

スパースモデリング

岡山大学 異分野基礎科学研究所

大槻純也



参考文献

Journal of the Physical Society of Japan **89**, 012001 (2020)

Invited Review Papers

<https://doi.org/10.7566/JPSJ.89.012001>

Sparse Modeling in Quantum Many-Body Problems

Junya Otsuki¹, Masayuki Ohzeki², Hiroshi Shinaoka³, and Kazuyoshi Yoshimi⁴

¹*Research Institute for Interdisciplinary Science, Okayama University, Okayama 700-8530, Japan*

²*Graduate School of Information Sciences, Tohoku University, Sendai 980-8579, Japan*

³*Department of Physics, Saitama University, Sakura, Saitama 338-8570, Japan*

⁴*Institute for Solid State Physics, University of Tokyo, Kashiwa, Chiba 277-8581, Japan*

open access

固体物理 Vol. 53, No. 4, pp. 173–188 (2018) 出版原稿

最大エントロピー法でいいの？

—スパースモデリングの量子多体論への応用—

東北大学大学院理学研究科	大槻純也
東北大学大学院情報科学研究科	大関真之
埼玉大学大学院理工学研究科	品岡 寛
東京大学物性研究所	吉見一慶

私のHPからダウンロード可

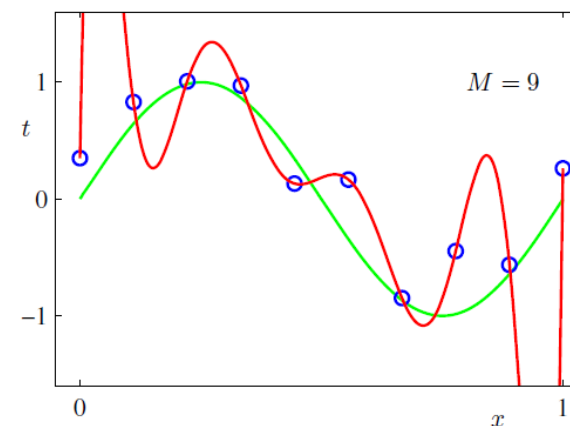
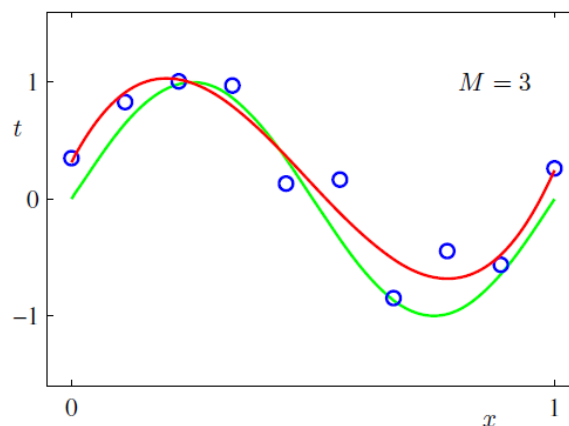
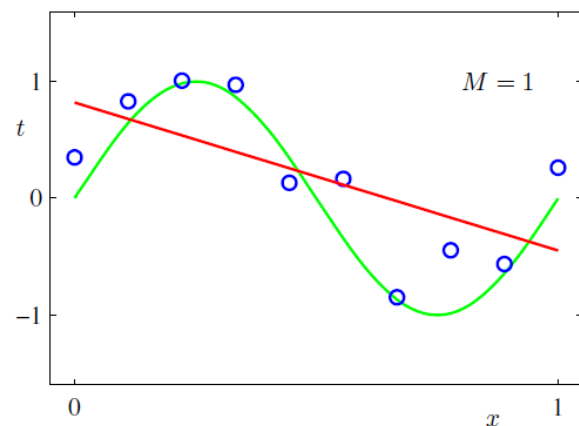
http://www.physics.okayama-u.ac.jp/~otsuki/activity_article.html

まずは一般的な話

小数のパラメーターで
実験データをそこそこ
フィットするモデル



たくさんのパラメーターで
実験データを完璧に
フィットするモデル



緑：真の曲線
赤：予測曲線
青点：観測データ

C. M. Bishop, PRML

知りたいのは「何が本質か」
不要なものを積極的に削りたい

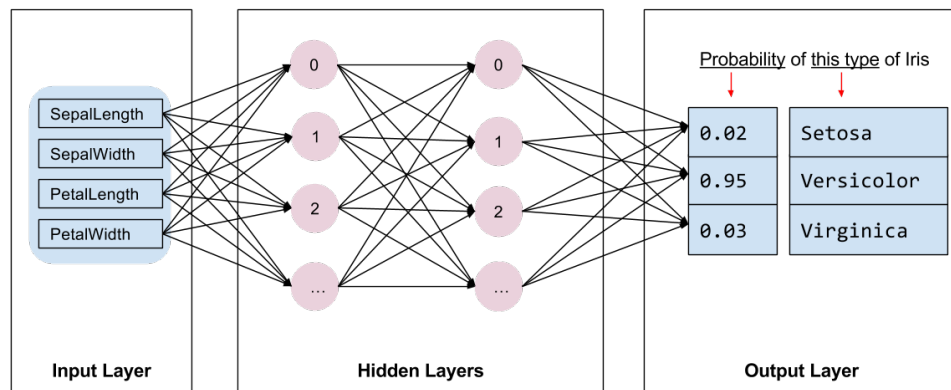


スパース性

機械学習 vs 数理統計 (データ科学)

機械学習

非線形変換による大規模最適化



tensorflow.org

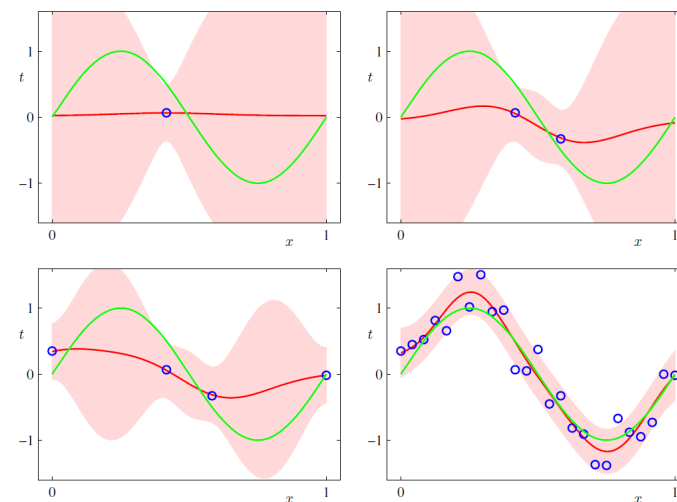
画像認識、予測、分類、囲碁・将棋、...

物理への応用：物質探索、変分波動関数、...

人間も予想しない結果

数理統計 (データ科学)

統計学に基づいた大規模数値計算法
(ベイズ統計学、多変量解析、回帰...)



