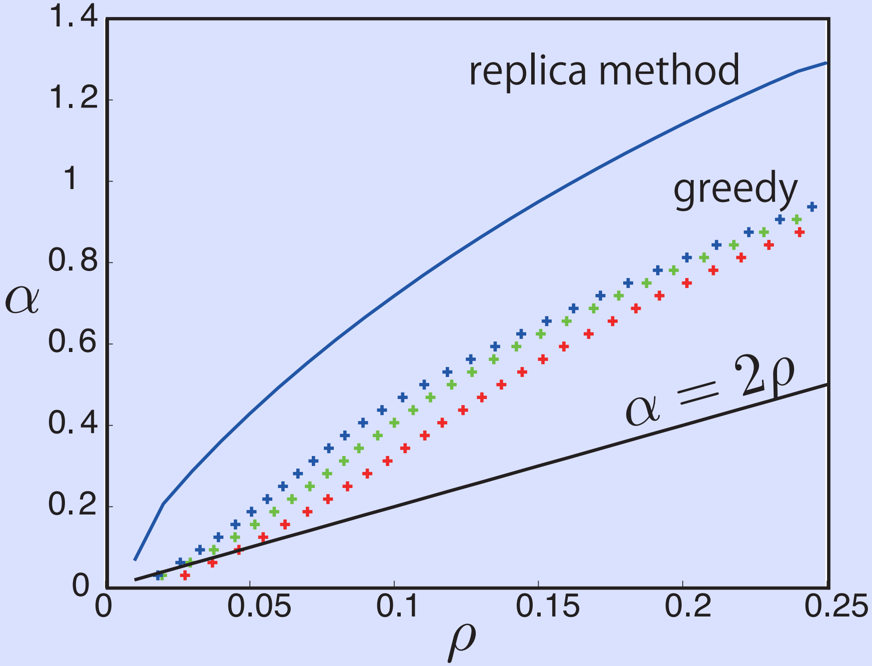


東京工業大学
大学院総合理工学研究科
樺島祥介

東京工業大学
大学院総合理工学研究科
許インイン

日本物理学会2015年秋季大会
関西大学千里山キャンパス19aCQ-6

次元選択の情報統計力学

京都大学大学院情報学研究科　大関　真之

Sparse Modeling

単純パーセプトロンの容量

E. Gardner and B. Derrida: J. Phys. A: Math. Gen. 21 (1988) 271

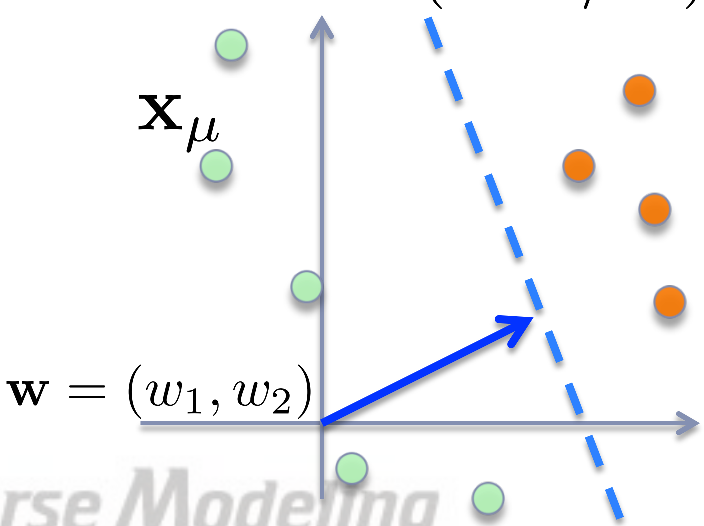
- Gardner-Cover容量
  - 単純パーセプトロン

$$y_\mu = \mathrm{sign}\left(\frac{1}{\sqrt{N}}\mathbf{x}_\mu^{\mathrm{T}}\mathbf{w}\right)$$

- ラベル（出力）：$y_\mu \in \{-1, +1\}$ 入力：$\mathbf{x}_\mu \in \mathbb{R}^N$ 結合：$\mathbf{w} \in \mathbb{R}^N$

