矩阵函数的（量子/经典）访问复杂度

邵长鹏

中科院数学院

arXiv:2311.06999, STOC 2024

joint work with Ashley Montanaro

2024 年 3 月 12 日

北京大学前沿计算研究中心

量子计算

- 以量子力学为基本原理的新型计算机模型，具有强大的并行计算能力，在一些问题上能够远比经典计算机快

量子计算

- 以量子力学为基本原理的新型计算机模型，具有强大的并行计算能力，在一些问题上能够远比经典计算机快
- Shor 大整数分解算法（指数加速 n vs $e^{n^{1/3}}$）

量子计算

- 以量子力学为基本原理的新型计算机模型，具有强大的并行计算能力，在一些问题上能够远比经典计算机快
- Shor 大整数分解算法（指数加速 n vs $e^{n^{1/3}}$）
- Grover 搜索算法（平方加速 $\sqrt{n}$ vs n）

量子计算

- 以量子力学为基本原理的新型计算机模型，具有强大的并行计算能力，在一些问题上能够远比经典计算机快
- Shor 大整数分解算法（指数加速 n vs $e^{n^{1/3}}$）
- Grover 搜索算法（平方加速 $\sqrt{n}$ vs n）
- Hamiltonian 模拟问题：计算 $e^{iHt}|\psi\rangle$（指数加速）

量子计算

- 以量子力学为基本原理的新型计算机模型，具有强大的并行计算能力，在一些问题上能够远比经典计算机快
- Shor 大整数分解算法（指数加速 n vs $e^{n^{1/3}}$）
- Grover 搜索算法（平方加速 $\sqrt{n}$ vs n）
- Hamiltonian 模拟问题：计算 $e^{iHt}|\psi\rangle$（指数加速）
- 在机器学习、优化问题、方程求解、图论问题等领域发挥了重要作用

量子计算

- 以量子力学为基本原理的新型计算机模型，具有强大的并行计算能力，在一些问题上能够远比经典计算机快
- Shor 大整数分解算法（指数加速 n vs $e^{n^{1/3}}$）
- Grover 搜索算法（平方加速 $\sqrt{n}$ vs n）
- Hamiltonian 模拟问题：计算 $e^{iHt}|\psi\rangle$（指数加速）
- 在机器学习、优化问题、方程求解、图论问题等领域发挥了重要作用
- 本报告：探讨矩阵函数问题上的量子优势

矩阵函数

定义

设 A 是一个厄米矩阵，特征值分解为 $A = UDU^\dagger$。设 $f(x)$ 是一个在 D 上有定义的函数，则[1]

$$f(A) := Uf(D)U^\dagger$$

其中若 $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$，则 $f(D) = \mathrm{diag}(f(\lambda_1), \ldots, f(\lambda_n))$

[1] Nicholas Higham, Functions of Matrices: Theory and Computations. SIAM 2008

矩阵函数

定义

设 A 是一个厄米矩阵，特征值分解为 $A = UDU^\dagger$。设 $f(x)$ 是一个在 D 上有定义的函数，则[1]

$$f(A) := Uf(D)U^\dagger$$

其中若 $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$，则 $f(D) = \mathrm{diag}(f(\lambda_1), \ldots, f(\lambda_n))$

例

- 设 $f = x^k$，则 $f(A) = A^k$

[1] Nicholas Higham, Functions of Matrices: Theory and Computations. SIAM 2008

矩阵函数

定义

设 A 是一个厄米矩阵，特征值分解为 $A = UDU^\dagger$。设 $f(x)$ 是一个在 D 上有定义的函数，则[1]

$$f(A) := Uf(D)U^\dagger$$

其中若 $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$，则 $f(D) = \mathrm{diag}(f(\lambda_1), \ldots, f(\lambda_n))$

例

- 设 $f = x^k$，则 $f(A) = A^k$
- 设 $f = e^x$，则 $f(A) = e^A (= \sum_k A^k/k!$ 当 $\|A\| \leq 1)$

[1] Nicholas Higham, Functions of Matrices: Theory and Computations. SIAM 2008

矩阵函数

定义

设 A 是一个厄米矩阵，特征值分解为 $A = UDU^\dagger$。设 $f(x)$ 是一个在 D 上有定义的函数，则[1]

$$f(A) := Uf(D)U^\dagger$$

其中若 $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$，则 $f(D) = \mathrm{diag}(f(\lambda_1), \ldots, f(\lambda_n))$

例

- 设 $f = x^k$，则 $f(A) = A^k$
- 设 $f = e^x$，则 $f(A) = e^A (= \sum_k A^k/k!$ 当 $\|A\| \leq 1)$
- 设 $f = 1/x$，则 $f(A) = A^{-1}$

[1] Nicholas Higham, Functions of Matrices: Theory and Computations. SIAM 2008

矩阵函数

定义

设 A 是一个厄米矩阵，特征值分解为 $A = UDU^T$。设 $f(x)$ 是一个在 D 上有定义的函数，则[2]

$$f(A) := Uf(D)U^T$$

其中若 $D = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$，则 $f(D) = \mathrm{diag}(f(\lambda_1), \ldots, f(\lambda_n))$

例

- 设 $f = x^k$，则 $f(A) = A^k$ (离散的随机游走)
- 设 $f = e^x$，则 $f(A) = e^A$ (连续的随机游走、ODE 的数值解)
- 设 $f = 1/x$，则 $f(A) = A^{-1}$ (线性方程组求解)

[2] Nicholas Higham, Functions of Matrices: Theory and Computations. SIAM 2008

矩阵函数在量子计算领域

在量子计算领域，一个非常重要的函数是 $f(x) = e^{ixt}$，则

$$f(A) = e^{iAt}.$$

对应于哈密顿模拟问题，是量子计算领域最基本也是最重要的问题之一。

矩阵函数在量子计算领域

在量子计算领域，一个非常重要的函数是 $f(x) = e^{ixt}$，则

$$f(A) = e^{iAt}.$$

对应于哈密顿模拟问题，是量子计算领域最基本也是最重要的问题之一。

定理 (Low & Chuang, 2017)