C. M. Bishop, PRML

情報抽出、確率に基づいて予測

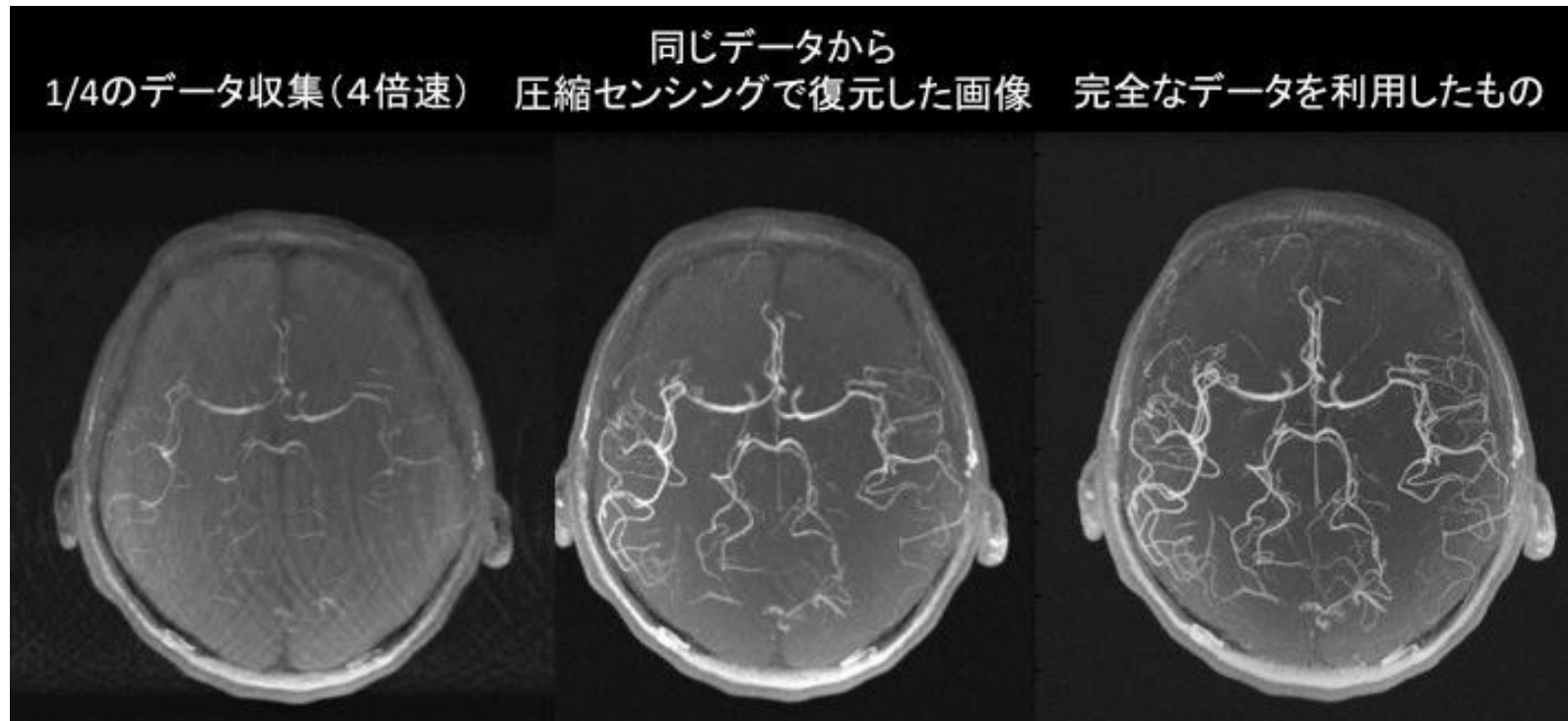
今日の話はこちら

圧縮センシング：MRIの例

Compressed sensing, Compressive sensing, Sparse sampling...

Candes, Romberg, Tao 2006
Donoho 2006
Lustig, Donoho, Pauly 2007

一部のフーリエ成分を計測 → 実空間像を生成



少ないデータからでも
完全に近い画像を復元

→ 計測の時間削減

解析精度向上

これまで計測精度不足で
見えなかったものが見える

圧縮センシングの数理

波数空間の測定 (例: MRI)

$$f(\mathbf{k}_i) = \int_V d\mathbf{r} \rho(\mathbf{r}) e^{-i\mathbf{k}_i \cdot \mathbf{r}}.$$

フーリエ信号(測定) 密度分布(知りたい)

離散化

$$f(\mathbf{k}_i) \rightarrow y_i, \quad e^{-i\mathbf{k}_i \cdot \mathbf{r}_j} \rightarrow A_{ij},$$

$$\rho(\mathbf{r}_j)(V/N) \rightarrow x_j,$$

$$\mathbf{y} = \mathbf{A}\mathbf{x} \quad \longrightarrow \quad \mathbf{x} = \mathbf{A}^{-1}\mathbf{y}$$

逆変換

一部のみ測定

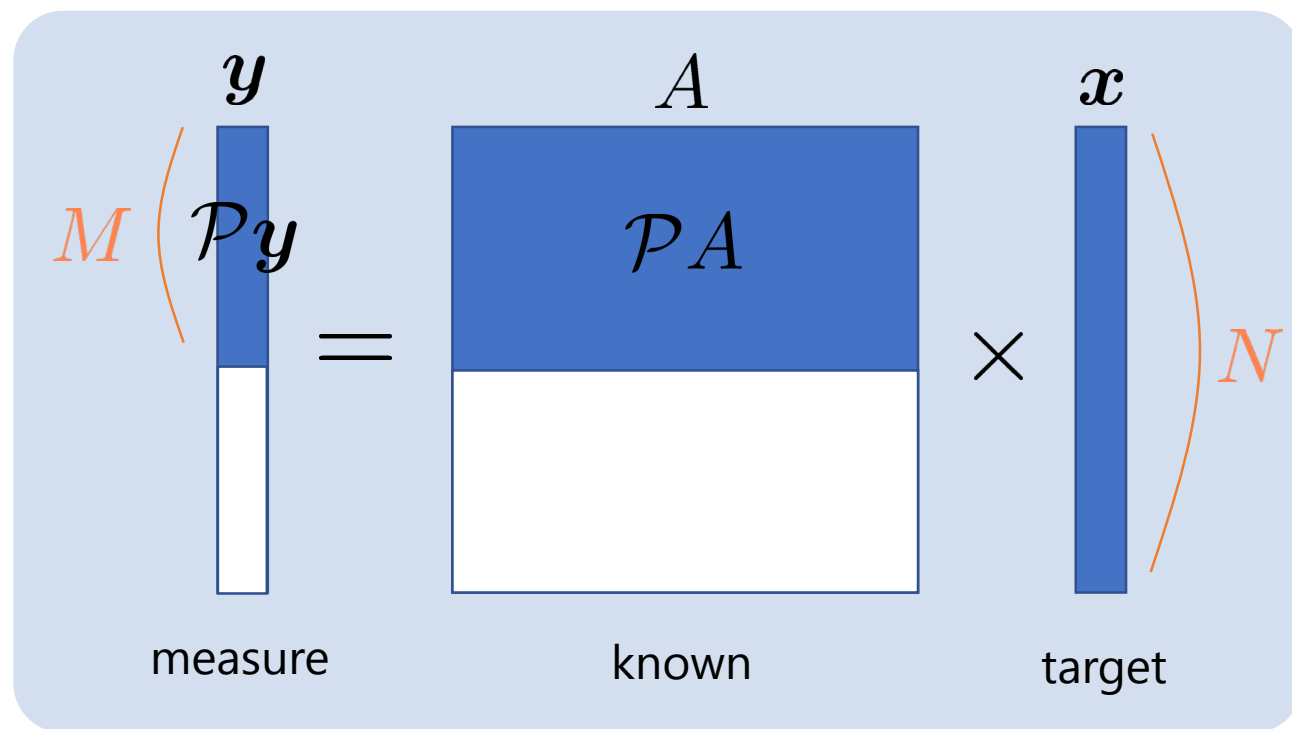
$$(\mathcal{P}\mathbf{y}) = (\mathcal{P}\mathbf{A})\mathbf{x} \quad \longrightarrow \quad \text{逆変換可能か?}$$

