame, and $y \in \mathcal{H}$. Assume that

$$y = \sum_{i \in I} c_i x_i \quad , \quad \sum_{i \in I} |c_i| < \infty \quad .$$

Then the n-term approximation errors fulfill

$$\sum_{n \in \mathbb{N}} n^{-1/2} \sigma_n(y)_X < \infty \quad .$$

Note

- Only one direction of the equivalence holds in the frame case.

For later use: Frame version

Theorem 6

Let $X = (x_i)_{i \in I} \subset \mathcal{H}$ *denote a frame, and* $y \in \mathcal{H}$*. Assume that*

$$y = \sum_{i \in I} c_i x_i \quad , \quad \sum_{i \in I} |c_i| < \infty \quad .$$

Then the n-term approximation errors fulfill

$$\sum_{n \in \mathbb{N}} n^{-1/2} \sigma_n(y)_X < \infty \quad .$$

Note

- Only one direction of the equivalence holds in the frame case.
- The ℓ^1-coefficients are not assumed to come from a dual frame.

Overview

Modulus of continuity

Modulus of continuity

Definition 7

For $h \in \mathbb{R}$, $f : \mathbb{R} \to \mathbb{C}$, the (forward) difference operator of step h is given by

$$\Delta_h f(x) = f(x+h) - f(x),$$

Modulus of continuity

Definition 7

For $h \in \mathbb{R}$, $f : \mathbb{R} \to \mathbb{C}$, the (forward) difference operator of step h is given by

$$\Delta_h f(x) = f(x+h) - f(x),$$

and for $r \in \mathbb{N}$, define the difference operator of order r, step h, inductively by

$$\Delta_h^r f(x) = \Delta_h(\Delta_h^{r-1} f(x)).$$

Modulus of continuity

Definition 7

For $h \in \mathbb{R}$, $f : \mathbb{R} \to \mathbb{C}$, the (forward) difference operator of step h is given by

$$\Delta_h f(x) = f(x+h) - f(x),$$

and for $r \in \mathbb{N}$, define the difference operator of order r, step h, inductively by

$$\Delta_h^r f(x) = \Delta_h(\Delta_h^{r-1} f(x)).$$

The $r-$th order modulus of smoothness of f in $L^2(\mathbb{R})$ is defined as the mapping

$$\omega_2^r(f,t) = \sup_{h \in \mathbb{R}, |h|<t} \|\Delta_h^r f(\cdot)\|_2.$$

Definition of Besov space

Definition of Besov space

Definition 8

Let f be a measurable function on $\mathbb{R}$.

Definition of Besov space

Definition 8

Let f be a measurable function on $\mathbb{R}$. For $1 < p < \infty$, $t \in \mathbb{R}_+$, the $r-$th order modulus of smoothness of f in $L^p(\mathbb{R})$ is defined by

$$\omega_p^r(f,t) = \sup_{h\in\mathbb{R},|h|<t} \|\Delta_h^r f(\cdot)\|_p.$$

Definition of Besov space

Definition 8

Let f be a measurable function on $\mathbb{R}$. For $1 < p < \infty$, $t \in \mathbb{R}_+$, the $r-$th order modulus of smoothness of f in $L^p(\mathbb{R})$ is defined by

$$\omega_p^r(f,t) = \sup_{h \in \mathbb{R}, |h| < t} \|\Delta_h^r f(\cdot)\|_p.$$

For $\alpha > 0, 1 \leq p, q < \infty$, $r = \lfloor \alpha \rfloor + 1$, a function f defined on $\mathbb{R}$ is said to be in the homogeneous Besov space $\dot{B}_{p,q}^{\alpha}(\mathbb{R})$, if

$$\|f\|_{\dot{B}_{p,q}^{\alpha}} := \left(\int_0^\infty (t^{-\alpha}\omega_p^r(f,t))^q \frac{dt}{t}\right)^{1/q} < \infty.$$

Definition of Besov space

Definition 8

Let f be a measurable function on $\mathbb{R}$. For $1 < p < \infty$, $t \in \mathbb{R}_+$, the $r-$th order modulus of smoothness of f in $L^p(\mathbb{R})$ is defined by

$$\omega_p^r(f,t) = \sup_{h\in\mathbb{R},|h|<t} \|\Delta_h^r f(\cdot)\|_p.$$

For $\alpha > 0, 1 \leq p, q < \infty$, $r = \lfloor \alpha \rfloor + 1$, a function f defined on $\mathbb{R}$ is said to be in the homogeneous Besov space $\dot{B}^\alpha_{p,q}(\mathbb{R})$, if

$$\|f\|_{\dot{B}^\alpha_{p,q}} := (\int_0^\infty (t^{-\alpha}\omega_p^r(f,t))^q \frac{dt}{t})^{1/q} < \infty.$$

Note

This definition extends the scale of homogeneous Sobolev spaces:
$\dot{W}^k_2(\mathbb{R}) = \dot{B}^k_{2,2}(\mathbb{R})$.