假定 A 稀疏，$\|A\| \leq 1$，则计算 $e^{iAt}|\psi\rangle$ 的量子复杂度为

$$\Theta\left(t + \frac{\log(1/\varepsilon)}{\log\log(1/\varepsilon)}\right).$$

矩阵函数在量子计算领域

在量子计算领域，一个非常重要的函数是 $f(x) = e^{ixt}$，则

$$f(A) = e^{iAt}.$$

对应于哈密顿模拟问题，是量子计算领域最基本也是最重要的问题之一。

定理 (Low & Chuang, 2017)

假定 A 稀疏，$\|A\| \leq 1$，则计算 $e^{iAt}|\psi\rangle$ 的量子复杂度为

$$\Theta\left(t + \frac{\log(1/\varepsilon)}{\log\log(1/\varepsilon)}\right).$$

该问题是 BQP-完全问题（也即能够在量子计算机上有效解决的最困难的一类问题）

矩阵函数在量子计算领域

- 另外一个BQP-完全问题是计算 $A^{-1}|b\rangle$。对应于求解线性方程组，这里的 $f(x) = 1/x$ [Harrow, Hassidim, Lloyd 2008]

矩阵函数在量子计算领域

- 另外一个BQP-完全问题是计算 $A^{-1}|b\rangle$。对应于求解线性方程组，这里的 $f(x) = 1/x$ [Harrow, Hassidim, Lloyd 2008]
- 在量子机器学习发挥着重要作用

矩阵函数在量子计算领域

- 另外一个BQP-完全问题是计算 $A^{-1}|b\rangle$。对应于求解线性方程组，这里的 $f(x) = 1/x$ [Harrow, Hassidim, Lloyd 2008]
- 在量子机器学习发挥着重要作用
- 假定 A 是稀疏的，则量子算法的复杂度为 $\Theta(\kappa)$，其中 κ 为 A 的条件数 [Chakraborty, Gilyén, Jeffery 2018]

矩阵函数在量子计算领域

- 另外一个BQP-完全问题是计算 $A^{-1}|b\rangle$。对应于求解线性方程组，这里的 $f(x) = 1/x$ [Harrow, Hassidim, Lloyd 2008]
- 在量子机器学习发挥着重要作用
- 假定 A 是稀疏的，则量子算法的复杂度为 $\Theta(\kappa)$，其中 κ 为 A 的条件数 [Chakraborty, Gilyén, Jeffery 2018]
- 本报告：针对一般矩阵函数问题，给出（量子/经典）算法的复杂度下界，估算量子计算优势的大小

矩阵函数问题

问题

设 A 是一个稀疏的厄米矩阵，$\|A\| \leq 1$，设 $f : [-1, 1] \to [-1, 1]$ 是一个函数，对任意给定的两个指标 i, j，研究计算 $f(A)_{i,j} \pm \varepsilon$ 的复杂度，尤其是复杂度下界。

矩阵函数问题

问题

设 A 是一个稀疏的厄米矩阵，$\|A\| \leq 1$，设 $f : [-1, 1] \to [-1, 1]$ 是一个函数，对任意给定的两个指标 i, j，研究计算 $f(A)_{i,j} \pm \varepsilon$ 的复杂度，尤其是复杂度下界。

定义 (访问/查询复杂度 (query complexity))

设 $A = (A_{i,j})_{n\times n}$ 是一个稀疏的厄米矩阵，给定两个 Oracle

$$(i,j) \longrightarrow \boxed{\mathcal{O}_1} \longrightarrow p_{i,j}$$

$$(i,j) \longrightarrow \boxed{\mathcal{O}_2} \longrightarrow A_{i,j}$$

其中 $p_{i,j}$ 是第 i 行第 j 个非零元素的位置。访问复杂度定义为逼近 $f(A)_{i,j}$ 所使用的 $\mathcal{O}_1, \mathcal{O}_2$ 的最少个数。

例子：矩阵幂次 A^d

经典算法：

- 设 A 的稀疏度为 s（即每行/列非零元素个数 $\leq s$）。按定义

$$(A^d)_{i,j} = \sum_{k_1} \sum_{k_2} \cdots \sum_{k_{d-1}} A_{i,k_1} A_{k_1,k_2} \cdots A_{k_{d-1},j}$$

由于稀疏性，复杂度为 $O(s^{d-1})$

例子：矩阵幂次 A^d

经典算法：

- 设 A 的稀疏度为 s（即每行/列非零元素个数 $\leq s$）。按定义

$$(A^d)_{i,j} = \sum_{k_1} \sum_{k_2} \cdots \sum_{k_{d-1}} A_{i,k_1} A_{k_1,k_2} \cdots A_{k_{d-1},j}$$

  由于稀疏性，复杂度为 $O(s^{d-1})$

- 对于 x^d，存在次数为 $\Theta(\sqrt{d})$ 的逼近多项式，因此结果可改进到 $s^{O(\sqrt{d})}$ [Sachdeva & Vishnoi, 2014]

例子：矩阵幂次 A^d

经典算法：

- 设 A 的稀疏度为 s（即每行/列非零元素个数 $\leq s$）。按定义

$$(A^d)_{i,j} = \sum_{k_1} \sum_{k_2} \cdots \sum_{k_{d-1}} A_{i,k_1} A_{k_1,k_2} \cdots A_{k_{d-1},j}$$

  由于稀疏性，复杂度为 $O(s^{d-1})$

- 对于 x^d，存在次数为 $\Theta(\sqrt{d})$ 的逼近多项式，因此结果可改进到 $s^{O(\sqrt{d})}$ [Sachdeva & Vishnoi, 2014]
- 该问题是 BQP-完全问题 [Janzing & Wocjan, 2007]