測定範囲への射影

Candes, Romberg, Tao 2006

Donoho 2006

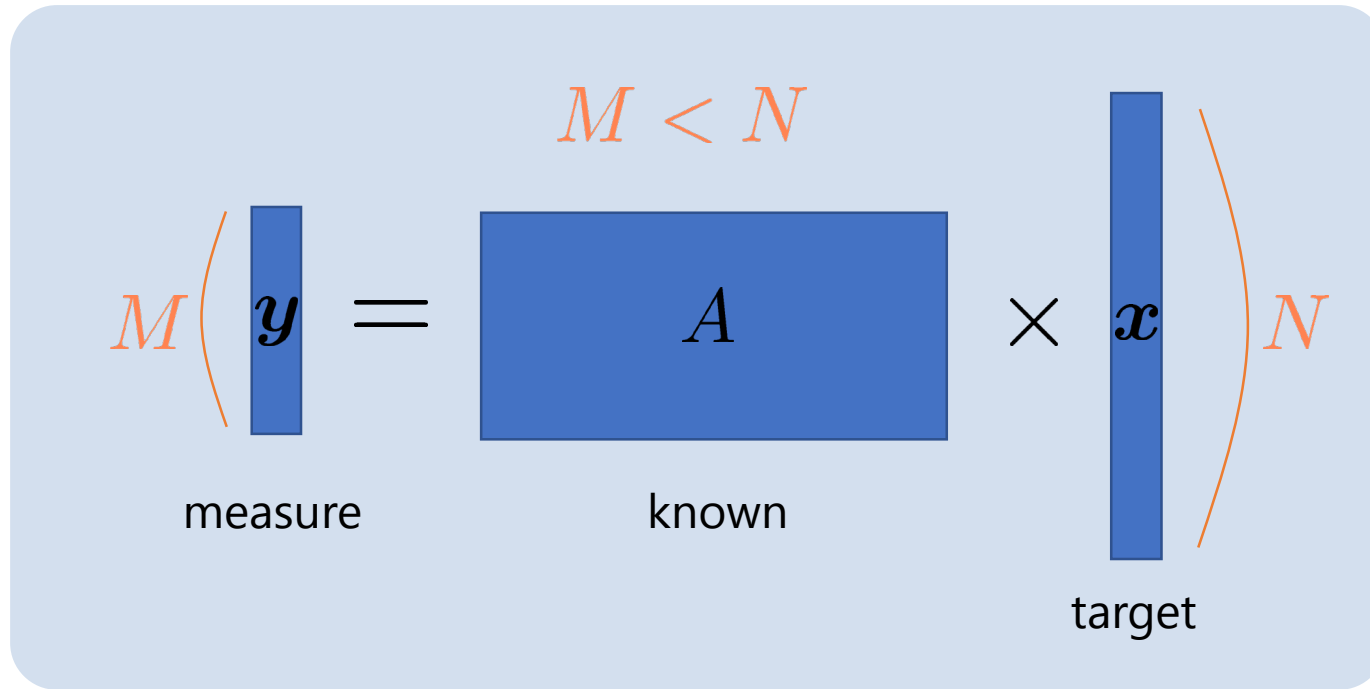
Lustig, Donoho, Pauly 2007



劣決定系逆問題

(Under-determined inverse problem)

劣決定系逆問題 (Under-determined inverse problem)



$\mathcal{P}y \rightarrow y, \mathcal{P}A \rightarrow A$ と定義しなおした

解が一意に定まらない
＝解けない（真の解は分からない）

しかし、 x の性質によっては
解ける場合がある（事前情報）

スパース性

x の成分のうち
の多くはゼロ



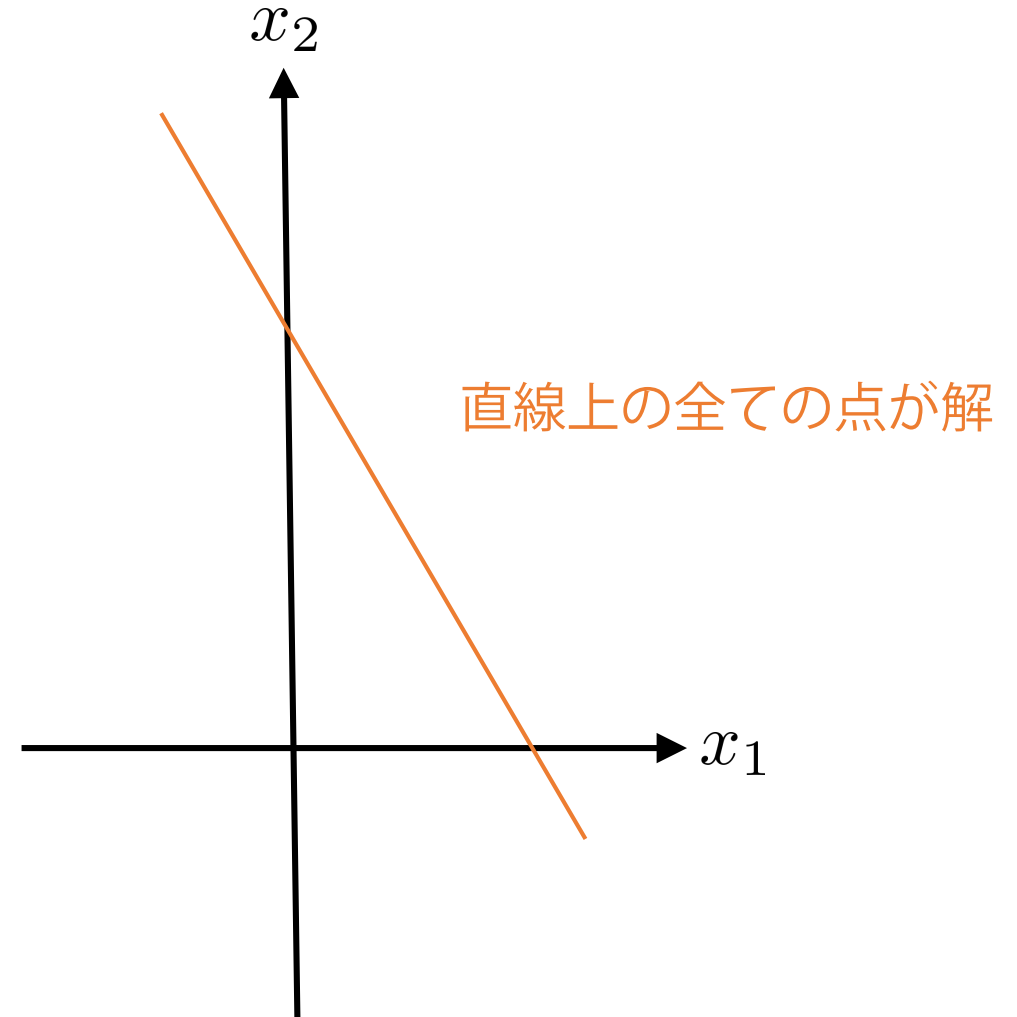
ただし、どこがゼロかはわからない

→ 非自明な問題
どうやってスパース解が得られるか？

劣決定系逆問題 (Under-determined inverse problem)

一番簡単な例： $M=1, N=2$

$$\begin{array}{ccccc} \boxed{y} & = & \boxed{A} & \times & \boxed{x} \\ \text{measure} & & \text{known} & & \text{target} \end{array}$$
$$y_1 = ax_1 + bx_2$$

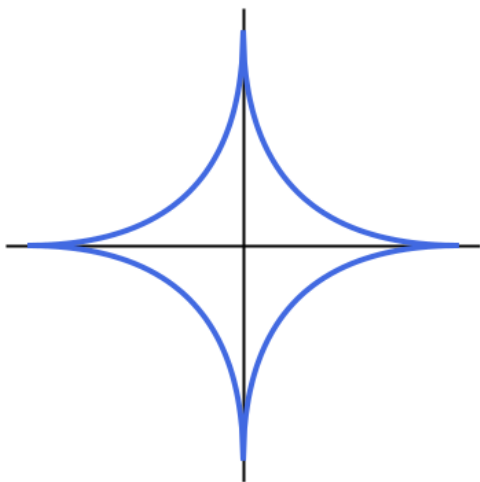


L_p ノルム

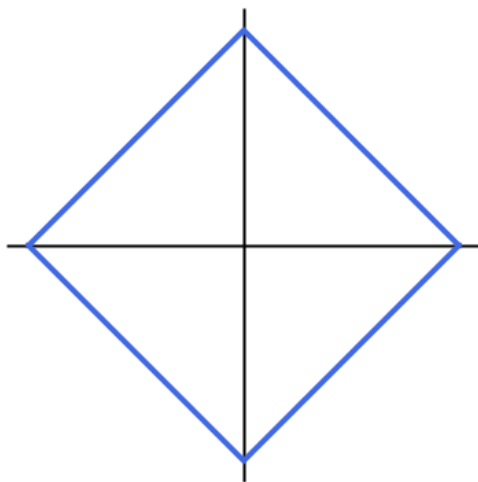
$$\|x\|_p = (|x_1|^p + |x_2|^p + \cdots + |x_N|^p)^{1/p}$$