Overview

Aims of this section

Setup

Aims of this section

Setup

- Given: Wavelet ONB $\Psi = (\psi_{j,k})_{j,k \in \mathbb{Z}}$ of $\mathrm{L}^2(\mathbb{R})$.

Aims of this section

Setup

- Given: Wavelet ONB $\Psi = (\psi_{j,k})_{j,k \in \mathbb{Z}}$ of $\mathrm{L}^2(\mathbb{R})$.
- Question 1: Can we identify the ℓ^1-sparse vectors with respect to Ψ?

Aims of this section

Setup

- Given: Wavelet ONB $\Psi = (\psi_{j,k})_{j,k \in \mathbb{Z}}$ of $\mathrm{L}^2(\mathbb{R})$.
- Question 1: Can we identify the ℓ^1-sparse vectors with respect to Ψ?
- Question 2: Different wavelets may have different sparse vectors!

Aims of this section

Setup

- Given: Wavelet ONB $\Psi = (\psi_{j,k})_{j,k \in \mathbb{Z}}$ of $\mathrm{L}^2(\mathbb{R})$.
- Question 1: Can we identify the ℓ^1-sparse vectors with respect to Ψ?
- Question 2: Different wavelets may have different sparse vectors! Can we control that, i.e., can we isolate a class of good wavelets which all have the same sparse vectors?

Aims of this section

Setup

- Given: Wavelet ONB $\Psi = (\psi_{j,k})_{j,k \in \mathbb{Z}}$ of $\mathrm{L}^2(\mathbb{R})$.
- Question 1: Can we identify the ℓ^1-sparse vectors with respect to Ψ?
- Question 2: Different wavelets may have different sparse vectors! Can we control that, i.e., can we isolate a class of good wavelets which all have the same sparse vectors?
- Answers: Good wavelets will be described in terms of smoothness, decay and vanishing moments;

Aims of this section

Setup

- Given: Wavelet ONB $\Psi = (\psi_{j,k})_{j,k \in \mathbb{Z}}$ of $\mathrm{L}^2(\mathbb{R})$.
- Question 1: Can we identify the ℓ^1-sparse vectors with respect to Ψ?
- Question 2: Different wavelets may have different sparse vectors! Can we control that, i.e., can we isolate a class of good wavelets which all have the same sparse vectors?
- Answers: Good wavelets will be described in terms of smoothness, decay and vanishing moments; the associated space of sparse vectors will be a suitable homogeneous Besov space.

Nice wavelets

Nice wavelets

Desirable properties of wavelets

A nice wavelet $\psi \in \mathrm{L}^2(\mathbb{R})$ typically has three properties

Nice wavelets

Desirable properties of wavelets

A nice wavelet $\psi \in \mathrm{L}^2(\mathbb{R})$ typically has three properties

(a) Fast decay, e.g. $|\psi(x)| \le C(1+|x|)^{-n}$;

Nice wavelets

Desirable properties of wavelets

A nice wavelet $\psi \in \mathrm{L}^2(\mathbb{R})$ typically has three properties

(a) Fast decay, e.g. $|\psi(x)| \le C(1+|x|)^{-n}$;

(b) Smoothness, e.g. $\psi^{(j)} \in \mathrm{L}^1(\mathbb{R})$, for all $1 \le j \le m$;

Nice wavelets

Desirable properties of wavelets

A nice wavelet $\psi \in \mathrm{L}^2(\mathbb{R})$ typically has three properties

(a) Fast decay, e.g. $|\psi(x)| \leq C(1+|x|)^{-n}$;

(b) Smoothness, e.g. $\psi^{(j)} \in \mathrm{L}^1(\mathbb{R})$, for all $1 \leq j \leq m$;

(c) Vanishing moments, e.g.

$$\forall 0 \leq j < k \; : \; \int_{\mathbb{R}} x^j \psi(x) dx = 0$$

with absolute convergence of the integral

Nice wavelets

Desirable properties of wavelets

A nice wavelet $\psi \in \mathrm{L}^2(\mathbb{R})$ typically has three properties

(a) Fast decay, e.g. $|\psi(x)| \leq C(1+|x|)^{-n}$;

(b) Smoothness, e.g. $\psi^{(j)} \in \mathrm{L}^1(\mathbb{R})$, for all $1 \leq j \leq m$;

(c) Vanishing moments, e.g.

$$\forall 0 \leq j < k \;:\; \int_{\mathbb{R}} x^j \psi(x) dx = 0$$

with absolute convergence of the integral

Shortly: Nice wavelets have good time-frequency localization.

