日本音響学会2015年秋季研究発表大会
ビギナーズセミナー

スパース信号表現って何？

東京大学大学院情報理工学系研究科
小山 翔一

スパース信号表現とは

- 信号を、ある基底の下で多くの成分の大きさがゼロ、またはゼロとみなせるほど小さくなる（**スパース**になる）ように表現すること。

- **圧縮センシング（Compressed sensing）**
  - スパースに表現可能な信号を少数の計測データから再構成を行う。
  - Donoho, Candés, Taoら（2003年頃）

- 応用例
  - トモグラフィ（X線CT、MRI）、画像復元（ノイズ除去、超解像など）

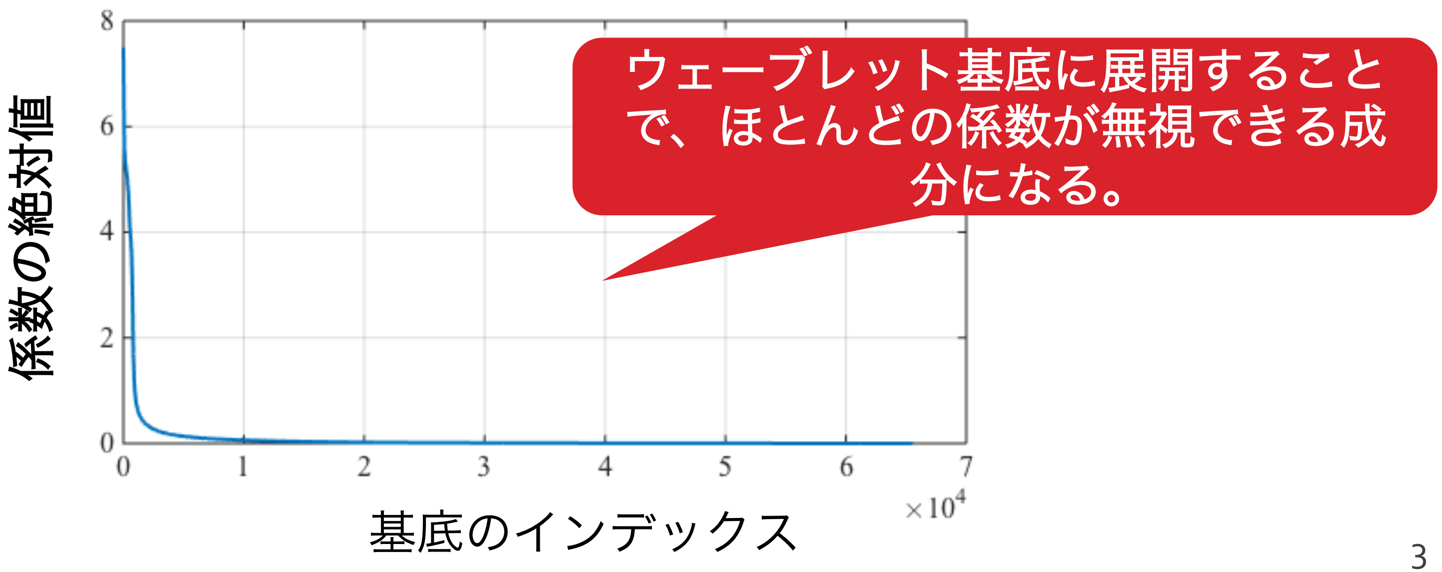
例：画像のスパース表現

- ウェーブレット基底での画像の表現

原画像

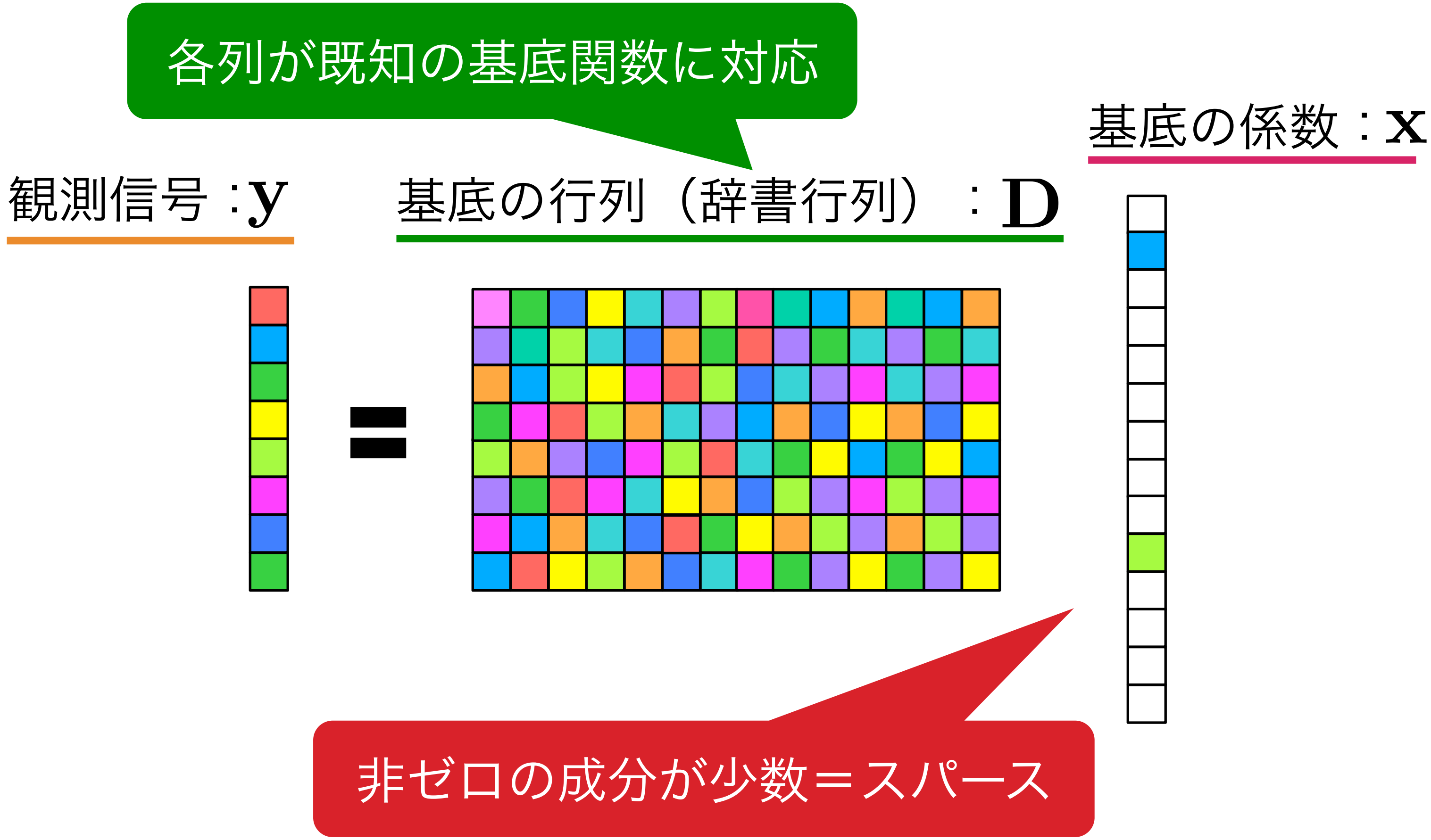
上位10%

上位3%



スパース信号表現とは