  - P個のランダムな入出力関係とN次元の結合で識別できる限界

$$\alpha\,(\equiv P/N) = 2$$



単純パーセプトロンの容量

E. Gardner and B. Derrida: J. Phys. A: Math. Gen. 21 (1988) 271

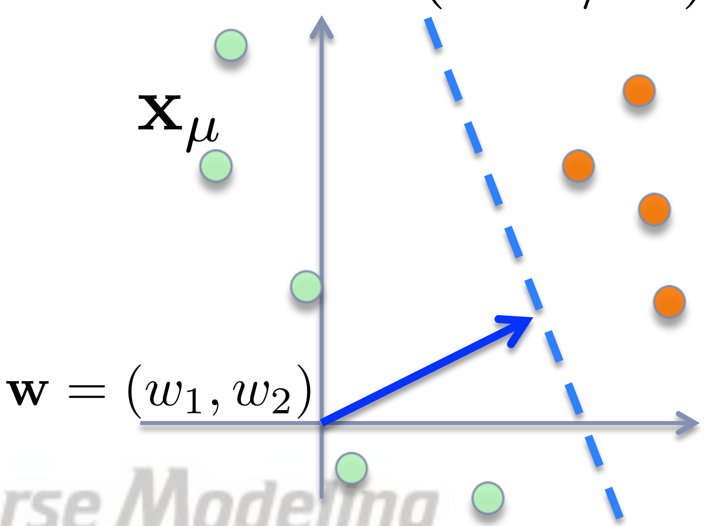
- Gardner-Cover容量
  - 単純パーセプトロン

$$y_\mu = \mathrm{sign}\left(\frac{1}{\sqrt{N}}\mathbf{x}_\mu^{\mathrm{T}}\mathbf{w}\right)$$

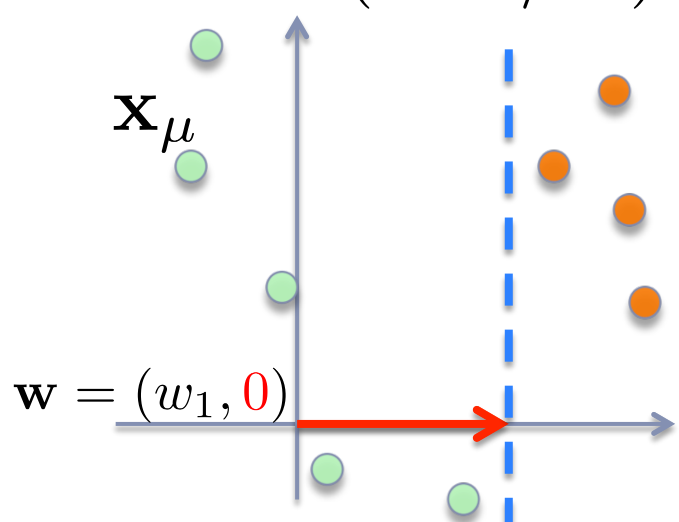
- ラベル（出力）：$y_\mu \in \{-1, +1\}$ 入力：$\mathbf{x}_\mu \in \mathbb{R}^N$ 結合：$\mathbf{w} \in \mathbb{R}^N$

