国立情報学研究所

今日からできる スパースモデリング

京都大学大学院情報学研究科システム科学専攻
大関　真之

QUANTUM ANNEALING

MACHINE LEARNING

Sparse Modeling

CREST

自己紹介

- 大関真之
  - 京都大学大学院情報学研究科システム科学専攻助教
    - 理論物理学：統計力学、量子アニーリング
    - 機械学習：深層学習、ボルツマン機械学習

自己紹介

- 大関真之
  - 京都大学大学院情報学研究科システム科学専攻助教
    - 理論物理学：統計力学、量子アニーリング
    - 機械学習：深層学習、ボルツマン機械学習
  - 東京工業大学大学院理工学研究科物性物理学専攻博士課程修了
    - 指導教官：西森秀稔「統計力学、量子アニーリング、etc」
  - ローマ大学物理学科プロジェクト雇用研究員
    - Leader：Giorgio Parisi「統計力学、最適化問題、etc」

引用：Google Scholar Citation，顔 東工大の研究者たち Vol.13 西森秀稔

自己紹介

- 大関真之
  - 京都大学大学院情報学研究科システム科学専攻助教
    - 理論物理学：統計力学、量子アニーリング
    - 機械学習：深層学習、ボルツマン機械学習

- カンニング大関？！
  - 朝日新聞（H27.1.16）、Newton（H27.2.26）、ケトル（H27.6.15）
  - TBSテレビ「あさチャン！」(H27.1.19)、NHK総合「おはよう日本」（H27.3.27）

昨今では、小型ワイヤレスイヤホンや文具型メッセージ受信機によるカンニングも横行。ちなみに中国では、カンニング防止のため、1200人の生徒たちが学校の校庭でテストを受けさせられたというニュースが話題を呼んだ

京都大学
情報学研究科
大関真之助教

日本最先端のアタマの中身

全世界の学生たちが戦々恐々！

史上初のカンニング検出システム

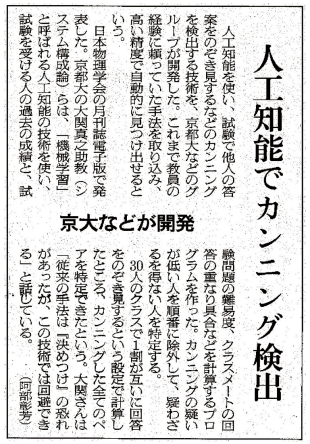
FOCUS Informatics

カンニングを見破るプログラムを開発

テストの正答データからカンニングの有無を検出

試験を受けて，答えがわからないときに，周りの人たちの答案を覗いてみたいと思うことはないだろうか。本人はバレないと思っていても，経験を積んだ教員が答案をチェックすると，不正が発覚する場合がある。カンニングの検出は，長年，人の勘に頼るところが大きかったが，京都大学大学院情報学研究科の大関真之助教らが，「機械学習」を利用した新しいプログラムを開発した。

採点データからカンニングを検出



人工知能でカンニング検出

人工知能を使い、試験で他人の答案をのぞき見するなどのカンニングを検出する技術を、京都大などのグループが開発した。これまで教員の経験に頼っていた手法を取り込み、高い精度で自動的に見つけ出せるという。

日本物理学会の月刊誌電子版で発表した。京都大の大関真之助教（システム構成論）らは、「機械学習」と呼ばれる人工知能の技術を使い、試験を受ける人の過去の成績と、試

京都大などが開発

験問題の難易度、クラスメートの回答の重なり具合などを計算するプログラムを作った。カンニングの疑いが低い人を順番に除外して、疑わしさを得ない人を特定する。

30人のクラスで1割が互いに回答をのぞき見するという設定で計算したところ、カンニングした全てのペアを特定できたという。大関さんは「従来の手法は『決めつけ』の恐れがあったが、この技術では回避できる」と話している。（阿部彰芳）

QUANTUM ANNEALING　MACHINE LEARNING　Sparse Modeling　CREST

自己紹介

- サイエンスライター？？

INTELLIGENCE DESIGNER

Intelligent Solutions for the Knowledge Society & Industry

original image: © drimi - Fotolia.com

知的情報処理の最前線：新規計算技術「量子アニーリング」

2015.11.17

大関 真之

1982年東京生まれ。2008年東京工業大学大学院理工学研究科物性物理学専攻博士課程早期修了。2010年から現職である京都大学大学院情報学研究科システム科学専攻助教。非常に複雑な多数の要素間の関係や集団としての性質を明らかにする統計力学と呼ばれる学問体系を切り口として、機械学習を始めとする現代のキーテクノロジーを独自の表現で理解して、広く社会に普及させることを目指している。大量の情報から本質的な部分を抽出する、または少数の情報から満足のいく精度で背後にある構造を明らかにすることができる「スパースモデリング」や、次世代コンピュータとして期待される量子コンピュータ、とりわけ「量子アニーリング」形式に関する研究活動を展開している。

今日からできるスパースモデリング

講義ノートあります.

