

量子アニーリングを用いたクラスタ分析

田中 宗[†] 佐藤 一誠^{††} 栗原 賢一^{†††} 宮下 精二^{††††} 中川 裕志^{†††††}

[†] 早稲田大学高等研究所 〒169-8050 東京都新宿区西早稲田1-6-1

^{††} 東京大学情報基盤センター 〒113-0033 東京都文京区本郷7-3-1

^{†††} グーグル株式会社 〒106-6150 東京都港区六本木6-10-1

^{††††} 東京大学大学院理学系研究科物理学専攻 〒113-0033 東京都文京区本郷7-3-1

^{†††††} 東京大学情報基盤センター 〒113-0033 東京都文京区本郷7-3-1

E-mail: †shu.tanaka@aoni.waseda.jp

あらまし 我々は、量子アニーリングを用いてクラスタ分析を行う方法を提案した。量子アニーリングは、解きたい最適化問題に対して量子揺らぎを導入し、それを徐々に弱めることにより最適化問題の解を自己組織化現象として得る手法である。クラスタ分析とは、膨大なデータを、そのデータの持つ意味によって分類することを指し、情報工学・機械学習における重要な技術である。我々は量子モンテカルロ法によるデモンストレーションの結果、熱揺らぎのみで行った場合に比べ、量子揺らぎを導入した場合に、より良い解を探索することに成功した。

キーワード 量子アニーリング, クラスタ分析, 量子モンテカルロ法

Clustering by Quantum Annealing

Shu TANAKA[†], Issei SATO^{††}, Kenichi KURIHARA^{†††}, Seiji MIYASHITA^{††††}, and Hiroshi
NAKAGAWA^{†††††}

[†] Waseda Institute for Advanced Study, Waseda University,
1-6-1 Nishi Waseda, Shinjuku-ku, Tokyo, Japan

^{††} Information Technology Center, The University of Tokyo,
7-3-1 Hongo, Bunkyo-ku, Tokyo, Japan

^{†††} Google Inc.,

6-10-1 Roppongi, Minato-ku, Tokyo, Japan

^{††††} Department of Physics, The University of Tokyo, 7-3-1 Hongo,
Bunkyo-ku, Tokyo, Japan

^{†††††} Information Technology Center, The University of Tokyo, 7-3-1 Hongo,
Bunkyo-ku, Tokyo, Japan

E-mail: †shu.tanaka@aoni.waseda.jp

Abstract We propose a new scheme for clustering by quantum annealing. The quantum annealing is a procedure to obtain a better or the best solution of optimization problems as self-assembly phenomena, where the quantum fluctuation effect is introduced and gradually decreases. The clustering is a method to divide huge dataset depending on the meaning of individual data, which is a key technology in information engineering and machine learning. We find that a better solution can be obtained when the quantum fluctuation effect is introduced.

Key words Quantum annealing, Clustering, Quantum Monte Carlo methods

1. はじめに

高度情報化社会やビッグデータ社会と呼ばれる現在、データに内在する意味によって膨大なデータを分類することの重要性

に対する認識が様々な分野で高まっている。膨大なデータ群の全体集合を、その潜在的意味によって部分集合に分類する方法を**クラスタ分析**と呼び、情報工学や機械学習における重要な技術である。クラスタ分析の適用範囲は非常に幅広い。自然科学

分野においては、バイオインフォマティクスやマテリアルズインフォマティクスといった分野が一例として挙げられる。また人文・社会科学分野においては、社会コミュニティの特徴付けや言語解析、言語学の一分野などが挙げられる。更に産業界に目を転じれば、画像認識や情報推薦の基礎技術として用いられている。クラスタ分析を行う際に求められるのは、**可能な限り短時間で、高品質なデータ分類を行うこと**である。データ分類の質を表す実関数が最大になるとき、与えられたモデルのもと、最も適切にデータ分類が行われているとする。すなわち、ある種の**組合せ最適化問題**と見ることができる。

