

大阪市立大学電子・物理工学特別講義

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



京都大学
KYOTO UNIVERSITY

情報統計力学編



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）



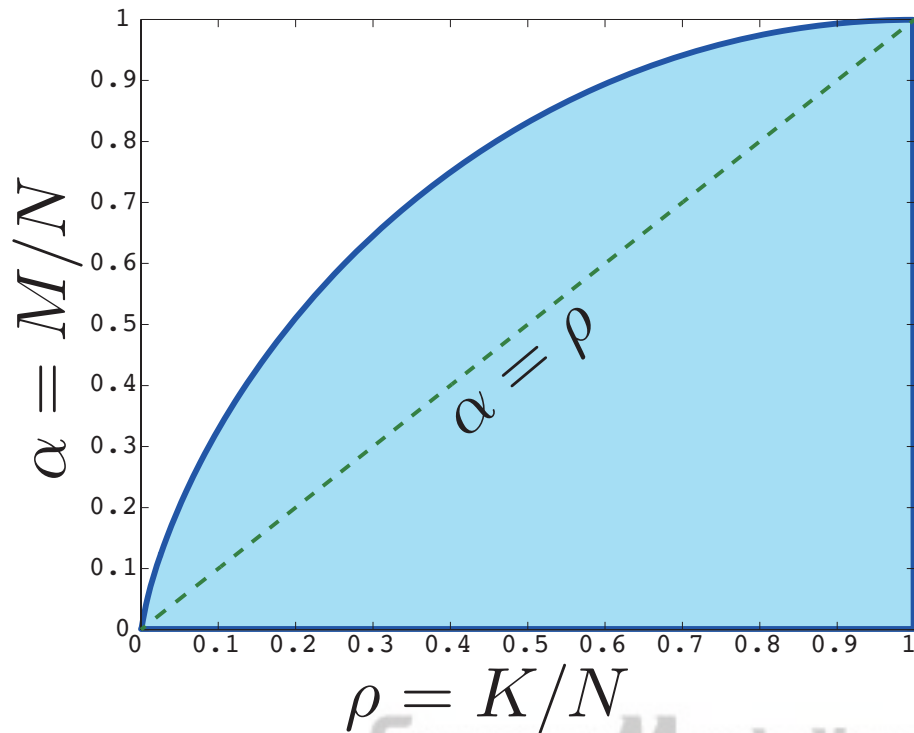
- ▶ **L1** ノルム最小化によるスパース解推定法
 - ▶ 基底追跡 (Basis Pursuit)

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

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

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

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



▶ 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

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

- ▶ 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



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

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

- ▶ 等式制約はもはや成立しない

- ▶ LASSO (Least Absolute Shrinkage and Selection Operators)

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

- ▶ ラグランジュ未定乗数法により等価な最適化問題

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





統計力学的アプローチの出発点

- ▶ ミクロな世界とマクロな世界を結ぶ処方箋
 - ▶ ハミルトニアン（エネルギー関数：ミクロ世界のルール）
 - ▶ 例：調和振動子（バネの運動）

$$H(\mathbf{x}, \mathbf{p}) = \frac{1}{2}\mathbf{p}^2 + \frac{1}{2}\mathbf{x}^2$$

- ▶ 例：Ising模型（磁性体のモデル）

$$H(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T J \mathbf{x}$$



統計力学的アプローチの出発点

- ▶ ミクロな世界とマクロな世界を結ぶ処方箋
 - ▶ ハミルトニアン（エネルギー関数：ミクロ世界のルール）
 - ▶ 例：調和振動子（バネの運動）

$$H(\mathbf{x}, \mathbf{p}) = \frac{1}{2}\mathbf{p}^2 + \frac{1}{2}\mathbf{x}^2$$

- ▶ 例：Ising模型（磁性体のモデル）

$$H(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T J \mathbf{x}$$

- ▶ 確率分布による記述

- ▶ Gibbs-Boltzmann分布

$$P(\mathbf{x}) = \frac{1}{Z} \exp\left(-\frac{H(\mathbf{x})}{T}\right)$$

- ▶ 分配関数

$$Z = \int d\mathbf{x} \exp\left(-\frac{H(\mathbf{x})}{T}\right)$$





統計力学的アプローチの中間層

- ▶ 分配関数の計算命!
 - ▶ 等エネルギー面にある状態の数を先に勘定する.

$$Z = \int dE \int d\mathbf{x} \delta(H(\mathbf{x}) - E) \exp\left(-\frac{E}{T}\right)$$



統計力学的アプローチの中間層

- ▶ 分配関数の計算命!
 - ▶ 等エネルギー面にある状態の数を先に勘定する.

$$Z = \int dE \int d\mathbf{x} \delta(H(\mathbf{x}) - E) \exp\left(-\frac{E}{T}\right)$$

- ▶ エントロピーの登場 $S(E) = \log \int d\mathbf{x} \delta(H(\mathbf{x}) - E)$

$$Z = \int dE \exp(S(E) - E/T)$$



- ▶ 分配関数の計算命!
 - ▶ 等エネルギー面にある状態の数を先に勘定する.

$$Z = \int dE \int d\mathbf{x} \delta(H(\mathbf{x}) - E) \exp\left(-\frac{E}{T}\right)$$

- ▶ エントロピーの登場 $S(E) = \log \int d\mathbf{x} \delta(H(\mathbf{x}) - E)$

$$Z = \int dE \exp(S(E) - E/T)$$

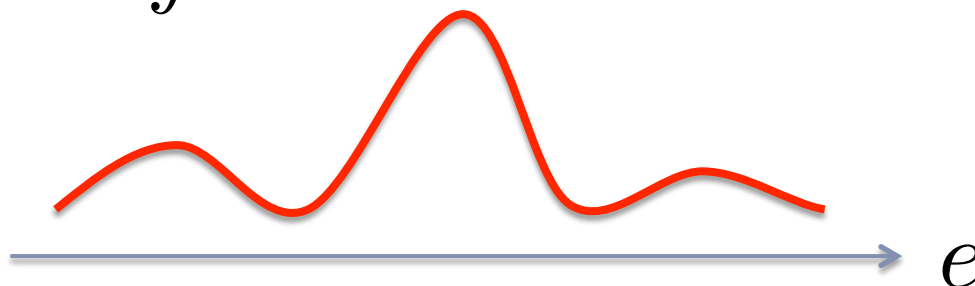
- ▶ とてつもない自由度の数であるということを思い出す. (熱力学的極限)

$$\mathbf{x} \in \mathbb{R}^N \quad N \rightarrow \infty$$



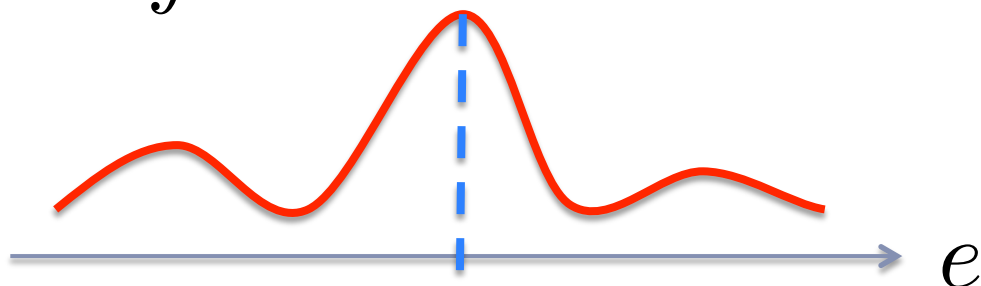
- ▶ 一気にマクロの世界へ!
- ▶ 積分の意味：全領域の足し上げ!