$N=2$ の場合

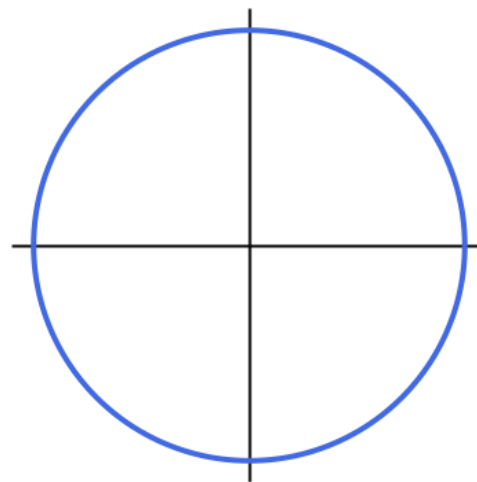
$p=0.5$



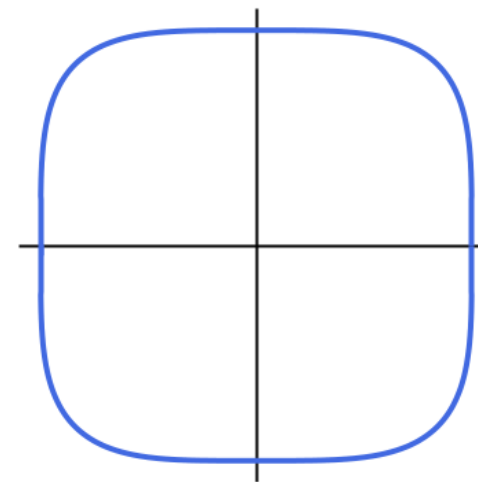
$p=1$



$p=2$



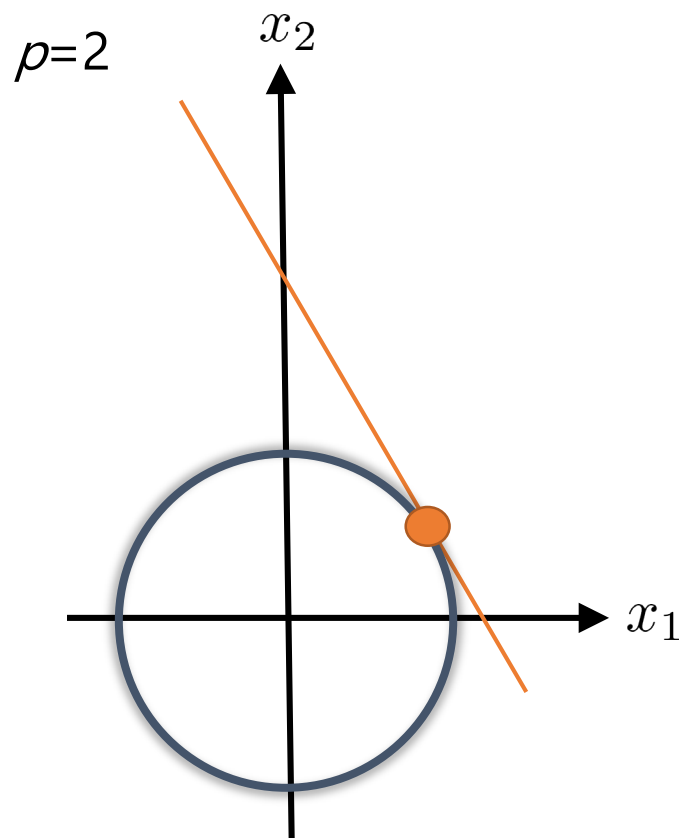
$p=4$



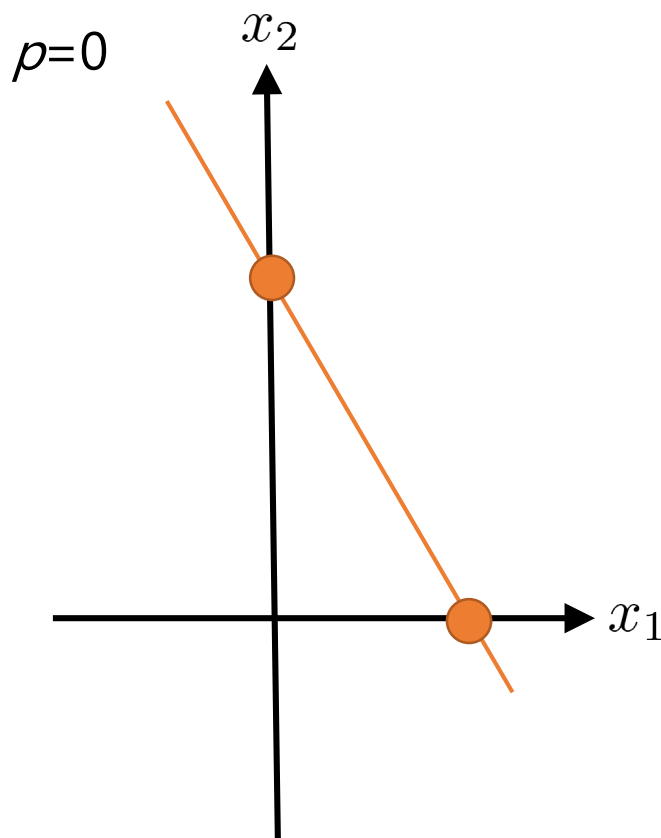
$p=0$:
非ゼロ要素の個数

L_p ノルム最小化

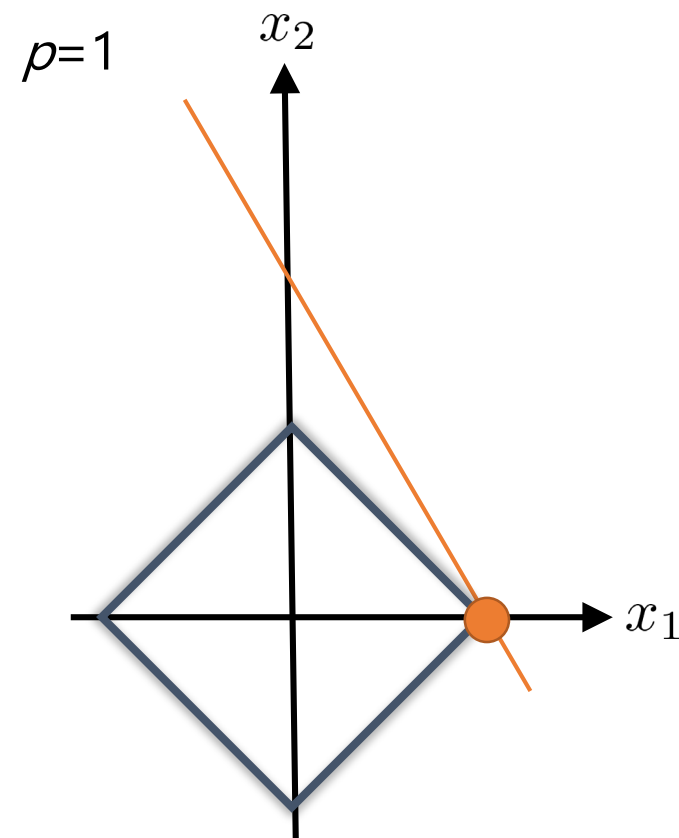
$$\min_x \|x\|_p \quad \text{subj. to} \quad y = Ax$$