  - P個のランダムな入出力関係とN次元の結合で識別できる限界

$\alpha\,(\equiv P/N) = 2$



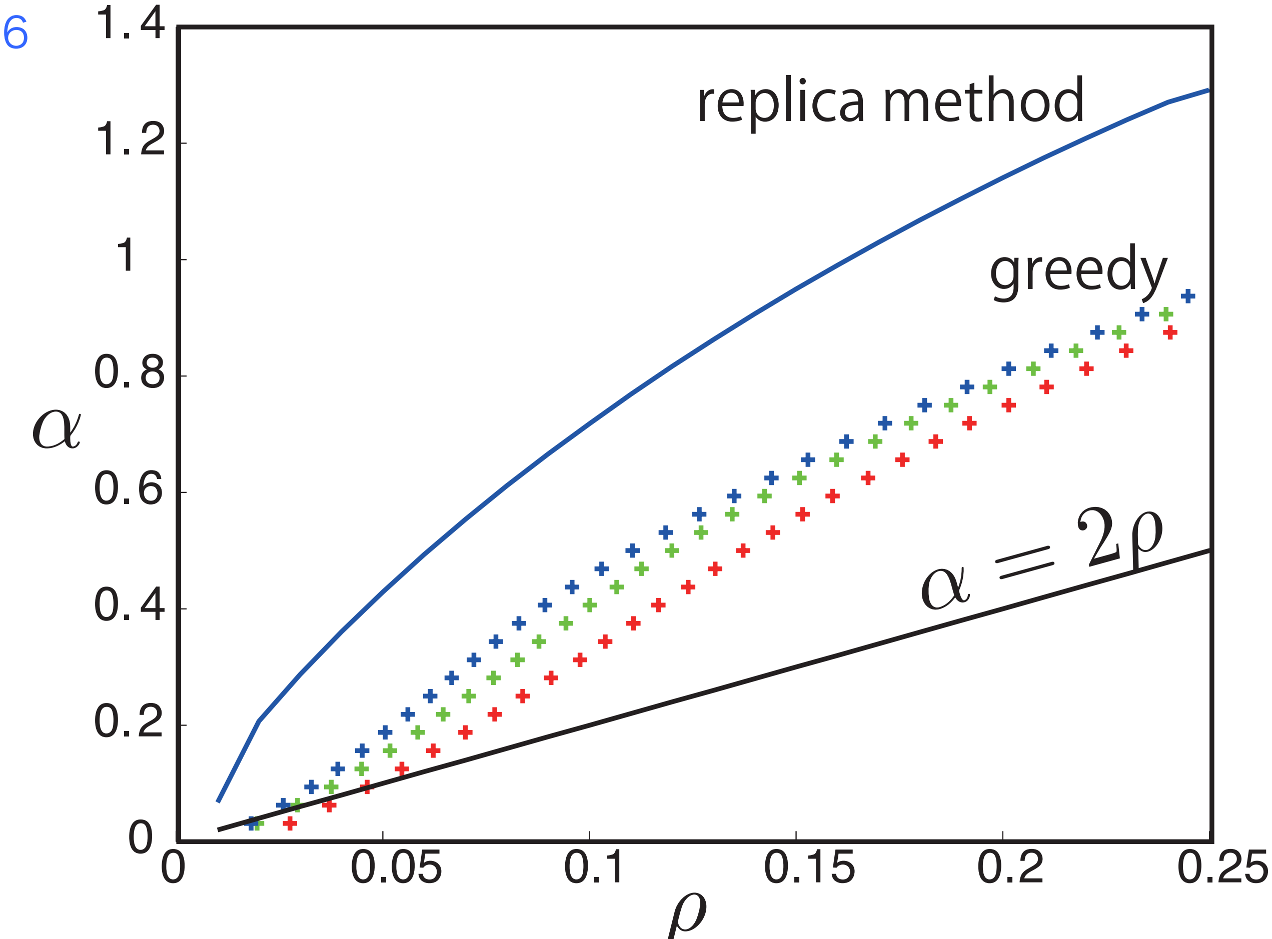
$\alpha\,(\equiv P/N) = 2?$



結果概要

M. Ohzeki, Xu Yingying and Y. Kabashima: to appear soon

- L0 norm （貪欲法）＋ Perceptron（非零割合：ρ ）
  - Gardner容量から増大！ $\alpha_c > 2\rho$
    - 貪欲法Binary Iterative Hard Thresholding
    - N=64、128、256



Sparse Modeling

解析概要 1

M. Ohzeki, Xu Yingying and Y. Kabashima: to appear soon

- どの結合成分をつかうかどうか？

$$P(\mathbf{w}|\mathbf{c}) \propto \delta\left(\sum_{i=1}^{N} c_i w_i^2 - N\rho\right) e^{-\frac{1-c_i}{2} w_i^2}$$

  - 非零の割合：ρ
  - 使うかどうか：$c_i \in \{0, 1\}$

- 非零の結合を指定された上でのGardner体積

$$V(\mathbf{c}|\xi^P) = \int d\mathbf{w} P(\mathbf{w}|\mathbf{c}) \prod_{\mu=1}^{P} \Theta\left(\frac{y_\mu}{\sqrt{N}} \sum_{i=1}^{N} c_i w_i x_{i\mu}\right)$$

  - ランダムな入出力関係：$\xi^P = \{(y_1, \mathbf{x}_1), (y_2, \mathbf{x}_2), \cdots, (y_P, \mathbf{x}_P)\}$
  - 指数関数でスケール仮定：$V(\mathbf{c}|\xi^P) \sim \exp(Ns)$
  - 識別可能なcが存在する：$s \sim O(1)$、存在しない：$s \longrightarrow -\infty$

Sparse Modeling

解析概要２

M. Ohzeki, Xu Yingying and Y. Kabashima: to appear soon

- Complexityの定義

$$\Sigma(s|\xi^P) = \frac{1}{N}\log\left[\sum_c \delta\left(Ns - \log V(\mathbf{c}|\xi^P)\right)\delta\left(\sum_{i=1}^N c_i - N\rho\right)\right]$$

- 母関数を経由して、complexityの計算をする.

$$\begin{aligned}\phi(m|\xi^P) &= \frac{1}{N}\log\left[\sum_{\mathbf{c}} V^m(\mathbf{c}|\xi^P)\delta\left(\sum_{i=1}^N c_i - N\rho\right)\right]\\ &\simeq \max_s\left\{ms + \Sigma(s|\xi^P)\right\}\end{aligned}$$

  - ルジャンドル変換

$$\Sigma(s^*|\xi^P) = \phi(m|\xi^P) - m\frac{\partial\phi(m|\xi^P)}{\partial m}$$

- 母関数をレプリカ法を用いて評価する.