Nice wavelets

Desirable properties of wavelets

A nice wavelet $\psi \in \mathrm{L}^2(\mathbb{R})$ typically has three properties

(a) Fast decay, e.g. $|\psi(x)| \le C(1+|x|)^{-n}$;

(b) Smoothness, e.g. $\psi^{(j)} \in \mathrm{L}^1(\mathbb{R})$, for all $1 \le j \le m$;

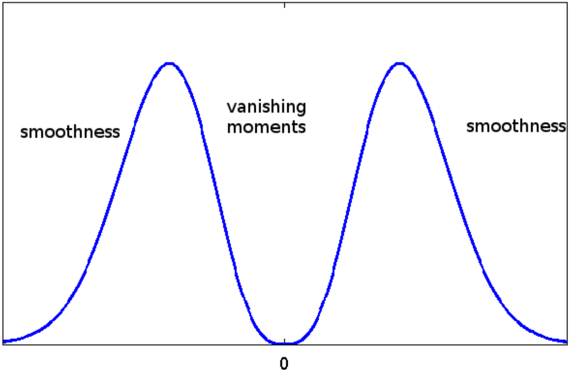
(c) Vanishing moments, e.g.

$$\forall 0 \le j < k \ : \ \int_{\mathbb{R}} x^j \psi(x) dx = 0$$

with absolute convergence of the integral

Shortly: Nice wavelets have good time-frequency localization.
Note: Frequency-side localization is understood away from zero.

Cartoon: Fourier side decay of wavelets



Plot of $|\widehat{\psi}|$.

Characterization of sparse vectors

Theorem 9 (DeVore/Jawerth/Popov)

Let $f \in \mathrm{L}^2(\mathbb{R})$.

Characterization of sparse vectors

Theorem 9 (DeVore/Jawerth/Popov)

Let $f \in \mathrm{L}^2(\mathbb{R})$. Assume that the wavelet ψ has suitably many vanishing moments, degrees of smoothness and decay, and let $\Psi = (\psi_{j,k})_{j,k}$ denote the associated wavelet ONB.

Characterization of sparse vectors

Theorem 9 (DeVore/Jawerth/Popov)

Let $f \in \mathrm{L}^2(\mathbb{R})$. Assume that the wavelet ψ has suitably many vanishing moments, degrees of smoothness and decay, and let $\Psi = (\psi_{j,k})_{j,k}$ denote the associated wavelet ONB. Then

$$f \text{ is } \ell^1\text{-sparse w.r.t. } \Psi \Longleftrightarrow f \in \dot{B}^{1/2}_{1,1}(\mathbb{R}) \quad .$$

Characterization of sparse vectors

Theorem 9 (DeVore/Jawerth/Popov)

Let $f \in \mathrm{L}^2(\mathbb{R})$. Assume that the wavelet ψ has suitably many vanishing moments, degrees of smoothness and decay, and let $\Psi = (\psi_{j,k})_{j,k}$ denote the associated wavelet ONB. Then

$$f \text{ is } \ell^1\text{-sparse w.r.t. } \Psi \iff f \in \dot{B}^{1/2}_{1,1}(\mathbb{R}) \quad .$$

Observations

Characterization of sparse vectors

Theorem 9 (DeVore/Jawerth/Popov)

Let $f \in \mathrm{L}^2(\mathbb{R})$. Assume that the wavelet ψ has suitably many vanishing moments, degrees of smoothness and decay, and let $\Psi = (\psi_{j,k})_{j,k}$ denote the associated wavelet ONB. Then

$$f \text{ is } \ell^1\text{-sparse w.r.t. } \Psi \Longleftrightarrow f \in \dot{B}^{1/2}_{1,1}(\mathbb{R}) \quad .$$

Observations

- Characterization of sparse vectors is independent of the wavelet.

Characterization of sparse vectors

Theorem 9 (DeVore/Jawerth/Popov)

Let $f \in \mathrm{L}^2(\mathbb{R})$. Assume that the wavelet ψ has suitably many vanishing moments, degrees of smoothness and decay, and let $\Psi = (\psi_{j,k})_{j,k}$ denote the associated wavelet ONB. Then

$$f \text{ is } \ell^1\text{-sparse w.r.t. } \Psi \Longleftrightarrow f \in \dot{B}^{1/2}_{1,1}(\mathbb{R}) \quad .$$

Observations

- Characterization of sparse vectors is independent of the wavelet.
- Characterization extends to arbitrary Besov spaces $\dot{B}^{\alpha}_{p,q} \rightsquigarrow$ weighted mixed summability condition on the coefficients.