$$Z = N \int de \exp \{ N (s(e) - e/T) \}$$

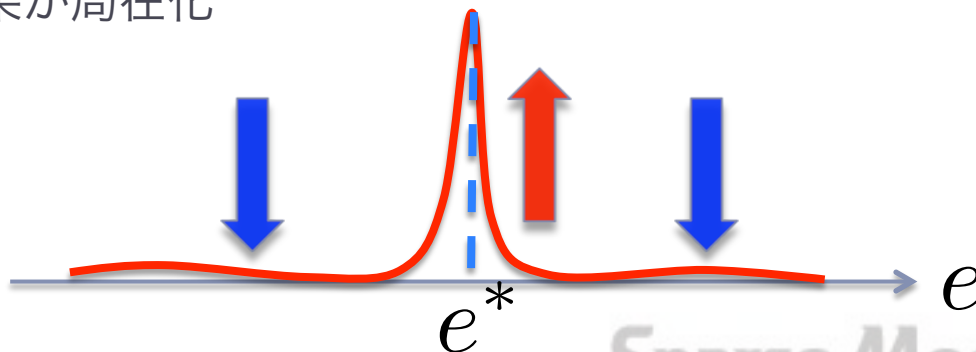


- ▶ 一気にマクロの世界へ!
- ▶ 積分の意味：全領域の足し上げ!

$$Z = N \int de \exp \{ N (s(e) - e/T) \}$$



- ▶ Nを大きくすると…効果が局在化





統計力学的アプローチの終着点

▶ 一気にマクロの世界へ!

▶ 鞍点評価

$$Z \sim \exp \{ N (s(e^*) - e^*/T) \}$$

▶ 自由エネルギーの登場

$$f = -\frac{T}{N} \log Z$$





統計力学的アプローチの終着点

- ▶ 一気にマクロの世界へ!

- ▶ 鞍点評価

$$Z \sim \exp \{ N (s(e^*) - e^*/T) \}$$

- ▶ 自由エネルギーの登場 $f = -\frac{T}{N} \log Z$

- ▶ 分配関数の評価を通じて、全ての必要な熱力学量が計算できる!

$$f = e^* - Ts(e^*)$$

「物質は、エネルギーが低く、エントロピーの大きい、平衡状態を実現する！」

$$f = \min_e \{ e - Ts(e) \}$$



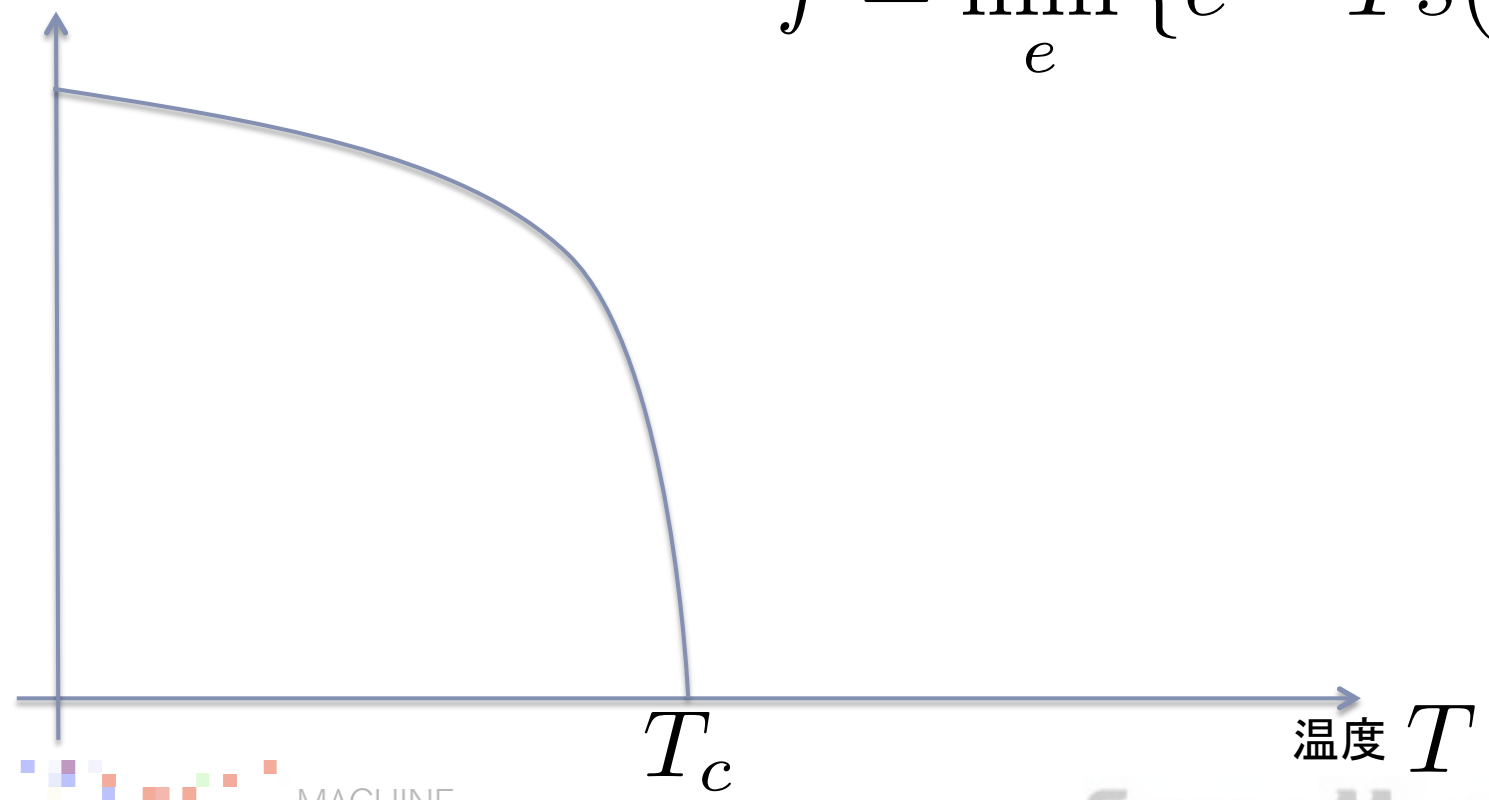


自由エネルギーの形状と相転移

▶ 磁性体の場合：Ising模型 $H(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T J \mathbf{x}$