Sparse Modeling

考察

M. Ohzeki, Xu Yingying and Y. Kabashima: to appear soon

- 典型的な非零の選び方でも識別可能なとき（ $\alpha < \alpha_{\mathrm{c}}^{\mathrm{typ}} = 2$ ）

$$\Sigma(s^*) = -\rho \log \rho - (1-\rho)\log(1-\rho)$$

Sparse Modeling

考察

M. Ohzeki, Xu Yingying and Y. Kabashima: to appear soon

- 典型的な非零の選び方でも識別可能なとき（ $\alpha < \alpha_{\mathrm{c}}^{\mathrm{typ}} = 2$ ）

$$\Sigma(s^*) \quad = \quad -\rho \log \rho - (1-\rho) \log(1-\rho)$$

- 典型的な非零の選び方では識別不可能なとき（$\alpha > \alpha_{\mathrm{c}}^{\mathrm{typ}} = 2$）
Complexityが連続関数なので

$$\Sigma(s^*) = \max_s \Sigma(s) > 0$$

である限り、$V(\mathbf{c}|\xi^P) > 0 \leftrightarrow s \sim O(1)$ である.

非零の選び方はある

Sparse Modeling

考察

M. Ohzeki, Xu Yingying and Y. Kabashima: to appear soon

- 典型的な非零の選び方でも識別可能なとき（ $\alpha < \alpha_{\rm c}^{\rm typ} = 2$ ）

$$\Sigma(s^*) \quad = \quad -\rho\log\rho - (1-\rho)\log(1-\rho)$$

- 典型的な非零の選び方では識別不可能なとき（$\alpha > \alpha_{\rm c}^{\rm typ} = 2$）
Complexityが連続関数なので

$$\Sigma(s^*) = \max_s \Sigma(s) > 0$$

である限り、$V(\mathbf{c}|\xi^P) > 0 \leftrightarrow s \sim O(1)$ である.

非零の選び方はある

$$\Sigma(s^*) < -\rho\log\rho - (1-\rho)\log(1-\rho)$$

となり、complexityが0となるところが求めたい容量である.

Sparse Modeling

解析概要 3

M. Ohzeki, Xu Yingying and Y. Kabashima: to appear soon

- 母関数（レプリカ対称仮定のもと）

$$
\begin{aligned}
\phi(m) \quad = \quad & \alpha \int Dz \log \left(\int Dy \left(\int Dx \Theta(\sqrt{\rho - q_1} x + \sqrt{q_1 - q_0} y + \sqrt{q_0} z) \right)^m \right) \\
& + \frac{m(\hat{Q} + \hat{q}_1)\rho}{2} - \frac{m^2}{2} (\hat{q}_1 q_1 - \hat{q}_0 q_0) - \\
& + \int Dz \log \left(1 + \frac{e^{-K}}{(\hat{Q} + \hat{q}_1)^{m/2}} \int Dy \exp \left(\frac{m(\sqrt{\hat{q}_1 - \hat{q}_0} y + \sqrt{\hat{q}_0} z)^2}{2(\hat{Q} + \hat{q}_1)} \right) \right) \\
& + K\rho
\end{aligned}
$$

Sparse Modeling

解析概要 3

M. Ohzeki, Xu Yingying and Y. Kabashima: to appear soon

- 母関数（レプリカ対称仮定のもと）

$$\begin{aligned}\phi(m) \quad = \quad & \alpha \int Dz \log \left(\int Dy \left(\int Dx \Theta(\sqrt{\rho - q_1} x + \sqrt{q_1 - q_0} y + \sqrt{q_0} z) \right)^m \right) \\ & + \frac{m(\hat{Q} + \hat{q}_1)\rho}{2} - \frac{m^2}{2} (\hat{q}_1 q_1 - \hat{q}_0 q_0) - \\ & + \int Dz \log \left(1 + \frac{e^{-K}}{(\hat{Q} + \hat{q}_1)^{m/2}} \int Dy \exp \left(\frac{m(\sqrt{\hat{q}_1 - \hat{q}_0} y + \sqrt{\hat{q}_0} z)^2}{2(\hat{Q} + \hat{q}_1)} \right) \right) \\ & + K\rho\end{aligned}$$

