Beating Brute Force for Systems of Polynomial Equations over Finite Fields
(有限体上の多変数連立代数方程式系に対する総当り探索の打破) [SODA 2017]

Daniel Lokshtanov (University of Bergen)

Ramamohan Paturi (University of California, San Diego)

玉置 卓 (京都大学)

Ryan Williams (MIT)

Huacheng Yu (Stanford University)

ERATO感謝祭 Season IV, 2017年8月4日, 国立情報学研究所, 千代田区

位置づけ

分野

理論計算機科学 → 計算・アルゴリズム理論 →

■ 厳密指数時間・媒介変数付アルゴリズム

■ Fine-Grained Complexity (精微な計算複雑性)

STOC・FOCS で1～2セッション

SODAで 2～3セッション

目的

P or NP困難, 多項式時間で解ける or 解けない

等より精密に必要十分な計算資源量を評価したい

($n^{1.9}$で十分 or n^2は必要, $2^{0.9n}$ で十分 or 2^nは必要等)

扱う問題

有限体F_q上のn変数連立d次方程式系

入力 d次多項式 $P_1, P_2, \ldots, P_m \in \mathrm{F}_q[x_1, x_2, \ldots, x_n]$

出力 $a \in \mathrm{F}_q^n$ s.t. $P_1(a) = P_2(a) = \cdots = P_m(a) = 0$

例 $[q = 3, n = 4, d = 5, m = 2]$

$$P_1 = 2x_1^2\, x_2^2\, x_3 + x_3^2\, x_4,\ P_2 = x_1 x_2 + x_2^2 + 1$$

$$a = (2,2,1,1)$$

扱う問題

有限体F_q上のn変数連立d次方程式系

入力 d次多項式 $P_1, P_2, \ldots, P_m \in \mathrm{F}_q[x_1, x_2, \ldots, x_n]$

出力 $a \in \mathrm{F}_q^n$ s.t. $P_1(a) = P_2(a) = \cdots = P_m(a) = 0$

例 $[q = 3, n = 4, d = 5, m = 2]$

$$P_1 = 2x_1^2\, x_2^2\, x_3 + x_3^2\, x_4,\ P_2 = x_1 x_2 + x_2^2 + 1$$

$$a = (2,2,1,1)$$

$d = 1$ のとき多項式時間で解ける (Gaußの消去法)

$d \geq 2$ のときNP困難, 総当り探索 q^n より速い

$\exists \varepsilon > 0, q^{(1-\varepsilon)n}$ 時間で解けるか未解決

本研究の結果1 [乱択, 探索]

有限体F_q上のn変数連立d次方程式系

条件	計算時間
$q = d = 2$	$2^{0.8765n}$
$q = 2, d > 2$	$2^{\left(1-\frac{1}{5d}\right)n}$
$q = p^k, \log p < 4edk$	$q^{\left(1-\frac{1}{200d}\right)n}$
$q = p^k, \log p \geq 4edk$	$q^n \left(\frac{\log q}{edk}\right)^{-kn}$

$e = 2.718\ldots$ はネイピア数

本研究の結果2 [決定性, 計数]

有限体F_q上のn変数連立d次方程式系

条件	計算時間
$q = p^k$	$q^{\left(1-\frac{1}{300dq^{\frac{6}{7k}}}\right)n}$

Cf. 本研究の結果1
[乱択, 探索]

$q^{\left(1-\frac{1}{200d}\right)n}$

問題の一般化

多項式の一般化 (ただし $q = 2$)

入力　ΣΠΣ算術回路 $P_1, P_2, \dots, P_m$

(P_i は変数と定数の和の積の和, 次数制限なし)

出力　$a \in \mathrm{F}_2^n$　s.t.　$P_1(a) = P_2(a) = \cdots = P_m(a) = 0$

例　$P_1 = (x_1 + x_2 + 1)(x_2 + x_3) + (x_1 + x_4)x_2 + 1$

$P_2 = x_1 x_2 \cdots x_n + x_2 + x_4$

s = 全体の項の数 (項=変数と定数の和の積)