検索：大関　真之 > TopないしはPublish

第一部：圧縮センシング

中学生の頃の思い出

中学生の頃の思い出

連立方程式

$$2x_1 + x_2 = 1$$

を満たす解を求めよ

$$2x_1 + x_2 = 1$$

を満たす解を求めよ

（一意には）できません

$$2x_1 + x_2 = 1$$

を満たす解を求めよ

（一意には）できません

式が足りないから

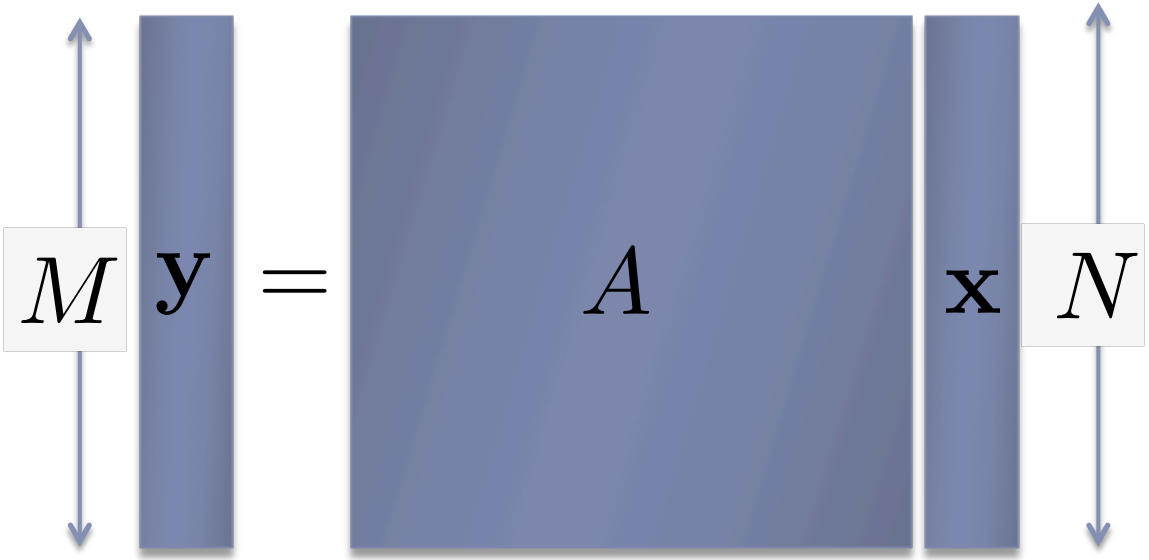
$$2x_1 + x_2 = 1$$

を満たす解を求めよ

(一意には) できません

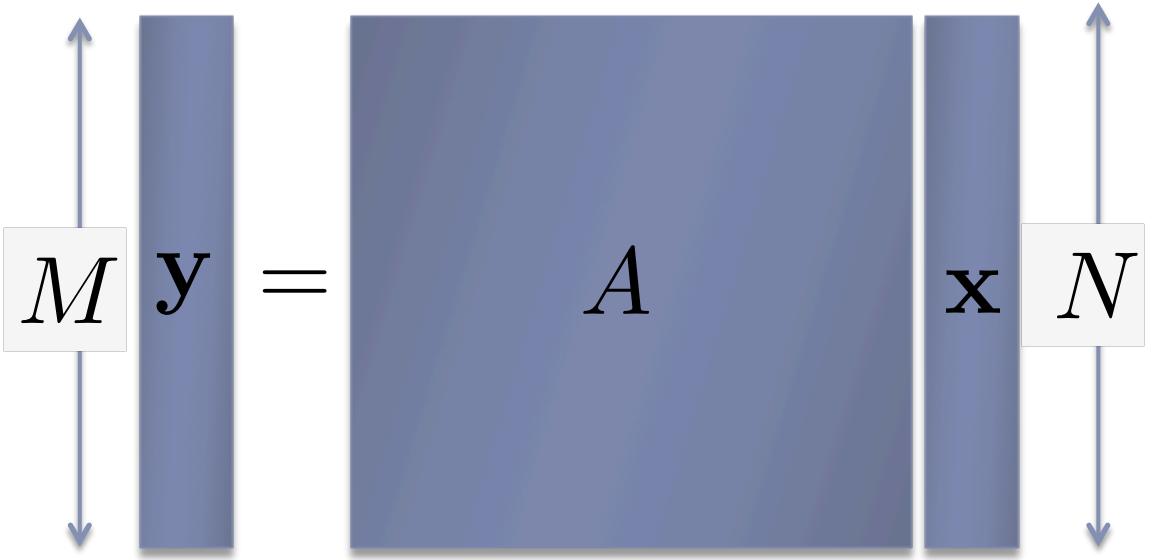
式が足りないから

情報が足りないから



$$\mathbf{y} = A\mathbf{x}$$

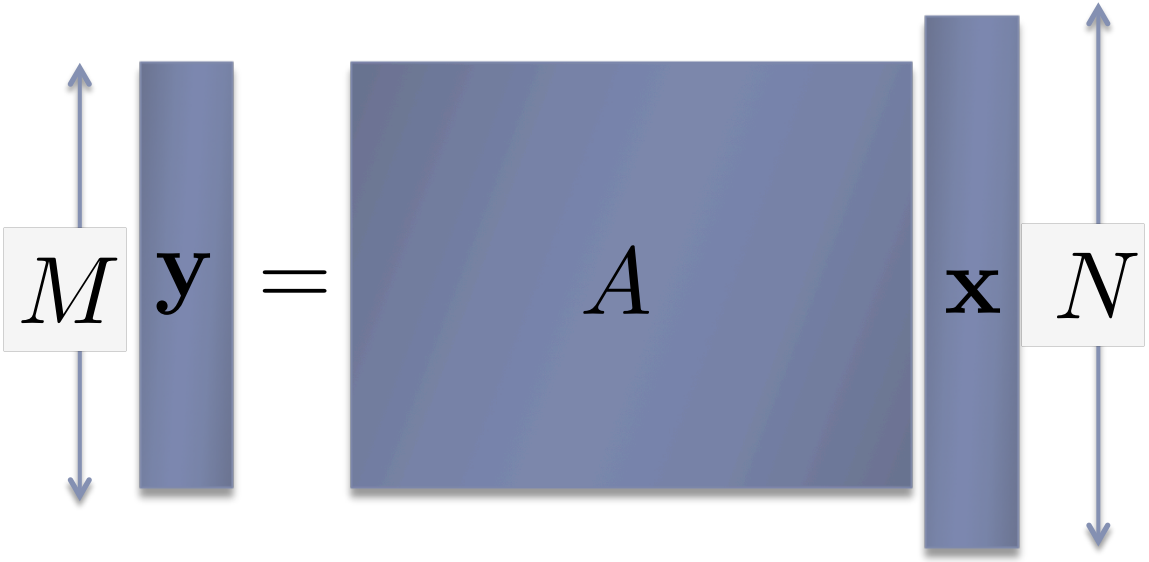
を満たす解を求めよ



$$\mathbf{y} = A\mathbf{x}$$

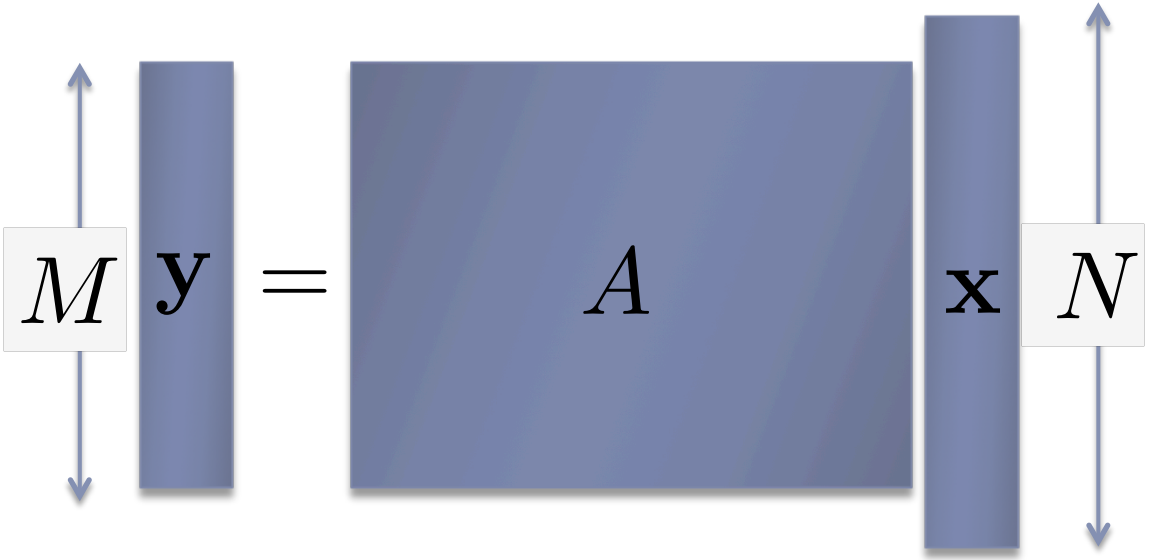
を満たす解を求めよ

逆行列があるからOK



$$\mathbf{y} = A\mathbf{x}$$

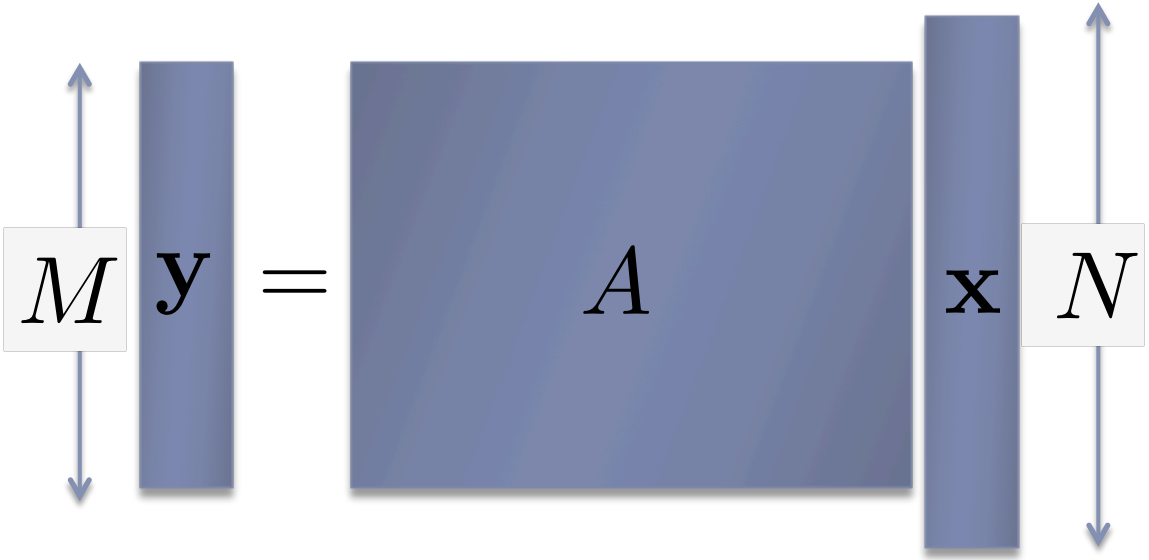
を満たす解を求めよ



$$\mathbf{y} = A\mathbf{x}$$

を満たす解を求めよ

（一意には）できません



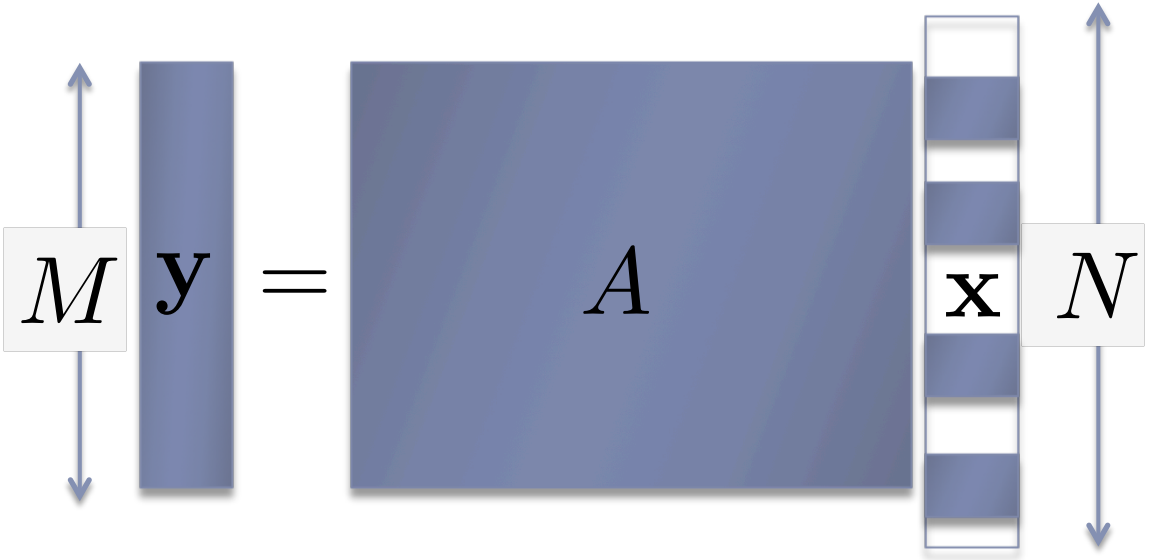
$$\mathbf{y} = A\mathbf{x}$$

を満たす解を求めよ

（一意には）できません

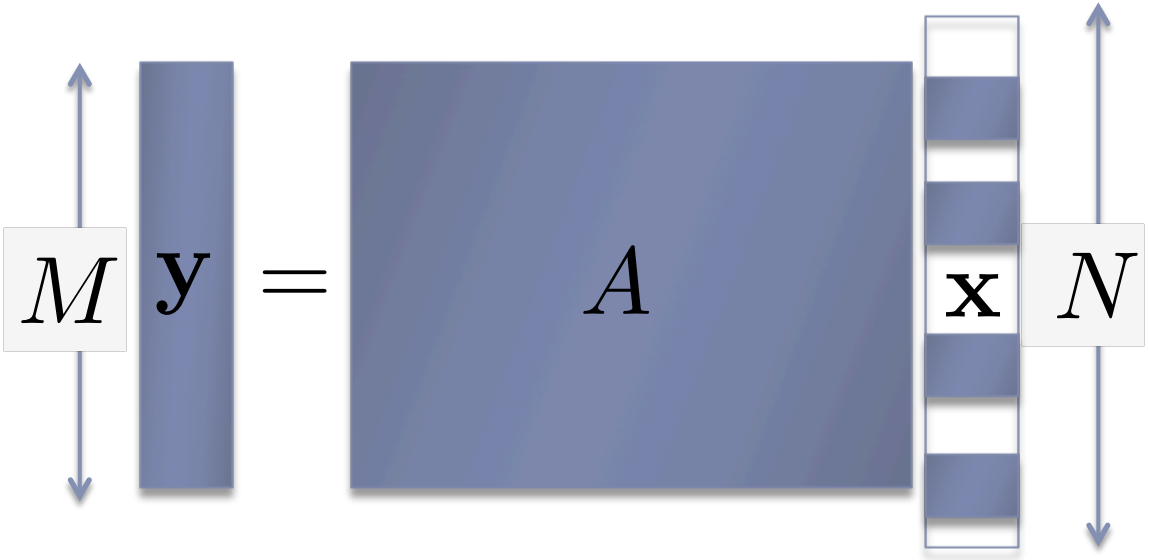
劣決定系

スパース性があるときは？



$$\mathbf{y} = A\mathbf{x}$$

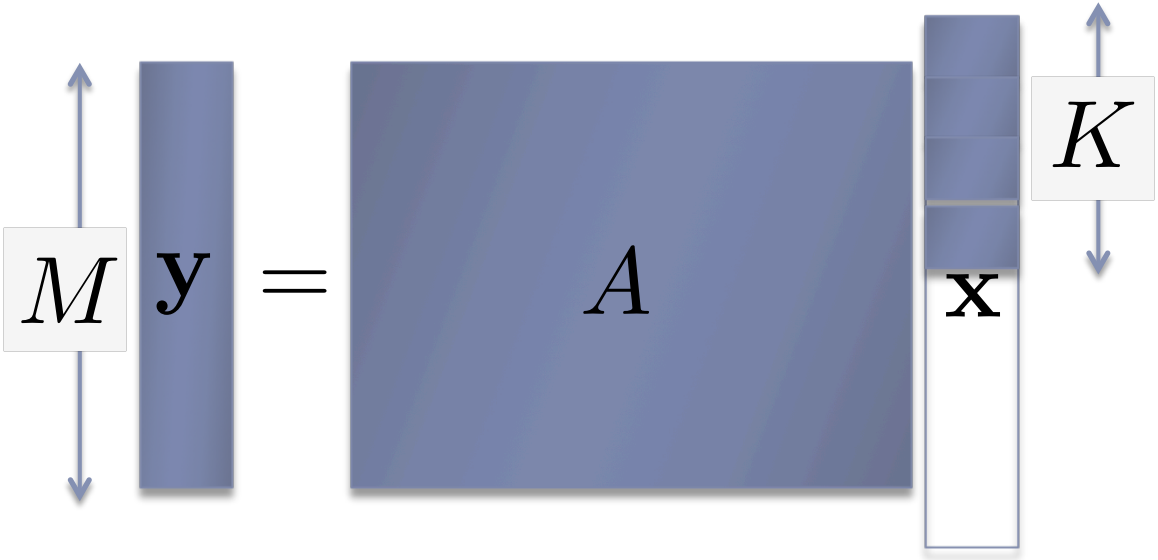
を満たす解を求めよ



$$\mathbf{y} = A\mathbf{x}$$

を満たす解を求めよ

できる！



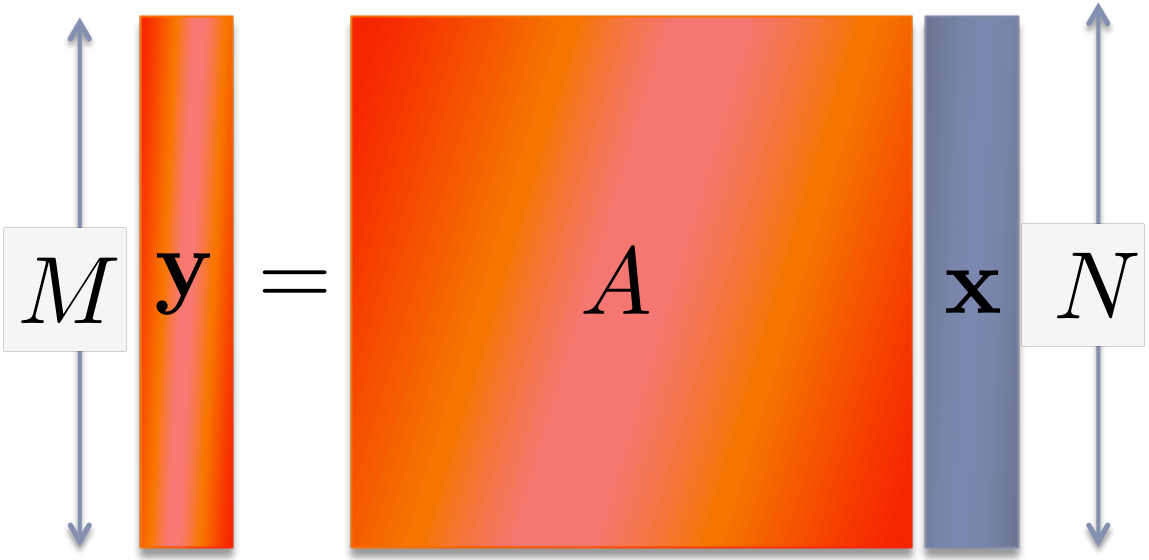
$$\mathbf{y} = A\mathbf{x}$$

を満たす解を求めよ

できる！

圧縮センシングの基礎

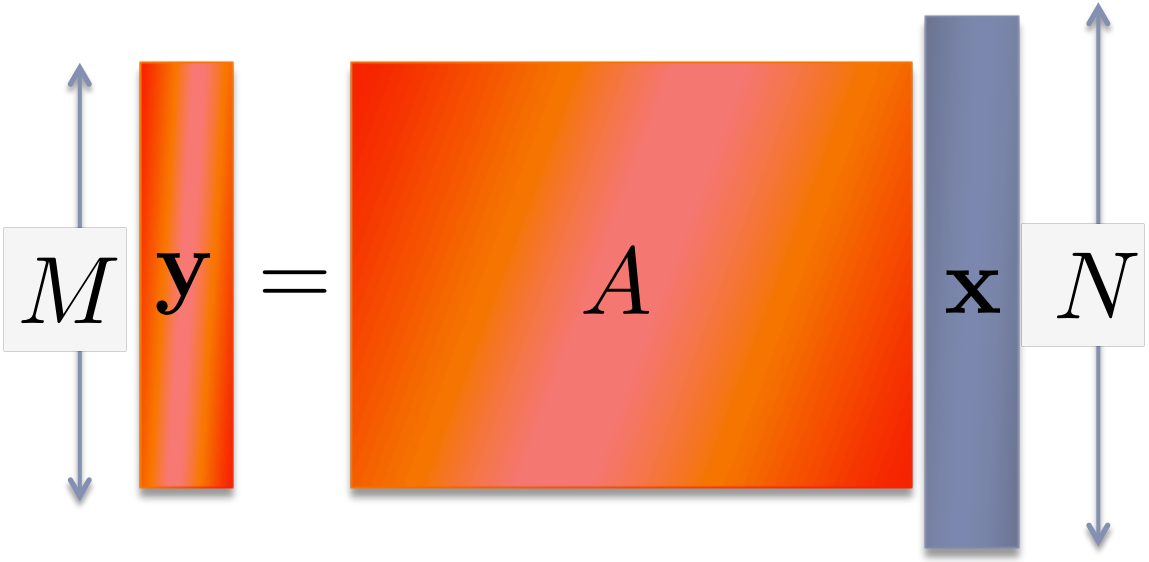
劣決定系の方程式も解がスパースなら解ける



観測過程

$$\mathbf{y} = A\mathbf{x}$$

の結果から入力を求めよ



観測縮減革命

$$\mathbf{y} = A\mathbf{x}$$

の結果から再現せよ

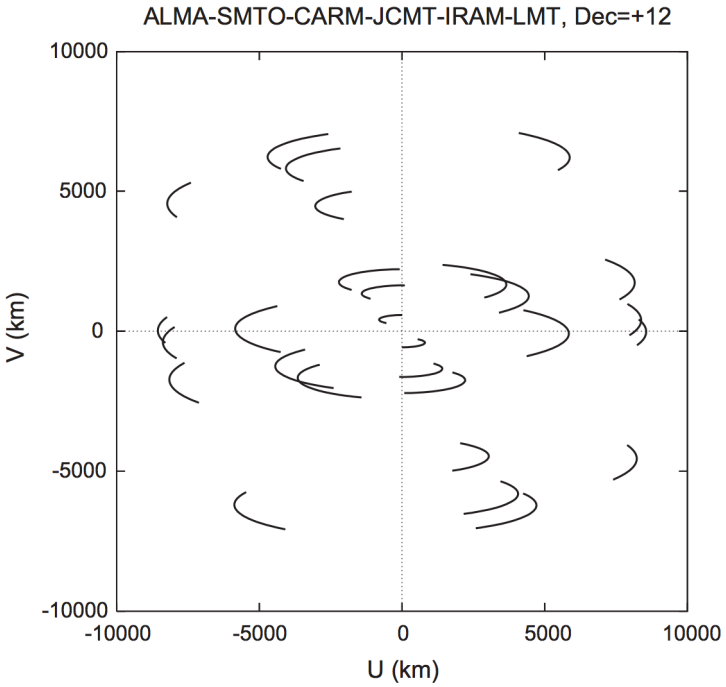
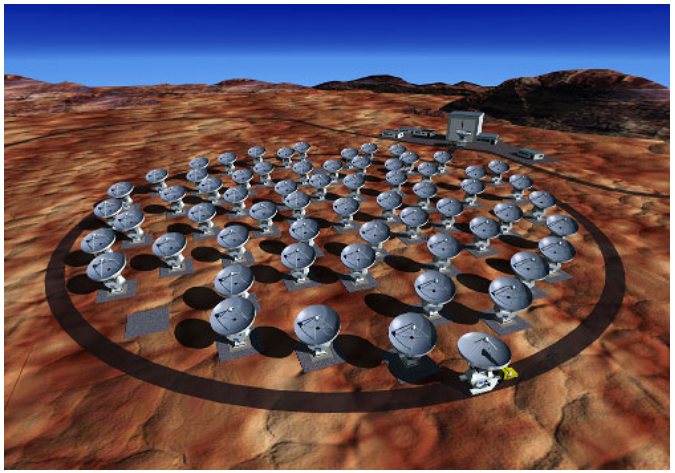
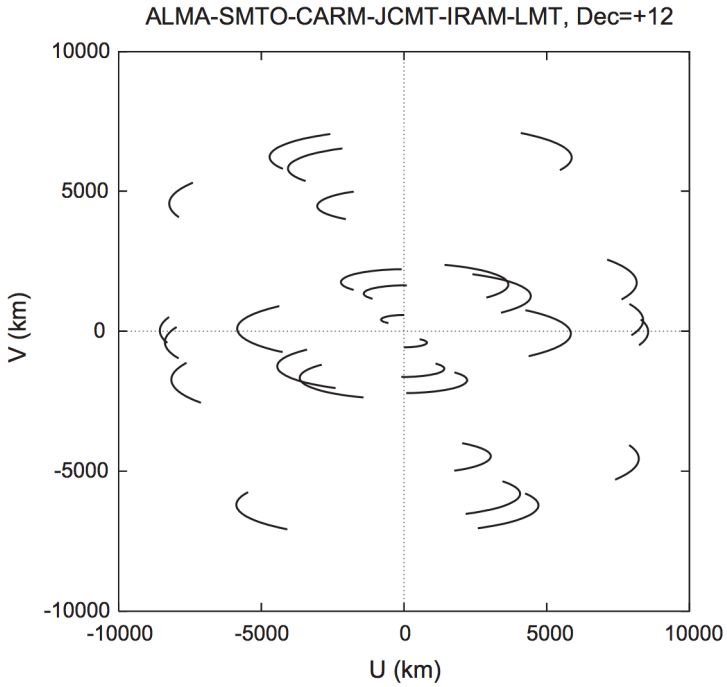
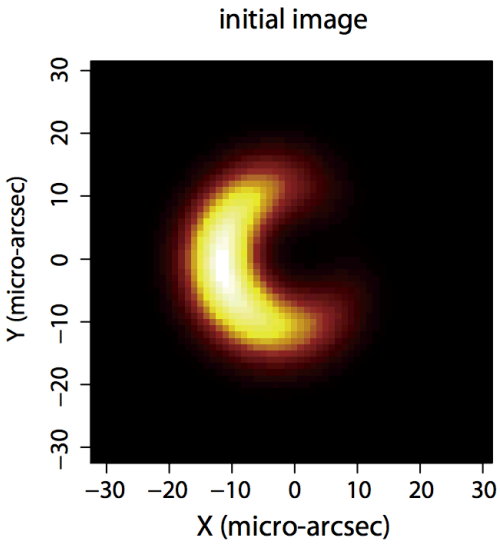


Fig. 5. Simulated UV coverage of M87 with six-station sub-mm VLBI array of EHT. Here it is assumed that observations are conducted at an elevation larger than 20° at each station.



ブラックホール撮像の例
M. Honma et al. Publ. Astron. Soc. Jon. (2014)





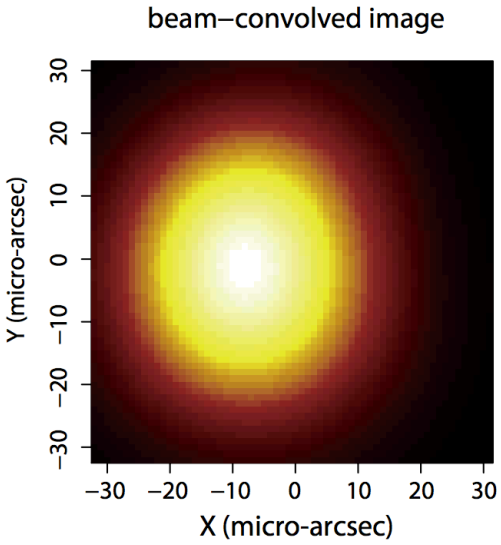
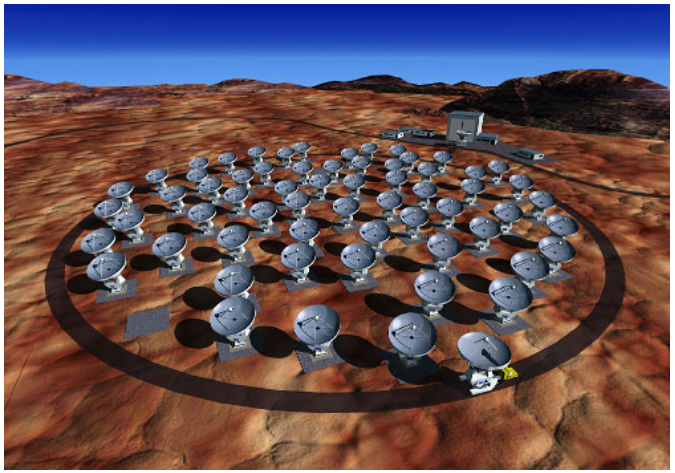


Fig. 5. Simulated UV coverage of M87 with six-station sub-mm VLBI array of EHT. Here it is assumed that observations are conducted at an elevation larger than 20° at each station.



ブラックホール撮像の例
M. Honma et al. Publ. Astron. Soc. Jon. (2014)

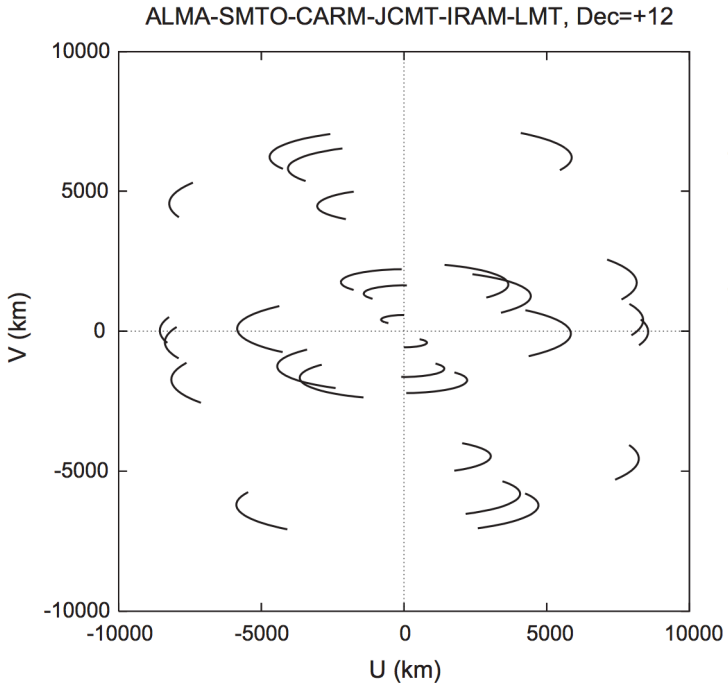
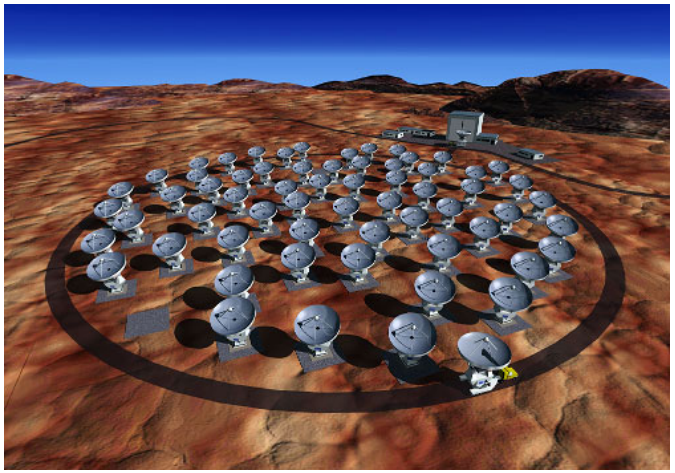
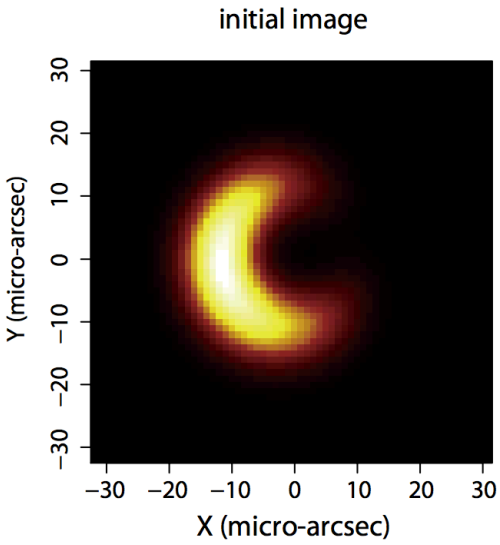
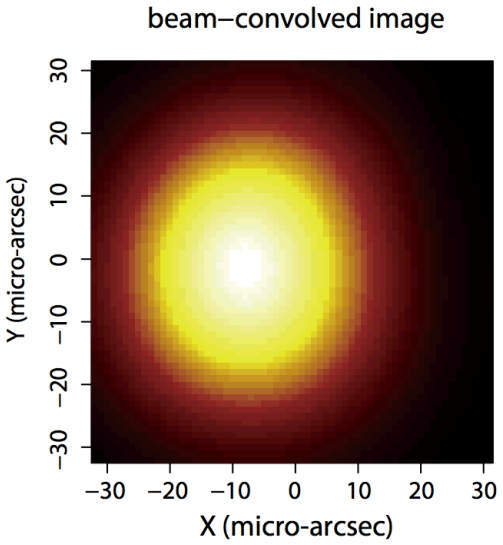
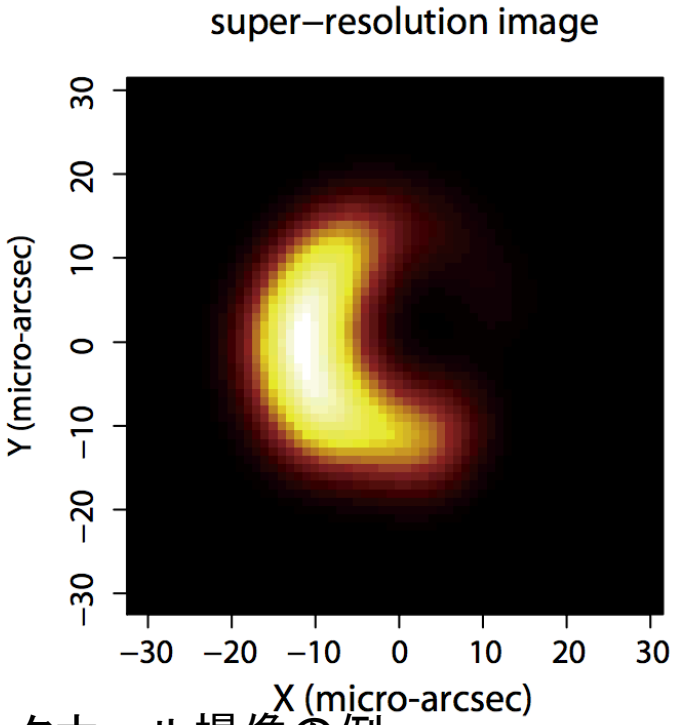