- 容量の条件
  - $\Sigma(s|\xi^P)$の最大値（$m \to 0$）=0

（どんなに頑張っても識別可能な結合はない）

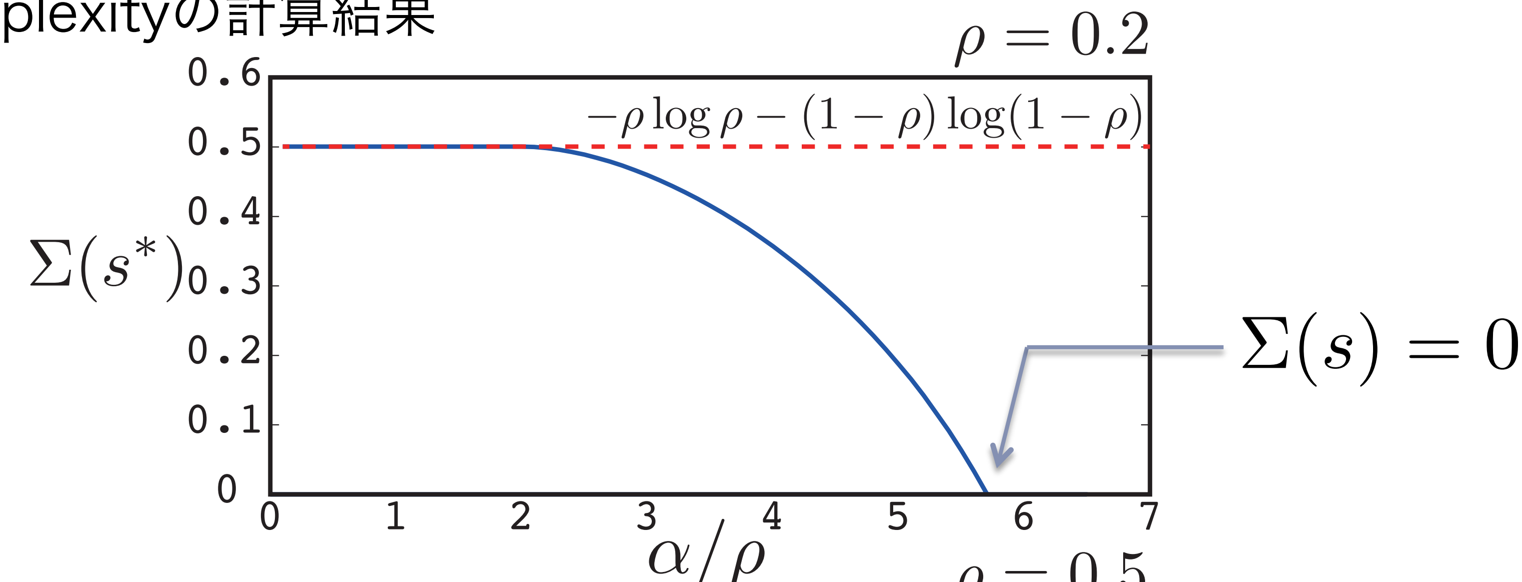
$$\max_s \left\{ \Sigma(s|\xi^P) \right\} = \lim_{m \to 0} \phi(m|\xi^P) = 0$$