本研究の結果3 [ΣΠΣ算術回路]

有限体F_2上のn変数方程式系

全体の項の数が s のとき

アルゴリズム	計算時間
乱択 探索	$2^{\left(1-\frac{1}{10\log\left(\frac{s}{n}\right)}\right)n}$
決定性 計数	$2^{\left(1-\frac{1}{1100\log\left(\frac{s}{n}\right)}\right)n}$

$s = O(n)$ のとき 2^n より指数的に高速

注意

F_2 上の n変数連立 d次方程式系は d-SATの一般化

e.g. $[d = 3]$

$C_1 = (\neg x_1 \vee x_2 \vee x_3) \Rightarrow p_1 = x_1(1 + x_2)(1 + x_3)$

$C_2 = (x_1 \vee \neg x_3 \vee \neg x_4) \Rightarrow p_2 = (1 + x_1)x_2x_2$

$C_3 = (x_2 \vee x_3 \vee x_4) \Rightarrow p_3 = (1 + x_1)(1 + x_2)(1 + x_3)$

$C_1 = C_2 = C_3 = 1 \Leftrightarrow p_1 = p_2 = p_3 = 0$

計算時間の最適性

■ **F_2上の n変数連立 d次方程式系** $2^{n(1-1/O(d))}$

Cf. **d-CNF SAT** $2^{n(1-1/d)}$

[Paturi-Pudlak-Zane'97,…]

■ **F_2上の n変数ΣΠΣ算術回路の方程式系**

全体の項の数が s のとき $2^{n(1-1/O(\log(s/n)))}$

Cf. **CNF SAT**

全体の節の数が s のとき $2^{n(1-1/(2\log(s/n)))}$

[Schuler'05,Calabro-Impagliazzo-Paturi'06,…]

先行・関連研究

アルゴリズム

■ 実用的には Gröbner基底 (最悪時は2重指数時間)
やSATソルバで解く試み

■ 特別な場合に総当り探索より速く解ける (十分条件)

問題の困難性を利用

■ 暗号の安全性
(Multivariate Cryptography ∈ 耐量子暗号)

■ 特定の問題に対するアルゴリズムの計算時間の最適性
(Fine-Grained Complexity)

アルゴリズムの概要

簡単のため $\mathrm{F}_2 = \{0,1\}$ の場合に限定
入力多項式は多重線形としてよい ($x_i^2 = x_i$より)