組合せ最適化問題は、データ数が大きくなると解の候補が指数関数もしくはそれ以上に激しく増大する、いわゆる**組合せ爆発**の問題が起こることが知られている。そのため最も安直な全探索は現実的な方法ではない。情報工学の分野では、困難な問題が存在する時、その問題に対する個別的な解法が次々に開発されてきており、多くの成功を取めている。一方物理学の観点から、汎用的なアルゴリズムを構築し、困難な問題に取り組むという立場がある。最も成功している例として、**シミュレーテッドアニーリング**と呼ばれる方法が挙げられる [1]。この方法を用いる前提として、解きたい問題をハミルトニアン形式で書く必要がある。この場合、解きたい問題の最適解が、ハミルトニアンの基底状態に対応する。多くの組合せ最適化問題の場合、イジングモデルと呼ばれる統計力学モデル：

$$\mathcal{H}_c = - \sum_{i,j} J_{ij} \sigma_i^z \sigma_j^z - \sum_i h_i \sigma_i^z, \quad \sigma_i^z = \pm 1 \quad (1)$$

で表すことができる。ここで第一項は相互作用を表し、第二項は磁場項を表す。シミュレーテッドアニーリングでは、温度を導入し、その温度を徐々に下げるといった操作を行う。それにより自発的に基底状態を生成する、いわゆる**自己組織化現象**を用いた方法である。温度は熱揺らぎ効果そのものであるため、揺らぎを制御する方法と読み替えることもできる。

物理学におけるもう一つの重要な揺らぎ効果 - 量子揺らぎを用いた基底状態探索アルゴリズムを**量子アニーリング**と呼ぶ [2]。量子アニーリングは、門脇と西森によって 1998 年に提案された方法である。解きたい問題をハミルトニアン形式で記述する段階までは、シミュレーテッドアニーリングと同じである。量子アニーリングでは、量子揺らぎを導入し、それを徐々に弱めるという操作を行い、自発的に解きたい問題の最適解を生成する。シミュレーテッドアニーリングでは古典系を考えているので、式 (2) は実数であるが、量子アニーリングを考える際は、パウリ行列を用いて行列表示する：

$$\hat{\mathcal{H}}_c = - \sum_{i,j} J_{ij} \hat{\sigma}_i^z \hat{\sigma}_j^z - \sum_i h_i \hat{\sigma}_i^z, \quad \hat{\sigma}_i^z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (2)$$

これは対角行列である。 $\hat{\sigma}_i^z$ と非可換な行列 $\hat{\sigma}_i^x$ を導入することにより、 $\hat{\sigma}_i^z$ の固有状態 $|\uparrow\rangle$ 及び $|\downarrow\rangle$ の間の遷移を引き起こすことを考える。そのため、

$$\hat{\mathcal{H}}_q = -\Gamma \sum_i \hat{\sigma}_i^x, \quad \hat{\sigma}_i^x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad (3)$$

なる項を導入し、この係数 Γ を徐々に小さくすることにより、量子アニーリングが実現する。

Farhi らによって、**断熱量子計算**という言葉が提唱された [3]。断熱量子計算では、システムを記述するハミルトニアンとして、組合せ最適化問題のハミルトニアンと、量子揺らぎ項のハミルトニアンの線形和を検討する。初期状態として、量子揺らぎ項のみのハミルトニアンの自明な基底状態を用意する。次に、量子揺らぎ項のハミルトニアンの係数を徐々に小さくし、同時に組合せ最適化問題のハミルトニアンの係数を徐々に大きくする。これにより初期状態から、組合せ最適化問題の非自明な基底状態へと**断熱定理**に基づいて遷移することが可能である。これが断熱量子計算である。また Santoro らによって、量子アニーリングをシミュレーションにより実装し、理論的な考察が行われた [4]。Santoro らが用いたシミュレーション方法は、量子力学の基礎方程式である**シュレディンガー方程式**と、確率遷移過程を表現する**マルコフ連鎖モンテカルロ法**の二つである。後者は必ずしも実際の量子系の実時間発展と対応するわけではないが、古典計算機を用いて量子アニーリングの「シミュレーション」を行うためには必要な手段の一つである。同様にモンテカルロ法を用いた量子アニーリングの研究として、組合せ最適化問題の代表例としてよく知られている**巡回セールスマン問題**や**3-SAT 問題**に対する量子アニーリングの性能評価が行われている [5], [6]。巡回セールスマン問題については、量子アニーリングのほうがシミュレーテッドアニーリングに比べて効率が良いことが示唆されているが、3-SAT 問題に対してはその逆の結果が報告されている。**どのようなタイプの最適化問題が量子アニーリングに適しているか、量子アニーリングに適する最適化問題のマッピング**については未開拓領域であり、今後の理論研究や実践的研究が必要とされている。