$$f = \min_e \{e - Ts(e)\}$$

磁化 m



QUANTUM ANNEALING



MACHINE LEARNING

Sparse Modeling

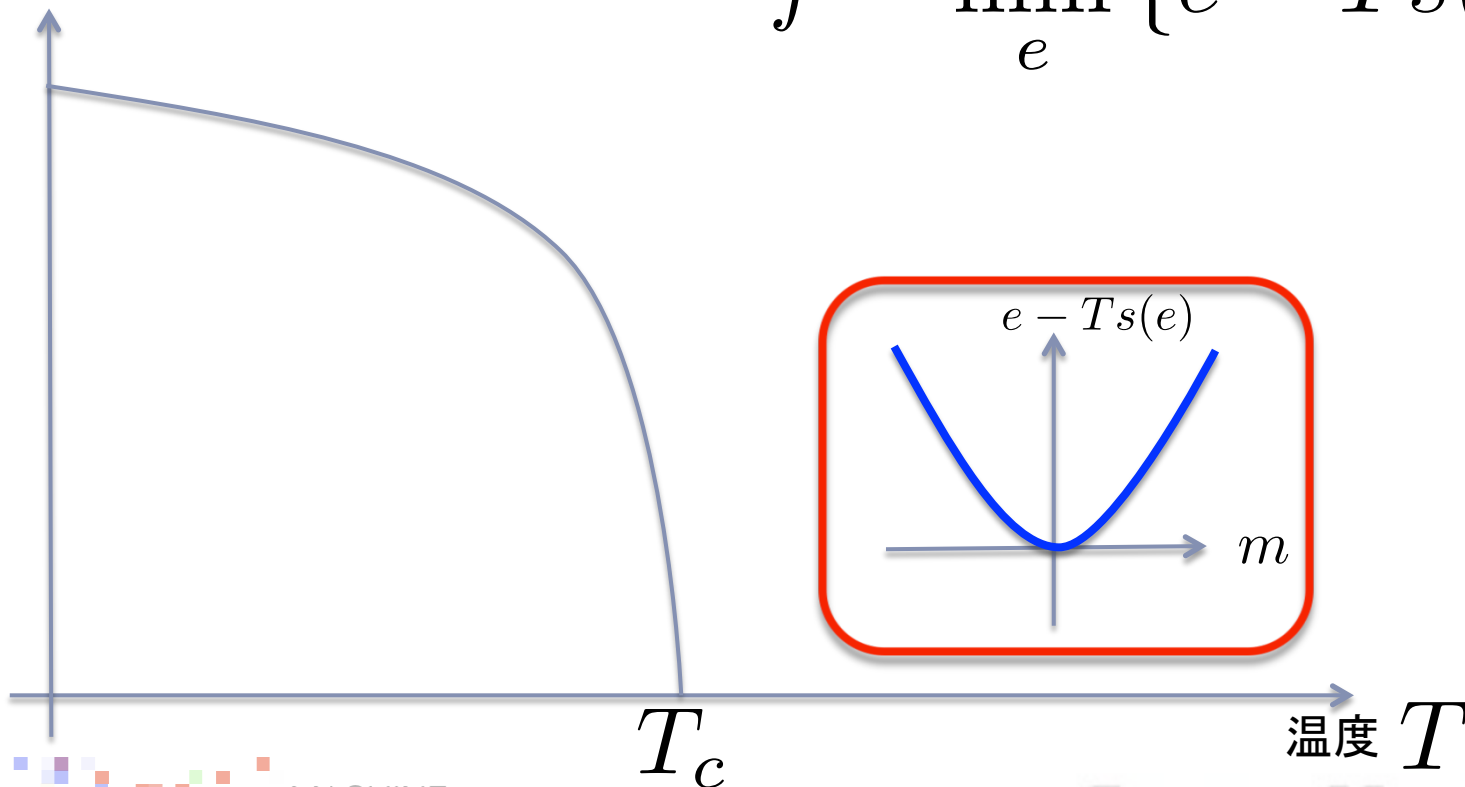


自由エネルギーの形状と相転移

▶ 磁性体の場合：Ising模型 $H(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T J \mathbf{x}$