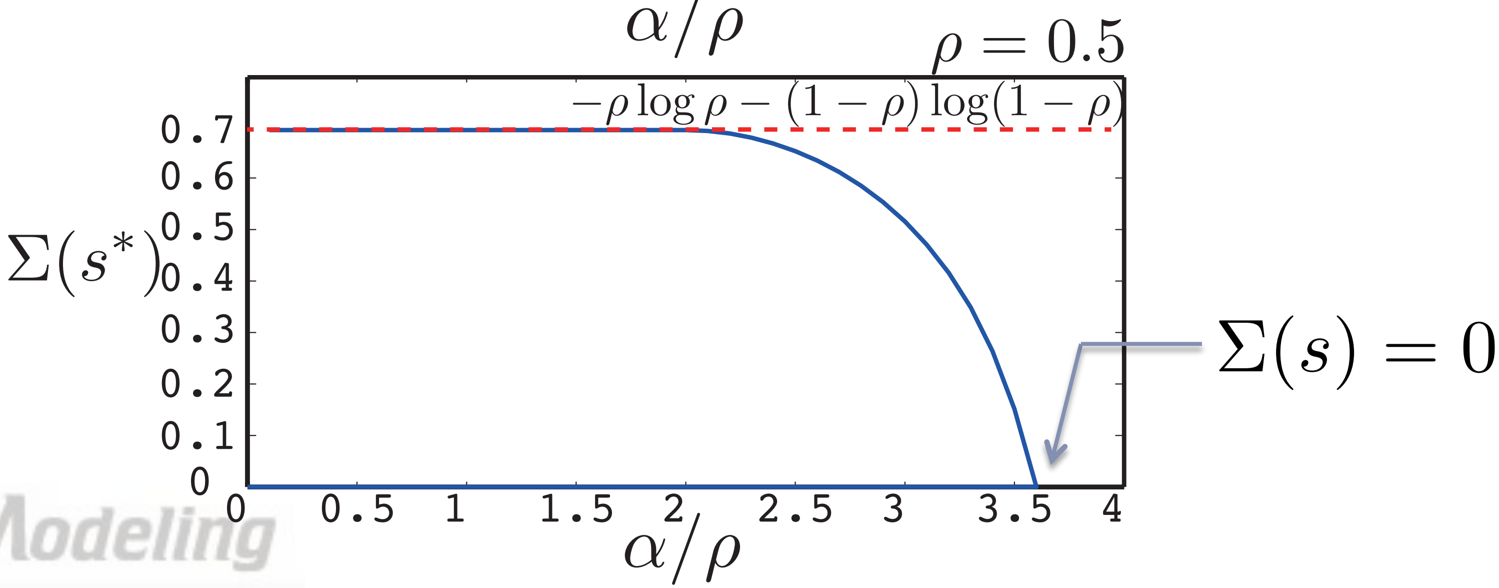
Sparse Modeling

いくつかの事例

M. Ohzeki, Xu Yingying and Y. Kabashima: to appear soon

- Complexityの計算結果

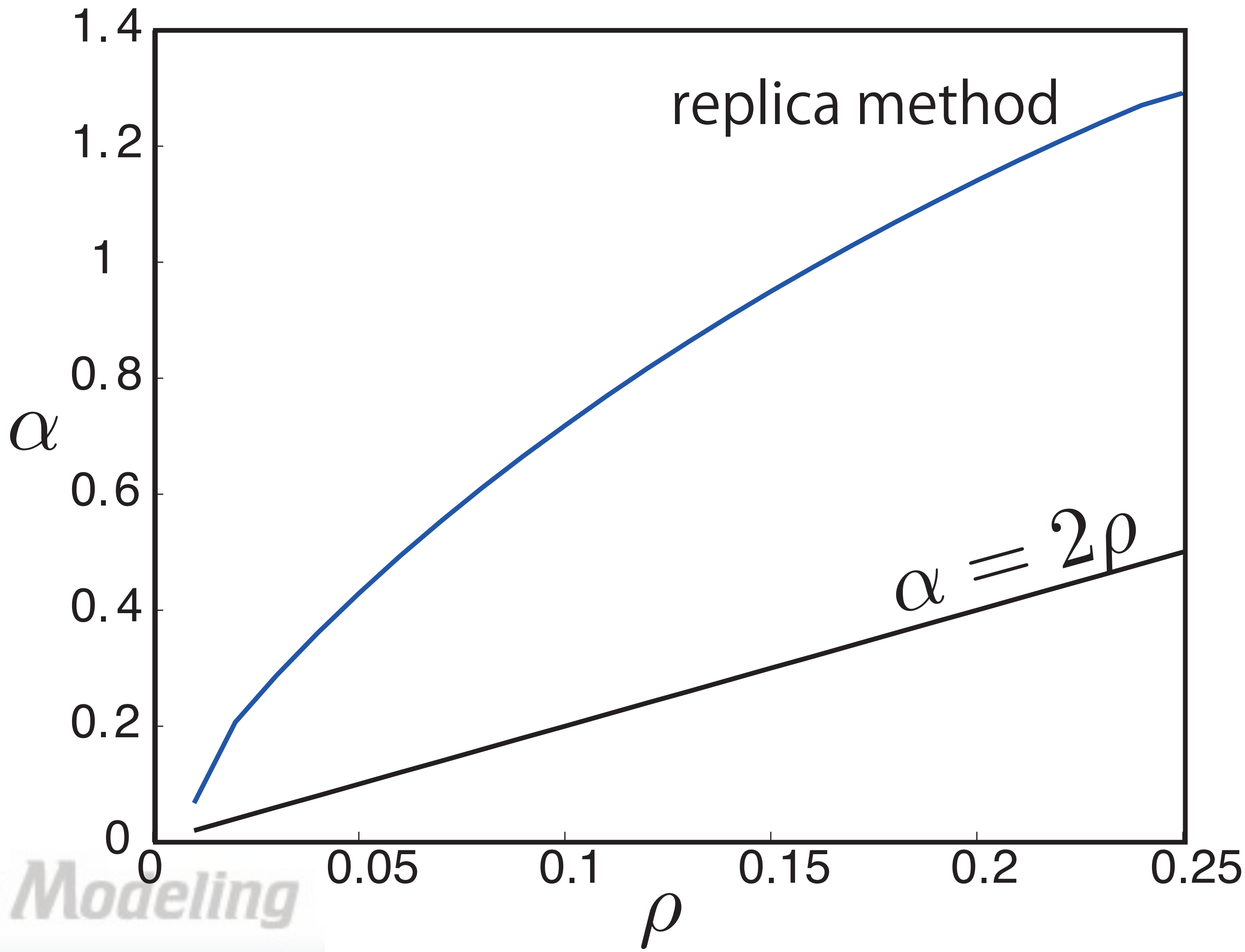




結果

M. Ohzeki, Xu Yingying and Y. Kabashima: to appear soon

- スパース結合による識別可能限界（容量）



数値実験

L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk: IEEE Trans, IT, 59, (2013) 2082.

- 貪欲法：Binary Iterative Hard Thresholding (BIHT)