その後、2011 年 5 月「**世界初の商用量子コンピュータ**」として、D-Wave Systems 社により開発された D-Wave が登場した [7]。128 個の量子ビットを搭載した量子アニーリングマシンであり、組合せ最適化問題や**機械学習**への応用が期待されている。更に 2013 年には、512 個の量子ビットを搭載した量子アニーリングマシン D-Wave Two が登場している。D-Wave を用いた組合せ最適化問題の検討例として代表的なものに、Ortiz らによるタンパク質の折り畳み安定構造の探索問題が挙げられる [8]。また、D-Wave における熱揺らぎ効果の実験的研究も進められている [9]。ある組合せ最適化問題の場合には、D-Wave が古典計算機よりも圧倒的な性能優位性を発揮するが、それとは別の組合せ最適化問題の場合には、D-Wave に優位性が見られないという結果も報告されている。このように、D-Wave そのものの性能評価は現在始まったばかりと言える。

我々はこのような現状の中、商用量子コンピュータとして D-Wave が登場する前の 2009 年から継続的に、機械学習に対する量子アニーリングの適用について、スーパーコンピュータを用いた大規模数値計算並びに理論的考察を行ってきた [10]~[13]。本講演では、量子アニーリングを用いてクラスタ分析を行う手法の提案、並びに、量子モンテカルロ法による大規模数値計算の結果について報告する。

2. 問題設定の概要

膨大なデータを、潜在的な意味により分類するクラスタ分析に対し、量子アニーリングを適用する。まずはじめに、クラスタ分析を数学的に表現する。 N 個のデータがあるとする。これを $\mathbf{x} = (x_1, x_2, \dots, x_N)$ と表現する。我々が検討する問題は、データ分類状態を σ としたとき、この対数尤度 $\log p(\mathbf{x}, \sigma)$ が最大となるような状況を探ることである。すなわち、 $\sigma^* = \operatorname{argmax}_{\sigma} \log p(\mathbf{x}, \sigma)$ を探索することである。

状態遷移過程を表すモデルとして、我々は**中華料理店過程 (Chinese Restaurant Process)**を検討する。これは以下のような手続きである。中華料理店を想像してみよう。中華料理店には幾つかのテーブルが用意されている。そこに次々と客が入店していき、あるルールに従ってテーブルを選択し、着席する様子を考える。ここで、 i 番目の客が居るテーブルの番号を z_i と表し、 $\mathbf{Z} = \{z_i\}_{i=1}^N$ で、 N 人のテーブルの位置情報を表現する。すなわち、 i 番目の客が k 番目のテーブルに居る場合、 $z_i = k$ と表現する。 i 番目の客をいったん退店させてからまた入店する際、1人以上の客が居るテーブルの総数を K とする。このとき以下の確率で、 i 番目の客はテーブルを選択する：

$$p(z_i | \mathbf{Z} \setminus z_i; \alpha) \propto \begin{cases} \frac{N_k}{\alpha + N - 1} & \text{(既に他の客が居るテーブル } k \text{ を選択)} \\ \frac{\alpha}{\alpha + N - 1} & \text{(誰も居ないテーブルを選択)} \end{cases} \quad (4)$$