$$f = \min_e \{e - Ts(e)\}$$

磁化 m



QUANTUM ANNEALING



MACHINE LEARNING

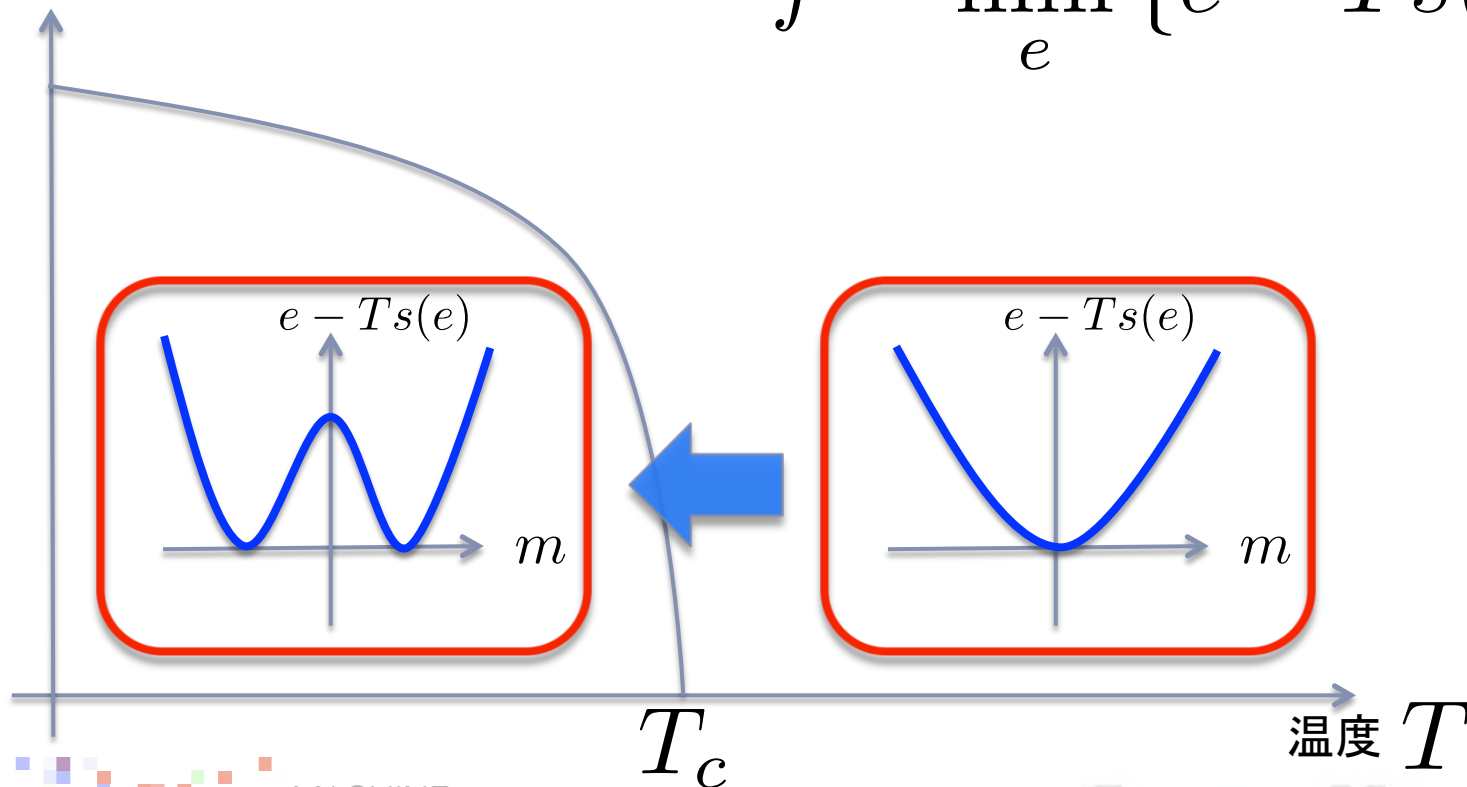
Sparse Modeling

自由エネルギーの形状と相転移

▶ 磁性体の場合：Ising模型 $H(\mathbf{x}) = \frac{1}{2} \mathbf{x}^T J \mathbf{x}$

$$f = \min_e \{e - Ts(e)\}$$

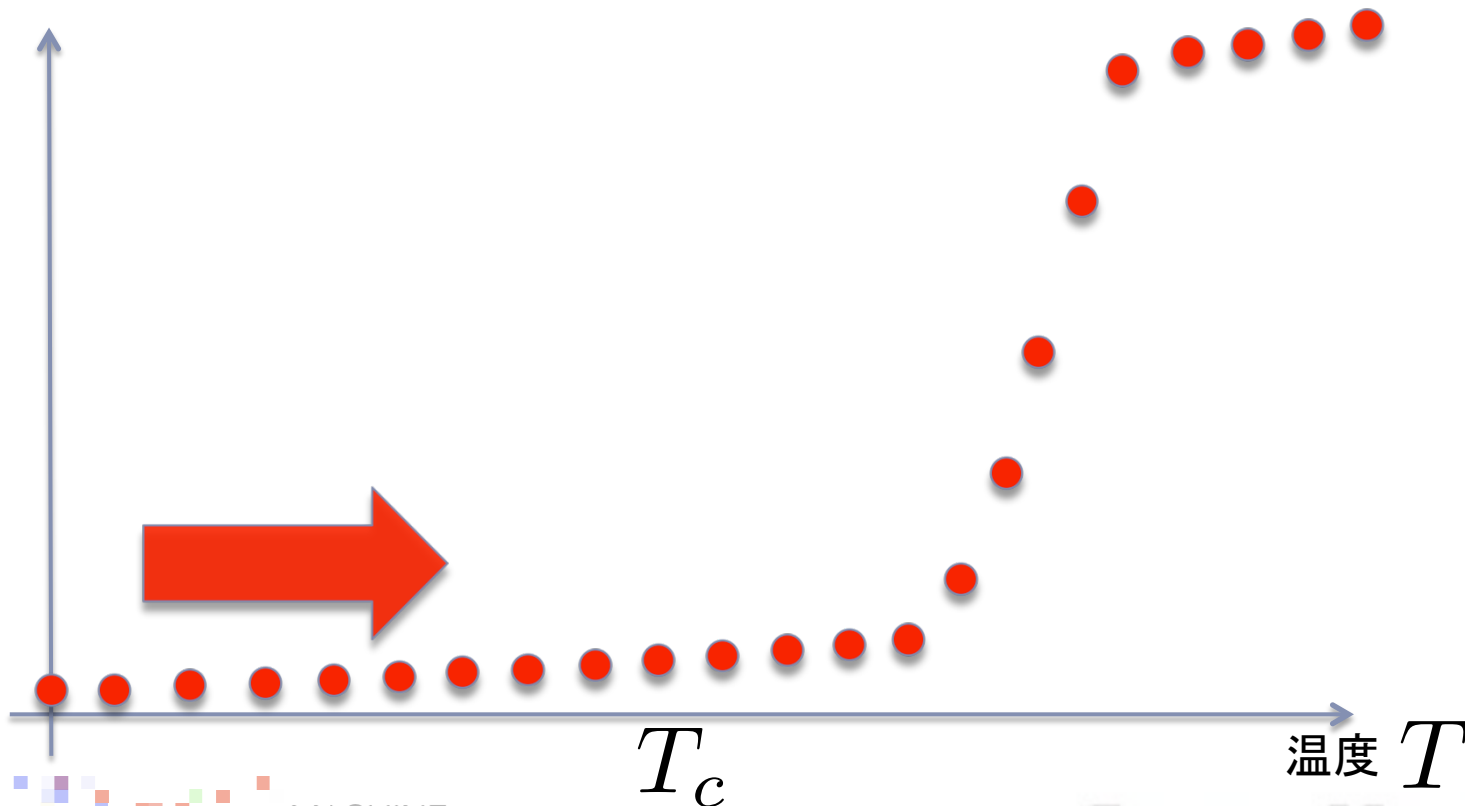
磁化 m



自由エネルギーの形状と相転移

▶ 変わった磁性体の場合：Potts模型 $f = \min_e \{e - Ts(e)\}$

内部エネルギー e