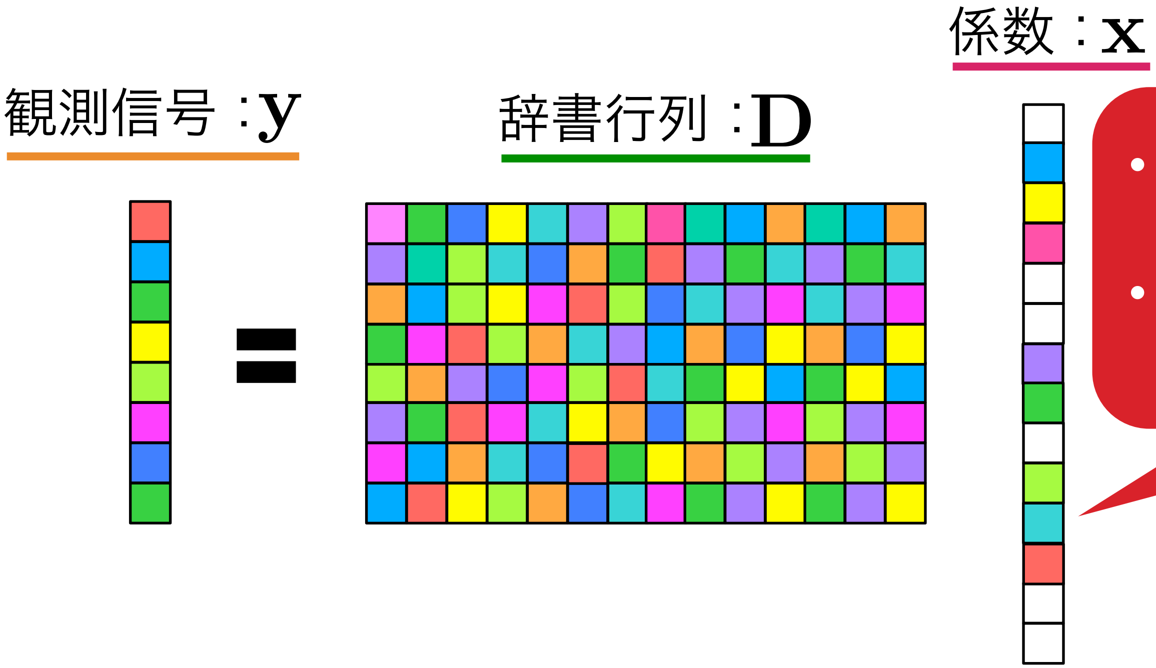
- 行列形式では、線形方程式として書ける。



$\mathbf{x}$ がスパースになるとなぜ嬉しいのか？
どうやってスパースな $\mathbf{x}$ を求めるのか？

圧縮センシング

➢ ゼロとなる要素が未知の場合 ➡ **圧縮センシング**の問題



- 未知数の数＞方程式の数の劣決定問題
- ゼロの成分が多いことはわかっている（場所は不明）

– 信号がある領域で多くのゼロ成分を持つものは多い

❑ どのようなアルゴリズムでスパースな解を求め、少ないサンプルから信号復元を行うか

❑ どのような条件において、信号復元の性能が保証されるか

信号復元アルゴリズム

- ベクトル $\mathbf{x} \in \mathbb{R}^N$ のノルム

$$\|\mathbf{x}\|_p = \left(\sum_{n=1}^{N} |x_n|^p \right)^{1/p} = \sqrt[p]{|x_1|^p + \cdots + |x_N|^p}$$