極端に大きな解は避けることができる
ただし、無難な解 = 特徴がない



全てのスパース解



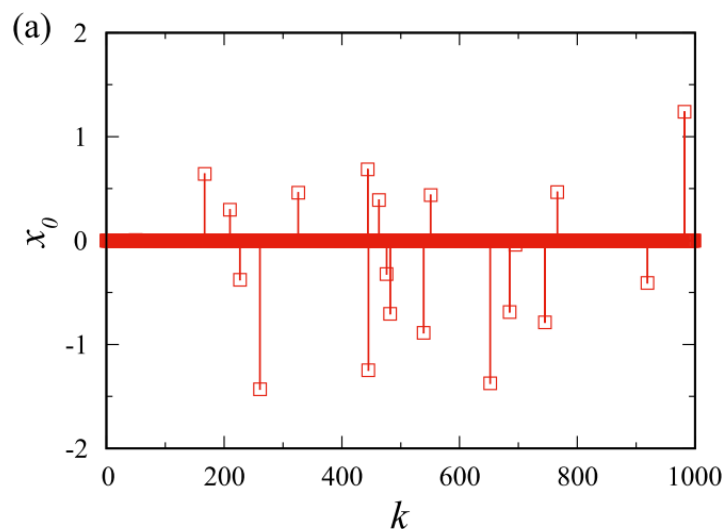
スパース解のうちのひとつ

例：ランダム行列

x_0 答え

次元 $N=1000$

非ゼロ要素数 $n=20$

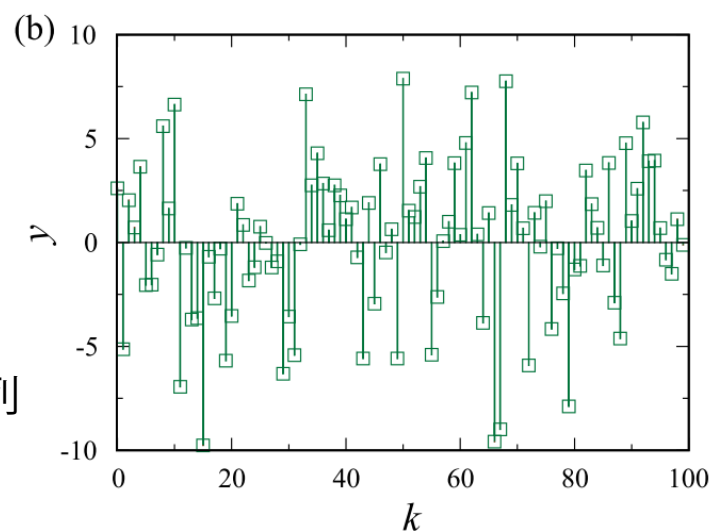


$y = Ax_0$

インプット

次元 $M=100$

A : ($M \times N$)ランダム行列

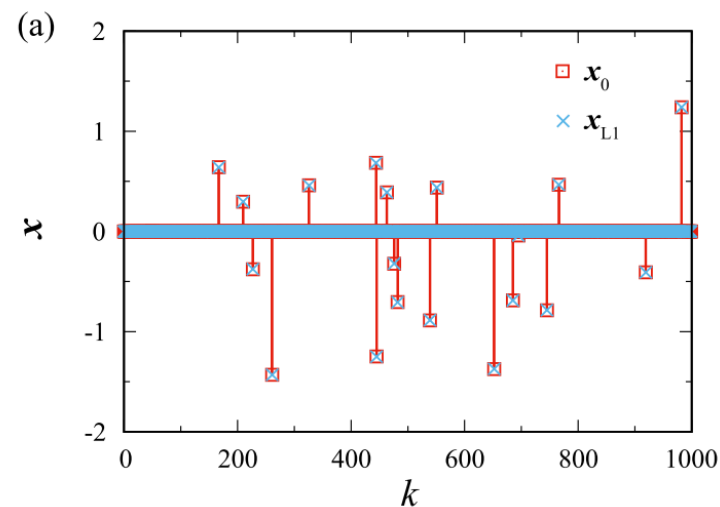


$$\min_x \|x\|_p \quad \text{subj. to} \quad y = Ax$$

$p=1$

L_1 ノルム最小解

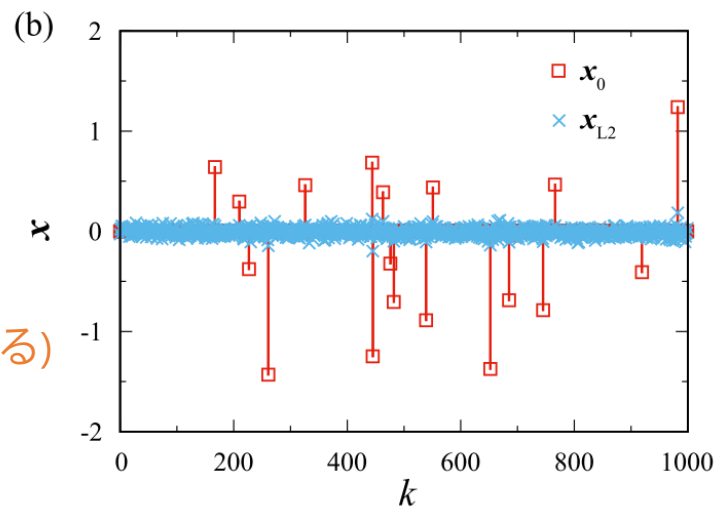
正解と一致



$p=2$

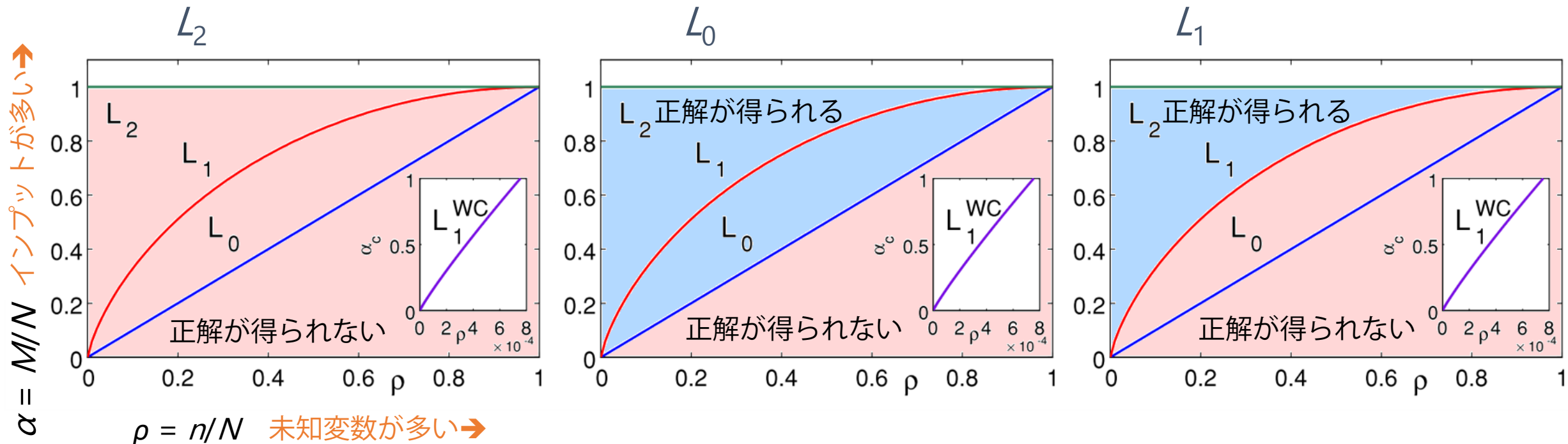
L_2 ノルム最小解

正解と一致しない
($y = Ax$ は満たしている)