例子：矩阵幂次 A^d

经典算法：

- 设 A 的稀疏度为 s（即每行/列非零元素个数 $\leq s$）。按定义

$$(A^d)_{i,j} = \sum_{k_1} \sum_{k_2} \cdots \sum_{k_{d-1}} A_{i,k_1} A_{k_1,k_2} \cdots A_{k_{d-1},j}$$

  由于稀疏性，复杂度为 $O(s^{d-1})$

- 对于 x^d，存在次数为 $\Theta(\sqrt{d})$ 的逼近多项式，因此结果可改进到 $s^{O(\sqrt{d})}$ [Sachdeva & Vishnoi, 2014]
- 该问题是 BQP-完全问题 [Janzing & Wocjan, 2007]
- 我们的结果: $\widetilde{\Omega}((s/2)^{(\sqrt{d}-1)/6})$

例子：矩阵幂次 A^d

经典算法：

- 设 A 的稀疏度为 s（即每行/列非零元素个数 $\leq s$）。按定义

$$(A^d)_{i,j} = \sum_{k_1} \sum_{k_2} \cdots \sum_{k_{d-1}} A_{i,k_1} A_{k_1,k_2} \cdots A_{k_{d-1},j}$$

  由于稀疏性，复杂度为 $O(s^{d-1})$

- 对于 x^d，存在次数为 $\Theta(\sqrt{d})$ 的逼近多项式，因此结果可改进到 $s^{O(\sqrt{d})}$ [Sachdeva & Vishnoi, 2014]
- 该问题是 BQP-完全问题 [Janzing & Wocjan, 2007]
- 我们的结果: $\widetilde{\Omega}((s/2)^{(\sqrt{d}-1)/6})$

例子：矩阵幂次 A^d

经典算法：

- 设 A 的稀疏度为 s（即每行/列非零元素个数 $\leq s$）。按定义

$$(A^d)_{i,j} = \sum_{k_1} \sum_{k_2} \cdots \sum_{k_{d-1}} A_{i,k_1} A_{k_1,k_2} \cdots A_{k_{d-1},j}$$

  由于稀疏性，复杂度为 $O(s^{d-1})$

- 对于 x^d，存在次数为 $\Theta(\sqrt{d})$ 的逼近多项式，因此结果可改进到 $s^{O(\sqrt{d})}$ [Sachdeva & Vishnoi, 2014]
- 该问题是 BQP-完全问题 [Janzing & Wocjan, 2007]
- 我们的结果: $\widetilde{\Omega}((s/2)^{(\sqrt{d}-1)/6})$

量子算法：

- 复杂度上界 $O(s\sqrt{d}/\varepsilon)$ [Gilyén, Su, Low, Wiebe, 2018]

例子：矩阵幂次 A^d

经典算法：

- 设 A 的稀疏度为 s（即每行/列非零元素个数 $\leq s$）。按定义

$$(A^d)_{i,j} = \sum_{k_1} \sum_{k_2} \cdots \sum_{k_{d-1}} A_{i,k_1} A_{k_1,k_2} \cdots A_{k_{d-1},j}$$

  由于稀疏性，复杂度为 $O(s^{d-1})$

- 对于 x^d，存在次数为 $\Theta(\sqrt{d})$ 的逼近多项式，因此结果可改进到 $s^{O(\sqrt{d})}$ [Sachdeva & Vishnoi, 2014]
- 该问题是 BQP-完全问题 [Janzing & Wocjan, 2007]
- 我们的结果: $\widetilde{\Omega}((s/2)^{(\sqrt{d}-1)/6})$

量子算法：

- 复杂度上界 $O(s\sqrt{d}/\varepsilon)$ [Gilyén, Su, Low, Wiebe, 2018]
- 我们的结果: 复杂度下界 $\Omega(\sqrt{d})$

一般性结果

设 f 连续，A 稀疏厄米，则计算 $f(A)_{i,j} \pm \varepsilon$ 的复杂度如下

	量子算法	经典算法
复杂度上界	$O(sd/\varepsilon)$	$O(s^{d-1})$
复杂度下界	$\Omega(d)$	$\Omega((s/2)^{(d-1)/6})$

其中逼近次数（approximate degree 标准记号 $\widetilde{\deg}_\varepsilon(f)$）

$$d = \min\{d' : |f(x) - g(x)| \le \varepsilon, \forall x \in [-1,1], \\ g(x) \text{ 是一个次数为 } d' \text{ 的多项式}\}.$$

逼近次数举例

前面提到了三个 BQP-完全问题对应于三个函数：e^{ixt}, $1/x$, x^d

逼近次数举例

前面提到了三个 BQP-完全问题对应于三个函数：e^{ixt}, $1/x$, x^d

- $\widetilde{\deg}_{1/3}(e^{ixt}) = \Theta(t)$,

逼近次数举例

前面提到了三个 BQP-完全问题对应于三个函数：e^{ixt}，$1/x$，x^d

- $\widetilde{\deg}_{1/3}(e^{ixt}) = \Theta(t)$，因此，量子 vs 经典 $= t$ vs c^t，其中 $(s/2)^{1/6} \leq c \leq s$

逼近次数举例

前面提到了三个 BQP-完全问题对应于三个函数：e^{ixt}，$1/x$，x^d

- $\widetilde{\deg}_{1/3}(e^{ixt}) = \Theta(t)$，因此，量子 vs 经典 $= t$ vs c^t，其中 $(s/2)^{1/6} \leq c \leq s$
- $\widetilde{\deg}_{1/3}(1/x) = \Theta(\kappa)$，这里假定区间为 $[-1, 1/\kappa] \cup [1/\kappa, 1]$，

逼近次数举例

前面提到了三个 BQP-完全问题对应于三个函数：e^{ixt}，$1/x$，x^d

- $\widetilde{\deg}_{1/3}(e^{ixt}) = \Theta(t)$，因此，量子 vs 经典 $= t$ vs c^t，其中 $(s/2)^{1/6} \le c \le s$
- $\widetilde{\deg}_{1/3}(1/x) = \Theta(\kappa)$，这里假定区间为 $[-1, 1/\kappa] \cup [1/\kappa, 1]$，因此，量子 vs 经典 $= \kappa$ vs c^κ