- 以下のコスト関数を最小化する $\mathbf{x}$ が、スパースな解として最良

$$\text{minimize } \|\mathbf{x}\|_0 \text{ subject to } \mathbf{y} = \mathbf{D}\mathbf{x}$$

非ゼロの要素の個数

非ゼロ要素に関するあらゆる数・位置の組み合わせが存在し、
解を求めることは計算量的に困難（NP困難な問題）

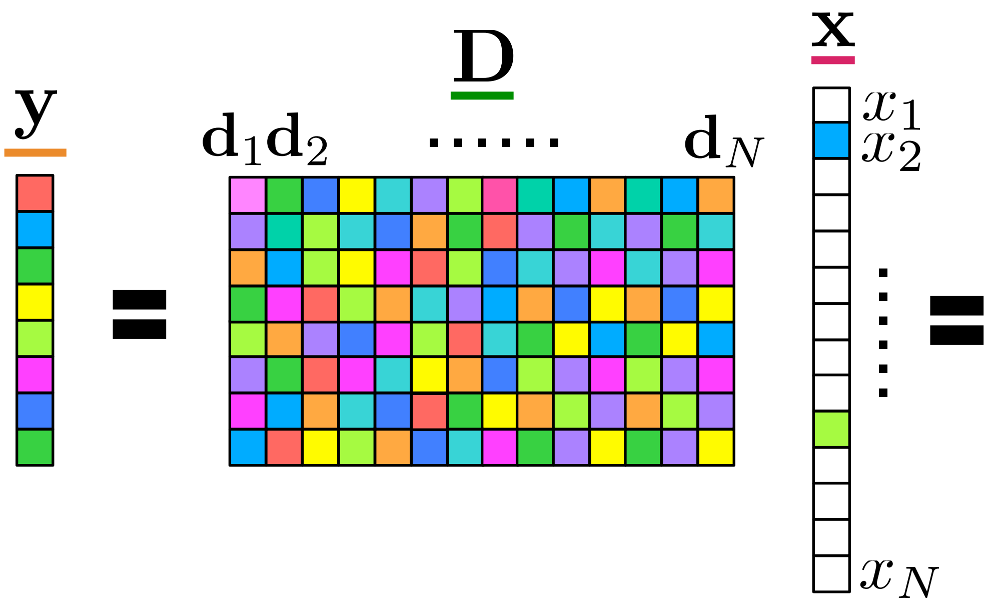
信号復元アルゴリズム

- 現実的な解法のいくつかを紹介

- **貪欲法**
  - 最良のサポートを順に付け加えることでサポートを構成
  - Orthogonal Matching Pursuit (OMP)など
- **凸緩和法**
  - コスト関数を凸関数として近似する
  - Basis Pursuit、Iterated-Reweighted-Least-Squares (IRLS)など
- **確率推論**
  - 統計的な推論による方法
  - Approximate message passing (AMP)など

貪欲法 (Greedy algorithms)

➢ 観測ベクトル $\mathbf{y}$ を最も良く近似する列ベクトルを順にサポートに付け加える



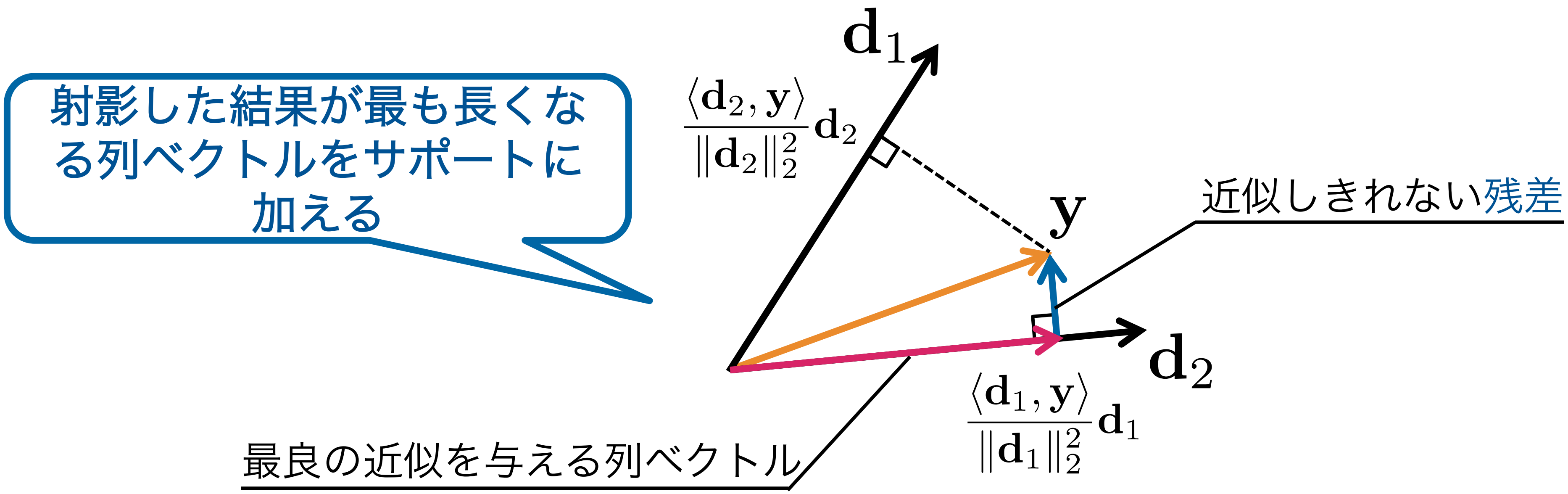
$$= x_1\mathbf{d}_1 + x_2\mathbf{d}_2 + \cdots + x_N\mathbf{d}_N$$