L_p ノルム最小化に関する厳密な結果

A: ($M \times N$) ランダム行列, M : y の次元, N : x の次元, n : x の非ゼロの個数



正解が得られるのは $\alpha=1$ の場合だけ

正解を得られる可能性が大きい
しかし、**計算量が膨大**(組み合わせ)

L_0 ほどではないが正解が得られる
かつ、**実用的な計算が可能**

ノイズのある場合： L_p ノルム正則化

測定ノイズを考慮

$$\mathbf{y} = A\mathbf{x} + \boldsymbol{\eta}$$

$\mathbf{y} = A\mathbf{x}$ を厳密に満たす必要はない

$$\|\mathbf{y} - A\mathbf{x}\|_2 \lesssim \eta$$

解の候補は無数に存在

正則化項付き最小二乗法

$$\min_{\mathbf{x}} \left\{ \frac{1}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 + \lambda \|\mathbf{x}\|_p \right\}$$

λ ：正則化の強さを決めるハイパーパラメーター
(値の決定方法は後程)

ベイズの定理による解釈

$$P(\mathbf{x}|\mathbf{y}) \propto P(\mathbf{y}|\mathbf{x})P(\mathbf{x})$$

$$P(\mathbf{y}|\mathbf{x}) \propto e^{-\frac{1}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2} \quad \text{尤度関数} = \text{二乗誤差}$$

$$P(\mathbf{x}) \propto e^{-\lambda \|\mathbf{x}\|_p} \quad \text{事前確率} = \text{正則化}$$

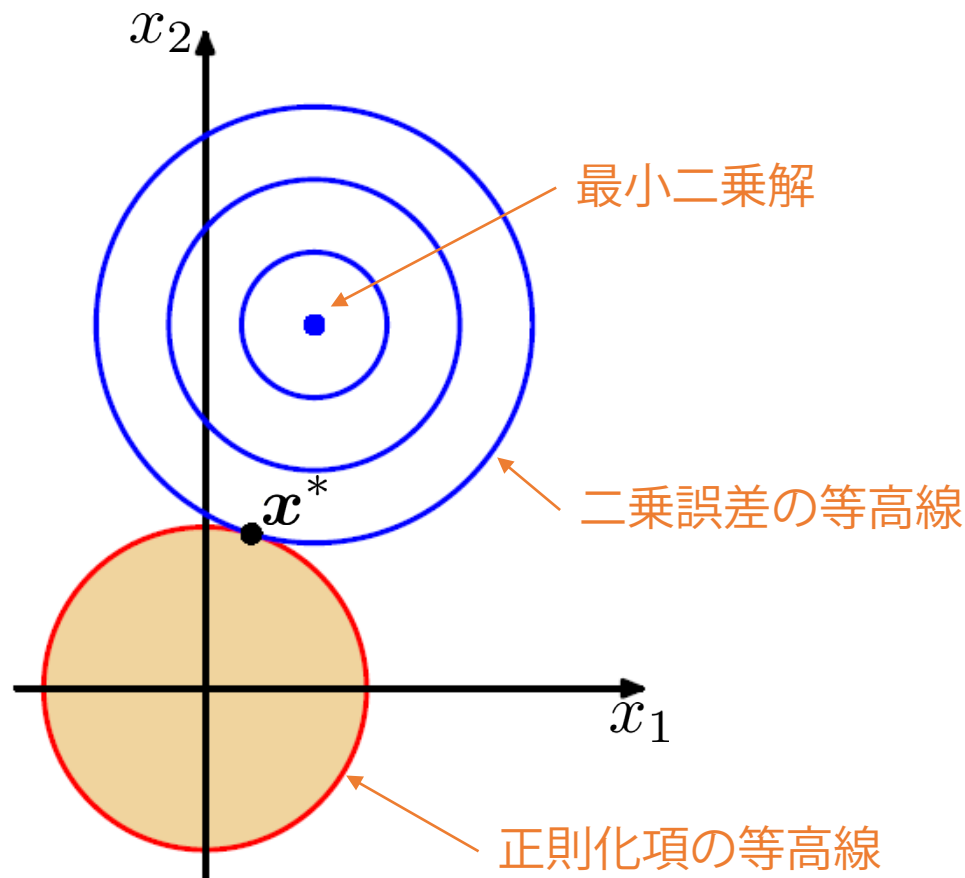
$p=1 \rightarrow$ ラプラス分布

事後確率を最大化
(MAP推定; maximum a posteriori)

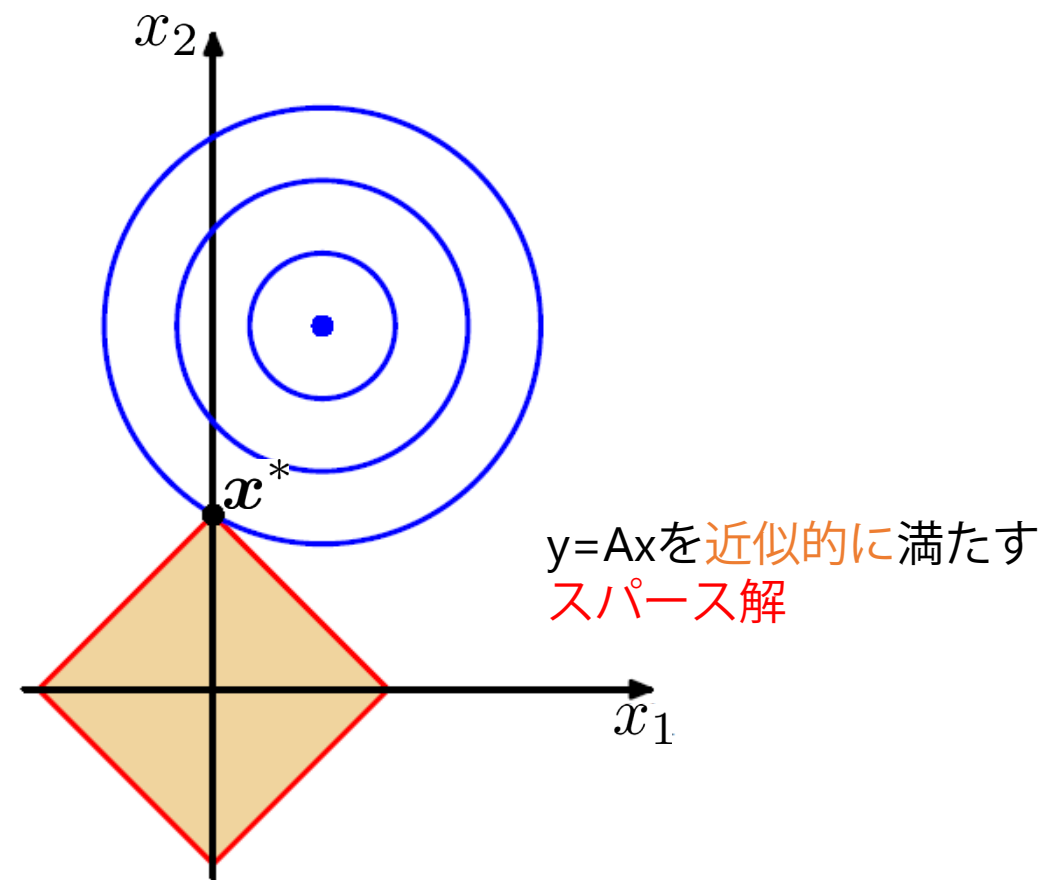
$$\max_{\mathbf{x}} P(\mathbf{x}|\mathbf{y})$$