Fig. 5. Simulated UV coverage of M87 with six-station sub-mm VLBI array of EHT. Here it is assumed that observations are conducted at an elevation larger than 20° at each station.





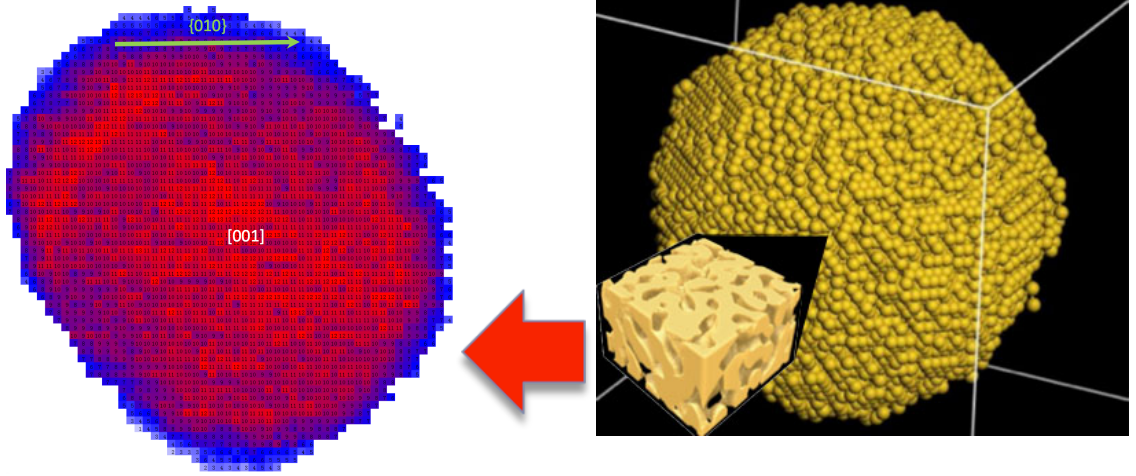




ブラックホール撮像の例

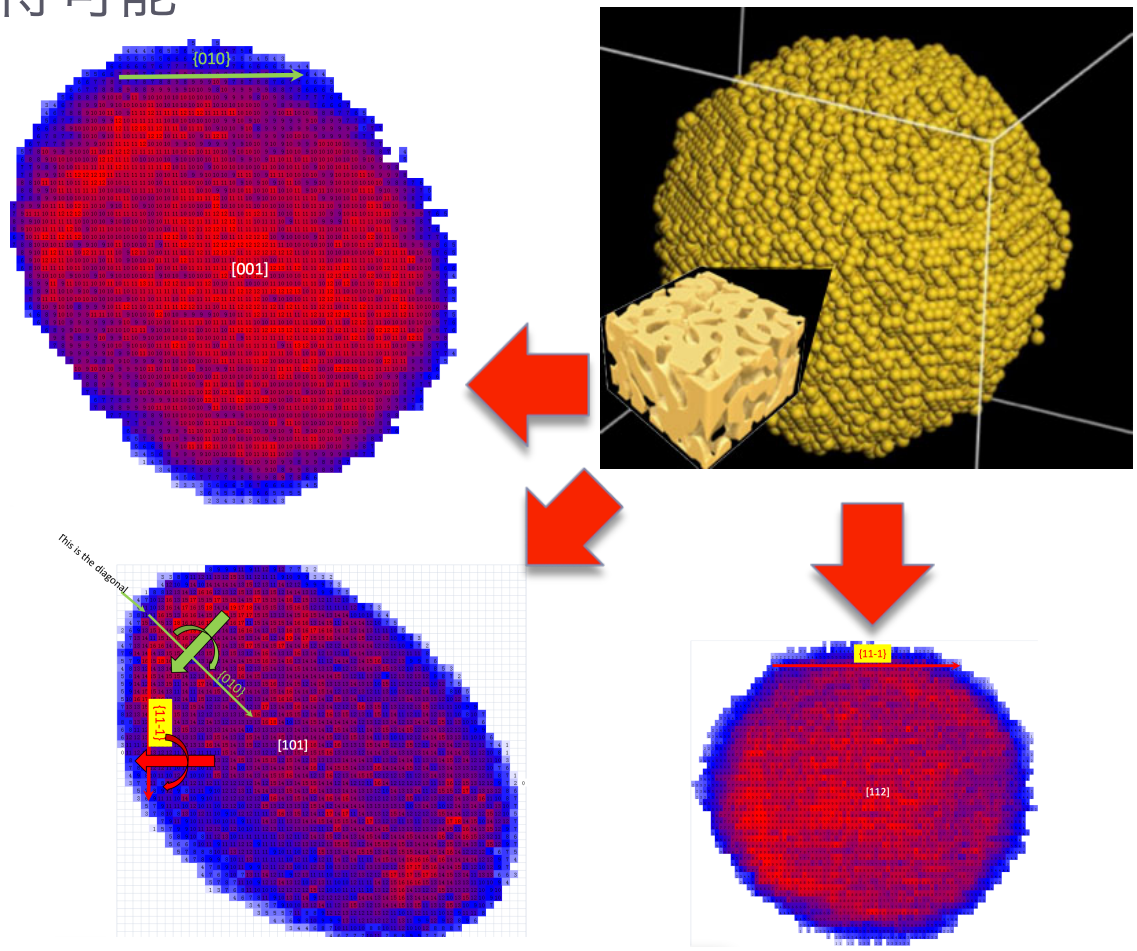
M. Honma et al. Publ. Astron. Soc. Jon. (2014)

- HAADF-STEMのよる撮像
  - 2次元射影図のみ取得可能



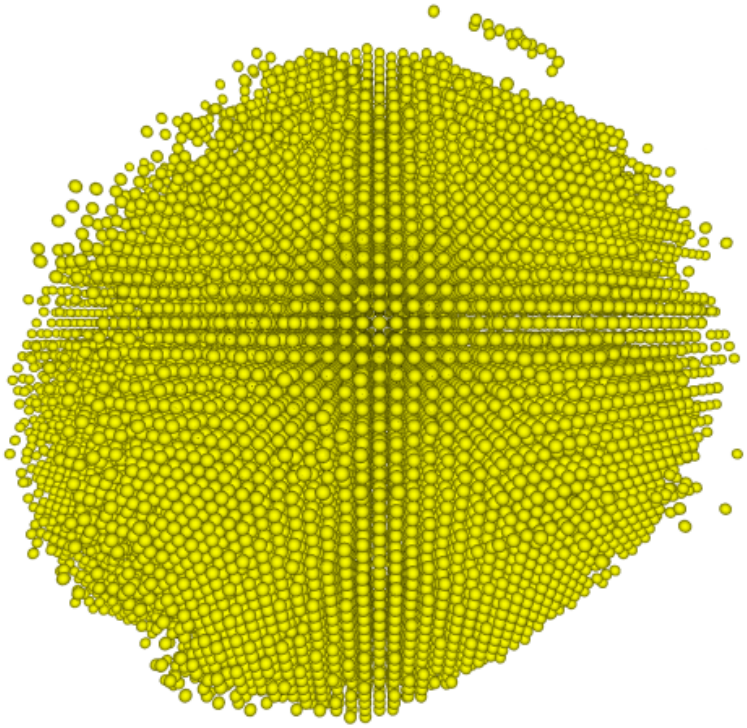
金のナノポーラスの例
Collaboration with C. Nakajima

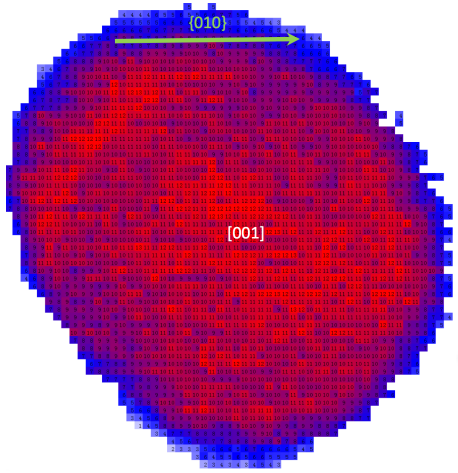
- HAADF-STEMのよる撮像
  - 2次元射影図のみ取得可能

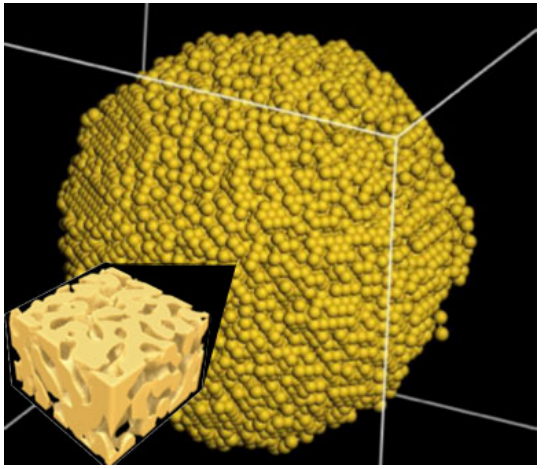


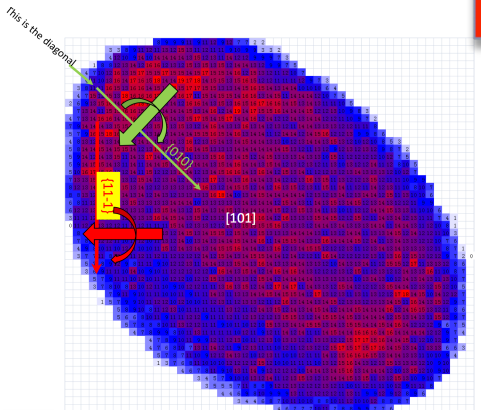
金のナノポーラスの例
Collaboration with C. Nakajima

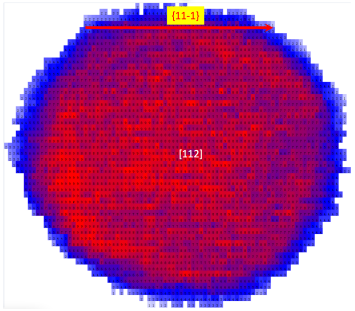
- HAADF-STEMのよる撮像
  - 2次元射影図のみ取得可能











金のナノポーラスの例
Collaboration with C. Nakajima

不足した観測情報＋推定方法

圧縮センシング

スパース解推定

- L0ノルム最小化によるスパース解推定
  - 以下の最小化問題でスパース解を探索しよう！

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad \mathbf{y} = A\mathbf{x}$$

  - 非零成分の個数が小さい（スパース）で方程式を満たすものを探す

スパース解推定

- L0ノルム最小化によるスパース解推定
  - 以下の最小化問題でスパース解を探索しよう！

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad \mathbf{y} = A\mathbf{x}$$

  - 非零成分の個数が小さい（スパース）で方程式を満たすものを探す
  - 残念ながら計算量的に困難（組み合わせ最適化問題）

スパース解推定

- L0ノルム最小化によるスパース解推定
  - 以下の最小化問題でスパース解を探索しよう！

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad \mathbf{y} = A\mathbf{x}$$

  - 非零成分の個数が小さい（スパース）で方程式を満たすものを探す
  - 残念ながら計算量的に困難（組み合わせ最適化問題）
- L1ノルム最小化によるスパース解推定
  - 代わりに次の最小化問題でスパース解を探索しよう！

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{y} = A\mathbf{x}$$

スパース解推定

- L0ノルム最小化によるスパース解推定
  - 以下の最小化問題でスパース解を探索しよう！

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad \mathbf{y} = A\mathbf{x}$$

  - 非零成分の個数が小さい（スパース）で方程式を満たすものを探す
  - 残念ながら計算量的に困難（組み合わせ最適化問題）
- L1ノルム最小化によるスパース解推定
  - 代わりに次の最小化問題でスパース解を探索しよう！

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{y} = A\mathbf{x}$$

  - 零が多ければ小さくなる＋大きさも小さくなりがち
  - 計算量は非常に軽い（Nの3乗程度）

ノルムについて注意

- Lpノルムの定義

$$\|\mathbf{x}\|_p = \sqrt[p]{|x_1|^p + |x_2|^p + \cdots + |x_N|^p}$$

  - L0「ノルム」はちょっと変
    - p->0の極限で定義.
  - L1ノルムは絶対値の和

$$\|\mathbf{x}\|_1 = |x_1| + |x_2| + \cdots + |x_N|$$

  - L2ノルムはユークリッド距離

$$\|\mathbf{x}\|_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_N^2}$$

スパース解推定?

- L2ノルム最小化によるスパース解推定は？
  - 以下の最小化問題でスパース解を探索しよう！

$$\min_{\mathbf{x}} \|\mathbf{x}\|_2 \quad \text{s.t.} \quad \mathbf{y} = A\mathbf{x}$$

  - 零が多ければ小さい＋大きさがかなり小さくなりがち+小さいかすが残る
  - 計算量は非常に軽い（Nの2〜3乗程度）

例題CS-I

以下のふたつの最小化問題を解いてみよう.

$$\min_{x_1,x_2} \{|x_1| + |x_2|\} \quad \text{s.t.} \quad y = A\mathbf{x}$$

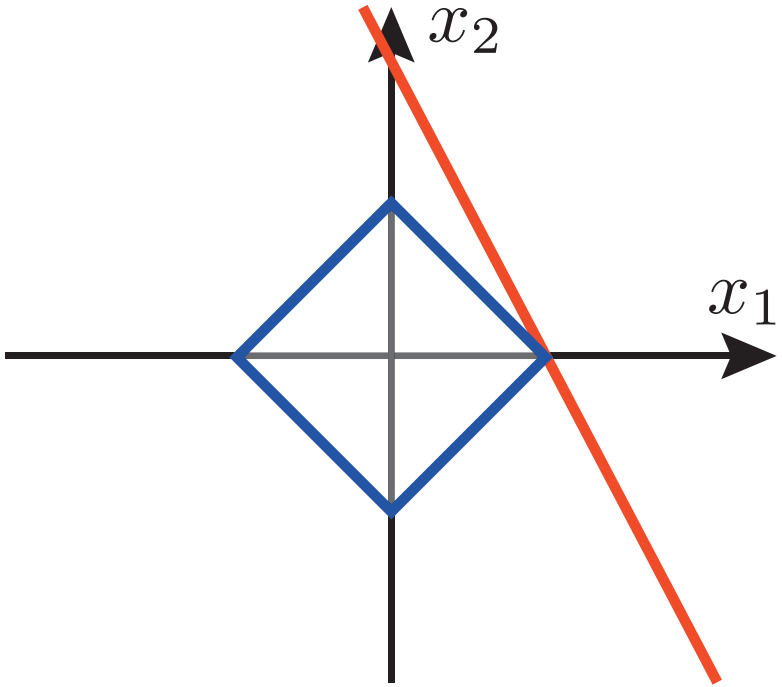
$$\min_{x_1,x_2} \left\{\sqrt{x_1^2 + x_2^2}\right\} \quad \text{s.t.} \quad y = A\mathbf{x}$$

ここで観測行列 A は、単なる行ベクトルに過ぎず、

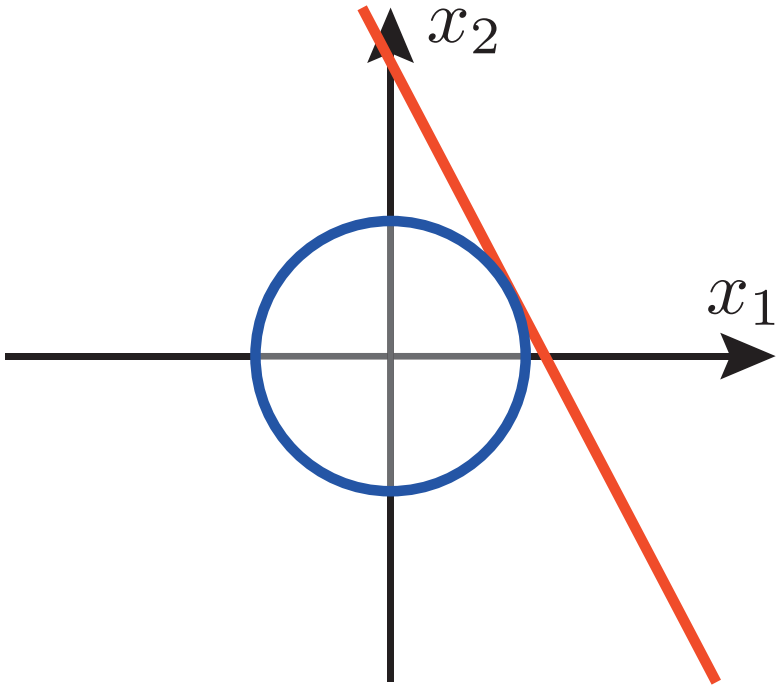
$$A = \begin{pmatrix} 2 & 1 \end{pmatrix}$$

とする. $y = 1$ とする.