逼近次数举例

前面提到了三个 BQP-完全问题对应于三个函数：e^{ixt}，$1/x$，x^d

- $\widetilde{\deg}_{1/3}(e^{ixt}) = \Theta(t)$，因此，量子 vs 经典 $= t$ vs c^t，其中 $(s/2)^{1/6} \leq c \leq s$
- $\widetilde{\deg}_{1/3}(1/x) = \Theta(\kappa)$，这里假定区间为 $[-1, 1/\kappa] \cup [1/\kappa, 1]$，因此，量子 vs 经典 $= \kappa$ vs c^κ
- $\widetilde{\deg}_{1/3}(x^d) = \Theta(\sqrt{d})$,

逼近次数举例

前面提到了三个 BQP-完全问题对应于三个函数：e^{ixt}，$1/x$，x^d

- $\widetilde{\deg}_{1/3}(e^{ixt}) = \Theta(t)$，因此，量子 vs 经典 $= t$ vs c^t，其中 $(s/2)^{1/6} \leq c \leq s$
- $\widetilde{\deg}_{1/3}(1/x) = \Theta(\kappa)$，这里假定区间为 $[-1, 1/\kappa] \cup [1/\kappa, 1]$，因此，量子 vs 经典 $= \kappa$ vs c^κ
- $\widetilde{\deg}_{1/3}(x^d) = \Theta(\sqrt{d})$，因此，量子 vs 经典 $= \sqrt{d}$ vs $c^{\sqrt{d}}$

量子算法的复杂度下界估计方法：多项式方法 [Beals, Buhrman, Cleve, Mosca, de Wolf 1998]

- 设 $f(x_1, \ldots, x_n) : \{\pm 1\}^n \to \{0, 1\}$ 是一个布尔函数，给定 $\mathcal{O} : i \to x_i$，使用尽可能少次数的 $\mathcal{O}$ 来计算 $f(x_1, \ldots, x_n)$

量子算法的复杂度下界估计方法：多项式方法 [Beals, Buhrman, Cleve, Mosca, de Wolf 1998]

- 设 $f(x_1,\ldots,x_n):\{\pm1\}^n\to\{0,1\}$ 是一个布尔函数，给定 $\mathcal{O}:i\to x_i$，使用尽可能少次数的 $\mathcal{O}$ 来计算 $f(x_1,\ldots,x_n)$
- 复杂度为 T 的量子算法形如：

$$U_T\mathcal{O}\cdots U_1\mathcal{O}U_0|0..0\rangle=\sqrt{1-p_1(x)^2}|0\rangle|\psi_0\rangle+p_1(x)|1\rangle|\psi_1\rangle,$$

其中 $U_0,U_1,\ldots,U_T$ 是和 $\mathcal{O}$ 无关的酉矩阵

量子算法的复杂度下界估计方法：多项式方法 [Beals, Buhrman, Cleve, Mosca, de Wolf 1998]

- 设 $f(x_1,\ldots,x_n):\{\pm1\}^n\to\{0,1\}$ 是一个布尔函数，给定 $\mathcal{O}:i\to x_i$，使用尽可能少次数的 $\mathcal{O}$ 来计算 $f(x_1,\ldots,x_n)$
- 复杂度为 T 的量子算法形如：

$$U_T\mathcal{O}\cdots U_1\mathcal{O}U_0|0..0\rangle=\sqrt{1-p_1(x)^2}|0\rangle|\psi_0\rangle+p_1(x)|1\rangle|\psi_1\rangle,$$

其中 $U_0,U_1,\ldots,U_T$ 是和 $\mathcal{O}$ 无关的酉矩阵

- 输出为 1，如果 $p_1(x)^2\geq 2/3$，否则输出为 0

量子算法的复杂度下界估计方法：多项式方法 [Beals, Buhrman, Cleve, Mosca, de Wolf 1998]

- 设 $f(x_1,\ldots,x_n):\{\pm1\}^n\to\{0,1\}$ 是一个布尔函数，给定 $\mathcal{O}:i\to x_i$，使用尽可能少次数的 $\mathcal{O}$ 来计算 $f(x_1,\ldots,x_n)$
- 复杂度为 T 的量子算法形如：

$$U_T\mathcal{O}\cdots U_1\mathcal{O}U_0|0..0\rangle=\sqrt{1-p_1(x)^2}|0\rangle|\psi_0\rangle+p_1(x)|1\rangle|\psi_1\rangle,$$

  其中 $U_0,U_1,\ldots,U_T$ 是和 $\mathcal{O}$ 无关的酉矩阵
- 输出为 1，如果 $p_1(x)^2\geq 2/3$，否则输出为 0
- 令 $g(x)=p_1(x)^2$，次数 $\leq 2T$ 且 $|f(x)-g(x)|\leq 1/3$，因此

$$T\geq\frac{1}{2}\deg(g)\geq\frac{1}{2}\widetilde{\deg}_{1/3}(f)$$

量子算法的复杂度下界估计方法：多项式方法 [Beals, Buhrman, Cleve, Mosca, de Wolf 1998]

- 设 $f(x_1,\ldots,x_n):\{\pm1\}^n\to\{0,1\}$ 是一个布尔函数，给定 $\mathcal{O}:i\to x_i$，使用尽可能少次数的 $\mathcal{O}$ 来计算 $f(x_1,\ldots,x_n)$
- 复杂度为 T 的量子算法形如：

$$U_T\mathcal{O}\cdots U_1\mathcal{O}U_0|0..0\rangle=\sqrt{1-p_1(x)^2}|0\rangle|\psi_0\rangle+p_1(x)|1\rangle|\psi_1\rangle,$$