– 辞書行列の各列 $\{\mathbf{d}_1, \mathbf{d}_2, \cdots, \mathbf{d}_N\}$ の中から、$\mathbf{y}$を最も良く近似するものを探索

– 最良の候補を$\mathbf{x}$のサポートS の要素として加えて行く

貪欲法 (Greedy algorithms)

- 単一の列ベクトルによる近似を繰り返す



最良近似を与える射影 $\hat{x}_n = \arg\min_{x_n} \|\mathbf{y} - x_n \mathbf{d}_n\|_2^2 = \frac{\langle \mathbf{d}_2, \mathbf{y} \rangle}{\|\mathbf{d}_2\|_2^2}\mathbf{d}_2$

近似誤差 $\epsilon_n = \min_{x_n} \|\mathbf{y} - x_n \mathbf{d}_n\|_2^2 = \|\mathbf{y}\|_2^2 - \frac{\langle \mathbf{d}_n, \mathbf{y} \rangle^2}{\|\mathbf{d}_n\|_2^2}$

Orthogonal Matching Pursuit (OMP)

- **Initialization**: Initialize k=0, and set

$$\mathbf{x}^0 = \mathbf{0}, \mathbf{r}^0 = \mathbf{y} - \mathbf{D}\mathbf{x}^0, S^0 = \emptyset$$

- **Main Iteration**: Increment k by 1 and perform the following steps
  - **Sweep**:

$$\epsilon_n = \min_{x_n} \|\mathbf{r}^{k-1} - x_n \mathbf{d}_n\|_2^2 = \|\mathbf{r}^{k-1}\|_2^2 - \frac{\langle \mathbf{d}_n, \mathbf{r}^{k-1} \rangle^2}{\|\mathbf{d}_n\|_2^2}$$

  - **Update support**:

$$\hat{n} = \arg \min_{n \notin S^{k-1}} \epsilon_n, \;\; S^k = S^{k-1} \cup \{\hat{n}\}$$

  - **Update provisional solution**:

$$\hat{\mathbf{x}}^k = \arg \min_{\mathbf{x}_{S^k}} \|\mathbf{y} - \mathbf{D}_{S^k} \mathbf{x}_{S^k}\|_2^2$$

  - **Update residual**:

$$\mathbf{r}^k = \mathbf{y} - \mathbf{D}_{S^k} \hat{\mathbf{x}}^k$$

  - **Stopping rule**: Stop if $\|\mathbf{r}^k\|_2 < \varepsilon$ holds.

凸緩和法 (Convex relaxation)

- ℓ_0ノルムは離散最適化となり困難であるため、連続関数で近似する

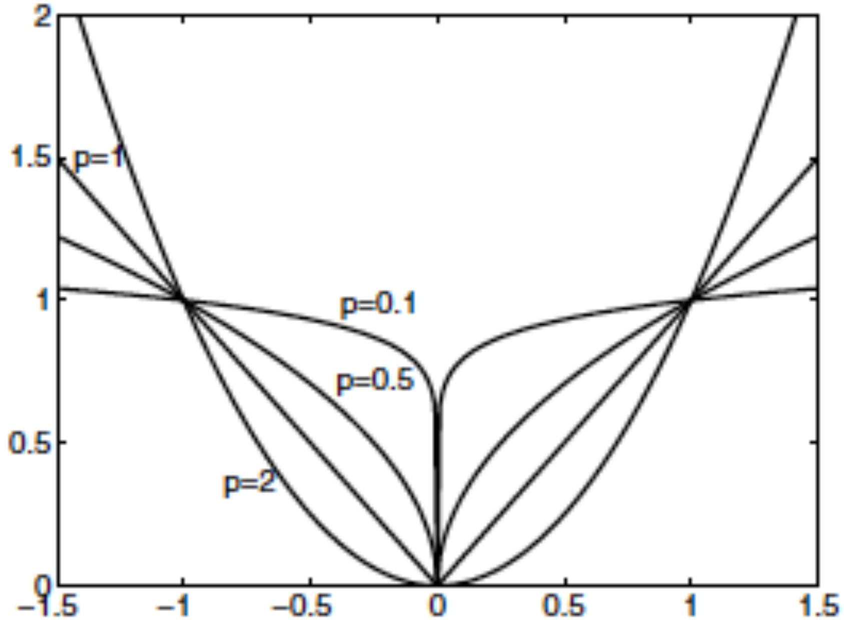
ℓ_0 ノルム

$$\|\mathbf{x}\|_0 = \lim_{p \to 0} \sum_{n=1}^{N} |x_n|^p$$

ℓ_p ノルム

$$\|\mathbf{x}\|_p = \left(\sum_{n=1}^{N} |x_n|^p \right)^{1/p}$$

$|x|^p$ のグラフ



M. Elad, "Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing," Springer, 2010.

凸緩和法 (Convex relaxation)