QUANTUM ANNEALING



MACHINE LEARNING

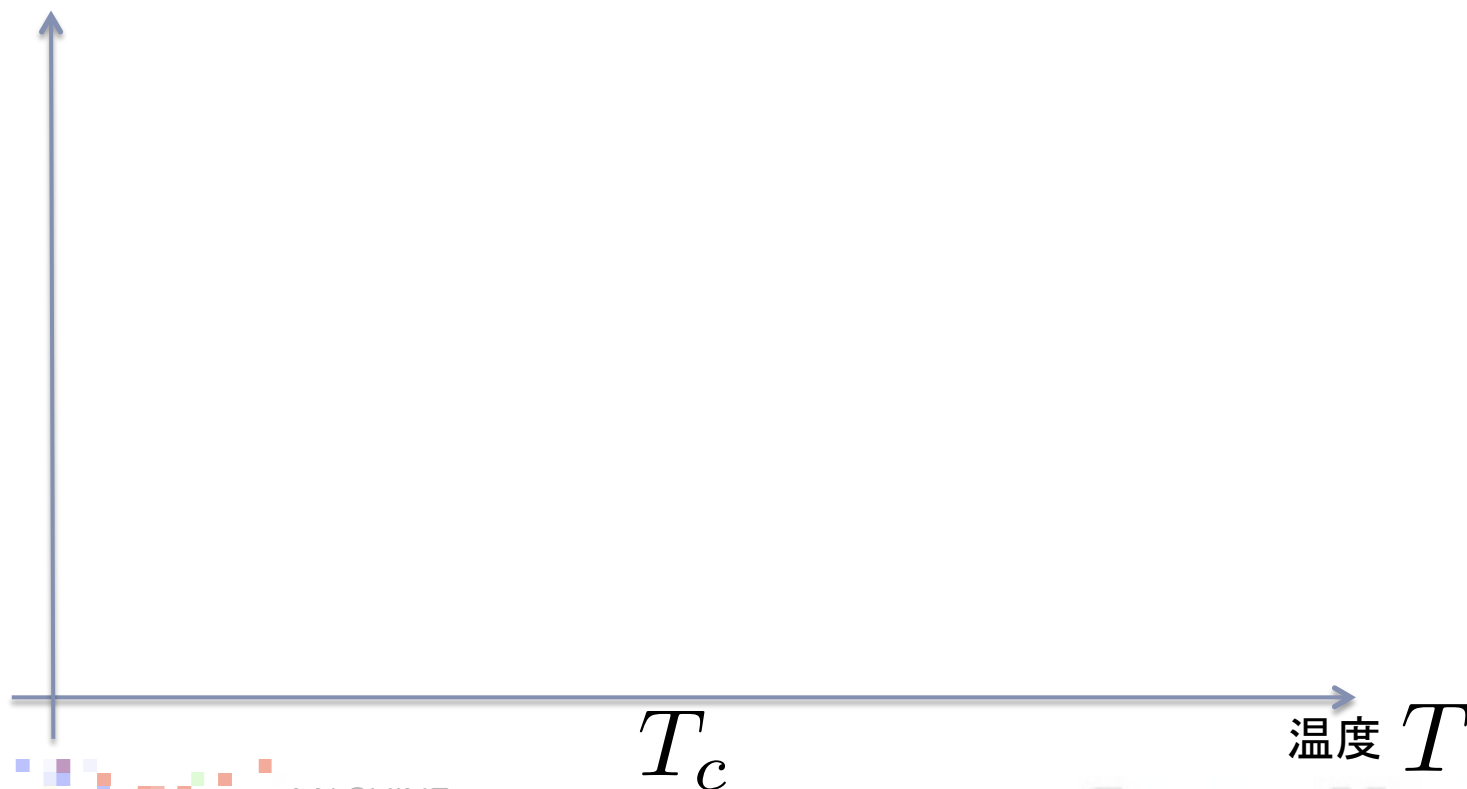
Sparse Modeling



自由エネルギーの形状と相転移

- ▶ 変わった磁性体の場合：Potts模型 $f = \min_e \{e - Ts(e)\}$

内部エネルギー e



QUANTUM ANNEALING

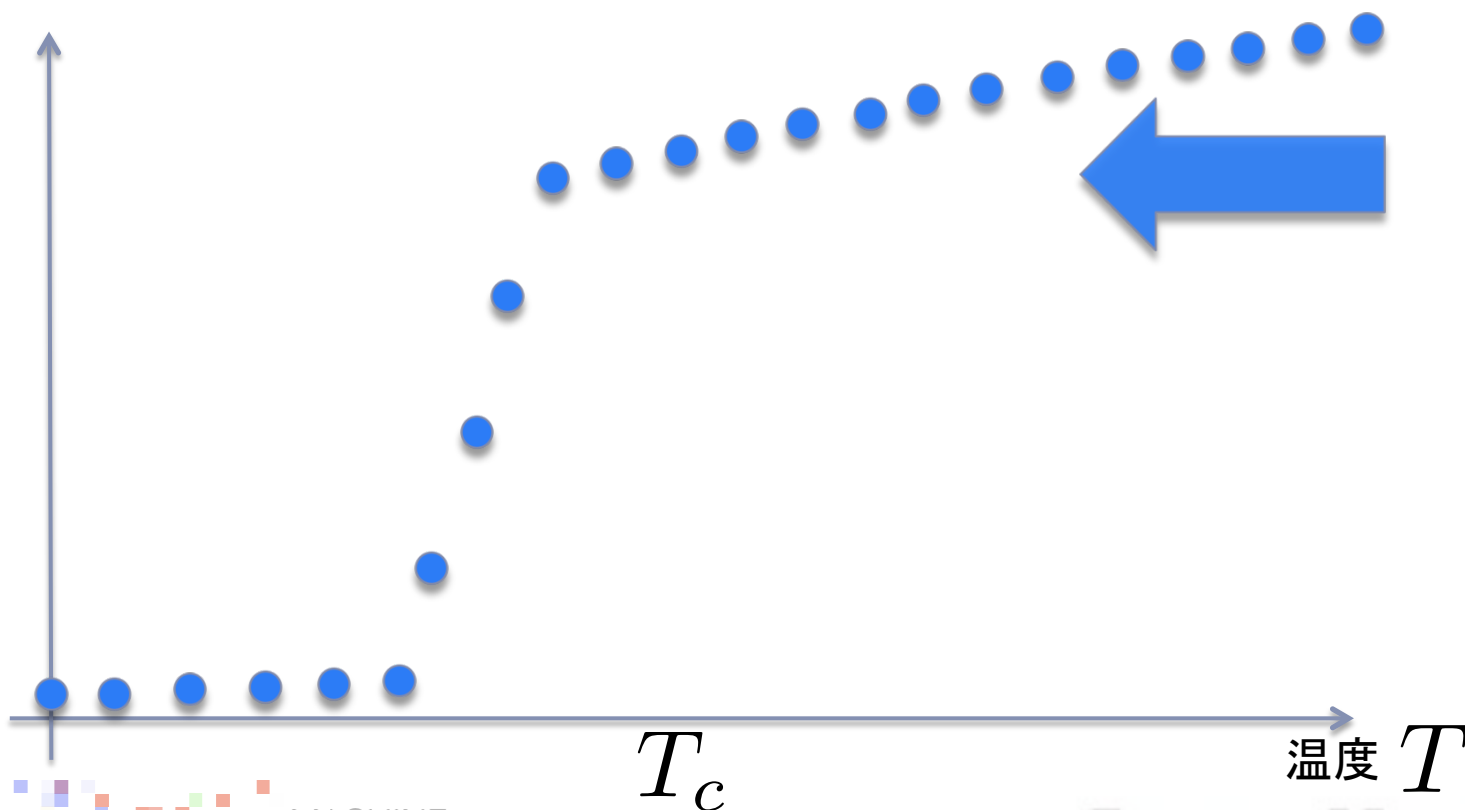
MACHINE LEARNING

Sparse Modeling

自由エネルギーの形状と相転移

- ▶ 変わった磁性体の場合：Potts模型 $f = \min_e \{e - Ts(e)\}$

内部エネルギー e



QUANTUM ANNEALING