  其中 $U_0,U_1,\ldots,U_T$ 是和 $\mathcal{O}$ 无关的酉矩阵
- 输出为 1，如果 $p_1(x)^2\geq 2/3$，否则输出为 0
- 令 $g(x)=p_1(x)^2$，次数 $\leq 2T$ 且 $|f(x)-g(x)|\leq 1/3$，因此

$$T\geq\frac{1}{2}\deg(g)\geq\frac{1}{2}\widetilde{\deg}_{1/3}(f)$$

- 该证明对于矩阵函数不可行，一个原因是 $\mathcal{O}_1$ 无多项式表示

关键定理

定理

设 $f(x):[-1,1]\to[-1,1]$ 是连续函数，设 $f_{\mathrm{odd}}(x), f_{\mathrm{even}}(x)$ 是奇、偶部分。则

- 存在对称的三对角阵 A 满足 $\|A\|\le 1$ 使得 $f(A)_{1,n}=\varepsilon$，其中 $n=\widetilde{\deg}_{\varepsilon}(f_{\mathrm{odd}})+O(1)$
- 存在对称的三对角阵 A 满足 $\|A\|\le 1$ 使得 $f(A)_{2,n-1}=\varepsilon$，其中 $n=\widetilde{\deg}_{\varepsilon}(f_{\mathrm{even}})+O(1)$

证明：线性半无限规划问题 + 三对角阵的性质 □

量子算法的复杂度下界

奇偶性判定问题（Parity problem）：给定 $x_1, \ldots, x_n \in \{0,1\}$，计算 $x_1 \oplus x_2 \oplus \cdots \oplus x_n$。经典/量子算法复杂度都为 $\Theta(n)$

量子算法的复杂度下界

奇偶性判定问题（Parity problem）：给定 $x_1, \ldots, x_n \in \{0,1\}$，计算 $x_1 \oplus x_2 \oplus \cdots \oplus x_n$。经典/量子算法复杂度都为 $\Theta(n)$

构造带权图 G：

- **顶点：** (i,t)，其中 $i \in \{0,1,\ldots,n\}, t \in \{0,1\}$
- **边：** $(i-1,t)$ 和 $(i, t \oplus x_i)$ 之间存在一条边

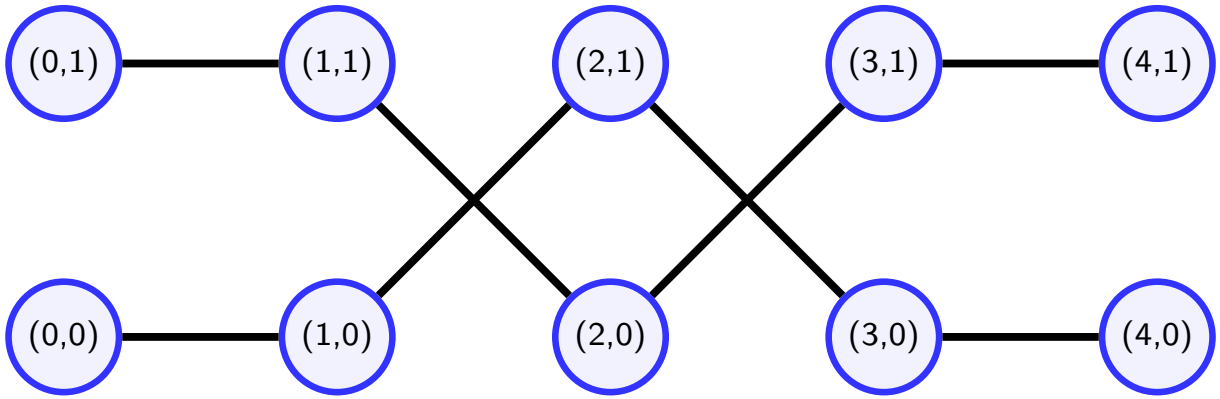
量子算法的复杂度下界

奇偶性判定问题（Parity problem）：给定 $x_1, \ldots, x_n \in \{0,1\}$，计算 $x_1 \oplus x_2 \oplus \cdots \oplus x_n$。经典/量子算法复杂度都为 $\Theta(n)$

构造带权图 G：

- **顶点：** (i,t)，其中 $i \in \{0,1,\ldots,n\}, t \in \{0,1\}$
- **边：** $(i-1,t)$ 和 $(i, t \oplus x_i)$ 之间存在一条边

例如 $(x_1,x_2,x_3,x_4) = (0,1,1,0)$，则 G 为



量子算法的复杂度下界

本质上，G 由两条道路构成

$$(0,0)-(1,x_0)-(2,x_0\oplus x_1)-\cdots-(n,x_0\oplus\cdots\oplus x_{n-1})$$
$$(0,1)-(1,1\oplus x_0)-(2,1\oplus x_0\oplus x_1)-\cdots-(n,1\oplus x_0\oplus\cdots\oplus x_{n-1})$$

量子算法的复杂度下界

本质上，G 由两条道路构成

$$(0,0)-(1,x_0)-(2,x_0\oplus x_1)-\cdots-(n,x_0\oplus\cdots\oplus x_{n-1})$$
$$(0,1)-(1,1\oplus x_0)-(2,1\oplus x_0\oplus x_1)-\cdots-(n,1\oplus x_0\oplus\cdots\oplus x_{n-1})$$

设 A 是图的邻接矩阵（本质上是两个三对角型矩阵），则

- **情形 1：** 若 $x_1\oplus x_2\oplus\cdots\oplus x_n=0$，则 $\langle 0,0|f(A)|n,1\rangle=0$

量子算法的复杂度下界

本质上，G 由两条道路构成

$$(0,0)-(1,x_0)-(2,x_0\oplus x_1)-\cdots-(n,x_0\oplus\cdots\oplus x_{n-1})$$
$$(0,1)-(1,1\oplus x_0)-(2,1\oplus x_0\oplus x_1)-\cdots-(n,1\oplus x_0\oplus\cdots\oplus x_{n-1})$$