正解

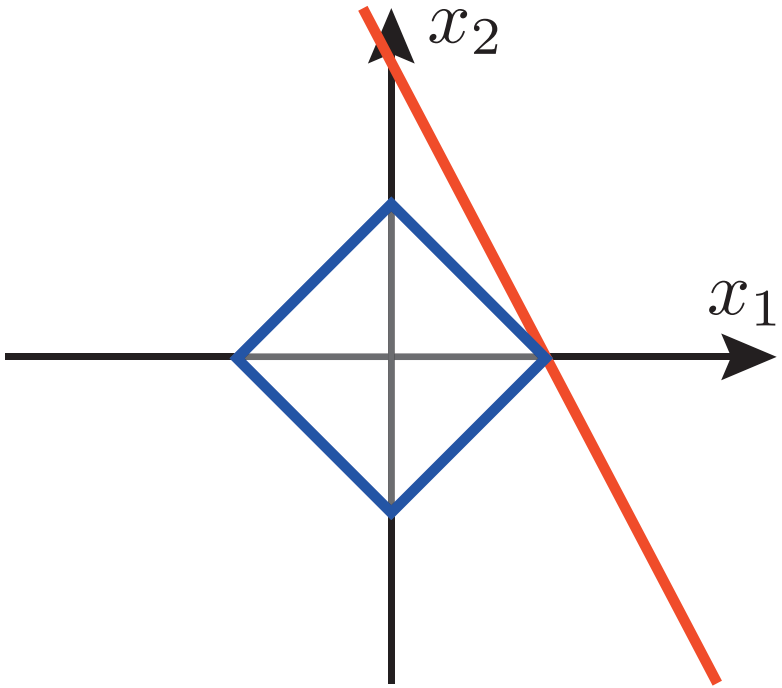


$(x_1, x_2) = (1/2, 0)$

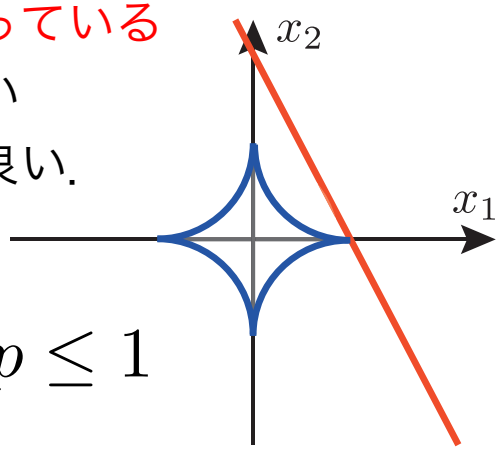


$(x_1, x_2) = (2/5, 1/5)$

吟味と注意

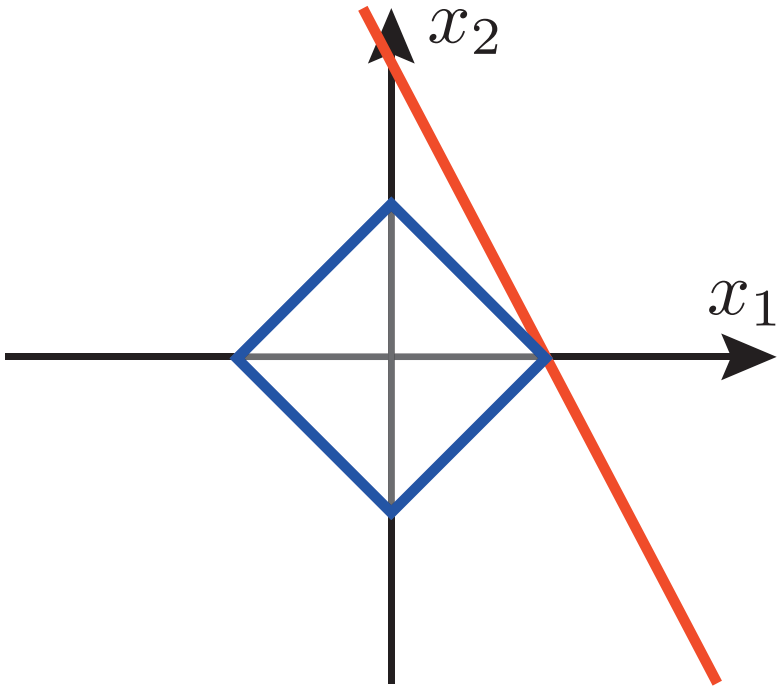


- L1ノルム最小化で
  スパース解は確かに得られる.
  - L1ノルムは尖っている
  - L2ノルムは丸い
  - Lpノルムでも良い.

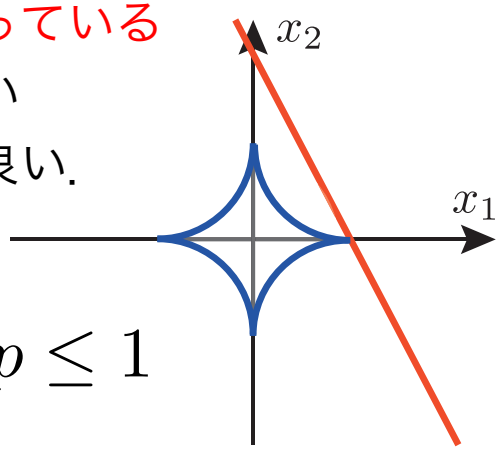


$0 < p \leq 1$

吟味と注意



- L1ノルム最小化で
  スパース解は確かに得られる.
  - L1ノルムは尖っている
  - L2ノルムは丸い
  - Lpノルムでも良い.



$$0 < p \leq 1$$

- L2ノルム最小化でも劣決定系の
  方程式の解が得られる.
  - ノルム最小化による解選択
    - L1ノルムはスパース解
    - L2ノルムは大きさが小さい解

QUANTUM ANNEALING MACHINE LEARNING Sparse Modeling CREST

基底追跡

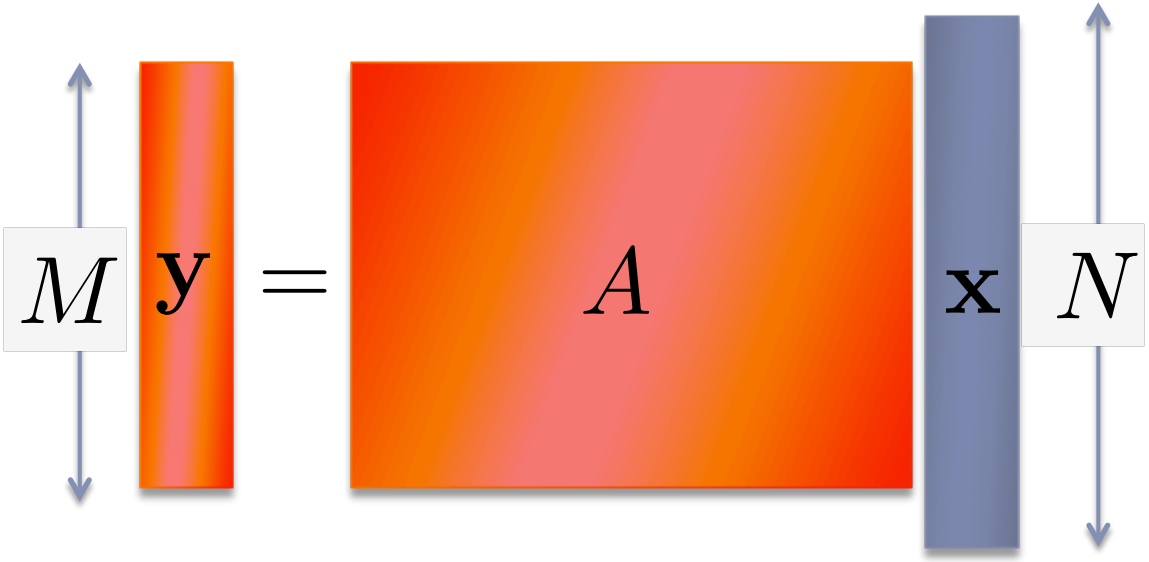
- L1ノルム最小化によるスパース解推定法
  - 基底追跡（Basis Pursuit）

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{y} = A\mathbf{x}$$

  - L1ノルムは零が多ければ小さくなる＋大きさも小さくなりがち
  - 計算量は非常に軽い（Nの3乗程度）

注意

圧縮センシング＝L1ノルム最小化ではない



足りない観測情報

$$\mathbf{y} = A\mathbf{x}$$

の結果から推定せよ

例題CS-II

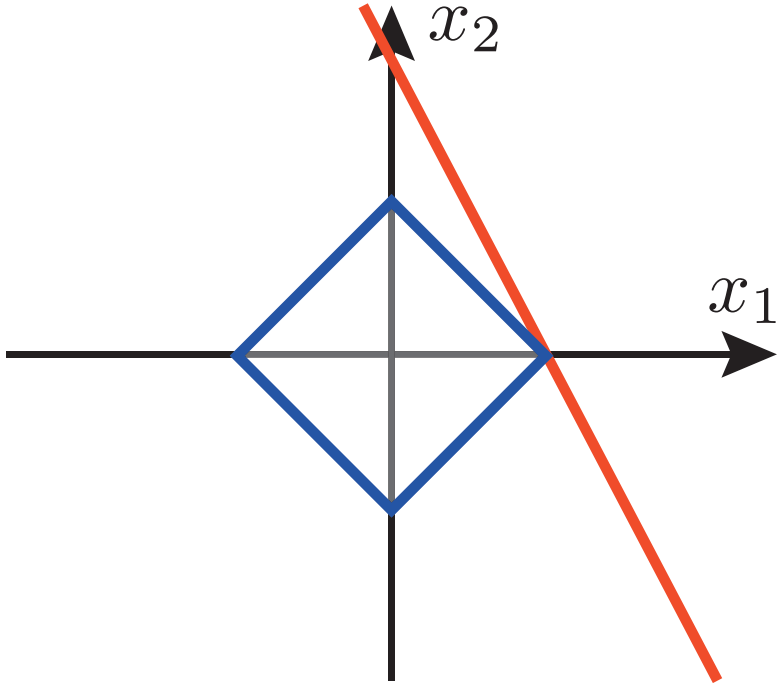
スパースな原信号ベクトル $\mathbf{x}_0^{\mathrm{T}} = (1/2, 0)$ に対して、$A = (2, 1)$ を作用して得られる $y = A\mathbf{x}_0 = 1$ のみが得られているとする．基底追跡により、スパース解を選択することにより原信号ベクトルを推定せよ．

例題CS-II

スパースな原信号ベクトル $\mathbf{x}_0^{\mathrm{T}} = (1/2, 0)$ に対して、$A = (2, 1)$ を作用して得られる $y = A\mathbf{x}_0 = 1$ のみが得られているとする．基底追跡により、スパース解を選択することにより原信号ベクトルを推定せよ．

- 正解はもちろん

$$(x_1, x_2) = (1/2, 0)$$

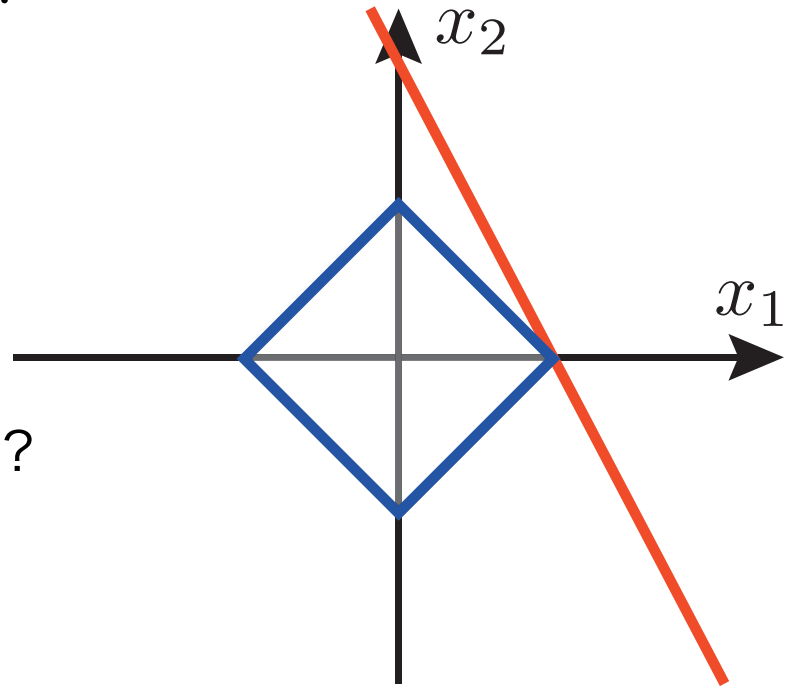


例題CS-II

スパースな原信号ベクトル $\mathbf{x}_0^{\mathrm{T}} = (1/2, 0)$ に対して、$A = (2, 1)$ を作用して得られる $y = A\mathbf{x}_0 = 1$ のみが得られているとする．基底追跡により、スパース解を選択することにより原信号ベクトルを推定せよ．

- 正解はもちろん

$$(x_1, x_2) = (1/2, 0)$$



- では原信号ベクトルが(0,1)だったら？？
  - 外れている….

基底追跡の性能

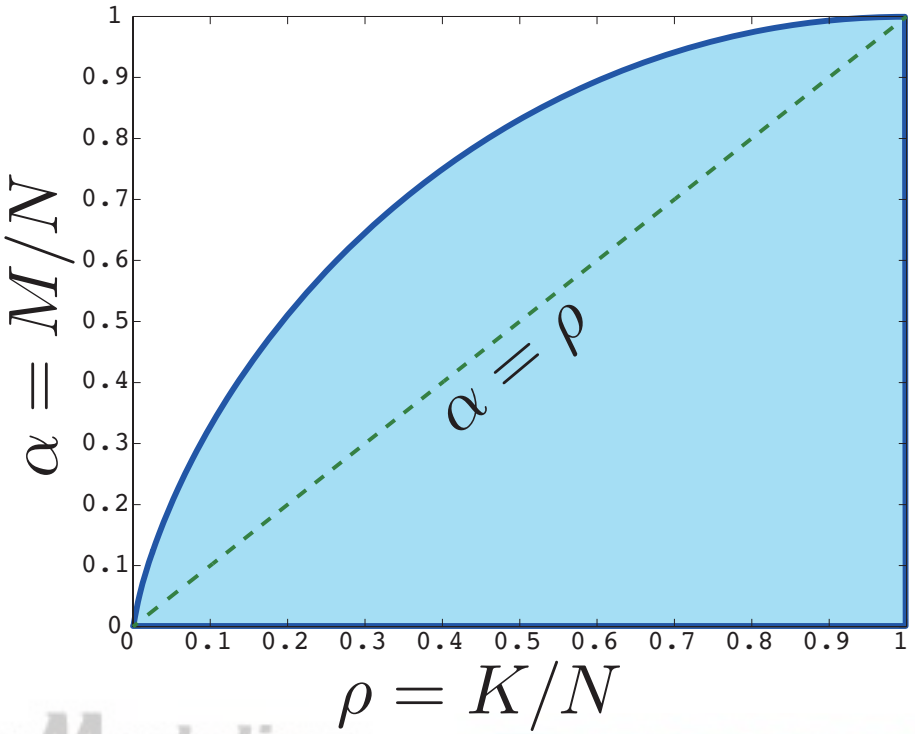
- L1ノルム最小化によるスパース解推定法
  - 基底追跡（Basis Pursuit）

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{y} = A\mathbf{x}$$

$$\frac{1}{\alpha} = 1 + \sqrt{\frac{\pi}{2}} t \mathrm{e}^{\frac{t^2}{2}} \{1 - 2Q(t)\}$$

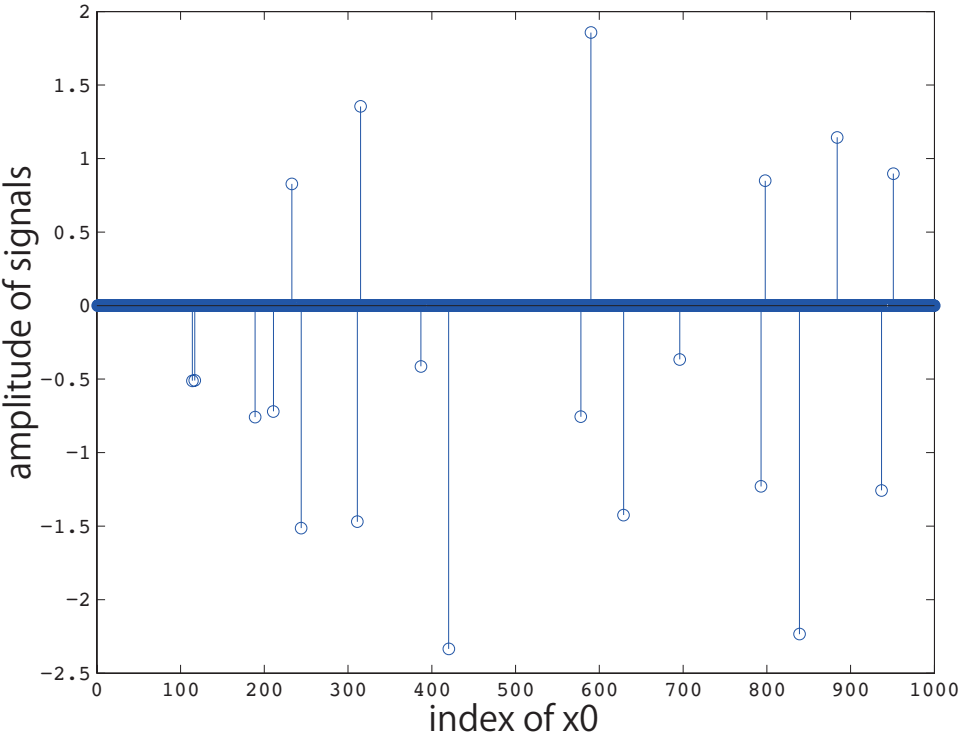
$$\frac{\rho}{1-\rho} = 2\left(\frac{\mathrm{e}^{\frac{t^2}{2}}}{t\sqrt{2\pi}} - Q(t)\right)$$