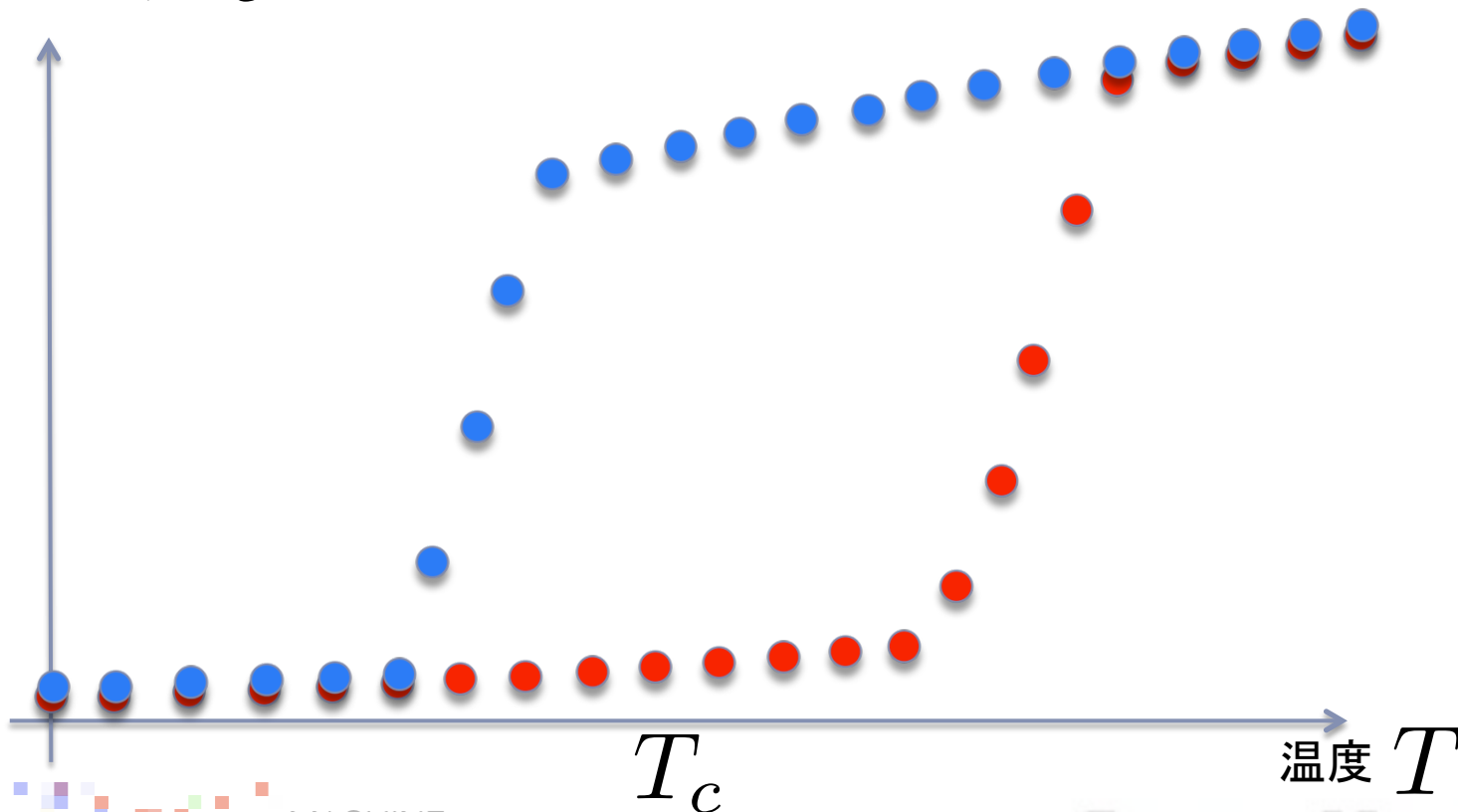
MACHINE LEARNING

Sparse Modeling

自由エネルギーの形状と相転移

- ▶ 変わった磁性体の場合：Potts模型 $f = \min_e \{e - Ts(e)\}$

内部エネルギー e



QUANTUM ANNEALING



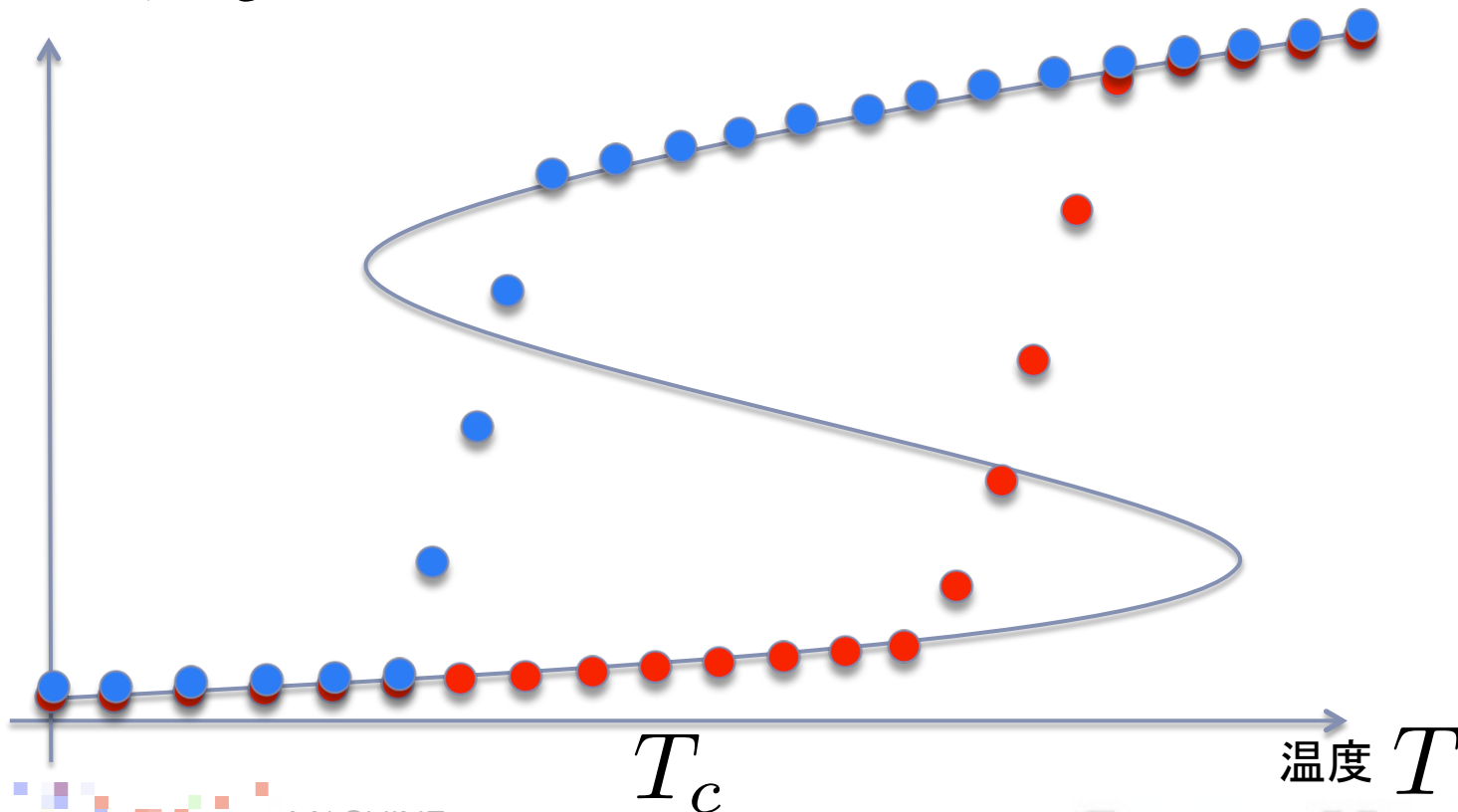
MACHINE LEARNING

Sparse Modeling

自由エネルギーの形状と相転移

- ▶ 変わった磁性体の場合：Potts模型 $f = \min_e \{e - Ts(e)\}$

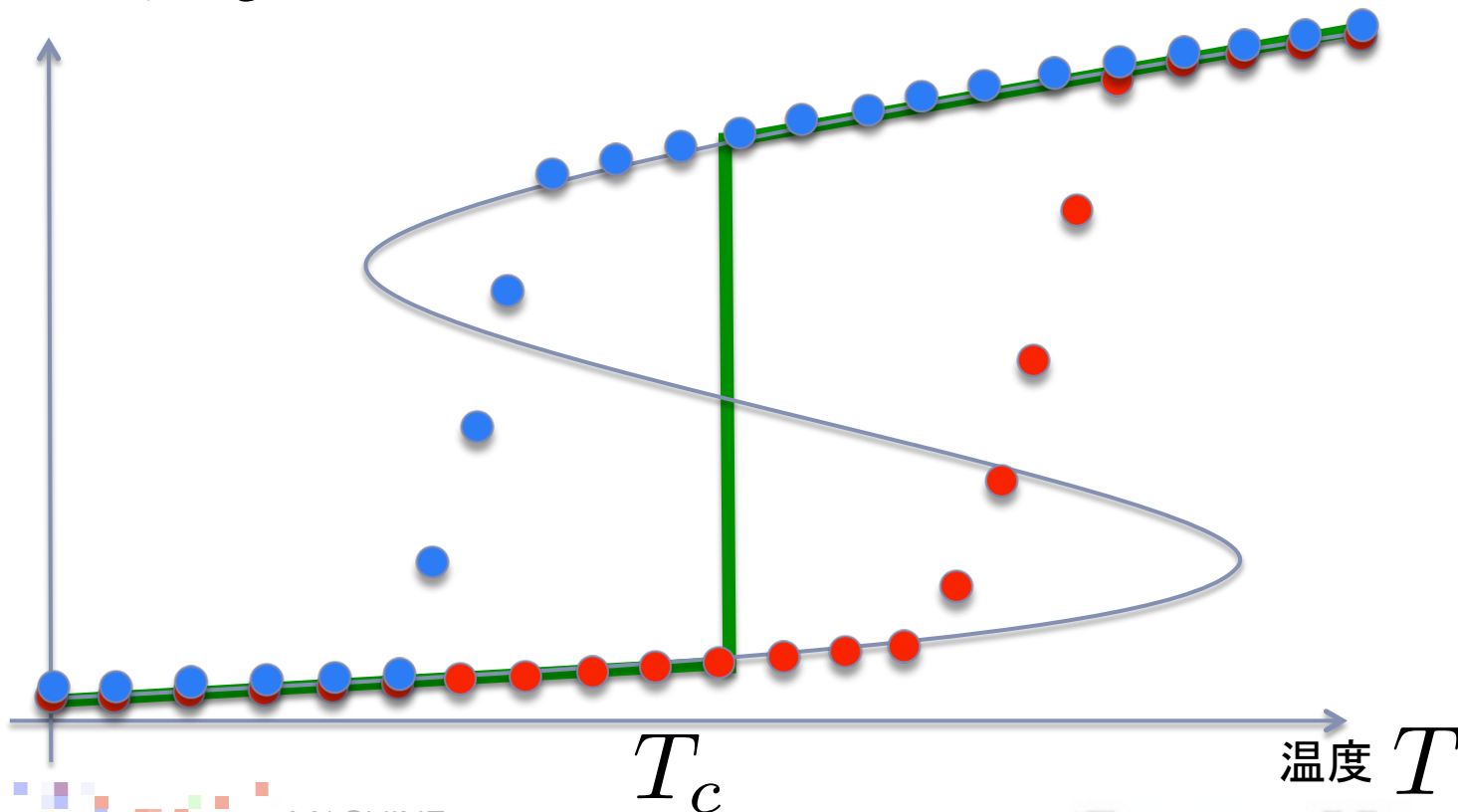
内部エネルギー e



自由エネルギーの形状と相転移