- 凸関数となれば最適化が容易となる ➡ $p \geq 1$
- 凸関数かつ最適解がスパースになることが望ましい

ℓ_p ノルム

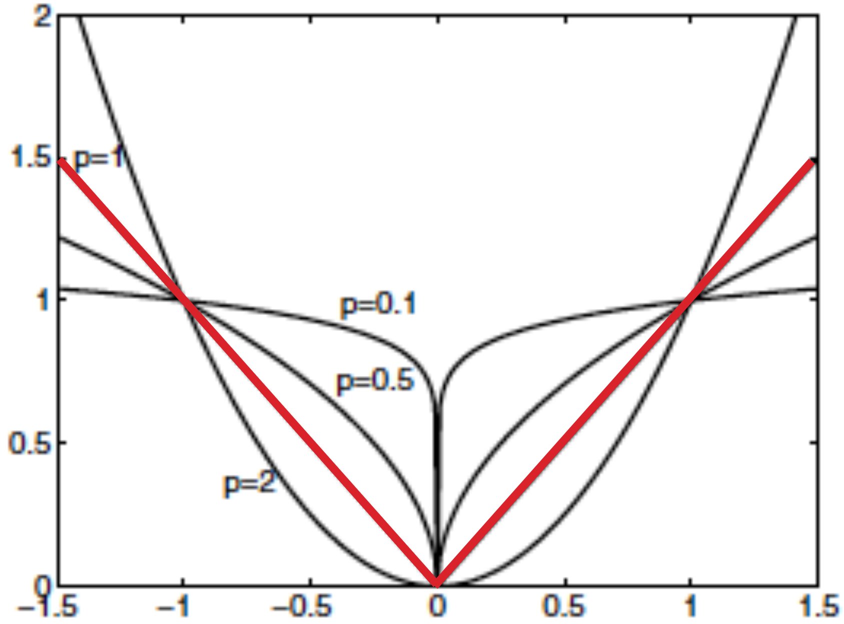
$$\|\mathbf{x}\|_p = \left(\sum_{n=1}^{N} |x_n|^p \right)^{1/p}$$