设 A 是图的邻接矩阵（本质上是两个三对角型矩阵），则

- **情形 1：** 若 $x_1\oplus x_2\oplus\cdots\oplus x_n=0$，则 $\langle 0,0|f(A)|n,1\rangle=0$
- **情形 2：** 若 $x_1\oplus x_2\oplus\cdots\oplus x_n=1$，则希望选取合适的权重，使得 $\langle 0,0|f(A)|n,1\rangle\geq\varepsilon$，权重由关键定理给出

Forrelation 问题 (Aaronson & Ambainis, 2015)

给定 $g_1, g_2 : \{0,1\}^n \to \{\pm 1\}$，令 $D_i = \mathsf{diag}(g_i(x) : x \in \{0,1\}^n)$，$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$。设

$$\begin{aligned} \Phi(g_1, g_2) &:= \langle 0^n | H^{\otimes n} D_1 H^{\otimes n} D_2 H^{\otimes n} | 0^n \rangle \\ &= \frac{1}{2^{3n/2}} \sum_{x,y \in \{0,1\}^n} (-1)^{x \cdot y} f(x) g(y). \end{aligned}$$

目标是计算 $\Phi(g_1, g_2) \pm 1/3$

Forrelation 问题 (Aaronson & Ambainis, 2015)

给定 $g_1, g_2 : \{0,1\}^n \to \{\pm 1\}$，令 $D_i = \mathsf{diag}(g_i(x) : x \in \{0,1\}^n)$，$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$。设

$$\begin{aligned} \Phi(g_1, g_2) &:= \langle 0^n | H^{\otimes n} D_1 H^{\otimes n} D_2 H^{\otimes n} | 0^n \rangle \\ &= \frac{1}{2^{3n/2}} \sum_{x,y \in \{0,1\}^n} (-1)^{x \cdot y} f(x) g(y). \end{aligned}$$

目标是计算 $\Phi(g_1, g_2) \pm 1/3$

关于该问题，经典算法的复杂度下界为 $\Omega(\sqrt{2^n}/n)$，量子算法的复杂度上界为 $O(1)$

Feynman 的 Clock 构造法

设 $U = U_{N-1} \cdots U_2 U_1$ 是一个酉算子，令

$$A = \begin{pmatrix} 0 & b_1 U_1^\dagger & & \\ b_1 U_1 & 0 & b_2 U_2^\dagger & \\ & b_2 U_2 & \ddots & \ddots \\ & & \ddots & \ddots \end{pmatrix}$$

Feynman 的 Clock 构造法

设 $U = U_{N-1} \cdots U_2 U_1$ 是一个酉算子，令

$$A = \begin{pmatrix} 0 & b_1 U_1^\dagger & & \\ b_1 U_1 & 0 & b_2 U_2^\dagger & \\ & b_2 U_2 & \ddots & \ddots \\ & & \ddots & \ddots \end{pmatrix}$$

设 $|\psi_t\rangle := |t\rangle \otimes U_t \cdots U_1 |0\rangle$，则

$$A|\psi_t\rangle = b_{t-1}|\psi_{t-1}\rangle + b_{t+1}|\psi_{t+1}\rangle$$

因此在子空间 $\{|\psi_t\rangle : t = 0, 1, \ldots, N-1\}$ 中，A 是一个三对角阵

经典算法的复杂度下界

在 Forrelation 问题中，$U = H^{\otimes n} D_1 H^{\otimes n} D_2 H^{\otimes n}$，为确保 A 的稀疏性，视

$$H^{\otimes n} = (H \otimes I \otimes \cdots \otimes I)(I \otimes H \otimes \cdots \otimes I) \cdots (I \otimes I \otimes \cdots \otimes H)$$

经典算法的复杂度下界

在 Forrelation 问题中，$U = H^{\otimes n} D_1 H^{\otimes n} D_2 H^{\otimes n}$，为确保 A 的稀疏性，视

$$H^{\otimes n} = (H \otimes I \otimes \cdots \otimes I)(I \otimes H \otimes \cdots \otimes I) \cdots (I \otimes I \otimes \cdots \otimes H)$$

这时 $N = 3n + 2$,

$$\begin{aligned} |\psi_0\rangle &= |0\rangle \otimes |0\rangle \\ |\psi_{N-1}\rangle &= |N-1\rangle \otimes H^{\otimes n} D_1 H^{\otimes n} D_2 H^{\otimes n} |0\rangle \end{aligned}$$

经典算法的复杂度下界

在 Forrelation 问题中，$U = H^{\otimes n} D_1 H^{\otimes n} D_2 H^{\otimes n}$，为确保 A 的稀疏性，视

$$H^{\otimes n} = (H \otimes I \otimes \cdots \otimes I)(I \otimes H \otimes \cdots \otimes I) \cdots (I \otimes I \otimes \cdots \otimes H)$$

这时 $N = 3n + 2$,

$$\begin{aligned} |\psi_0\rangle &= |0\rangle \otimes |0\rangle \\ |\psi_{N-1}\rangle &= |N-1\rangle \otimes H^{\otimes n} D_1 H^{\otimes n} D_2 H^{\otimes n} |0\rangle \end{aligned}$$

令 $|\phi_{N-1}\rangle = |N-1\rangle \otimes |0\rangle$，则

$$\langle \phi_{N-1} | f(A) | \psi_0 \rangle = \langle \psi_{N-1} | f(A) | \psi_0 \rangle \cdot \Phi(g_1, g_2)$$

经典算法的复杂度下界

在 Forrelation 问题中，$U = H^{\otimes n} D_1 H^{\otimes n} D_2 H^{\otimes n}$，为确保 A 的稀疏性，视

$$H^{\otimes n} = (H \otimes I \otimes \cdots \otimes I)(I \otimes H \otimes \cdots \otimes I) \cdots (I \otimes I \otimes \cdots \otimes H)$$