L_2 ノルム正則化 と L_1 ノルム正則化 の比較

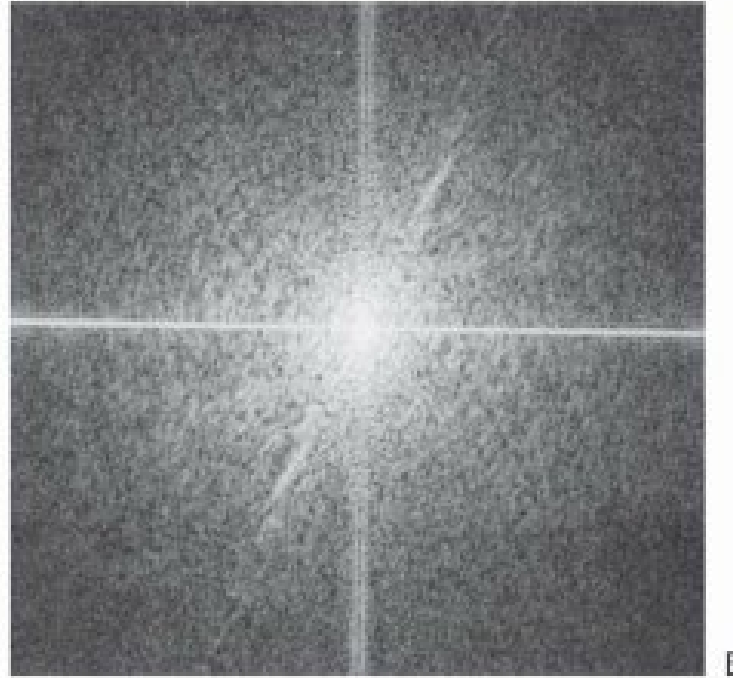
$$L_2 \quad x^* = \min_x \left\{ \frac{1}{2} \|y - Ax\|_2^2 + \lambda \|x\|_2 \right\}$$



$$L_1 \quad x^* = \min_x \left\{ \frac{1}{2} \|y - Ax\|_2^2 + \lambda \|x\|_1 \right\}$$



x は本当にスパースか？



<https://radiologykey.com/fourier-transform/>

Images → Fourier, wavelet, ...
Sound → Fourier

基底変換を通してスパースになる場合がある

= データが圧縮可能

→ 圧縮基底を利用して
測定時間量を減らす

一般化LASSO

LASSO (Least Absolute Shrinkage of Selection Operators)

R. Tibshirani, J. R. Stat. Soc. B 58, 267 (1996)

$$\min_{\boldsymbol{x}} \left\{ \frac{1}{2} \|\boldsymbol{y} - A\boldsymbol{x}\|_2^2 + \lambda \|\boldsymbol{x}\|_1 \right\}$$

$\boldsymbol{x}$ がスパースであることを仮定

一般化LASSO

R. Tibshirani, J. Taylor, Ann. Stat. 39, 1335 (2011)

$$\min_{\boldsymbol{x}} \left\{ \frac{1}{2} \|\boldsymbol{y} - A\boldsymbol{x}\|_2^2 + \lambda \|B\boldsymbol{x}\|_1 \right\}$$

$B\boldsymbol{x}$ がスパースであることを仮定
(B は非正方行列)

例：MRI

$$\min_x \left\{ \frac{1}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 + \lambda \|B\mathbf{x}\|_1 \right\}$$

$\mathbf{y}$: MRI信号 (測定データ)

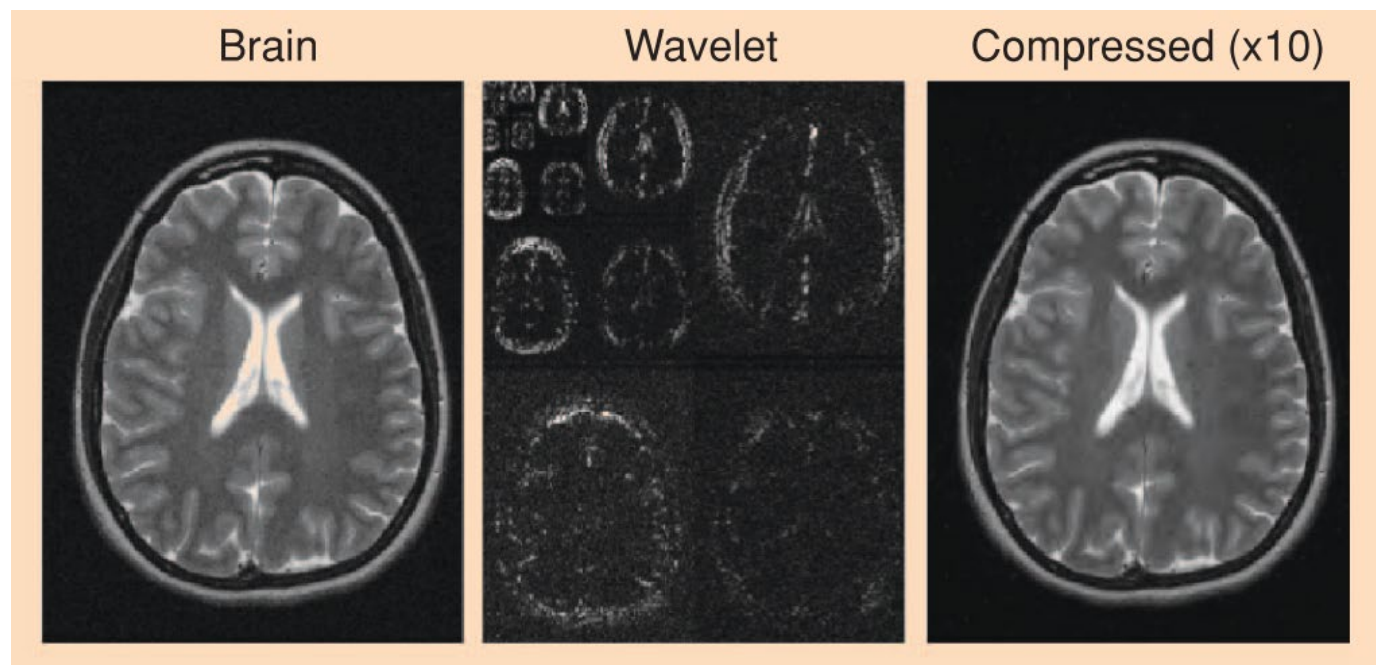
A : フーリエ基底

$\mathbf{x}$: 2D画像 (ほしいもの)

B : Wavelet変換 (JPEG2000)

and/or

Total variation (差分)

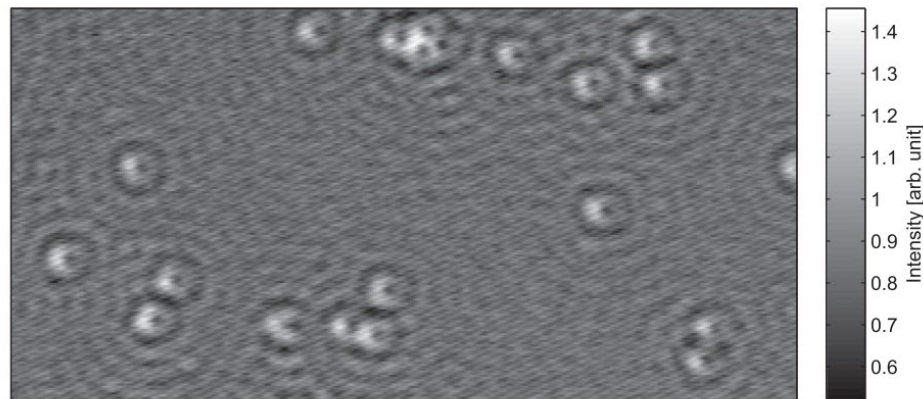


例：STS

実空間dI/dV画像 Ag(111)表面

(a)