$p = 1$

ℓ_1 ノルム

$$\|\mathbf{x}\|_1 = \sum_{n=1}^{N} |x_n|$$

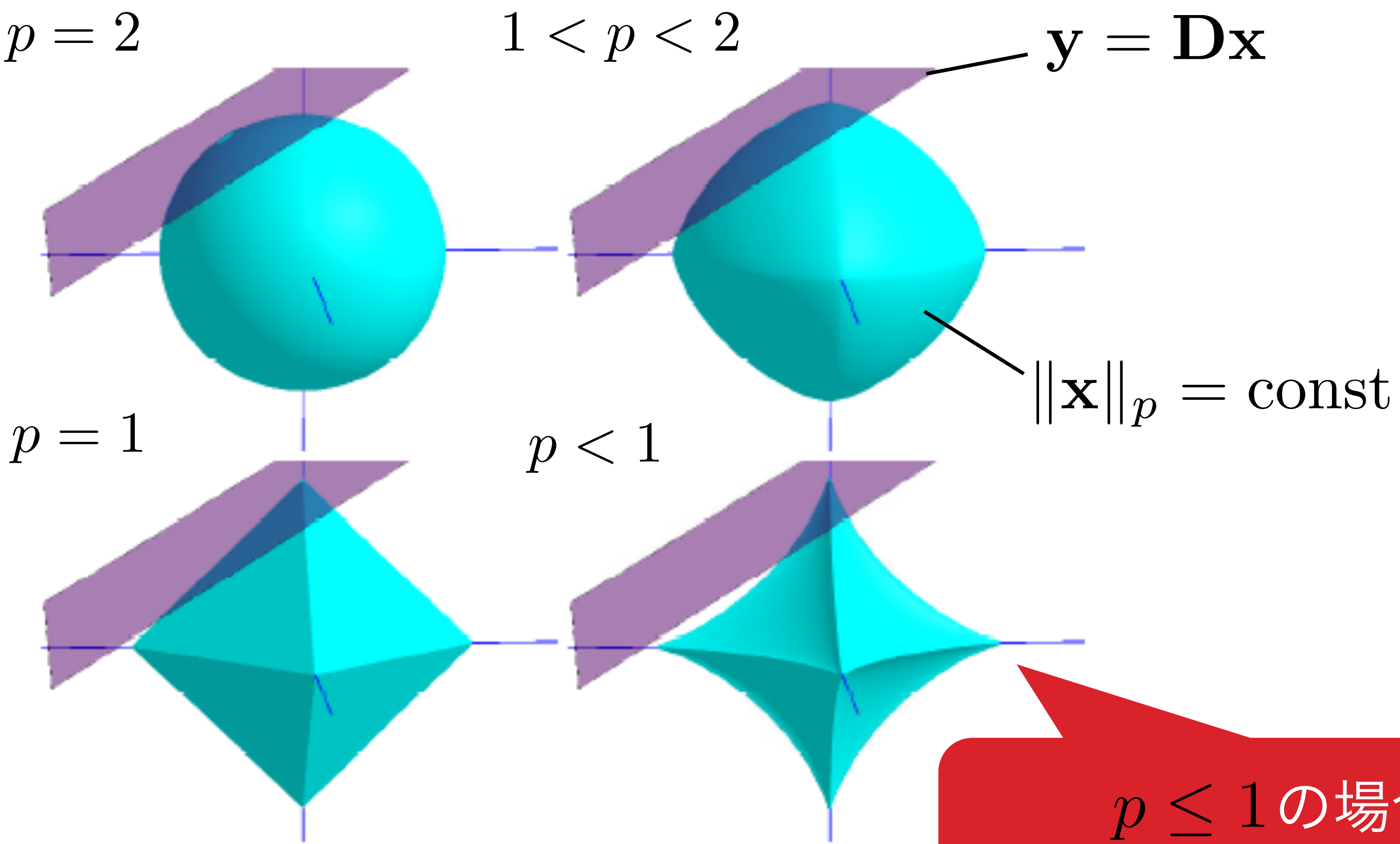
$|x|^p$ のグラフ



M. Elad, "Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing," Springer, 2010.

ノルムと解との関係

$$\boxed{\text{minimize } \|\mathbf{x}\|_p \quad \text{subject to } \ \mathbf{y} = \mathbf{D}\mathbf{x}}$$



M. Elad, “Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing,” Springer, 2010.

$p \leq 1$の場合に
解がスパースになりやすい

Basis Pursuit

- minimize $\|\mathbf{x}\|_1$ subject to $\mathbf{y} = \mathbf{D}\mathbf{x}$

- 線形計画問題へ書き換える

$$\mathbf{x} = \mathbf{u} - \mathbf{v} \quad \begin{cases} \mathbf{u} : \mathbf{x} \text{ の正の成分はそのまま、負の成分をゼロ} \\ \mathbf{v} : \mathbf{x} \text{ の正の成分はゼロ、負の成分をその絶対値} \end{cases}$$

$$\mathbf{z} = \left[\mathbf{u}^T, \mathbf{v}^T\right]^T \in \mathbb{R}^{2N}$$

全ての要素が1のベクトル

$$\Rightarrow \|\mathbf{x}\|_1 = \mathbf{1}^T(\mathbf{u} + \mathbf{v}),\ \mathbf{D}\mathbf{x} = \mathbf{D}(\mathbf{u} - \mathbf{v}) = [\mathbf{D}, -\mathbf{D}]\,\mathbf{z}$$

$$\Rightarrow \boxed{\text{minimize } \mathbf{1}^T\mathbf{z} \text{ subject to } \mathbf{y} = [\mathbf{D}, -\mathbf{D}]\,\mathbf{z} \text{ and } \mathbf{z} \succeq \mathbf{0}}$$

線形計画問題を解くためのアルゴリズムが使える

Basis Pursuit

➢ MATLAB code w/ Optimization Toolbox

線形計画問題を解く関数

```
1.z = linprog(ones(2*N,1), [], [],
   [D,-D], y, zeros(2*N,1), []);

2.x = z(1:N)-z(N+1:2*N);
```

z から x に変換

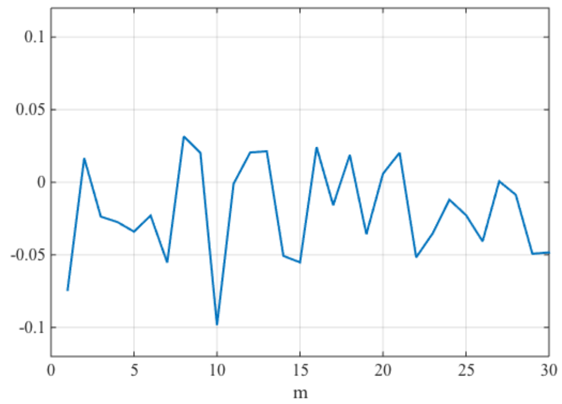
MATLABだと２行で書けます。

数値例

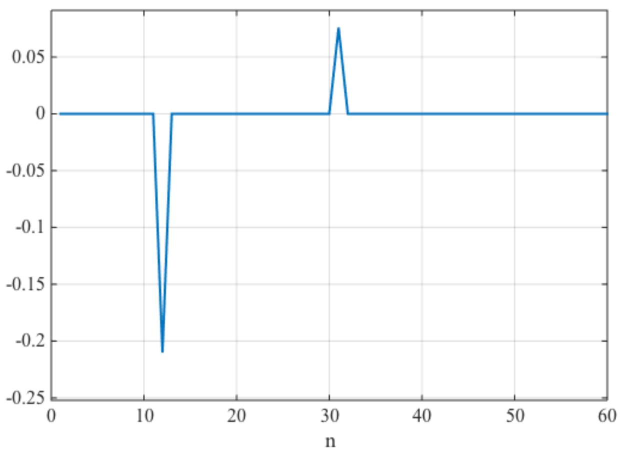
- $\mathbf{y} \in \mathbb{R}^{30},\ \mathbf{x} \in \mathbb{R}^{60},\ \mathbf{D} \in \mathbb{R}^{30\times 60}$ の場合の実験結果