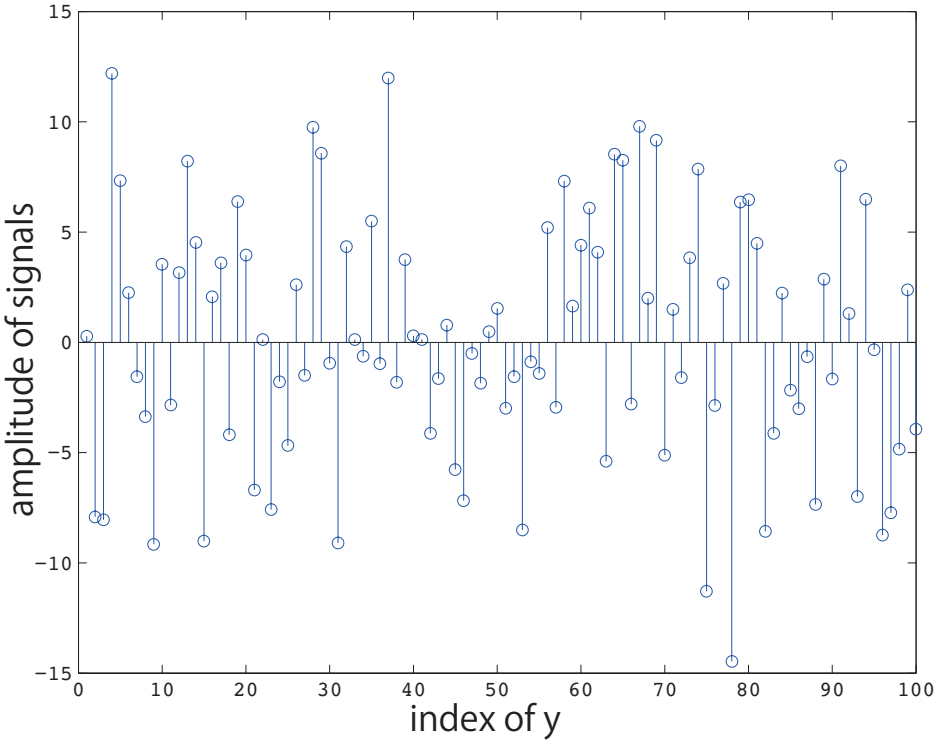
$$Q(t) = \int_t^\infty \frac{\mathrm{e}^{\frac{x^2}{2}}}{\sqrt{2\pi}} dx$$



実践例

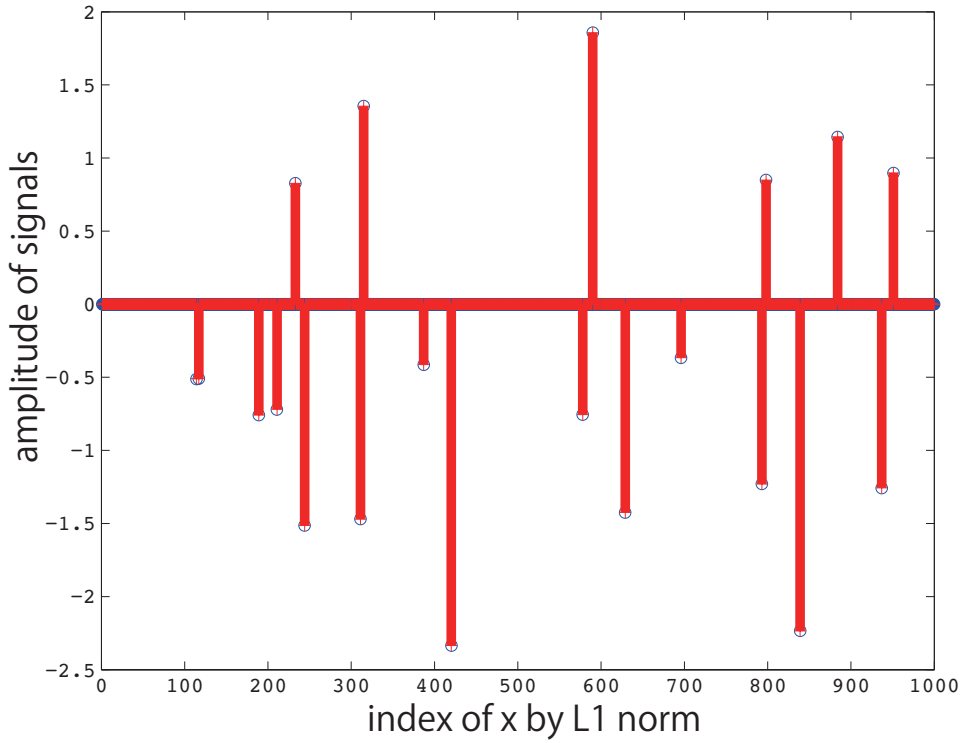
- 基底追跡（M=100,N=1000,K=20,ガウスランダム観測行列）

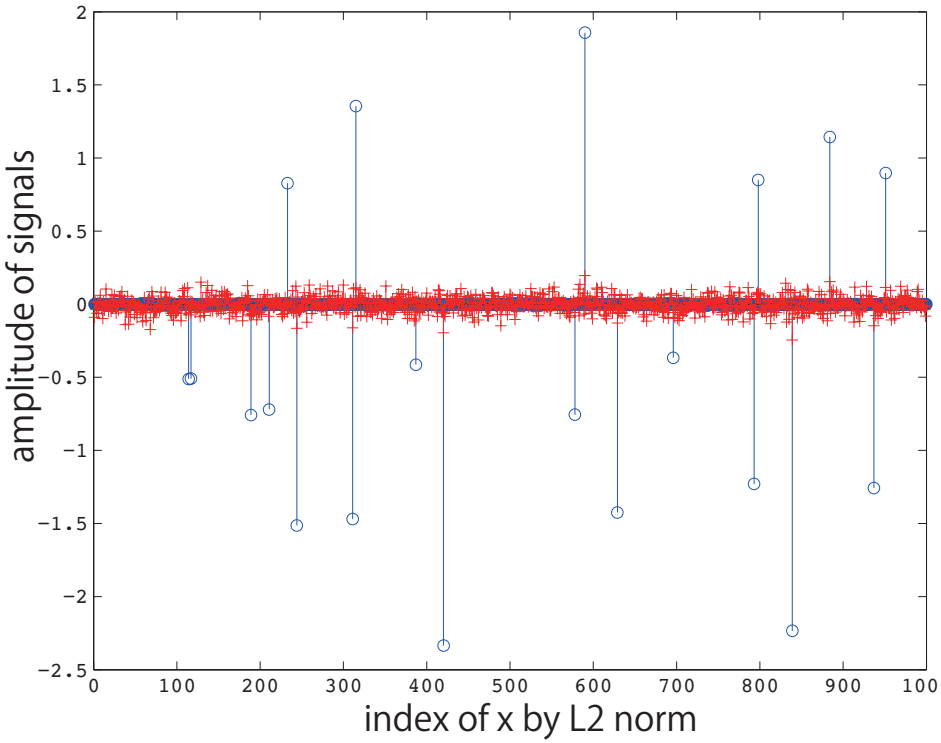




実践例

- 基底追跡（M=100,N=1000,K=20,ガウスランダム観測行列）





ノイズ有り圧縮センシング

- 観測結果にはノイズがつきもの？！
  - 加法的ノイズが混入した場合

$$\mathbf{y} = A\mathbf{x} + \sigma_0 \mathbf{w}$$

ノイズ有り圧縮センシング

- 観測結果にはノイズがつきもの？！
  - 加法的ノイズが混入した場合

$$\mathbf{y} = A\mathbf{x} + \sigma_0 \mathbf{w}$$

  - 等式制約はもはや成立しない
    - LASSO（Least Absolute Shrinkage and Selection Operators）

$$\min_{\mathbf{x}} \left\{ \|\mathbf{y} - A\mathbf{x}\|_2^2 \right\} \quad \text{s.t.} \quad \|\mathbf{x}\|_1 \leq a$$

ノイズ有り圧縮センシング

- 観測結果にはノイズがつきもの？！
  - 加法的ノイズが混入した場合

$$\mathbf{y} = A\mathbf{x} + \sigma_0 \mathbf{w}$$

  - 等式制約はもはや成立しない
    - LASSO（Least Absolute Shrinkage and Selection Operators）

$$\min_{\mathbf{x}} \left\{ \|\mathbf{y} - A\mathbf{x}\|_2^2 \right\} \quad \text{s.t.} \quad \|\mathbf{x}\|_1 \leq a$$