変種問題を考える

整数上で方程式が1本のとき解を計数

入力 d次多項式 $P \in Z[x_1, x_2, \ldots, x_n]$

s.t. $\frac{n}{2} \gg d = \Omega(n), \forall x \in \{0,1\}^n, P(x) \in \{0,1\}$

出力 $\sum_{x \in \{0,1\}^n} P(x) = \#\{x \in \{0,1\}^n \mid P(x) = 1\}$

例 $[n = d = 3]$

$P = x_1\, x_2\, x_3 - x_1\, x_3 - x_2\, x_3 + x_3 + 1$

$\sum_{x \in \{0,1\}^3} P(x) = 7$

変種問題を考える

整数上で方程式が1本のとき解を計数

入力 d次多項式 $P \in Z[x_1, x_2, \ldots, x_n]$

s.t. $\frac{n}{2} \gg d = \Omega(n), \forall x \in \{0,1\}^n, P(x) \in \{0,1\}$

出力 $\sum_{x \in \{0,1\}^n} P(x) = \#\{x \in \{0,1\}^n \mid P(x) = 1\}$

例 $[n = d = 3]$

$P = x_1\, x_2\, x_3 - x_1\, x_3 - x_2\, x_3 + x_3 + 1$

$\sum_{x \in \{0,1\}^3} P(x) = 7$

単純な全点評価: $\mathrm{poly}(n) \binom{n}{d} 2^n = 2^n \times 2^{\Omega(n)}$時間

Yatesの全点評価: $\mathrm{poly}(n)\ 2^n$時間

変種問題を考える

整数上で方程式が1本のとき解を計数

入力 d次多項式 $P \in Z[x_1, x_2, \ldots, x_n]$

s.t. $\frac{n}{2} \gg d = \Omega(n), \forall x \in \{0,1\}^n, P(x) \in \{0,1\}$

出力 $\sum_{x \in \{0,1\}^n} P(x) = \#\{x \in \{0,1\}^n \mid P(x) = 1\}$

Yatesの全点評価: $\mathrm{poly}(n)\ 2^n$時間

実は $\mathrm{poly}(n) 2^{\frac{1}{2}\{n+\log\binom{n}{d}\}} \ll 2^n \times 2^{-\Omega(n)}$ 時間で可能

入力 d次多項式 $P \in Z[x_1, x_2, \ldots, x_n]$

s.t. $\frac{n}{2} \gg d = \Omega(n), \forall x \in \{0,1\}^n, P(x) \in \{0,1\}$

出力 $\sum_{x \in \{0,1\}^n} P(x) = \#\{x \in \{0,1\}^n \mid P(x) = 1\}$

$n' := \frac{1}{2}\{n - \log \binom{n}{d}\}$とおいて部分和を考える

$Q(x_1, \ldots, x_{n-n'}) := \sum_{y \in \{0,1\}^{n'}} P(x_1, \ldots, x_{n-n'}, y_1, \ldots, y_{n'})$

入力 d次多項式 $P \in Z[x_1, x_2, \ldots, x_n]$

s.t. $\frac{n}{2} \gg d = \Omega(n), \forall x \in \{0,1\}^n, P(x) \in \{0,1\}$

出力 $\sum_{x \in \{0,1\}^n} P(x) = \#\{x \in \{0,1\}^n \mid P(x) = 1\}$

$n' := \frac{1}{2}\{n - \log \binom{n}{d}\}$とおいて部分和を考える

$Q(x_1, \ldots, x_{n-n'}) := \sum_{y \in \{0,1\}^{n'}} P(x_1, \ldots, x_{n-n'}, y_1, \ldots, y_{n'})$

- Q は $n - n'$変数d次多項式
- Q を求める時間 $\mathrm{poly}(n) \binom{n}{d} 2^{n'} \sim 2^{\frac{1}{2}\{n + \log \binom{n}{d}\}}$
- $\sum_{x \in \{0,1\}^n} P(x) = \sum_{x \in \{0,1\}^{n-n'}} Q(x_1, \ldots, x_{n-n'})$

入力 d次多項式 $P \in Z[x_1, x_2, \ldots, x_n]$

s.t. $\frac{n}{2} \gg d = \Omega(n), \forall x \in \{0,1\}^n, P(x) \in \{0,1\}$

出力 $\sum_{x \in \{0,1\}^n} P(x) = \#\{x \in \{0,1\}^n \mid P(x) = 1\}$

$n' := \frac{1}{2}\{n - \log \binom{n}{d}\}$とおいて部分和を考える

$Q(x_1, \ldots, x_{n-n'}) := \sum_{y \in \{0,1\}^{n'}} P(x_1, \ldots, x_{n-n'}, y_1, \ldots, y_{n'})$

- Q は $n - n'$変数d次多項式
- Q を求める時間 $\mathrm{poly}(n) \binom{n}{d} 2^{n'} \sim 2^{\frac{1}{2}\{n + \log \binom{n}{d}\}}$
- $\sum_{x \in \{0,1\}^n} P(x) = \sum_{x \in \{0,1\}^{n-n'}} Q(x_1, \ldots, x_{n-n'})$

Yatesの全点評価をQに適用

$\mathrm{poly}(n)\, 2^{n-n'} \sim 2^{\frac{1}{2}\{n + \log \binom{n}{d}\}}$時間で出力が求まる