Characterization of sparse vectors

Theorem 9 (DeVore/Jawerth/Popov)

Let $f \in \mathrm{L}^2(\mathbb{R})$. Assume that the wavelet ψ has suitably many vanishing moments, degrees of smoothness and decay, and let $\Psi = (\psi_{j,k})_{j,k}$ denote the associated wavelet ONB. Then

$$f \text{ is } \ell^1\text{-sparse w.r.t. } \Psi \Longleftrightarrow f \in \dot{B}^{1/2}_{1,1}(\mathbb{R}) \quad .$$

Observations

- Characterization of sparse vectors is independent of the wavelet.
- Characterization extends to arbitrary Besov spaces $\dot{B}^{\alpha}_{p,q} \rightsquigarrow$ weighted mixed summability condition on the coefficients.
- Wavelet basis is a joint unconditional basis of a whole scale of Besov spaces.

Significance of wavelet characterization

Significance of wavelet characterization

In mathematics:

Significance of wavelet characterization

In mathematics:

- Approximation theory and its applications (numerical analysis, etc.);

Significance of wavelet characterization

In mathematics:

- Approximation theory and its applications (numerical analysis, etc.);
- Functional analysis: Study of operators;

Significance of wavelet characterization

In mathematics:

- Approximation theory and its applications (numerical analysis, etc.);
- Functional analysis: Study of operators;
- Statistics, etc.

Significance of wavelet characterization

In mathematics:

- Approximation theory and its applications (numerical analysis, etc.);
- Functional analysis: Study of operators;
- Statistics, etc.

In applications

Important source of heuristics for signal processing applications, most prominently

- compression;
- denoising.

Summary

Summary

- Systematic way of predicting wavelet coefficient decay from analytic properties of the signal, i.e., a method of spotting the sparse signals.

Summary

- Systematic way of predicting wavelet coefficient decay from analytic properties of the signal, i.e., a method of spotting the sparse signals.
- Important concept: *Consistency*, i.e., one needs to develop a theory of sparse signals that is independent of the wavelet, as long as the latter is chosen from a suitably defined and accessible class of nice wavelet.

Summary

- Systematic way of predicting wavelet coefficient decay from analytic properties of the signal, i.e., a method of spotting the sparse signals.
- Important concept: *Consistency*, i.e., one needs to develop a theory of sparse signals that is independent of the wavelet, as long as the latter is chosen from a suitably defined and accessible class of nice wavelet.
- Here: Nice wavelets defined in terms of decay, smoothness, vanishing moments.

Summary

- Systematic way of predicting wavelet coefficient decay from analytic properties of the signal, i.e., a method of spotting the sparse signals.
- Important concept: *Consistency*, i.e., one needs to develop a theory of sparse signals that is independent of the wavelet, as long as the latter is chosen from a suitably defined and accessible class of nice wavelet.
- Here: Nice wavelets defined in terms of decay, smoothness, vanishing moments.
- Plan for the following: Develop an analogous theory for wavelets in higher dimensions.

Summary

- Systematic way of predicting wavelet coefficient decay from analytic properties of the signal, i.e., a method of spotting the sparse signals.
- Important concept: *Consistency*, i.e., one needs to develop a theory of sparse signals that is independent of the wavelet, as long as the latter is chosen from a suitably defined and accessible class of nice wavelet.
- Here: Nice wavelets defined in terms of decay, smoothness, vanishing moments.
- Plan for the following: Develop an analogous theory for wavelets in higher dimensions.
  Challenge: The notion will have to depend on the dilation group!

Overview

Ronald A. DeVore.
Nonlinear approximation.
In *Acta numerica, 1998*, volume 7 of *Acta Numer.*, pages 51–150. Cambridge Univ. Press, Cambridge, 1998.

Ronald A. DeVore, Björn Jawerth, and Bradley J. Lucier.
Image compression through wavelet transform coding.
IEEE Trans. Inform. Theory, 38(2, part 2):719–746, 1992.

Ronald A. DeVore, Björn Jawerth, and Vasil Popov.
Compression of wavelet decompositions.
Amer. J. Math., 114(4):737–785, 1992.

David L. Donoho and Iain M. Johnstone.
Adapting to unknown smoothness via wavelet shrinkage.
J. Amer. Statist. Assoc., 90(432):1200–1224, 1995.

Michael Frazier and Björn Jawerth.
Decomposition of Besov spaces.
Indiana Univ. Math. J., 34(4):777–799, 1985.

Michael Frazier, Björn Jawerth, and Guido Weiss.
Littlewood-Paley theory and the study of function spaces, volume 79 of *CBMS Regional Conference Series in Mathematics*.
Published for the Conference Board of the Mathematical Sciences, Washington, DC, 1991.

Yves Meyer.
Wavelets and operators, volume 37 of *Cambridge Studies in Advanced Mathematics*.
Cambridge University Press, Cambridge, 1992.
Translated from the 1990 French original by D. H. Salinger.

P. Wojtaszczyk.
A mathematical introduction to wavelets, volume 37 of *London Mathematical Society Student Texts*.
Cambridge University Press, Cambridge, 1997.