    - ラグランジュ未定乗数法により等価な最適化問題

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

そもそも解はスパースなの？

- 観測したいものは本当にスパースだろうか？

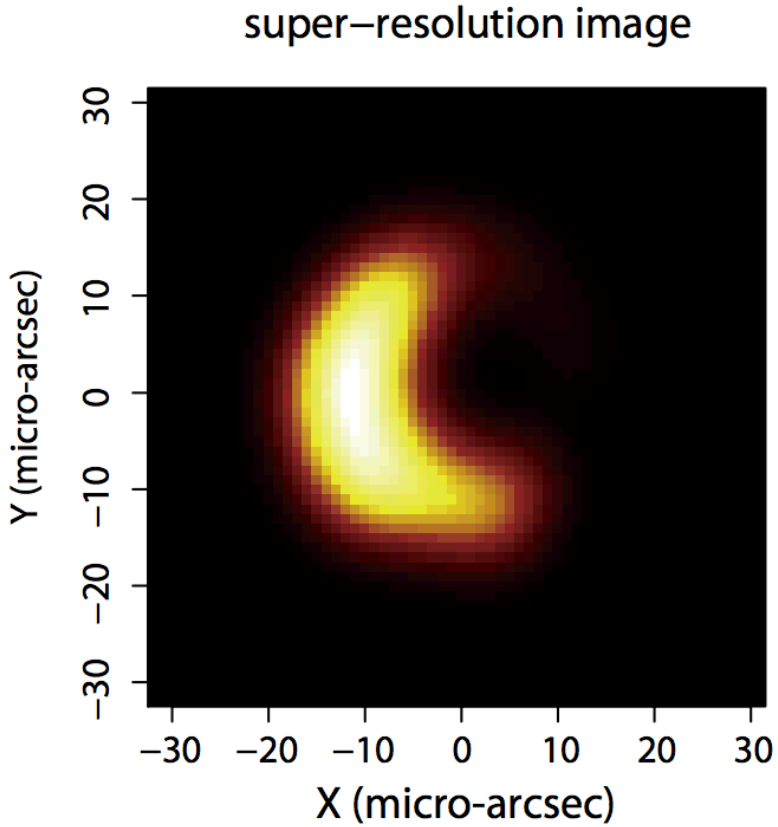
そもそも解はスパースなの？

- 観測したいものは本当にスパースだろうか？
  - 実空間でのスパース性

$$\|\mathbf{x}\|_1$$

  - 隣接間差分のスパース性

$$\|\mathrm{TV}(\mathbf{x})\|_1$$



ブラックホール撮像の例

M. Honma et al. Publ. Astron. Soc. Jon. (2014)

そもそも解はスパースなの？

- 観測したいものは本当にスパースだろうか？
  - 実空間でのスパース性

$$\|\mathbf{x}\|_1$$

  - 隣接間差分のスパース性

$$\|\mathrm{TV}(\mathbf{x})\|_1$$

original image

自然画像の例

そもそも解はスパースなの？

- 観測したいものは本当にスパースだろうか？
  - 実空間でのスパース性

$$\|\mathbf{x}\|_1$$

  - 隣接間差分のスパース性

$$\|\mathrm{TV}(\mathbf{x})\|_1$$

  - ウェーブレット変換によるスパース性

$$\|W(\mathbf{x})\|_1$$

自然画像の例

そもそも解はスパースなの？

- 観測したいものは本当にスパースだろうか？
  - 実空間でのスパース性

$$\|\mathbf{x}\|_1$$

  - 隣接間差分のスパース性

$$\|\mathrm{TV}(\mathbf{x})\|_1$$

  - ウェーブレット変換によるスパース性

$$\|W(\mathbf{x})\|_1$$

reconstructed image

自然画像の例

そもそも解はスパースなの？

- 観測したいものは本当にスパースだろうか？
  - 実空間でのスパース性

$$\|\mathbf{x}\|_1$$

  - 隣接間差分のスパース性

$$\|\mathrm{TV}(\mathbf{x})\|_1$$

  - ウェーブレット変換によるスパース性

$$\|W(\mathbf{x})\|_1$$

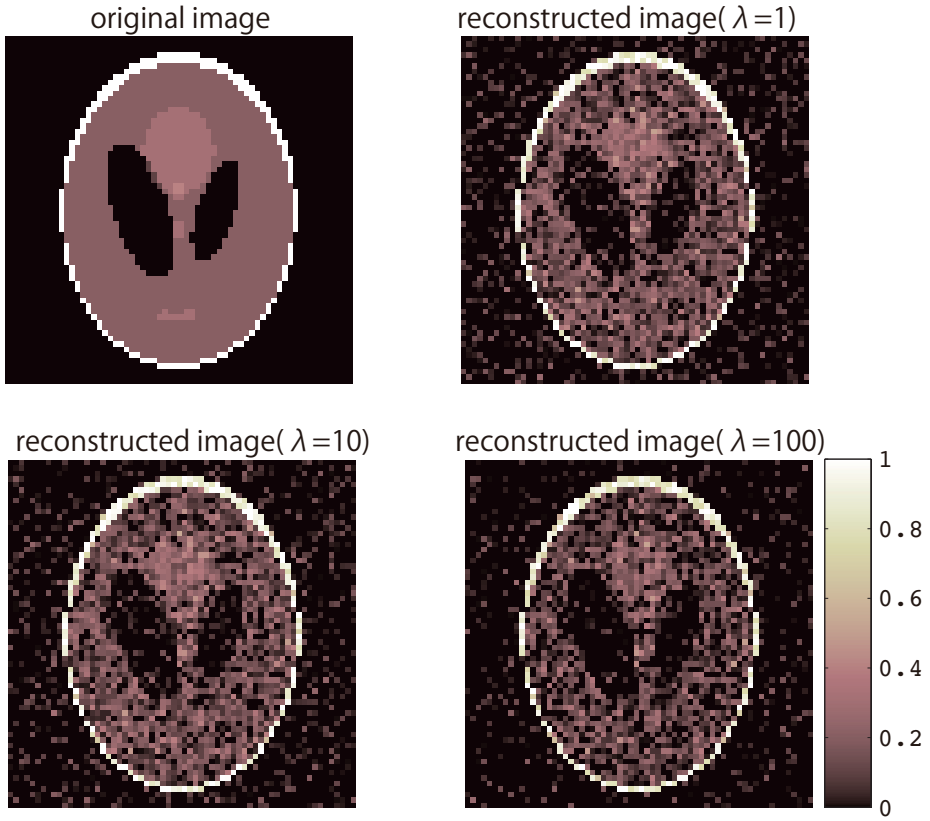
  - スパースモデリングで獲得するスパース性

reconstructed image

自然画像の例

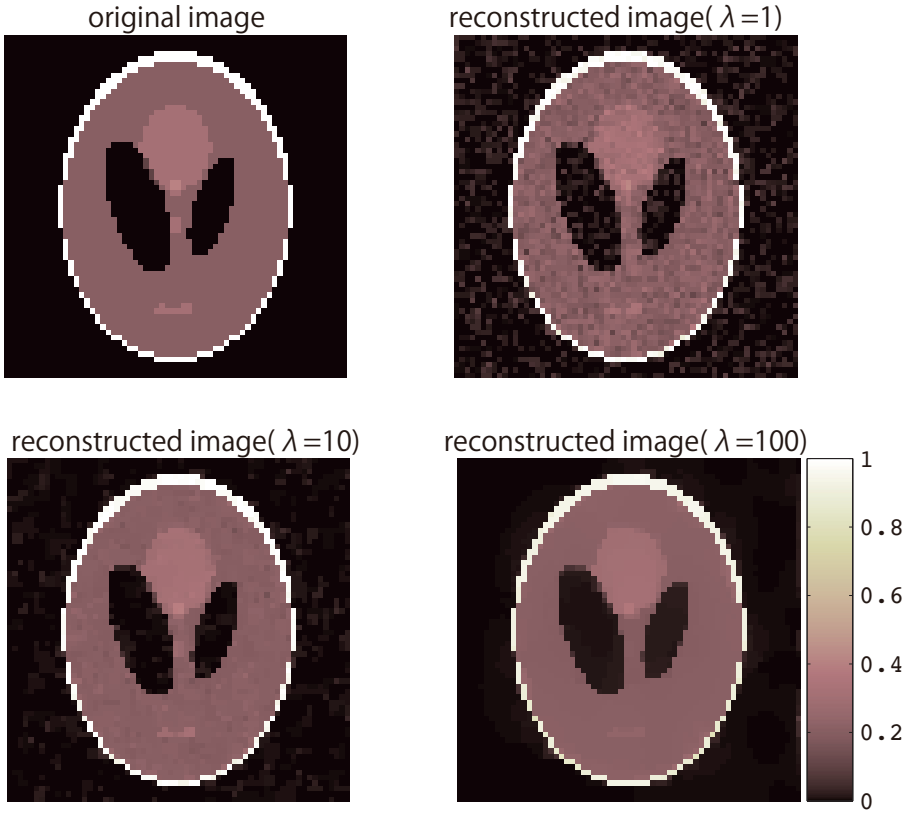
そもそも解はスパースなの？

- 実践例



そもそも解はスパースなの？

- 実践例



そもそも線形観測なの？

- 観測過程は線形変換で書かれるのか？

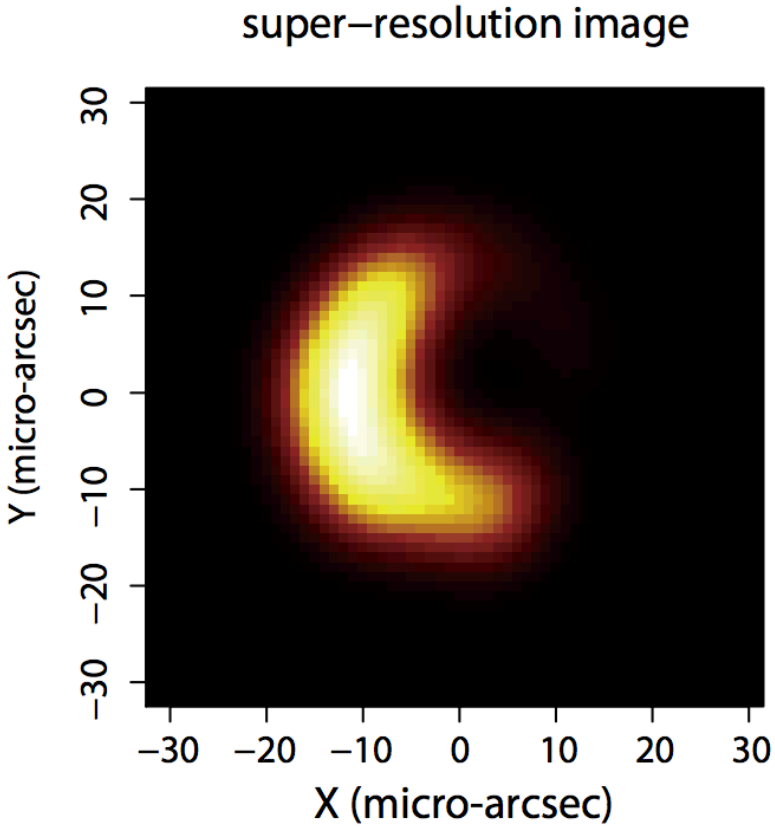
そもそも線形観測なの？

- 観測過程は線形変換で書かれるのか？
  - フーリエ変換部分行列+コイル強度

$$\mathbf{y} = FR\mathbf{x}_0 + \sigma_0\mathbf{w}$$

  - フーリエ変換部分行列＋天候要素

$$\mathbf{y} = FR\mathbf{x}_0 + \sigma_0\mathbf{w}$$

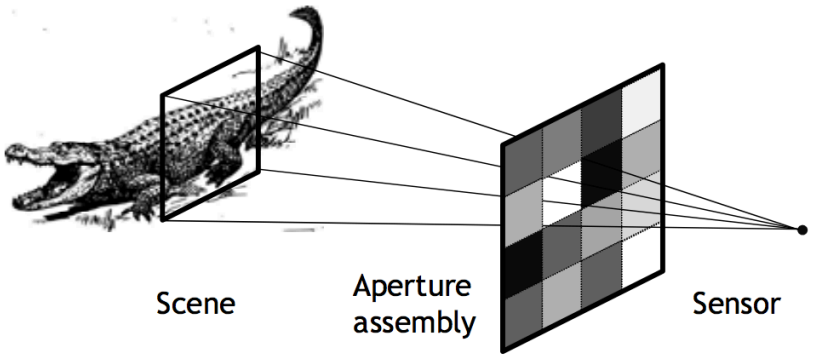


ブラックホール撮像の例

そもそも線形観測なの？

- 観測過程は線形変換で書かれるのか？
  - フーリエ変換部分行列+コイル強度

$$\mathbf{y} = FR\mathbf{x}_0 + \sigma_0\mathbf{w}$$



  - フーリエ変換部分行列＋天候要素

$$\mathbf{y} = FR\mathbf{x}_0 + \sigma_0\mathbf{w}$$

  - デザインした観測

$$\mathbf{y} = A\mathbf{x}_0 + \sigma_0\mathbf{w}$$

Lenseless camera
G. Huang et al: arXiv:1305.7181

圧縮センシング実践編

スパース解推定

- L0ノルム最小化によるスパース解推定
  - 以下の最小化問題でスパース解を探索しよう！

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad \mathbf{y} = A\mathbf{x}$$

  - 非零成分の個数が小さい（スパース）で方程式を満たすものを探す
  - 残念ながら計算量的に困難
- L1ノルム最小化によるスパース解推定
  - 代わりに次の最小化問題でスパース解を探索しよう！

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{y} = A\mathbf{x}$$

  - 零が多ければ小さくなる＋大きさも小さくなりがち
  - 計算量は非常に軽い（Nの3乗程度）

お手軽に圧縮センシング？！

- 各種ライブラリが勢揃い
  - Matlab、Python用の最適化ライブラリ
    - CVX for Matlab, CVXPY for python, Convex.jl for Julia

```
M = 20;
N = 100;
K = 5;%非ゼロ要素の数
A = randn(M,N);%Aを適当に生成
x0 = [ones(K,1); zeros(N-K,1)];
x0 = x0(randperm(N));%どこかわからないようにばらまく
y = A*x0;%出力データの生成

cvx_begin%ここからcvxのお作法
variable x(N)%解きたい変数をかく
minimize(norm(x,1))%最小化したい関数を書く
subject to
norm(A*x-y,2) <= 0.00001;%制約条件．ここでは不等式で押さえた．
cvx_end
```

    - http://stanford.edu/~boyd/papers/cvx_short_course.html

ノイズ有り圧縮センシング

- 観測結果にはノイズがつきもの？！
  - 加法的ノイズが混入した場合

$$\mathbf{y} = A\mathbf{x} + \sigma_0 \mathbf{w}$$

  - 等式制約はもはや成立しない
    - LASSO（Least Absolute Shrinkage and Selection Operators）

$$\min_{\mathbf{x}} \left\{ \|\mathbf{y} - A\mathbf{x}\|_2^2 \right\} \quad \text{s.t.} \quad \|\mathbf{x}\|_1 \leq a$$

    - ラグランジュ未定乗数法により等価な最適化問題

ノイズ有り圧縮センシング

- 観測結果にはノイズがつきもの？！
  - 加法的ノイズが混入した場合

$$\mathbf{y} = A\mathbf{x} + \sigma_0 \mathbf{w}$$

  - 等式制約はもはや成立しない
    - LASSO（Least Absolute Shrinkage and Selection Operators）

$$\min_{\mathbf{x}} \left\{ \|\mathbf{y} - A\mathbf{x}\|_2^2 \right\} \quad \text{s.t.} \quad \|\mathbf{x}\|_1 \leq a$$

    - ラグランジュ未定乗数法により等価な最適化問題

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

例題

以下の最小化問題を解け.

$$\min_x \left\{ |x| + \frac{1}{2\lambda}(y-x)^2 \right\}$$

ここで λ, y は適当な実数である.

正解

- 絶対値関数は場合分けすれば良い
  - X>0のときについて平方完成すると以下のようになる.

$$\frac{1}{2\lambda}\{x-(y-\lambda)\}^2+y-\frac{\lambda}{2}$$

  - ２次関数の最小値は頂点！？
    - 頂点の正負により最小値が変わる.

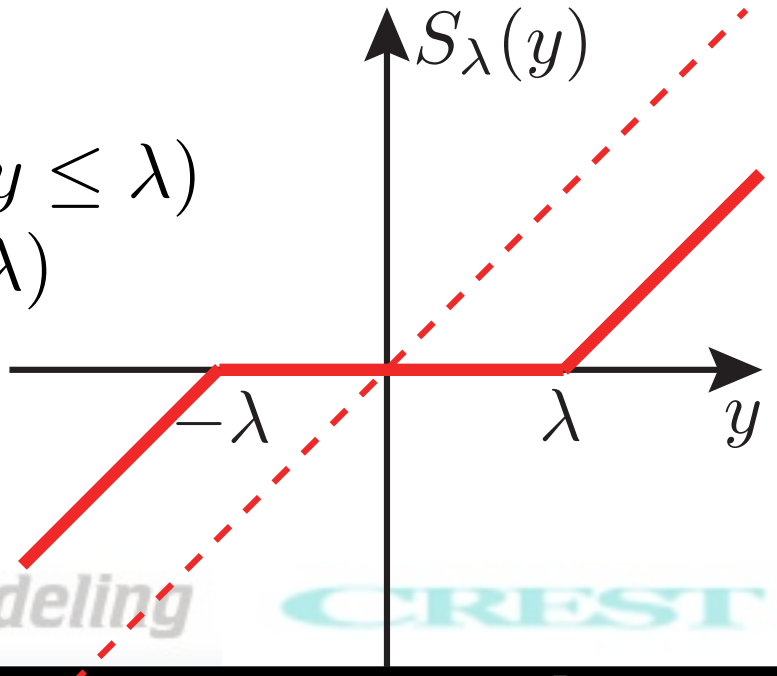
結果

- 軟判定しきい値関数の導入

$$\min_x \left\{ |x| + \frac{1}{2\lambda}(y-x)^2 \right\}$$

  - この最小化問題の解は、以下の軟判定しきい値関数で与えられる.

$$S_\lambda(y) = \begin{cases} y-\lambda & (y > \lambda) \\ 0 & (-\lambda \le y \le \lambda) \\ y+\lambda & (y < -\lambda) \end{cases}$$



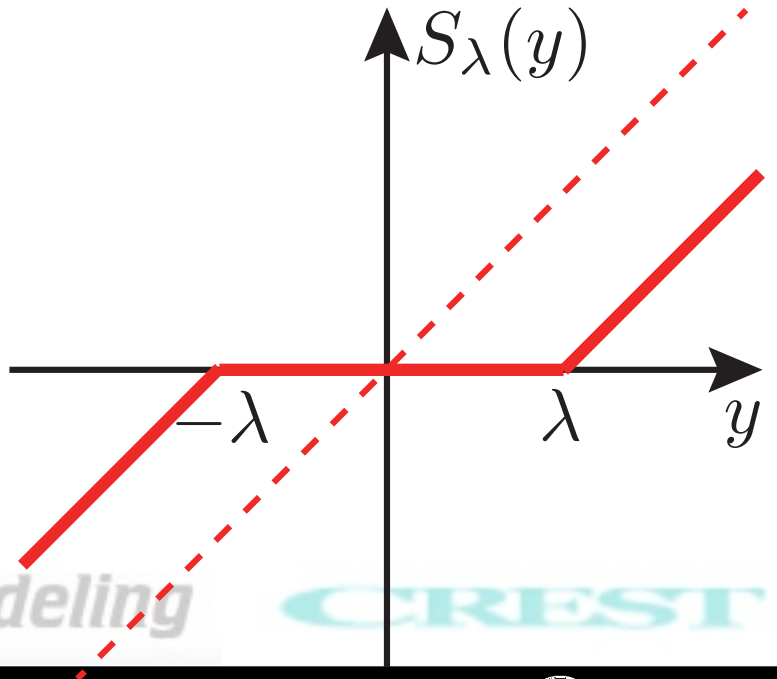
多変数であっても同様

- L1ノルム、L2ノルムの分離性により

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

  - この最小化問題の解は、軟判定しきい値関数で与えられる.

$$\mathbf{x}^* = S_\lambda(\mathbf{y})$$



多変数であっても同様

- L1ノルム、L2ノルムの分離性により

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

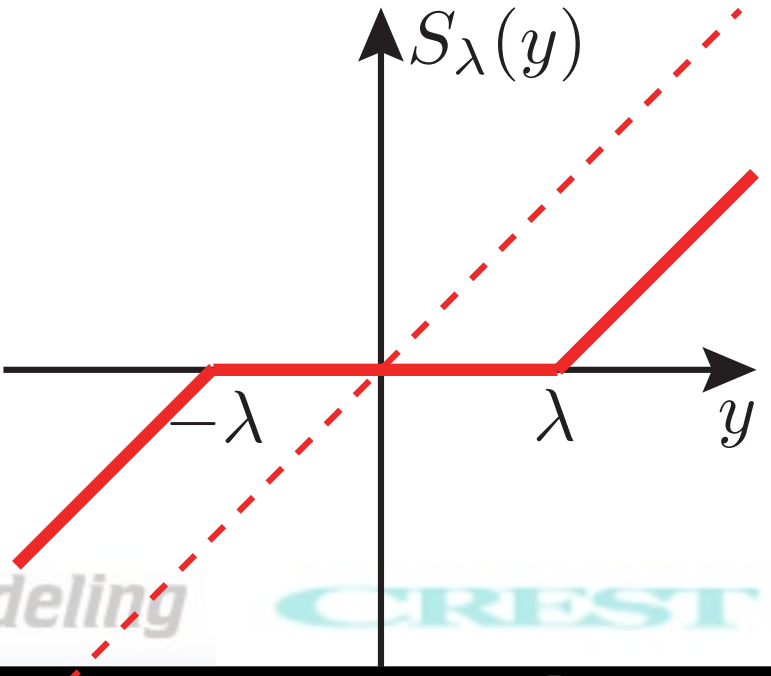
  - この最小化問題の解は、軟判定しきい値関数で与えられる.

$$\mathbf{x}^* = S_\lambda(\mathbf{y})$$

  - 分離性は、成分毎の和に分かれること

$$\|\mathbf{x}\|_1 = \|x_1\| + \|x_2\| + \cdots + \|x_N\|$$

$$\|\mathbf{x}\|_2^2 = x_1^2 + x_2^2 + \cdots + x_N^2$$



問題点

- 解くべき問題はLASSO型最小化問題

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

  - ちょっとそのまま軟判定しきい値関数は利用できない

問題点

- 解くべき問題はLASSO型最小化問題

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

  - ちょっとそのまま軟判定しきい値関数は利用できない
  - なんとかして分離性を持つ形にしたい.

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

問題点

- 解くべき問題はLASSO型最小化問題

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

  - ちょっとそのまま軟判定しきい値関数は利用できない
  - なんとかして分離性を持つ形にしたい.

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

  - そうすれば軟判定しきい値関数を適用できる

$$\mathbf{x}^* = S_\lambda(\mathbf{v})$$

メジャライザー最小化

- 最小化したい関数 $g(\mathbf{x})$ を上からふたするメジャライザーを導入

$$q_L(\mathbf{x}, \mathbf{v}) = g(\mathbf{v}) + (\nabla g(\mathbf{v}))^{\mathrm{T}} (\mathbf{x} - \mathbf{v}) + \frac{L}{2} \|\mathbf{x} - \mathbf{v}\|_2^2$$

$$g(\mathbf{x}) \leq q_L(\mathbf{x}, \mathbf{v})$$

メジャライザー最小化

- 最小化したい関数 $g(\mathbf{x})$ を上からふたするメジャライザーを導入

$$q_L(\mathbf{x}, \mathbf{v}) = g(\mathbf{v}) + (\nabla g(\mathbf{v}))^{\mathrm{T}} (\mathbf{x} - \mathbf{v}) + \frac{L}{2} \|\mathbf{x} - \mathbf{v}\|_2^2$$

$$g(\mathbf{x}) \leq q_L(\mathbf{x}, \mathbf{v})$$

  - メジャライザーは２次関数なので平方完成すれば最小値は即座にわかる

メジャライザー最小化

- 最小化したい関数 $g(\mathbf{x})$ を上からふたするメジャライザーを導入

$$q_L(\mathbf{x}, \mathbf{v}) = g(\mathbf{v}) + (\nabla g(\mathbf{v}))^{\mathrm{T}} (\mathbf{x} - \mathbf{v}) + \frac{L}{2} \|\mathbf{x} - \mathbf{v}\|_2^2$$

$$g(\mathbf{x}) \leq q_L(\mathbf{x}, \mathbf{v})$$

  - メジャライザーは２次関数なので平方完成すれば最小値は即座にわかる

- 逐次メジャライザーの最小化をすると…

$$g(\mathbf{x}[t+1]) \leq q_L(\mathbf{x}[t+1], \mathbf{x}[t]) \leq g(\mathbf{x}[t])$$

メジャライザー最小化

- 最小化したい関数 $g(\mathbf{x})$ を上からふたするメジャライザーを導入

$$q_L(\mathbf{x}, \mathbf{v}) = g(\mathbf{v}) + (\nabla g(\mathbf{v}))^{\mathrm{T}} (\mathbf{x} - \mathbf{v}) + \frac{L}{2} \|\mathbf{x} - \mathbf{v}\|_2^2$$

$$g(\mathbf{x}) \leq q_L(\mathbf{x}, \mathbf{v})$$

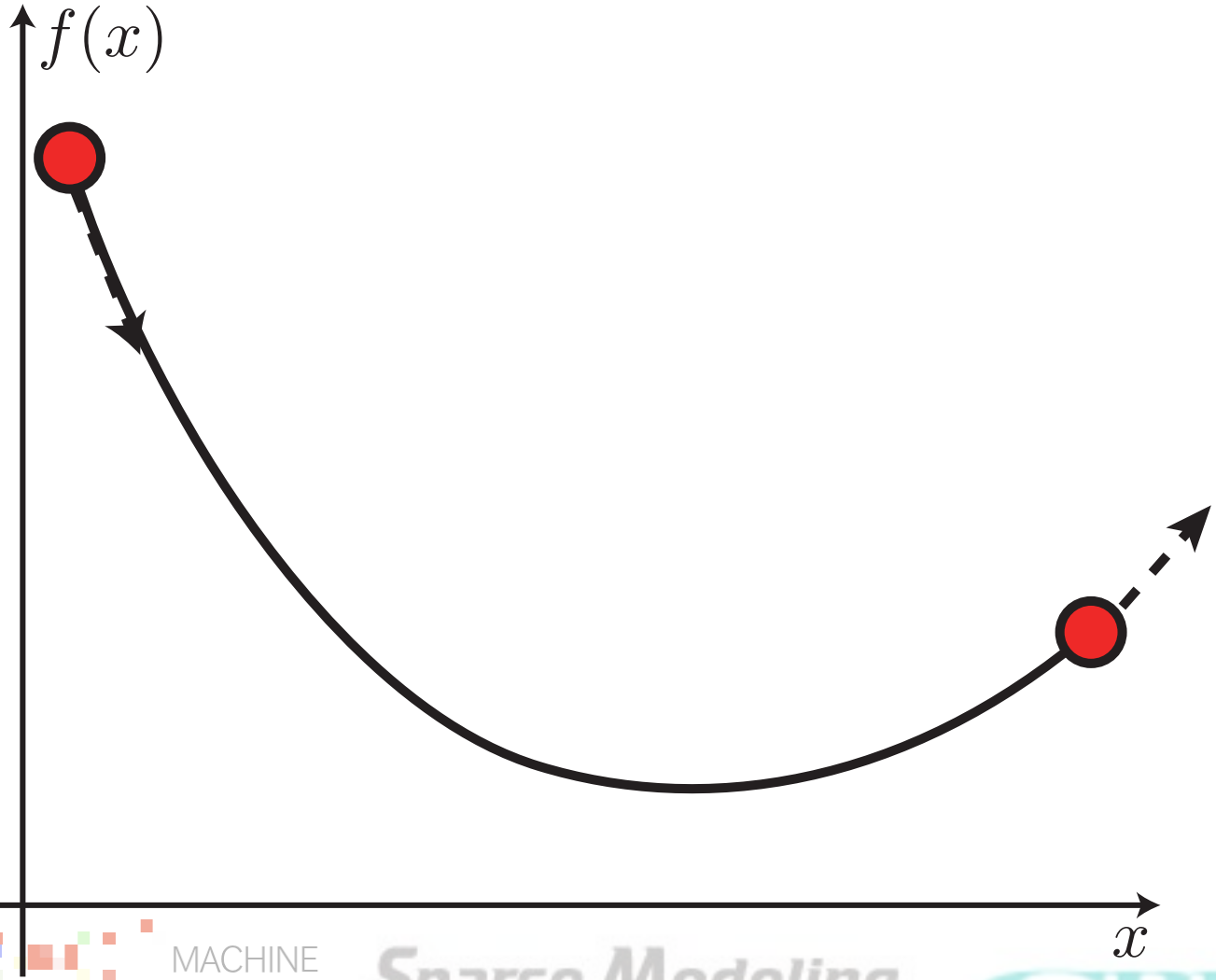
  - メジャライザーは２次関数なので平方完成すれば最小値は即座にわかる

- 逐次メジャライザーの最小化をすると…

$$g(\mathbf{x}[t+1]) \leq q_L(\mathbf{x}[t+1], \mathbf{x}[t]) \leq g(\mathbf{x}[t])$$