元の問題に戻る

入力 d次多項式 $P_1, P_2, \ldots, P_m \in \mathrm{F}_q[x_1, x_2, \ldots, x_n]$
出力 $a \in \mathrm{F}_2^n$ s.t. $P_1(a) = P_2(a) = \cdots = P_m(a) = 0$

1. 方程式を1本にまとめる

$$p_1(x) = p_2(x) = \cdots = p_m(x) = 0 \Leftrightarrow P(x) = 1$$

元の問題に戻る

入力 d次多項式 $P_1, P_2, \ldots, P_m \in \mathrm{F}_q[x_1, x_2, \ldots, x_n]$
出力 $a \in \mathrm{F}_2^n$ s.t. $P_1(a) = P_2(a) = \cdots = P_m(a) = 0$

1. 方程式を1本にまとめる

$$p_1(x) = p_2(x) = \cdots = p_m(x) = 0 \Leftrightarrow P(x) = 1$$

2. (有限体に戻ったので) 部分論理和を考える

$$Q(x_1, \ldots, x_{n-n'}) := \mathrm{V}_{y \in \{0,1\}^{n'}} P(x_1, \ldots, x_{n-n'}, y_1, \ldots, y_{n'})$$

元の問題に戻る

入力 d次多項式 $P_1, P_2, \ldots, P_m \in \mathrm{F}_q[x_1, x_2, \ldots, x_n]$
出力 $a \in \mathrm{F}_2^n$ s.t. $P_1(a) = P_2(a) = \cdots = P_m(a) = 0$

1. 方程式を1本にまとめる

$$p_1(x) = p_2(x) = \cdots = p_m(x) = 0 \Leftrightarrow P(x) = 1$$

2. (有限体に戻ったので) 部分論理和を考える

$$Q(x_1, \ldots, x_{n-n'}) := \bigvee_{y \in \{0,1\}^{n'}} P(x_1, \ldots, x_{n-n'}, y_1, \ldots, y_{n'})$$

3. Qを多項式Rで確率的に近似 ← **論文で最も技術的な部分**
(正確な多項式表現を求めるには2^n時間超)

元の問題に戻る

入力 d次多項式 $P_1, P_2, \ldots, P_m \in \mathrm{F}_q[x_1, x_2, \ldots, x_n]$
出力 $a \in \mathrm{F}_2^n$ s.t. $P_1(a) = P_2(a) = \cdots = P_m(a) = 0$

1. 方程式を1本にまとめる

$$p_1(x) = p_2(x) = \cdots = p_m(x) = 0 \Leftrightarrow P(x) = 1$$

2. (有限体に戻ったので) 部分論理和を考える

$$Q(x_1, \ldots, x_{n-n'}) := \bigvee_{y \in \{0,1\}^{n'}} P(x_1, \ldots, x_{n-n'}, y_1, \ldots, y_{n'})$$

3. Qを多項式Rで確率的に近似
(正確な多項式表現を求めるには2^n時間超)

4. RにYatesの全点評価を適用
1～4の繰り返しでQの全点評価が高い確率で求まる

ステップ1

有限体F_2上のn変数連立d次方程式系

入力 d次多項式 $P_1, P_2, \ldots, P_m \in \mathrm{F}_q[x_1, x_2, \ldots, x_n]$
出力 $a \in \mathrm{F}_2^n$ s.t. $P_1(a) = P_2(a) = \cdots = P_m(a) = 0$

1. 方程式を1本にまとめる

$P \coloneqq (1 + P_1) \cdots (1 + P_m)$ とおくと

$$p_1(x) = p_2(x) = \cdots = p_m(x) = 0 \Leftrightarrow P(x) = 1$$

ステップ2

有限体F_2上のn変数連立d次方程式系

入力 d次多項式 $P_1, P_2, \dots, P_m \in \mathrm{F}_q[x_1, x_2, \dots, x_n]$
出力 $a \in \mathrm{F}_2^n$ s.t. $P_1(a) = P_2(a) = \cdots = P_m(a) = 0$