各要素を乱数で生成

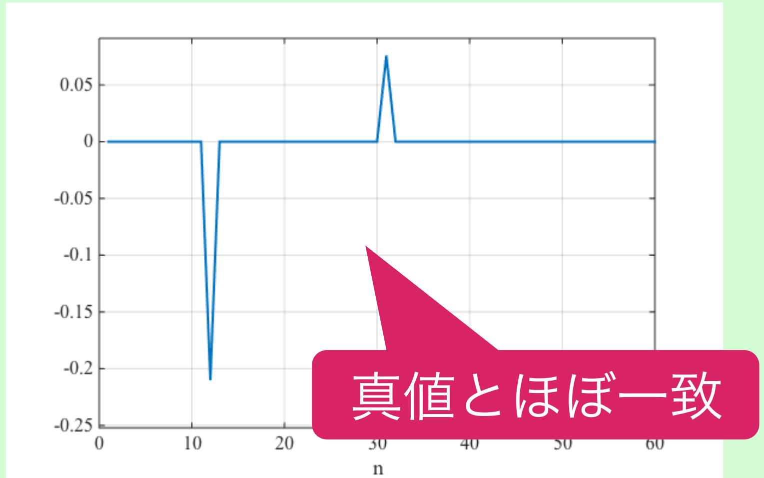
観測信号 $\mathbf{y}$



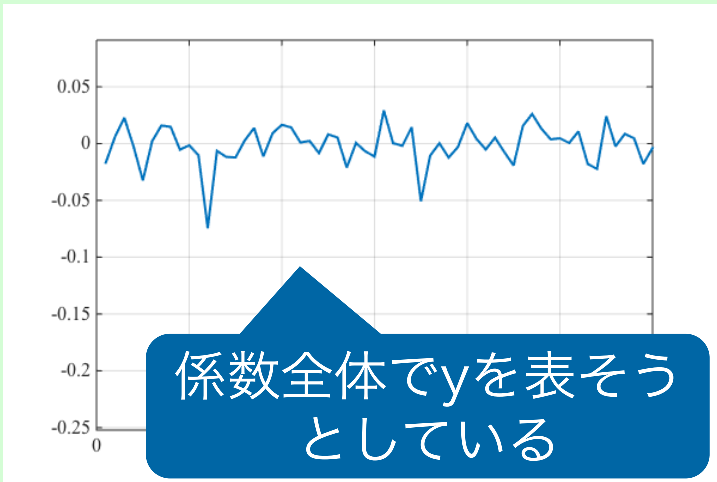
真値 $\mathbf{x}$



Basis Pursuit（p=1）

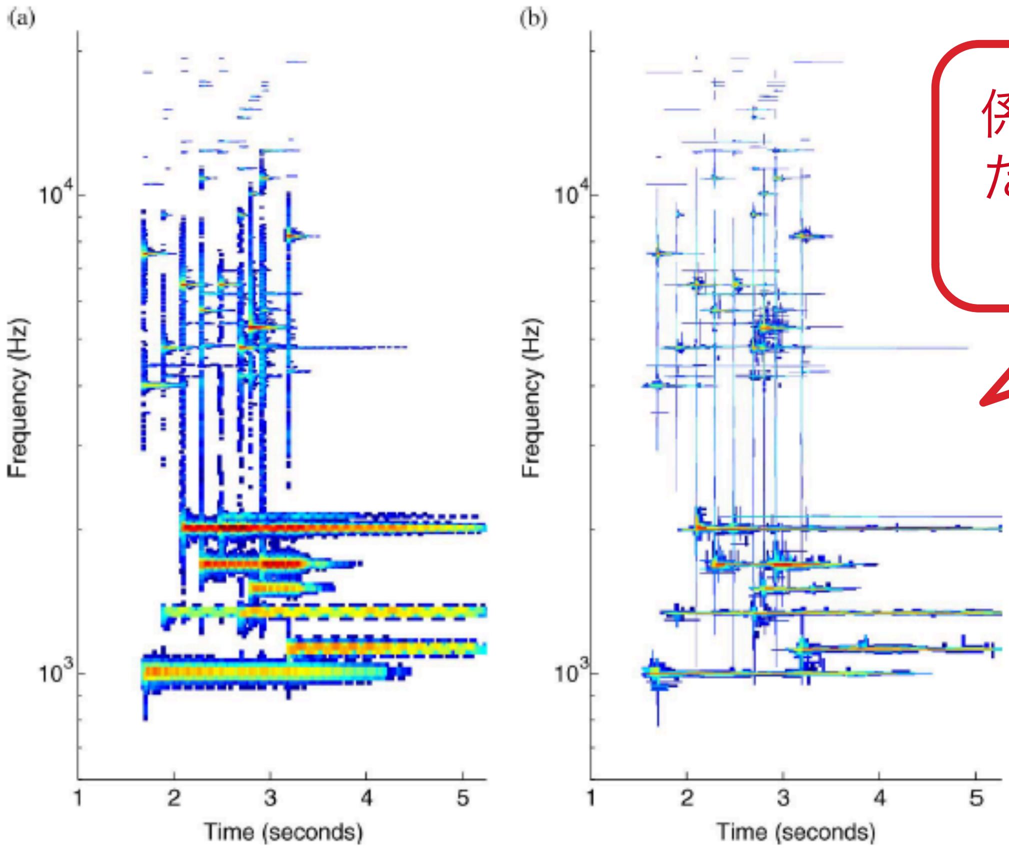


最小ノルム解（p=2）



音声・音響信号処理への応用①

➢ オーディオ符号化 [Plumbley+ Proc. IEEE 2010]