実空間



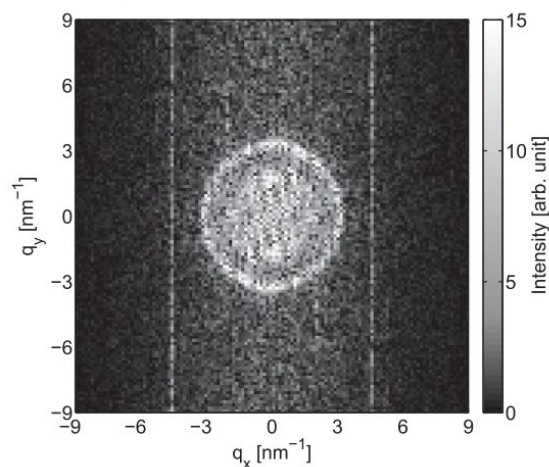
フーリエ変換



(b)

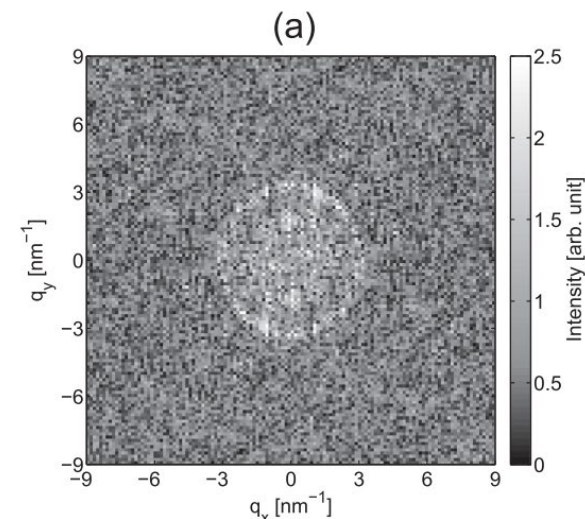
準粒子干渉
パターン

波数空間

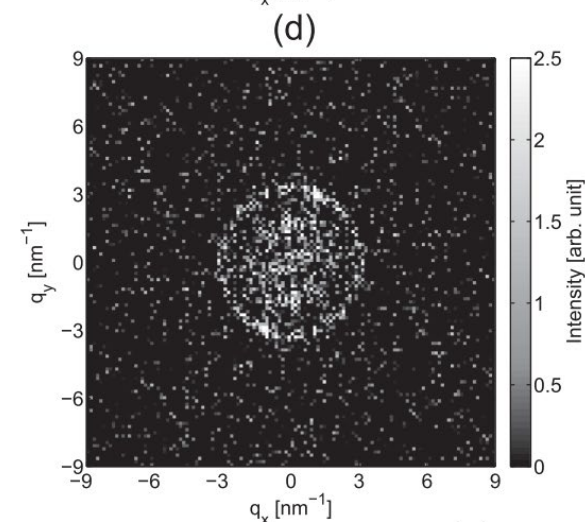


1/9の画像データのみ使用

通常の方法で
フーリエ変換



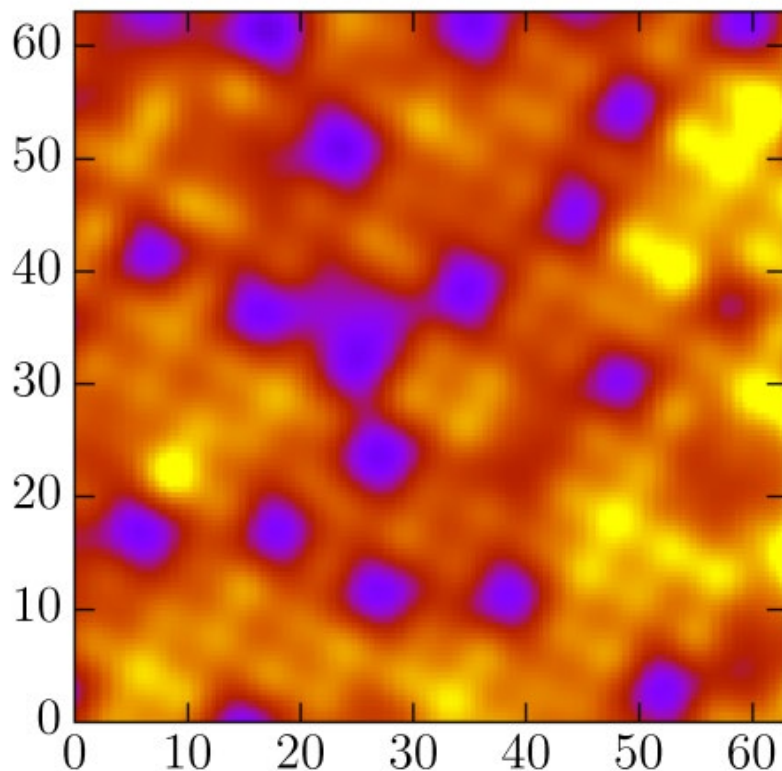
LASSO



STM: 原子位置の判定

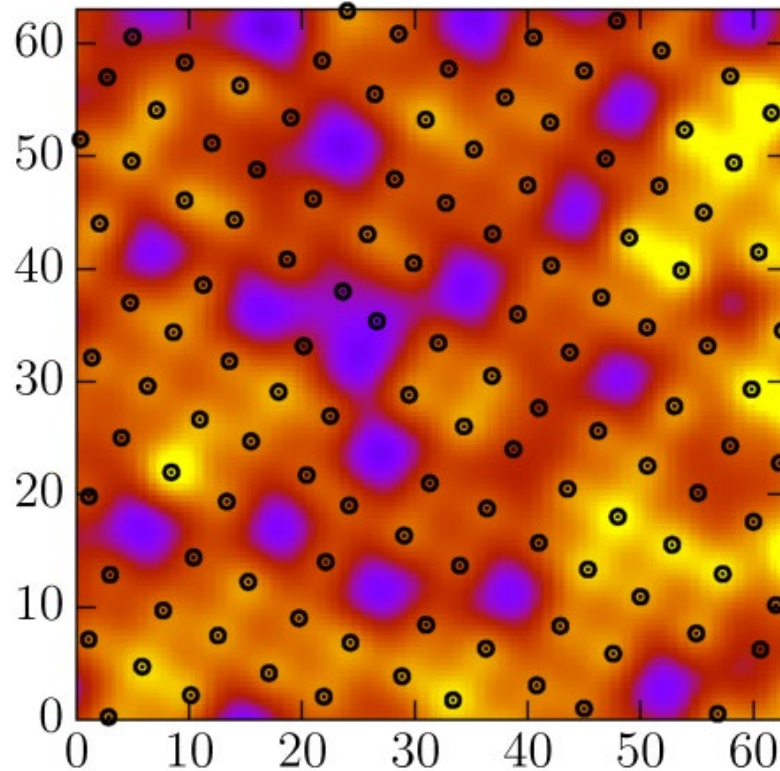
SrVO₃(001) thin-film

STM実空間像



(a) Original data y_i

黒丸：原子位置



(c) Estimated peak positions r_k^*

原子をたくさん配置しておき、
LASSOで不要な原子を削除

次回：LASSO最適化問題の解法

- 凸最適化問題だが、微分が不連続なため、一般的な方法は使えない
- LASSO
 - Coordinate descent (scikit-learnで使われている)
 - Majorization minimization (MM) algorithm; ISTA, FISTA
- generalized LASSO
 - Scikit-learnにはない
 - 行列Bの形式によりアルゴリズム・ライブラリがわかれている
 - 汎用的なものとしてADMM (Alternating Direction Method of Multiplier)がある
Boyd et al. *Foundations and Trends in Machine Learning* 3, 1 (2011)
拡張ラグランジュ法を利用した方法で非常に汎用的。拘束条件（総和則、非負性など）も課すことができる。ただし、収束は早くはない。