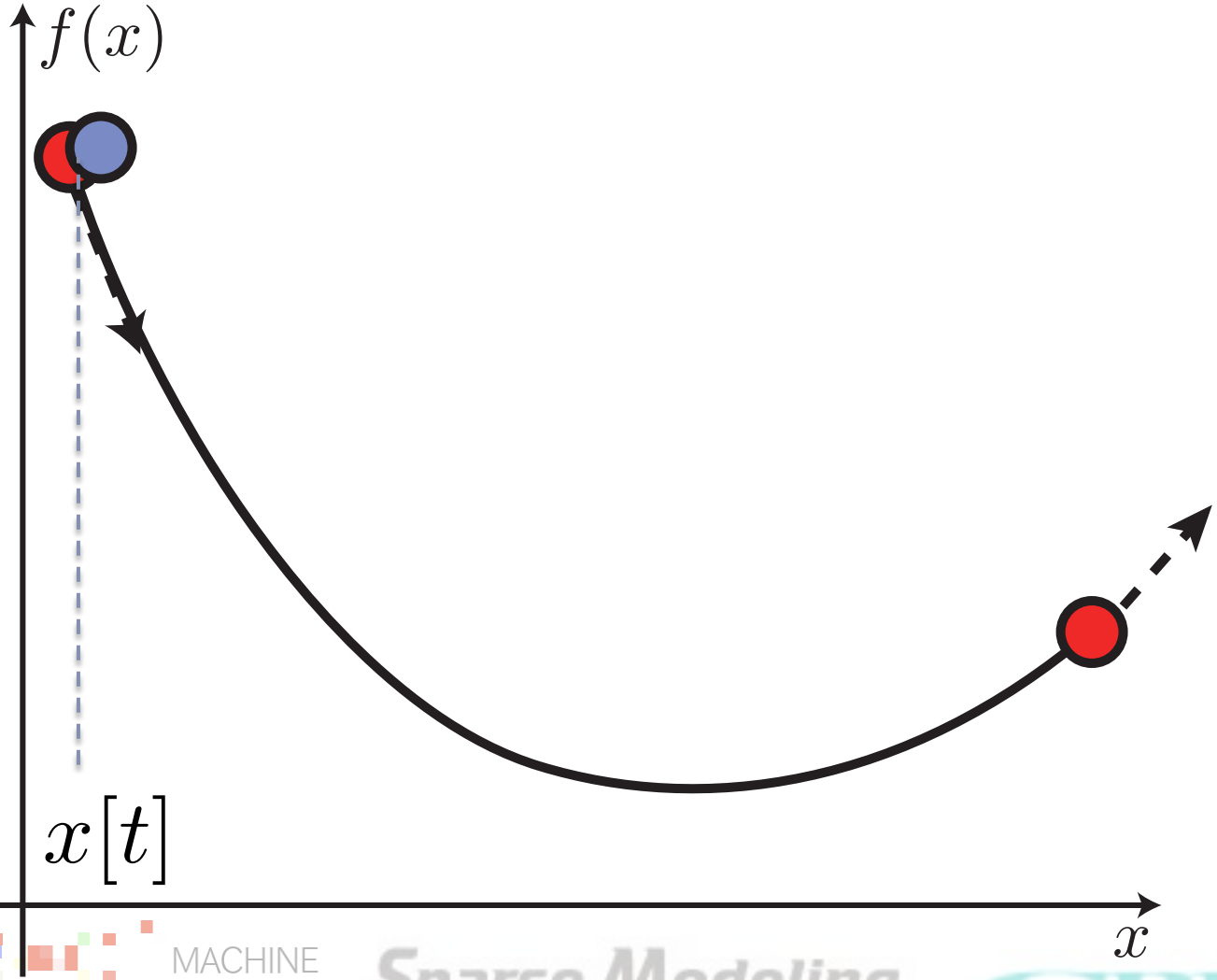
  - 本来の最小化したい関数も小さくできる！

メジャライザー逐次最小化のイメージ



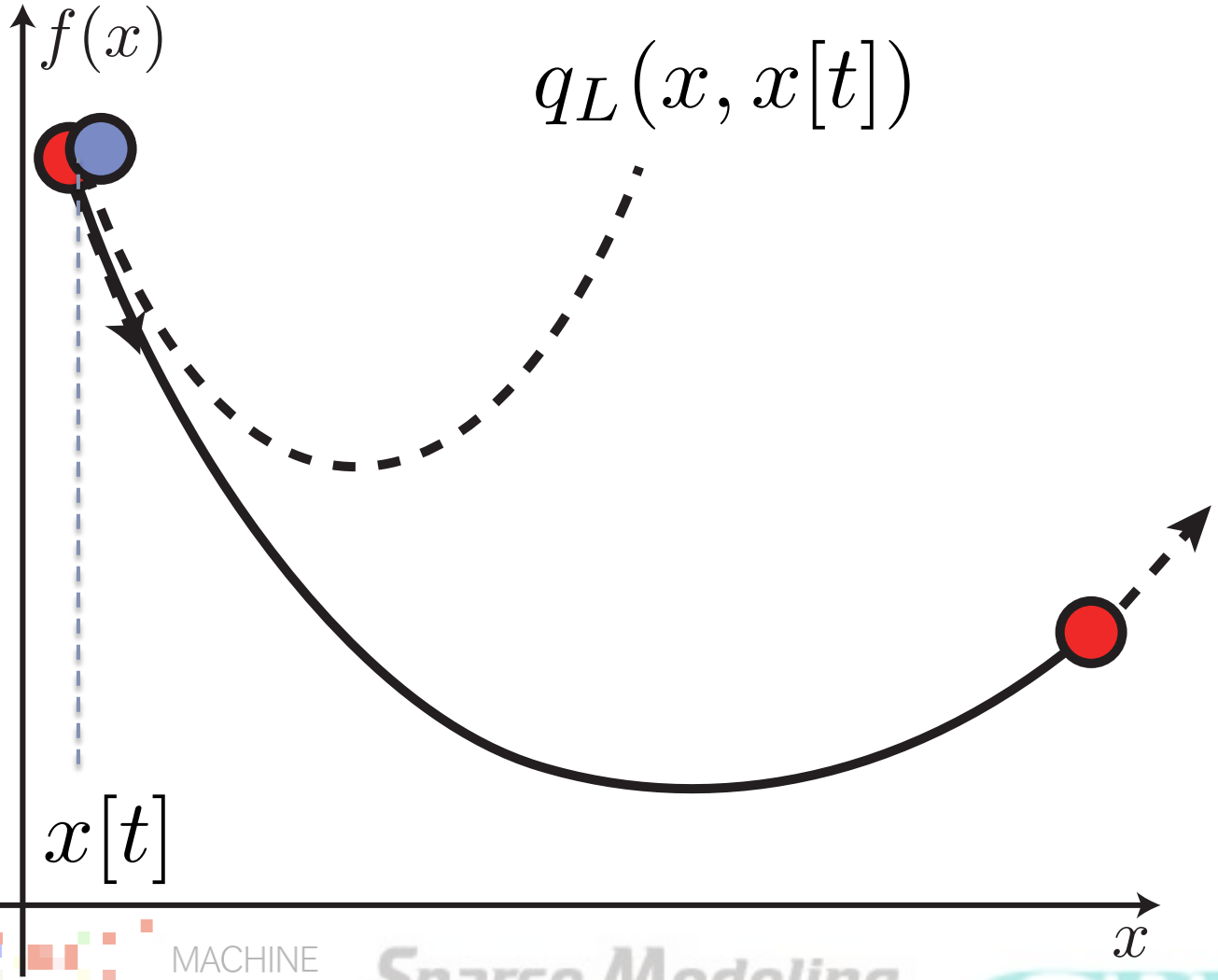
QUANTUM ANNEALING
MACHINE LEARNING
Sparse Modeling
CREST

メジャライザー逐次最小化のイメージ



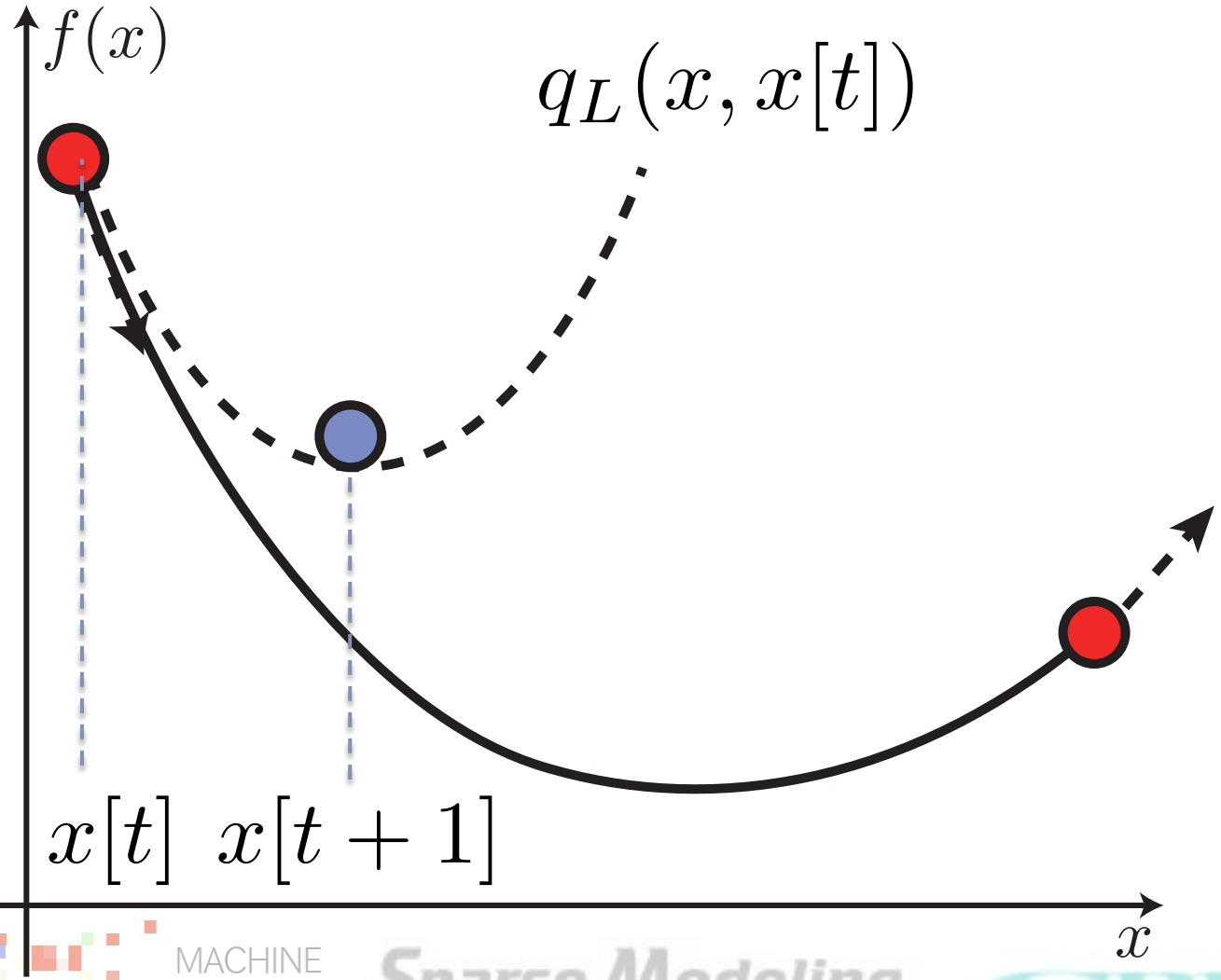
QUANTUM ANNEALING
MACHINE LEARNING
Sparse Modeling
CREST

メジャライザー逐次最小化のイメージ



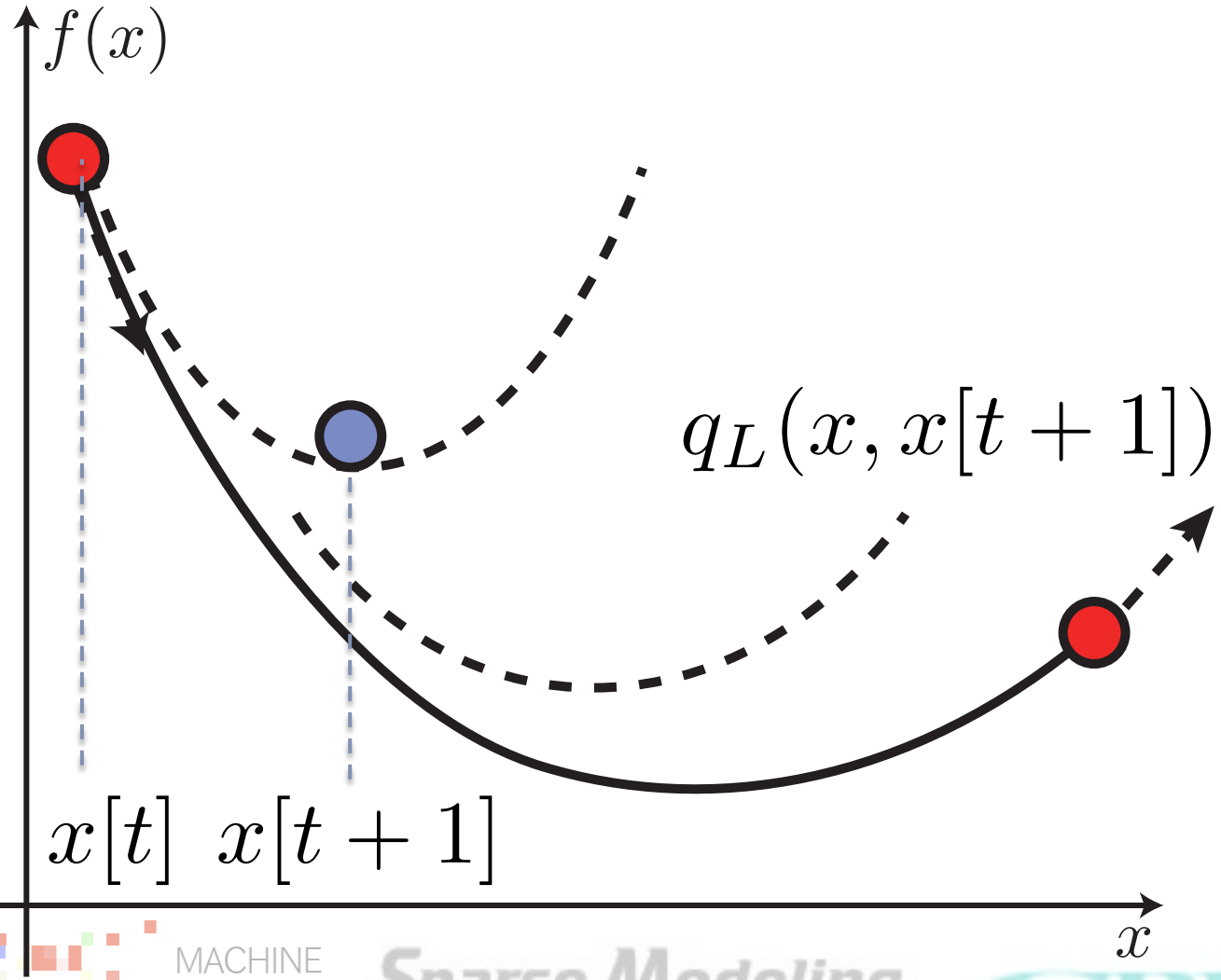
QUANTUM ANNEALING
MACHINE LEARNING
Sparse Modeling
CREST

メジャライザー逐次最小化のイメージ



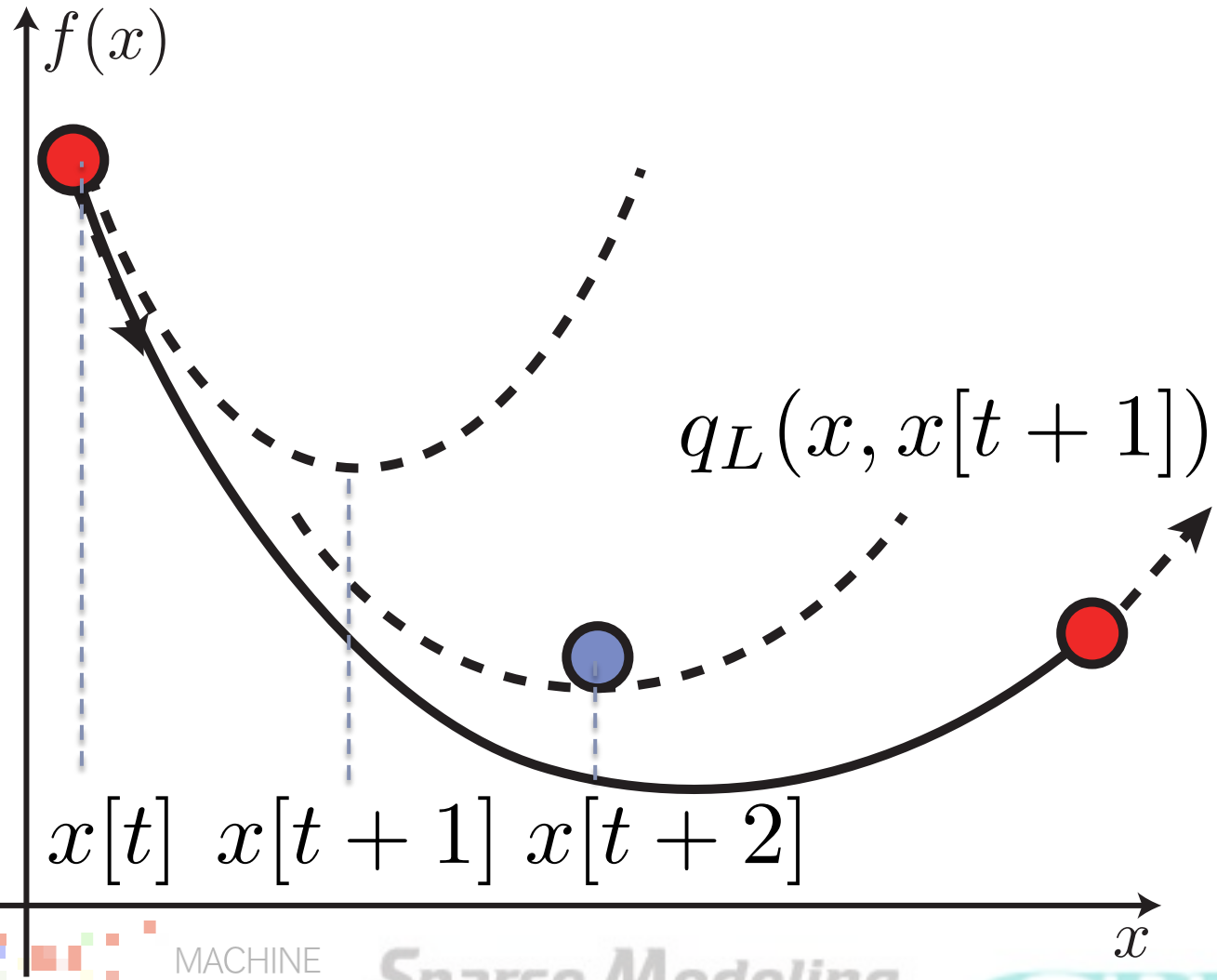
QUANTUM ANNEALING
MACHINE LEARNING
Sparse Modeling
CREST