- ▶ 変わった磁性体の場合：Potts模型 $f = \min_e \{e - Ts(e)\}$

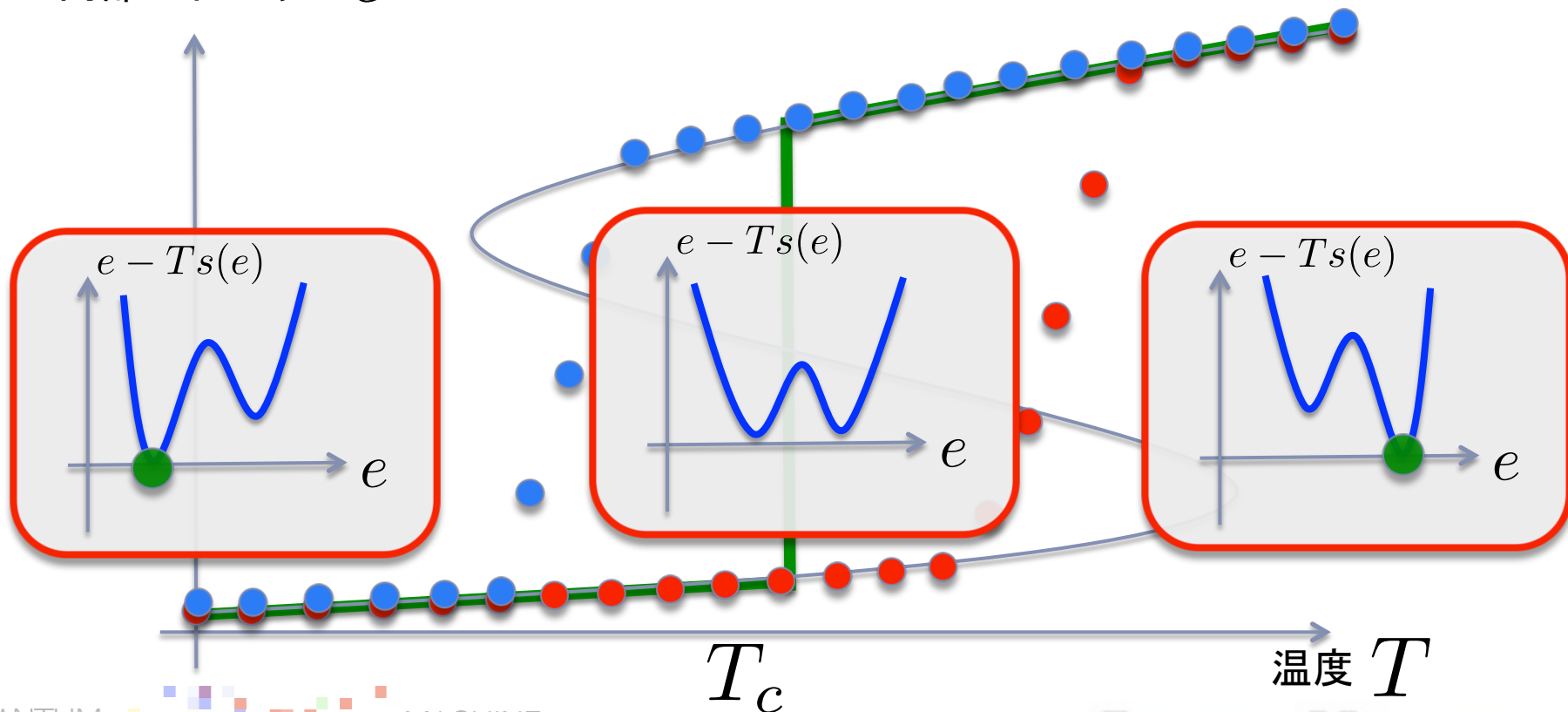
内部エネルギー e



自由エネルギーの形状と相転移

- ▶ 変わった磁性体の場合：Potts模型 $f = \min_e \{e - Ts(e)\}$

内部エネルギー e



QUANTUM ANNEALING



MACHINE LEARNING

Sparse Modeling

- ▶ 基底追跡型の最適化問題の解析
 - ▶ どんな \mathbf{x} が生成されるか？

$$P(\mathbf{x}|A, \mathbf{y}) \propto \exp(-\beta \|\mathbf{x}\|_1) \delta(\mathbf{y} - A\mathbf{x})$$