ここで N_k はテーブル k に既に居る客の人数、 α は中華料理店過程のハイパーパラメータを表す。 α が大きいほど、客 i は新しいテーブルを選ぶようになる。この場合 \mathbf{Z} の尤度は、 \mathbf{Z} における、1人以上の客がいるテーブルの総数を $K(\mathbf{Z})$ と書くと、

$$p(\mathbf{Z}) = \frac{\alpha^{K(\mathbf{Z})}}{\prod_{l=1}^N (N - l + \alpha)} \prod_{k=1}^{K(\mathbf{Z})} (N_k - 1)! \quad (5)$$

で表される。

ビット変数で記述した方が取り扱いが容易なので、ビット行列を以下のように導入する。例として、5人の客がレストランに居る状況を考える。すなわち $N = 5$ である。このとき、客 1,2,4 が1番目のテーブル、客 3,5 が2番目のテーブルに居るとする。つまり、 $z_1 = 1, z_2 = 1, z_3 = 2, z_4 = 1, z_5 = 2$ である。グラフ理論の**隣接行列**を用いて、この状態を表現する。同じテーブルを共有している客同士は結合しているとして、

$$\begin{pmatrix} 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{pmatrix} \quad (6)$$

というビット行列を考える。中華料理店過程を考える際、このビット行列 B は以下の性質を満たす必要がある。

- ビット行列は対称行列である。すなわち、任意の i, j に対し、 $B_{i,j} = B_{j,i}$ である。

- 対角成分は1である。すなわち、任意の i に対し、 $B_{i,i} = 1$ である。

- 任意の i, j に対し、 $B_i / |B_i| \cdot B_j / |B_j| = 0$ or 1 である。ただしここで、 B_i は i 行目のベクトルを表す。つまり、 $B_i = (B_{i1}, B_{i2}, \dots, B_{iN})$ である。

上記の性質を満たしながら、 \mathbf{Z} そのものではなく、ビット行列を更新していく過程を検討する。これにより、イジングモデルの統計力学を用いることが可能になり、量子アニーリングの実装が容易になる。ただしここで、データ点の個数 (客の人数) が N であるとき、イジングスピンの個数は N^2 になっている。ビット行列で表現される状態を式 (2) で表現されるイジングモデルの状態とみなし、量子揺らぎを式 (3) によって導入する。

データ数 N が少ない場合には、数値計算を用いた厳密対角化や、時間に依存する Schrödinger 方程式を解くことにより、量子アニーリングを古典コンピュータで実装することができるが、データ数が大きい時にはこの方法を取ることはできない。そのため、量子モンテカルロ法を使った擬似的なシミュレーションを行うことにより、量子アニーリングを実装する。量子系に対しモンテカルロ法を適用する場合は、演算子の指数関数の取り扱いに対して近似を行う。代表的なものとして、鈴木トロッター公式 [14], [15] を用いた経路積分表示が挙げられる。本研究ではこれを用いて量子アニーリングを実装した。経路積分表示による量子アニーリングについては、文献 [12] 及び、[13] を参照されたい。

3. 検討したデータセット及び量子アニーリングのスケジュール

我々は、データセットとして以下の三つのネットワークを検討した^(注1)。

- ネットワーク科学者の共著関係を表すデータセット
Netscience (1589 名の研究者; $N = 1589$)
- 引用関係を表すデータセット
Citeseer (2110 論文; $N = 2110$)
- Wikipedia における管理者選挙データセット

Wikivote (7115 名の Wikipedia ユーザー; $N = 7115$)
それぞれネットワーク構造の特徴が異なることが経験則として知られている。

次に、我々が用いた量子アニーリングのスケジュールについて述べる。経路積分表示をする際、我々は虚時間方向を離散的に取った。虚時間方法のスライス数 (トロッタ数) を、 $m = 16$ とした。まずはじめに、シミュレーテッドアニーリングをいくつかの場合について実行した。逆温度 β について、 $\beta(t) = \beta_0 \log(1 + t)$, $\beta_0 \sqrt{t}$, $\beta_0 t$ のスケジュールを考え、 $\beta_0/m = 0.2, 0.4, 0.6$ のそれぞれについて検討した。その結果、