这时 $N = 3n + 2$,

$$\begin{aligned} |\psi_0\rangle &= |0\rangle \otimes |0\rangle \\ |\psi_{N-1}\rangle &= |N-1\rangle \otimes H^{\otimes n} D_1 H^{\otimes n} D_2 H^{\otimes n} |0\rangle \end{aligned}$$

令 $|\phi_{N-1}\rangle = |N-1\rangle \otimes |0\rangle$，则

$$\begin{array}{ccccc} \langle \phi_{N-1}|f(A)|\psi_0\rangle & = & \langle \psi_{N-1}|f(A)|\psi_0\rangle \cdot \Phi(g_1, g_2) \\ \downarrow & & \downarrow & & \downarrow \\ f(A) \text{ 的一个分量} & & \text{易} & & \text{难} \end{array}$$

关键定理（回顾）

定理

设 $f(x):[-1,1]\to[-1,1]$ 是连续函数，设 $f_{\mathrm{odd}}(x), f_{\mathrm{even}}(x)$ 是奇、偶部分。则

- 存在对称的三对角阵 A 满足 $\|A\|\le 1$ 使得 $f(A)_{1,n}=\varepsilon$，其中 $n=\widetilde{\deg}_{\varepsilon}(f_{\mathrm{odd}})+O(1)$
- 存在对称的三对角阵 A 满足 $\|A\|\le 1$ 使得 $f(A)_{2,n-1}=\varepsilon$，其中 $n=\widetilde{\deg}_{\varepsilon}(f_{\mathrm{even}})+O(1)$

对偶多项式方法

回顾（approximate degree）：

$$\widetilde{\deg}_\varepsilon(f) = \min\{d : |f(x) - g(x)| \le \varepsilon, \forall x \in [-1,1], \\ g(x) \text{ 是一个次数为} d \text{ 的多项式}\}.$$

对偶多项式方法

回顾（approximate degree）：

$$\widetilde{\deg}_\varepsilon(f) = \min\{d : |f(x)-g(x)| \le \varepsilon, \forall x \in [-1,1], \\ g(x) \text{ 是一个次数为} d \text{ 的多项式}\}.$$

考虑线性半无限规划（linear semi-infinite program）：

$$\begin{aligned} \min_{g,\delta} \quad & \delta \\ \text{s.t.} \quad & |f(x)-g(x)| \le \delta, \quad \forall x \in [-1,1], \\ & \deg(g) \le d. \end{aligned}$$

对偶多项式方法

回顾（approximate degree）：

$$\widetilde{\deg}_\varepsilon(f) = \min\{d : |f(x)-g(x)| \le \varepsilon, \forall x \in [-1,1], \\ g(x) \text{ 是一个次数为} d \text{ 的多项式}\}.$$

考虑线性半无限规划（linear semi-infinite program）：

$$\begin{aligned} \min_{g,\delta} \quad & \delta \\ \text{s.t.} \quad & |f(x)-g(x)| \le \delta, \quad \forall x \in [-1,1], \\ & \deg(g) \le d. \end{aligned}$$

最优值记成 $\mathrm{Val}(f(x),d)$，容易验证

$$\widetilde{\deg}_\varepsilon(f) = \min\{d : \mathrm{Val}(f(x),d) = \varepsilon\}$$

对偶多项式方法

对偶形式

$$\max_{h} \quad \int_{-1}^{1} f(x)h(x)\,dx \tag{1}$$

$$\text{s.t.} \quad \int_{-1}^{1} h(x)x^k\,dx = 0, \quad k \in \{0,1,\ldots,d\}, \tag{2}$$

$$\int_{-1}^{1} |h(x)|\,dx = 1. \tag{3}$$

对偶多项式方法

对偶形式

$$\max_{h} \quad \int_{-1}^{1} f(x)h(x)\,dx \tag{1}$$

$$\text{s.t.} \quad \int_{-1}^{1} h(x)x^k\,dx = 0, \quad k \in \{0,1,\ldots,d\}, \tag{2}$$

$$\int_{-1}^{1} |h(x)|\,dx = 1. \tag{3}$$

- (2) 说明 $h(x)$ 和低次数的单项式正交 $\Rightarrow \deg(h)$ 高

对偶多项式方法

对偶形式

$$\max_{h} \quad \int_{-1}^{1} f(x)h(x)\,dx \tag{1}$$

$$\text{s.t.} \quad \int_{-1}^{1} h(x)x^k\,dx = 0, \quad k \in \{0,1,\ldots,d\}, \tag{2}$$

$$\int_{-1}^{1} |h(x)|\,dx = 1. \tag{3}$$

- (2) 说明 $h(x)$ 和低次数的单项式正交 $\Rightarrow \deg(h)$ 高
- (3) 说明 $h(x)$ 的 L_1 范数为 1

对偶多项式方法

对偶形式

$$\max_{h} \quad \int_{-1}^{1} f(x)h(x)\,dx \tag{1}$$

$$\text{s.t.} \quad \int_{-1}^{1} h(x)x^k\,dx = 0, \quad k \in \{0,1,\ldots,d\}, \tag{2}$$

$$\int_{-1}^{1} |h(x)|\,dx = 1. \tag{3}$$

- (2) 说明 $h(x)$ 和低次数的单项式正交 $\Rightarrow \deg(h)$ 高
- (3) 说明 $h(x)$ 的 L_1 范数为 1
- 最优值为 ε，(1) $\Rightarrow$ 说明 $f(x)$ 和 $h(x)$ 的相关性弱，也即 $\deg(f)$ 低

求解线性半无限规划的离散化方法

存在 $\{\lambda_1, \ldots, \lambda_n\}$，$n \leq d+2$，使得原问题等价于

$$\begin{aligned} \min_{g,\delta} \quad & \delta \\ \text{s.t.} \quad & |f(\lambda_i) - g(\lambda_i)| \leq \delta, \quad \forall i = 1, \ldots, n, \\ & \deg(g) \leq d. \end{aligned}$$