$$\mathbf{w}[t+1] = H_K\left(\mathbf{w}[t] - \frac{1}{M}\sum_{\mu=1}^{M}\mathbf{x}_\mu\left(\mathrm{sign}\left(\mathbf{x}_\mu^{\mathrm{T}}\mathbf{w}[t]\right) - y_\mu\right)\right)$$

  - 硬判定しきい値関数（上位K成分だけを残す）：$H_K(\mathbf{v})$

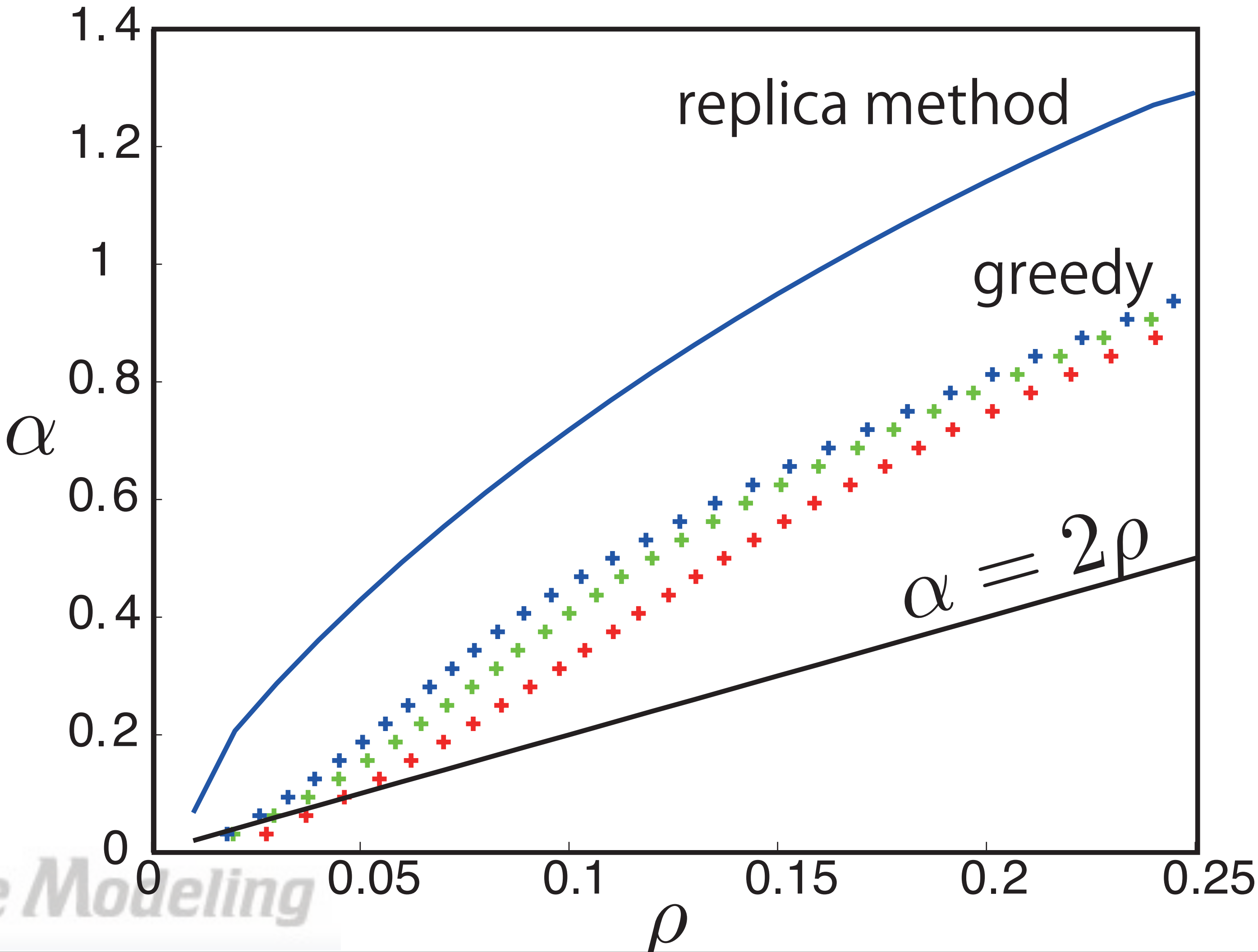
- 手順
  - 1 : K=1から徐々に非零成分を増やす.
  - 2 : BIHTを実行する.
  - 3 : 識別できる入出力関係数Pcを記録する.
- 計算条件
  - サンプル数：Nsam = 10000
  - 次元数：N=64,128,256