(注1) : 上から順に、

http://www.casos.cs.cmu.edu/computational_tools/datasets/external/netscience/

<http://www.cs.umd.edu/projects/linqs/projects/lbc/index.html>

<http://snap.stanford.edu/data/wiki-Vote.htm>
のデータを用いた。

$\beta_0/m = 0.4$ とし、 $\beta(t) = \beta_0\sqrt{t}$ とした場合が、最もシミュレーテッドアニーリングが良い解を出すことがわかった。我々は熱揺らぎ効果と量子揺らぎ効果を同時に弱める方法を、文献 [10], [12] の知見を活かして検討した。熱揺らぎ効果は先程の条件、すなわち、

$$\frac{\beta(t)}{m} = 0.4\sqrt{t} \quad (7)$$

を用い、更に量子揺らぎ項を

$$\frac{\beta\Gamma}{m} = \Gamma_0 \frac{\tau}{t} \quad (8)$$

のスケジュールで弱めるというスケジュールを考えた。ここで τ はアニーリング時間を表す。ここで Γ_0 は調整可能なパラメータである。

4. 結果・今後の展望

3章で述べたデータセットについて、量子モンテカルロ法を用いた擬似的な量子アニーリングでは、シミュレーテッドアニーリングに比べて効率よく良い解を得られる場合があることが分かった。 Γ_0 の値が小さすぎる場合は、虚時間方向の相互作用が強くなるため、状態更新がうまくいかないという問題点がある。また Γ_0 の値が大きすぎる場合は、独立なシミュレーテッドアニーリングが m 個並列で行われているという状況になり、実質的にシミュレーテッドアニーリングと変わらない。そのため、中間の Γ_0 を用いる必要がある。この方法を用いる場合には、適切な Γ_0 を決定する方法が必要である。

ただしここで行った大規模数値計算は、量子モンテカルロ法に基づくものであり、実際の量子ダイナミクスとは異なる。鈴木トロッタ公式による、虚時間方向の離散化に伴う人為的な効果も無視できないため、クラスタ分析に対する量子アニーリングの有用性をより詳細に解明するためには、D-Wave などの量子アニーリング機を用いた実験や、大規模系の時間発展をシミュレートする方法の開発を行う必要がある。

文 献

- [1] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, *Science*, **220**, 671 (1983).
- [2] T. Kadowaki and H. Nishimori, *Phys. Rev. E*, **58**, 5355 (1998).
- [3] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. pred, *Science*, **292**, 472 (2001).
- [4] G. E. Santoro, R. Martoňák, E. Tosatti, and R. Car, *Science*, **295**, 2427 (2002).
- [5] R. Martoňák, G. E. Santoro, and E. Tosatti, *Phys. Rev. E*, **70**, 057701 (2004).
- [6] D. A. Battaglia, G. E. Santoro, and E. Tosatti, *Phys. Rev. E*, **71**, 066707 (2005).
- [7] <http://www.dwavesys.com>
- [8] A. P. Ortiz, N. Dickson, M. D. Brook, G. Rose, and A. A. Guzik, *Scientific Reports* **2**, 571 (2012).
- [9] N. G. Dickson *et al.*, *Nature Communications*, **4**, 1903 (2013).
- [10] K. Kurihara, S. Tanaka, and S. Miyashita, *Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence (UAI2009)*.
- [11] I. Sato, K. Kurihara, S. Tanaka, H. Nakagawa, and S. Miyashita, *Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence (UAI2009)*.
- [12] 田中宗、栗原賢一、宮下精二、科研費特定領域研究「情報統計力学の深化と展開」研究会「情報統計力学の広がり：量子・画像・そして展開」講演概要集（http://www.shutanaka.com/papers_files/ShuTanaka_DEXSML10.pdf よりダウンロード可能。）
- [13] I. Sato, S. Tanaka, K. Kurihara, S. Miyashita, and H. Nakagawa, *Neurocomputing*, **121**, 523 (2013).
- [14] H. F. Trotter, *Proc. Am. Math. Soc.*, **10**, 545 (1959).
- [15] M. Suzuki, *Prog. Theor. Phys.*, **56**, 1454 (1976).