求解线性半无限规划的离散化方法

存在 $\{\lambda_1, \ldots, \lambda_n\}$，$n \le d+2$，使得原问题等价于

$$\begin{aligned} \min_{g,\delta} \quad & \delta \\ \text{s.t.} \quad & |f(\lambda_i) - g(\lambda_i)| \le \delta, \quad \forall i = 1, \ldots, n, \\ & \deg(g) \le d. \end{aligned}$$

对偶形式

$$\begin{aligned} \max_{h_i} \quad & \sum_{i=1}^{n} f(\lambda_i) h_i \\ \text{s.t.} \quad & \sum_{i=1}^{n} h_i \lambda_i^k = 0, \quad \forall k \in \{0, 1, \ldots, d\}, \\ & \sum_{i=1}^{n} |h_i| = 1. \end{aligned}$$

求解线性半无限规划的离散化方法

- 约束

$$\sum_{i=1}^{n} h_i \lambda_i^k = 0 \quad \forall k \in \{0, 1, \ldots, d\}$$

是关于 h_i 的一个范德蒙线性方程组，秩为 $\min(d+1, n)$，因此要存在非平凡解，只能有 $n = d+2$。方程解满足

$$h_i = \frac{\alpha}{\prod_{j \neq i}(\lambda_i - \lambda_j)}, \quad i \in \{1, 2, \ldots, n\}.$$

求解线性半无限规划的离散化方法

- 约束

$$\sum_{i=1}^{n} h_i \lambda_i^k = 0 \quad \forall k \in \{0, 1, \ldots, d\}$$

是关于 h_i 的一个范德蒙线性方程组，秩为 $\min(d+1, n)$，因此要存在非平凡解，只能有 $n = d+2$。方程解满足

$$h_i = \frac{\alpha}{\prod_{j \neq i}(\lambda_i - \lambda_j)}, \quad i \in \{1, 2, \ldots, n\}.$$

- 约束

$$\sum_{i=1}^{n} |h_i| = 1$$

确定了参数 α 的值

三对角型矩阵

$$A = \begin{pmatrix} 0 & b_1 & & \\ b_1 & \ddots & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ & & b_{n-1} & 0 \end{pmatrix}$$

设特征值为 $\{\lambda_i\}_{i=1}^n$，则 $f(A)_{1,n} = b_1 b_2 \cdots b_{n-1} \sum_{i=1}^{n} \frac{f(\lambda_i)}{\prod_{j\neq i}(\lambda_i - \lambda_j)}$.

三对角型矩阵

$$A = \begin{pmatrix} 0 & b_1 & & \\ b_1 & \ddots & \ddots & \\ & \ddots & \ddots & b_{n-1} \\ & & b_{n-1} & 0 \end{pmatrix}$$

设特征值为 $\{\lambda_i\}_{i=1}^n$，则 $f(A)_{1,n} = b_1 b_2 \cdots b_{n-1} \sum_{i=1}^n \frac{f(\lambda_i)}{\prod_{j\neq i}(\lambda_i - \lambda_j)}$.

引理 (关键引理)

对任意的 $x_m > \cdots > x_1 > 0$, 存在唯一的 A 使得 $n = 2m$, $b_1, \ldots, b_{n-1} > 0$，其特征值为 $\{\lambda_i\}_{i=1}^n = \{\pm x_i\}_{i=1}^m$，且

$$\sum_{i=1}^m \frac{b_1 b_2 \cdots b_{2m-1}}{|\prod_{j\neq i}(\lambda_i - \lambda_j)|} = 1.$$

对偶多项式方法与三对角阵的关系

对偶多项式方法	三对角阵
$\sum_{i=1}^{n} \frac{\alpha}{\lvert\prod_{j\neq i}(\lambda_i-\lambda_j)\rvert} = 1$	$\sum_{i=1}^{m} \frac{b_1\cdots b_{2m-1}}{\lvert\prod_{j\neq i}(\lambda_i-\lambda_j)\rvert} = 1$
（约束条件）	
$\sum_{i=1}^{n} \frac{\alpha f(\lambda_i)}{\prod_{j\neq i}(\lambda_i-\lambda_j)}$	$f(A)_{1,2m} = \sum_{i=1}^{2m} \frac{b_1\cdots b_{2m-1} f(\lambda_i)}{\prod_{j\neq i}(\lambda_i-\lambda_j)}$
（最优值）	

总结

- 对于矩阵函数问题，我们有

	量子算法	经典算法
复杂度上界	$O(sd/\varepsilon)$	$O(s^{d-1})$
复杂度下界	$\Omega(d)$	$\Omega((s/2)^{(d-1)/6})$

总结

- 对于矩阵函数问题，我们有

	量子算法	经典算法
复杂度上界	$O(sd/\varepsilon)$	$O(s^{d-1})$
复杂度下界	$\Omega(d)$	$\Omega((s/2)^{(d-1)/6})$

- 关于逼近次数 d，证明了量子算法的最优性

总结

- 对于矩阵函数问题，我们有

	量子算法	经典算法
复杂度上界	$O(sd/\varepsilon)$	$O(s^{d-1})$
复杂度下界	$\Omega(d)$	$\Omega((s/2)^{(d-1)/6})$

- 关于逼近次数 d，证明了量子算法的最优性
- 对于矩阵函数问题，证明了量子和经典算法的差距是指数级

一些问题

- 对精度 ε 的下界估计，$\Omega(1/\varepsilon)$?

一些问题

- 对精度 ε 的下界估计，$\Omega(1/\varepsilon)$?
- 经典算法的上、下界问题?

一些问题

- 对精度 ε 的下界估计，$\Omega(1/\varepsilon)$?
- 经典算法的上、下界问题?
- 应用?

一些问题

- 对精度 ε 的下界估计，$\Omega(1/\varepsilon)$?
- 经典算法的上、下界问题?
- 应用?

谢 谢!