Sparse Modeling

数値実験結果

L. Jacques, J. N. Laska, P. T. Boufounos, and R. G. Baraniuk: IEEE Trans, IT, 59, (2013) 2082.

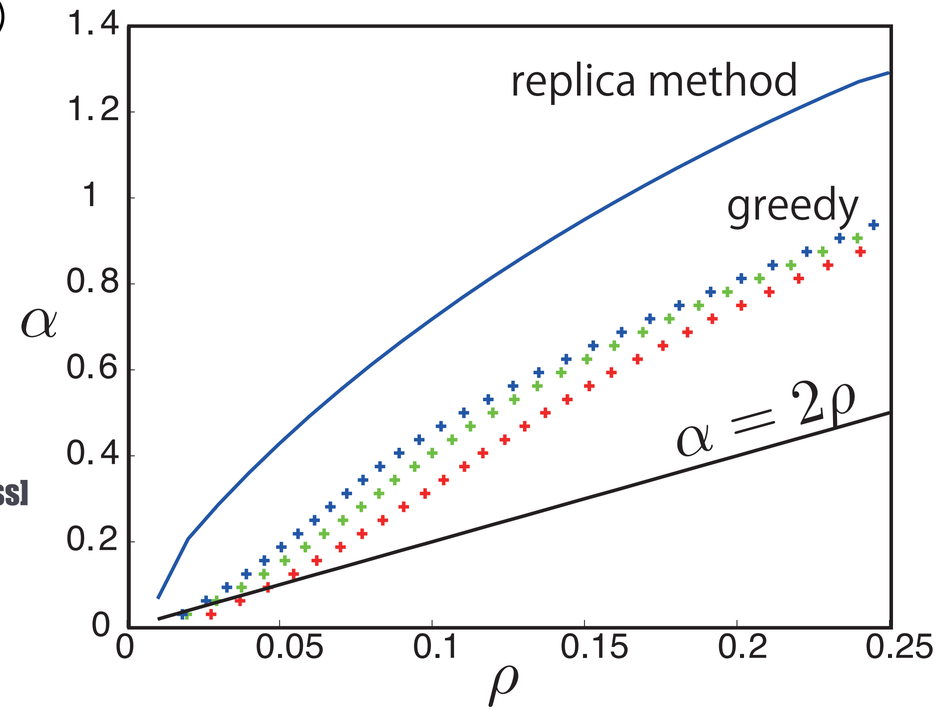
- 次元数：N=64（赤）,128（緑）,256（青）



まとめ

M. Ohzeki, Xu Yingying and Y. Kabashima: to appear soon

- 実はL1 norm + Perceptronでも
  - Gardner容量から変わらず… $\alpha_c = 2\rho$
- L0 norm （貪欲法）+ Perceptron
  - Gardner容量から増大！ $\alpha_c > 2\rho$
    - 貪欲法Binary Iterative Hard Thresholding
    - N=64（赤），128（緑），256（青）



- 貪欲法の良好なアルゴリズム追求
- Ising Perceptronの場合は？
  - [M. Ohzeki, Xu Yingying, and Y. Kabashima: work in progress]

Sparse Modeling