係数のほとんどがゼロとなるため、より低ビットレートでデータ圧縮が可能となる

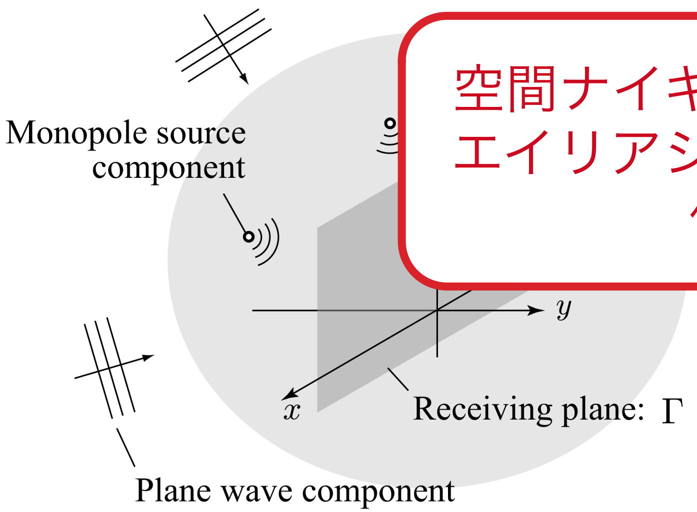
一般的なMDCT基底

窓長を変えた8つのMDCT基底の組み合わせ

音声・音響信号処理への応用②

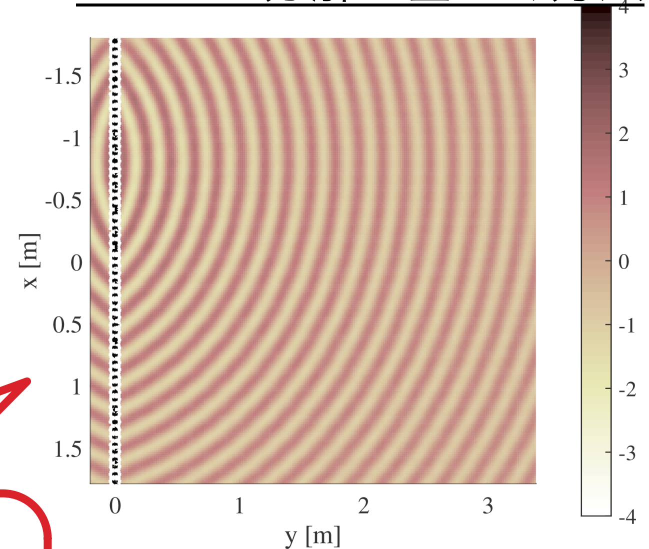
➢ 音場収音・再現の超解像化 [Koyama+ ICASSP 2014]

- 音場をモノポール音源成分と平面波成分の和としてモデル化
- モノポール音源成分が空間的にスパースに分布していることを仮定
- 各成分に分解した上でスピーカ駆動信号に変換し、空間エイリアシングの誤差を低減

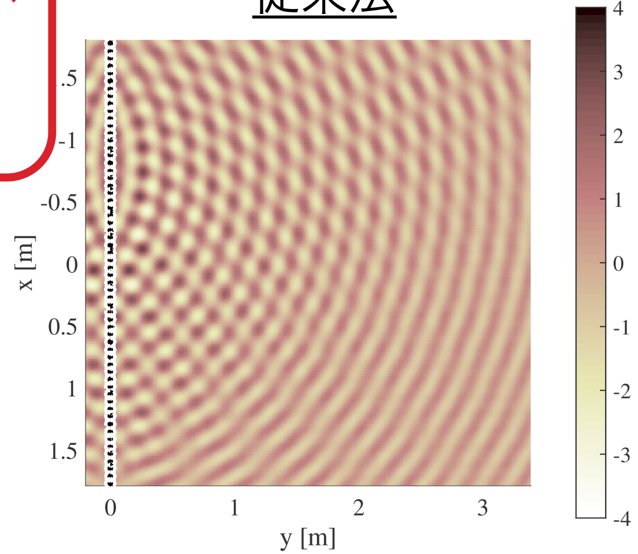


空間ナイキスト周波数以上で、エイリアシングによる誤差を低減可能

スパース分解に基づく方法



従来法



まとめ

- スパース信号表現って何？
  - 信号をある基底の下でスパースになるように表現する。
  - 少ないサンプルから信号を復元。
  - 貪欲法、凸緩和法など、現実的な計算時間で解けるアルゴリズムが提案されている。
  - 音声・音響信号処理に対しても広く応用可能。
  - 「L1正則化は21世紀の最小二乗法」 by E. Candés
- 音響学会誌, 71巻, 11号にて、「**スパース表現に基づく音響信号処理**」の小特集があります！

関連文献

- M. Elad, "Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing," Springer, 2010.
- 樺島, "圧縮センシングの数理的側面〜Compressed Sensing / Compressive Sensing〜," セミナー資料, 2013.
- M. D. Plumbley, T, Blumensath, L. Daudet, R. Gribonval, and M. E. Davies, "Sparse Representations in Audio and Music: From Coding to Source Separation," Proc. IEEE, vol. 98, no. 6, 2010.
- E. Ravelli, G. Richard, and L. Daudet, "Union of MDCT Bases for Audio Coding," IEEE Trans. ASLP, vol. 16, no. 8, 2008.
- S. Koyama, S. Shimauchi, H. Ohmuro, "Sparse sound field representation in recording and reproduction for reducing spatial aliasing artifacts," Proc. IEEE ICASSP, 2014.
- B. A. Cipra, "L1 magic," SIAM News, vol. 39, no. 9, 2006.