$P \coloneqq (1 + P_1) \cdots (1 + P_m)$ とおいた

2. (有限体に戻ったので) 部分論理和を考える
$Q(x_1, \dots, x_{n-n'}) \coloneqq \mathrm{V}_{y \in \{0,1\}^{n'}} P(x_1, \dots, x_{n-n'}, y_1, \dots, y_{n'})$

理由 和を取ると解が偶数個の場合に0になるのを回避

ステップ3

有限体F_2上のn変数連立d次方程式系

入力 d次多項式 $P_1, P_2, \dots, P_m \in \mathrm{F}_q[x_1, x_2, \dots, x_n]$
出力 $a \in \mathrm{F}_2^n$ s.t. $P_1(a) = P_2(a) = \cdots = P_m(a) = 0$

$P \coloneqq (1 + P_1) \cdots (1 + P_m)$
$Q(x_1, \dots, x_{n-n'}) \coloneqq \bigvee_{y \in \{0,1\}^{n'}} P(x_1, \dots, x_{n-n'}, y_1, \dots, y_{n'})$

3. Qを多項式Rで確率的に近似

理由 Pの次数 $\gg n$
論理和の多項式表現の次数 $= 2^{n'}$
Qを素直に展開して項の和の形で書くと 2^n 時間を超える

ステップ3

有限体F_2上のn変数連立d次方程式系

入力 d次多項式 $P_1, P_2, \dots, P_m \in \mathrm{F}_q[x_1, x_2, \dots, x_n]$
出力 $a \in \mathrm{F}_2^n$ s.t. $P_1(a) = P_2(a) = \cdots = P_m(a) = 0$

$P \coloneqq (1 + P_1) \cdots (1 + P_m)$
$Q(x_1, \dots, x_{n-n'}) \coloneqq \bigvee_{y \in \{0,1\}^{n'}} P(x_1, \dots, x_{n-n'}, y_1, \dots, y_{n'})$

3. Qを多項式Rで確率的に近似

⇒ 計算の理論の有名結果 [Razborov'87] が使える
「定数段数回路 (計算モデル) が多数決関数を計算できない」を示すために導入された手法
否定的な結果は肯定的な結果にも有用!

アルゴリズム (再掲)

入力 d次多項式 $P_1, P_2, \ldots, P_m \in \mathrm{F}_q[x_1, x_2, \ldots, x_n]$
出力 $a \in \mathrm{F}_2^n$ s.t. $P_1(a) = P_2(a) = \cdots = P_m(a) = 0$

1. 方程式を1本にまとめる

$$P_1(x) = P_2(x) = \cdots = P_m(x) = 0 \Leftrightarrow P(x) = 1$$

2. (有限体に戻ったので) 部分論理和を考える

$$Q(x_1, \ldots, x_{n-n'}) := \bigvee_{y \in \{0,1\}^{n'}} P(x_1, \ldots, x_{n-n'}, y_1, \ldots, y_{n'})$$

3. Qを多項式Rで確率的に近似
(正確に求めようとすると2^n時間かかる)

4. RにYatesの全点評価を適用
1～4の繰り返しでQの全点評価が高い確率で求まる

まとめ

■ **有限体F_q上のn変数連立d次方程式系**に対して
総当り探索 q^n より速いアルゴリズムを与えた
■ 論理関数の疎な近似表現
(低次数多項式/低ランク行列) が重要な役割

課題
■ 指数領域アルゴリズムのため実用的ではない
⇒ 多項式領域で $\exists \varepsilon > 0, q^{(1-\varepsilon)n}$ 時間で解けるか

■ 総当り打破が未解決な他の問題を,
本問題への帰着により解決できるか
■ 他に有用な論理関数の疎な近似表現はあるか