メジャライザー逐次最小化のイメージ



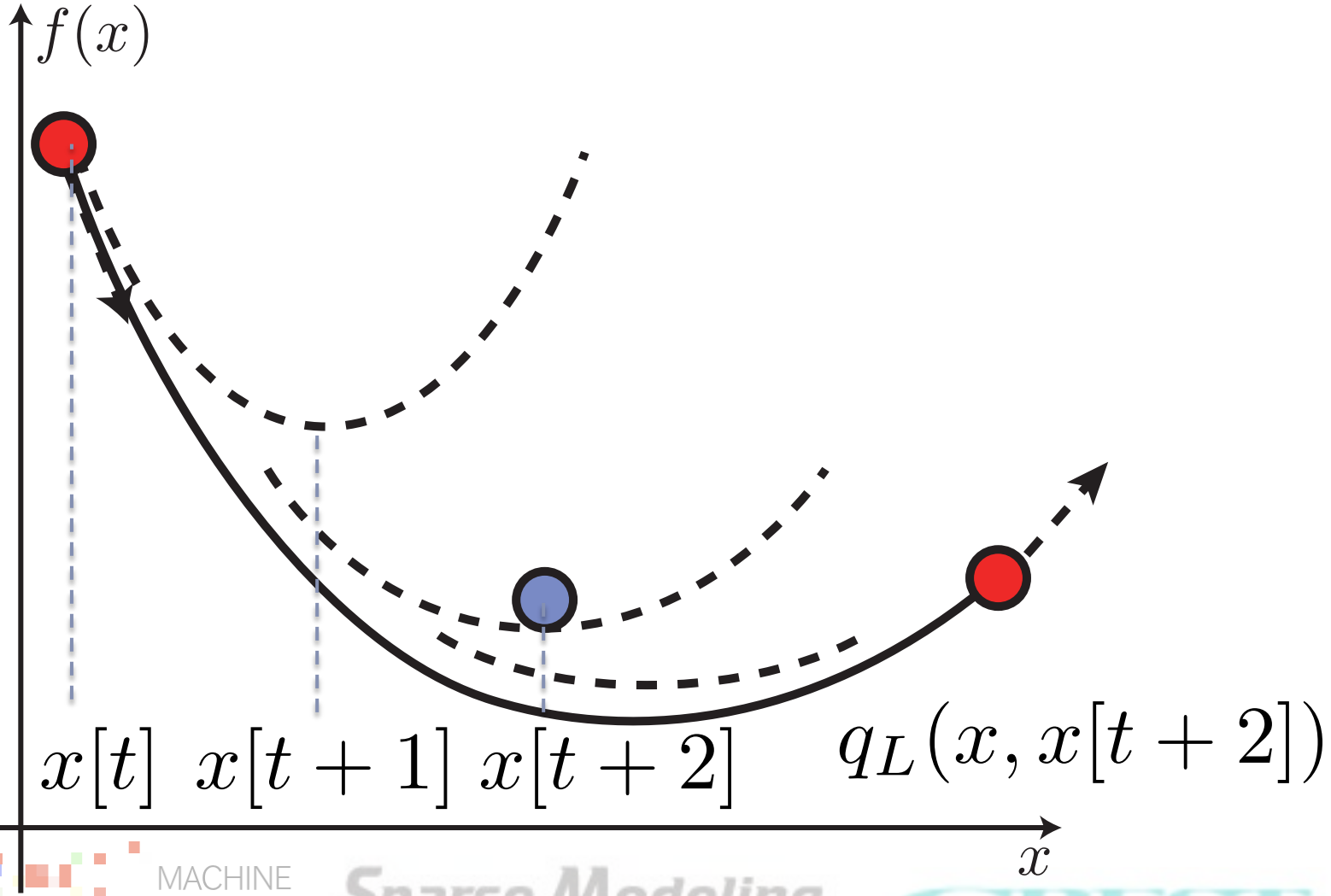
QUANTUM ANNEALING
MACHINE LEARNING
Sparse Modeling
CREST

メジャライザー逐次最小化のイメージ



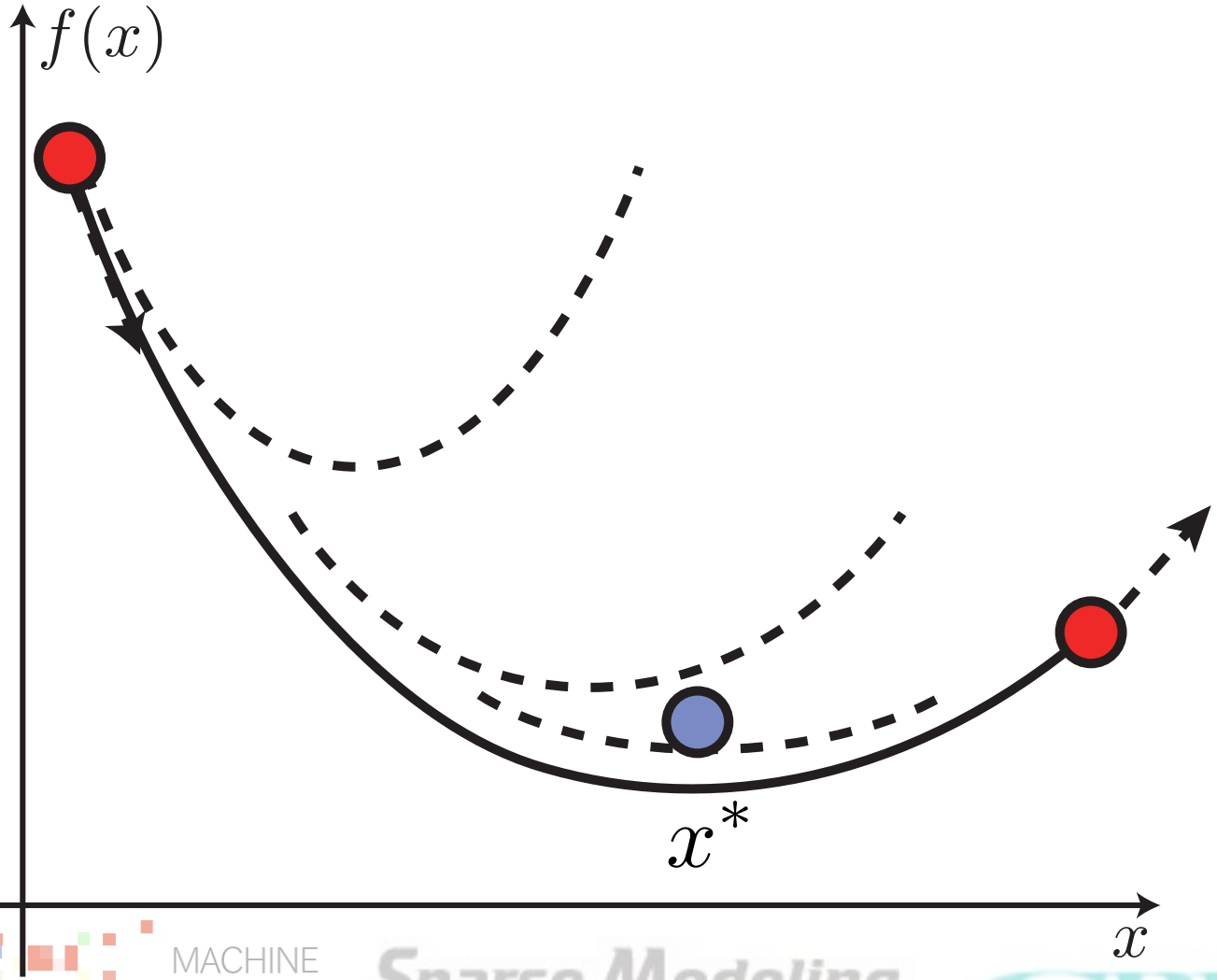
QUANTUM ANNEALING
MACHINE LEARNING
Sparse Modeling
CREST

メジャライザー逐次最小化のイメージ



QUANTUM ANNEALING
MACHINE LEARNING
Sparse Modeling
CREST

メジャライザー逐次最小化のイメージ



QUANTUM ANNEALING
MACHINE LEARNING
Sparse Modeling
CREST

ISTA: Iterative Shrinkage Thresholding Algorithm

1. $t = 0$ とする．初期化 $\mathbf{x}[0]$. (例えば $\mathbf{x}[0] = A^{\mathrm{T}}\mathbf{y}$)
2. 平方完成により $g(\mathbf{x})$ の 2 次関数近似の頂点を求める．

$$\mathbf{v}[t] = \mathbf{x}[t] + \frac{1}{L\lambda} A^{\mathrm{T}}(\mathbf{y} - A\mathbf{x}[t])$$

3. 軟判定しきい値関数を適用する．

$$\mathbf{x}[t+1] = S_{1/L}\left(\mathbf{v}[t]\right)$$

4. 終了基準を満たすまでステップ 2-4 を繰り返す．

ISTAの別解釈

- 罰金法
  - 古典的制約つき最適化問題の解法

$$\min_{\mathbf{x}} \left\{ \|\mathbf{x}\|_1 + g(\mathbf{x}) + \frac{\mu}{2} \|\mathbf{x} - \mathbf{x}[t]\|_2^2 \right\}$$

  - 暫定解に近いところで最適解を探索させる.
  - 罰金項の係数μが大きいと、制約条件に忠実

ISTAの別解釈

- 罰金法
  - 古典的制約つき最適化問題の解法

$$\min_{\mathbf{x}} \left\{ \|\mathbf{x}\|_1 + g(\mathbf{x}) + \frac{\mu}{2} \|\mathbf{x} - \mathbf{x}[t]\|_2^2 \right\}$$

  - 暫定解に近いところで最適解を探索させる.
  - 罰金項の係数μが大きいと、制約条件に忠実

- ISTA＝罰金法＋線形近似
  - コスト関数を暫定解付近で展開

$$\min_{\mathbf{x}} \left\{ \|\mathbf{x}\|_1 + g(\mathbf{x}[t]) + (\nabla g(\mathbf{x}[t]))^{\mathrm{T}} (\mathbf{x} - \mathbf{x}[t]) + \frac{\mu}{2} \|\mathbf{x} - \mathbf{x}[t]\|_2^2 \right\}$$

基底追跡の解法

- ノイズなし圧縮センシング
  - 基底追跡

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{y} = A\mathbf{x}$$

  - 制約条件ありの最適化問題
    - ラグランジュ未定乗数

$$\min_{\mathbf{x}} \left\{ \|\mathbf{x}\|_1 + (\mathbf{h}[t])^{\mathrm{T}} (\mathbf{y} - A\mathbf{x}) \right\}$$

      - 最小化のときは正の乗数、最大化のときは負の乗数
      - 未定乗数の更新と最適化にかかる変数の更新でやや不安定
    - 罰金法

$$\min_{\mathbf{x}} \left\{ \|\mathbf{x}\|_1 + \frac{\mu}{2} \|\mathbf{y} - A\mathbf{x}\|_2^2 \right\}$$

      - 罰金係数を徐々に大きくしていかなければならない

基底追跡の解法

- 拡張ラグランジュ法
  - ラグランジュ未定乗数と罰金法の組み合わせ

$$\min_{\mathbf{x}} \left\{ \|\mathbf{x}\|_1 + \left(\mathbf{h}^{\mathrm{T}}\right)(\mathbf{y} - A\mathbf{x}) + \frac{\mu}{2}\|\mathbf{y} - A\mathbf{x}\|_2^2 \right\}$$

  - ラグランジュ未定乗数の更新則

$$\mathbf{h}[t+1] = \mathbf{h}[t] + \mu[t](\mathbf{y} - A\mathbf{x})$$

    - 制約条件の式による変分を更新に利用
    - 罰金係数をそこまで大きくしなくてもよい
    - 未定乗数の計算が安定化する

基底追跡の解法

- 拡張ラグランジュ法
  - ラグランジュ未定乗数と罰金法の組み合わせ

$$\min_{\mathbf{x}} \left\{ \|\mathbf{x}\|_1 + \left(\mathbf{h}^{\mathrm{T}}\right)(\mathbf{y} - A\mathbf{x}) + \frac{\mu}{2}\|\mathbf{y} - A\mathbf{x}\|_2^2 \right\}$$

  - ラグランジュ未定乗数の更新則

$$\mathbf{h}[t+1] = \mathbf{h}[t] + \mu[t](\mathbf{y} - A\mathbf{x})$$

    - 制約条件の式による変分を更新に利用
    - 罰金係数をそこまで大きくしなくてもよい
    - 未定乗数の計算が安定化する
- Bregman反復法と等価である

最近の進展

- ADMM(Alternating Direction Method of Multipliers)
  - ふたつのコスト関数の最適化問題を解く方法

$$\min_{\mathbf{x}} \{f(\mathbf{x}) + g(\mathbf{x})\}$$

  - 制約条件ありの最適化問題へ敢えて変更する

$$\min_{\mathbf{x}} \{f(\mathbf{z}) + g(\mathbf{x})\} \quad \text{s.t.} \quad \mathbf{x} - \mathbf{z} = \mathbf{0}$$

  - LASSOであれば、

$$f(\mathbf{x}) = \|\mathbf{x}\|_1 \quad g(\mathbf{x}) = \frac{1}{2\lambda} \|\mathbf{y} - A\mathbf{x}\|_2^2$$

ADMMの実際

- 拡張ラグランジュ法をちょっと適当にやる

$$\min_{\mathbf{x},\mathbf{z}} \left\{ f(\mathbf{z}) + g(\mathbf{x}) + \left(\mathbf{h}^{\mathrm{T}}\right)(\mathbf{x}-\mathbf{z}) + \frac{\mu}{2}\|\mathbf{x}-\mathbf{z}\|_2^2 \right\}$$

$$\mathbf{h}[t+1] = \mathbf{h}[t] + \mu(\mathbf{x}-\mathbf{z})$$

  - 拡張ラグランジュ法との違い
    - 罰金係数は適当な値で固定
    - ふたつの変数についての最適化を交互にやる

$$\mathbf{x}[t+1] = \arg\min_{\mathbf{x}} \left\{ f(\mathbf{z}[t]) + g(\mathbf{x}) + (\mathbf{h}[t])^{\mathrm{T}}(\mathbf{x}-\mathbf{z}[t]) + \frac{\mu}{2}\|\mathbf{x}-\mathbf{z}[t]\|_2^2 \right\}$$

$$\mathbf{z}[t+1] = \arg\min_{\mathbf{z}} \left\{ f(\mathbf{z}) + g(\mathbf{x}[t+1]) + (\mathbf{h}[t])^{\mathrm{T}}(\mathbf{x}[t+1]-\mathbf{z}) + \frac{\mu}{2}\|\mathbf{x}[t+1]-\mathbf{z}\|_2^2 \right\}$$

圧縮センシングのまとめ

- 観測過程のモデリング＋推定のアルゴリズム
  - L0ノルム最小化

$$\min_{\mathbf{x}} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad \mathbf{y} = A\mathbf{x}$$

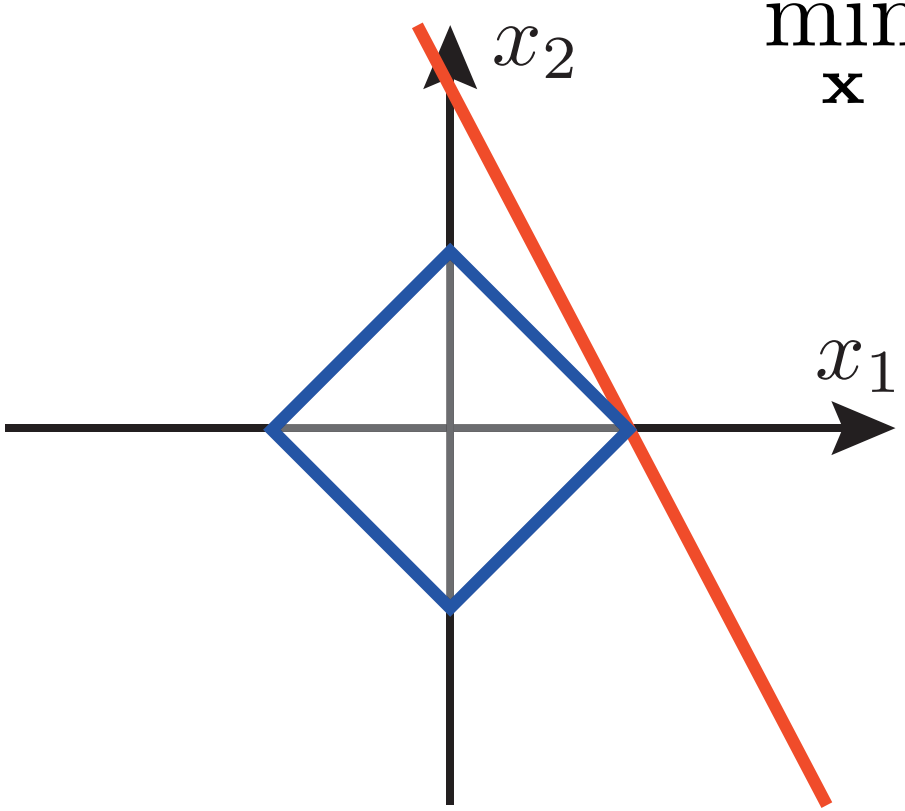
    - 非零成分の個数が小さい（スパース）で方程式を満たすものを探す
    - 残念ながら計算量的に困難（組み合わせ最適化問題）

  - L1ノルム最小化

$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{y} = A\mathbf{x}$$

    - 零が多ければ小さくなる＋大きさも小さくなりがち
    - 計算量は非常に軽い（メジャライザー最小化）

最適化問題と正則化



$$\min_{\mathbf{x}} \|\mathbf{x}\|_1 \quad \text{s.t.} \quad \mathbf{y} = A\mathbf{x}$$

- 制約条件を満たす解は無数
  - 正則化で解を選択
  - 正則化で解をしぼる
  - どんな正則化？＝アルゴリズム