- ▶ 逆温度 β 無限大で、

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

- ▶ 様々な A と \mathbf{y} での評価を必要とするので、レプリカ法を利用する。

$$[\log Z(A, \mathbf{y})]_{A, \mathbf{y}} = \lim_{n \rightarrow 0} \frac{[Z^n(A, \mathbf{y})]_{A, \mathbf{y}} - 1}{n}$$





罰金法の利用

▶ 分配関数の評価

▶ nレプリカの系における多重積分の評価

$$[Z^n(A, \mathbf{y})]_{A, \mathbf{y}} = \left[\int \prod_{a=1}^n d\mathbf{x}_a \exp(-\beta \|\mathbf{x}_a\|_1) \delta(\mathbf{y} - A\mathbf{x}^a) \right]_{A, \mathbf{y}}$$

▶ 罰金法

$$\delta(\mathbf{y} - A\mathbf{x}^a) = \exp\left(-\frac{1}{2\lambda} \|\mathbf{y} - A\mathbf{x}\|_2^2\right) \quad (\lambda \rightarrow +0)$$

▶ 仮定

- ノイズなし観測過程 $\mathbf{y} = A\mathbf{x}^0$
- Aは各成分ガウスランダム行列 (平均0、分散1/N)
- \mathbf{x}^0 は各成分

$$P_0(x) = (1 - \rho)\delta(x) + \rho \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}x^2\right)$$



ひたすらガウス積分

▶ 自由エネルギー (nの1次部分)

▶ レプリカ対称性の下、秩序パラメータ Q, χ, m と共役パラメータの鞍点評価

$$-f = \max_{Q, \tilde{Q}} \left\{ -\frac{\alpha}{2\chi} (\rho - 2m + Q) + \frac{1}{2} (Q\tilde{Q} - \chi\tilde{\chi}) - m\tilde{m} \right. \\ \left. + \frac{1-\rho}{\tilde{Q}} G(\tilde{\chi}) + \frac{\rho}{\tilde{Q}} G(\tilde{\chi} + \tilde{m}^2) \right\}$$

▶ 関数の定義

$$G(a) = (a + 1) H \left(\frac{1}{\sqrt{a}} \right) - \sqrt{\frac{a}{2\pi}} \exp \left(-\frac{1}{2a} \right)$$

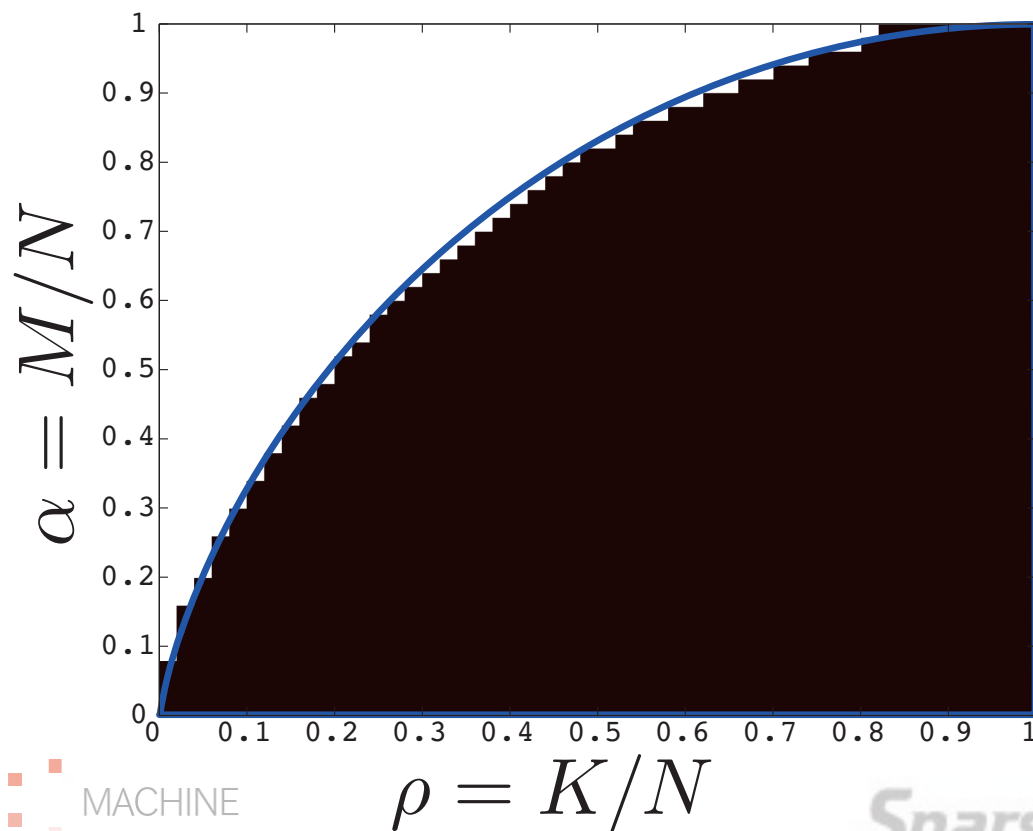
$$H(x) = \int_x^\infty Dz$$



▶ 相転移

▶ 性能指標

$$\text{MSE} = \frac{1}{N} \|\mathbf{x} - \mathbf{x}_0\|_2^2$$



- ▶ 基底追跡型の最適化問題の解析
 - ▶ どんな \mathbf{x} が生成されるか？

$$P(\mathbf{x}|A, \mathbf{y}) \propto \exp(-\beta \|\mathbf{x}\|_1) \delta(\mathbf{y} - A\mathbf{x})$$

- ▶ 罰金法

$$\delta(\mathbf{y} - A\mathbf{x}^a) = \exp\left(-\frac{1}{2\lambda} \|\mathbf{y} - A\mathbf{x}\|_2^2\right) (\lambda \rightarrow +0)$$

- ▶ ラグランジュ未定乗数法

$$\delta(\mathbf{y} - A\mathbf{x}^a) = \exp(i\mathbf{h}^T (\mathbf{y} - A\mathbf{x}))$$

- ▶ 拡張ラグランジュ法

$$\delta(\mathbf{y} - A\mathbf{x}^a) = \exp\left(i\mathbf{h}^T (\mathbf{y} - A\mathbf{x}) - \frac{1}{2\lambda} \|\mathbf{y} - A\mathbf{x}\|_2^2\right)$$

- ▶ どれも等価な結果を導く



黒板でひたすら計算….