



Algoritmo de Grover Aplicado ao Problema de Geometria de Distâncias Moleculares

Andrês Rodrigues Oliveira

Resumo

Utilizando conceitos básicos de Álgebra Linear, podemos criar uma “linguagem” para trabalharmos em Computação Quântica. Apesar de ser ainda uma promessa para o futuro, já há resultados teóricos importantes, mostrando algoritmos capazes de superar os computadores clássicos, como o Algoritmo de Grover. O problema de Geometria de Distâncias Moleculares é um importante problema na área de bioquímica computacional, que consiste em determinar a estrutura 3D de uma molécula conhecendo algumas distâncias entre seus átomos. O presente trabalho discute como podemos aplicar o Algoritmo de Grover em tal problema.

Palavras-chave: Computação Quântica, Algoritmo de Grover, Geometria de Distâncias Moleculares

1 Introdução

O cálculo da estrutura 3D de uma molécula de proteína é um problema fundamental da biologia computacional, pois o desenvolvimento de novos medicamentos é baseado nas estruturas das proteínas envolvidas nesse complexo e altamente custoso procedimento da indústria farmacêutica.

Uma das maneiras de se fazer isso é utilizando informações coletadas por experimentos de Ressonância Magnética Nuclear (RMN). Esses dados coletados da RMN fornecem distâncias entre átomos próximos da molécula (Donald, 2011). Então é esse o nosso problema: determinar a estrutura 3D de uma molécula de proteína conhecendo algumas distâncias entre seus átomos, problema este conhecido na literatura como *Molecular Distance Geometry Problem* (MDGP) (Crippen et al., 1988; Liberti et al., 2014), um problema NP-Difícil (Saxe, 1979).

Para tentar resolver esse problema de geometria molecular, será introduzido um algoritmo de computação quântica chamado Algoritmo de Grover. Esse procedimento é baseado no texto de Lavor et al. (2005).

2 Algoritmo de Grover

Em computação, um dos principais problemas são os de busca, ou seja, dada uma lista de elementos, queremos encontrar um elemento específico desta lista. Seja o conjunto $\{0, 1, \dots, N -$



1}, em que $N = 2^n$ com $n \in \mathbb{N}$ e seja a função

$$f : \{0, 1, \dots, N - 1\} \longrightarrow \{0, 1\}$$

dada por

$$f(i) = \begin{cases} 1, & \text{se } i = i_0, \\ 0, & \text{caso contrário,} \end{cases} \quad (2.1)$$

em que $i_0 \in \{0, 1, \dots, N - 1\}$ é o elemento procurado. Vamos assumir aqui que esse elemento é único.

Para encontrar tal i_0 , classicamente, teríamos que chamar a função f para cada elemento de $\{0, 1, \dots, N - 1\}$ e verificar se aquele elemento é o nosso i_0 , ou seja, sua complexidade é $O(N)$. Apresentaremos aqui o Algoritmo de Grover (Grover, 1996, 1997), um algoritmo quântico que resolve esse problema com complexidade $O(\sqrt{N})$.

Para isso precisaremos de dois registradores quânticos: o primeiro terá n qubits e o segundo tem apenas um qubit. O estado inicial do primeiro registrador será dado por

$$|\psi\rangle = \frac{1}{\sqrt{N}} \sum_{i=0}^{N-1} |i\rangle,$$

enquanto o segundo registrador terá seu estado inicial dado por

$$|-\rangle = \frac{|0\rangle - |1\rangle}{\sqrt{2}}.$$

O operador de grover G será a composição dos operadores,

$$G = ((H^{\otimes n}(2|0\rangle\langle 0| - I)H^{\otimes n}) \otimes I) U_f, \quad (2.2)$$

em que H é o operador de Hadamard, I é o operador de Identidade e U_f é chamado de *oráculo* (a versão quântica da função f), dado por

$$U_f(|i\rangle|j\rangle) = |i\rangle|j \oplus f(i)\rangle. \quad (2.3)$$

A complexidade deste algoritmo será a quantidade t de vezes que teremos que aplicar este operador G no estado inicial $|\psi\rangle|-\rangle$. Sendo assim, com $t = \left\lfloor \frac{\pi}{4}\sqrt{N} \right\rfloor$, após uma medição no primeiro registrador do estado resultante $G^t(|\psi\rangle|-\rangle)$, obteremos $|i_0\rangle$ com uma probabilidade maior ou igual a $1 - \frac{1}{\sqrt{N}}$.

3 Geometria de Distâncias Moleculares

Definição 1 (*Molecular Distance Geometry Problem*). Seja $G = (V, E, d)$ um grafo simples, conectado, não direcionado e ponderado nas arestas por $d : E \longrightarrow (0, \infty)$, em que os vértices V representam os átomos da molécula e as arestas E , com peso d , representam as distâncias conhecidas entre os átomos da molécula. Queremos encontrar uma imersão desse grafo no \mathbb{R}^3 , ou seja, encontrar uma função

$$\begin{aligned} \mathbf{x} : V &\rightarrow \mathbb{R}^3 \\ v &\mapsto \mathbf{x}(v) \end{aligned}$$



tal que

$$\forall(u, v) \in E, \|\mathbf{x}_u - \mathbf{x}_v\| = d_{u,v} \quad (3.1)$$

onde $\|\cdot\|$ é a norma euclidiana (utilizaremos a notação \mathbf{x}_v para $\mathbf{x}(v)$ e $d_{u,v}$ para $d(u, v)$). Chamaremos \mathbf{x}_v de uma realização para o vértice v e a função \mathbf{x} apenas de realização.

Sendo \mathbf{x} uma realização, podemos definir uma função que associa essa realização a um número real:

$$g(\mathbf{x}) = \sum_{(u,v) \in E} (\|\mathbf{x}_u - \mathbf{x}_v\|^2 - d_{u,v}^2)^2. \quad (3.2)$$

É fácil ver que \mathbf{x} é uma realização que satisfaz (3.1) se, e somente se, $g(\mathbf{x}) = 0$.

Se tivermos neste grafo $G = (V, E, d)$ de um MDGP, com $n = |V|$, uma ordem total em V tal que:

- Para todo $k \in \{4, \dots, n\}$ temos

$$\{(v_{k-3}, v_k), (v_{k-2}, v_k), (v_{k-1}, v_k)\} \subseteq E;$$

- Para todo $k \in \{3, \dots, n\}$ os vértices $\{v_{k-2}, v_{k-1}, v_k\}$ formam uma clique e vale a desigualdade

$$d_{v_{k-2}, v_k} < d_{v_{k-2}, v_{k-1}} + d_{v_{k-1}, v_k},$$

então estaremos em um caso particular chamado *Discretizable Molecular Distance Geometry Problem* (DMDGP) (Lavor et al., 2012) pois com essa ordem conseguimos garantir que existem finitas soluções. Para facilitar a notação, vamos denotar o vértice v_k simplesmente por k , bem como a realização de v_k por apenas \mathbf{x}_k , em que $k \in \{1, 2, \dots, n\}$. Podemos então entender uma realização \mathbf{x} como sendo uma sequência de n pontos no \mathbb{R}^3 , ou seja, $\mathbf{x} = (\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n)$.

O que queremos fazer é usar o Algoritmo de Grover para encontrar soluções do DMDGP, para isso temos que associar cada estado do primeiro registrador $\{|0\rangle, |1\rangle, \dots, |N-1\rangle\}$ a uma solução do DMDGP. Tomando $N = 2^{n-3}$, onde n é a quantidade de átomos no DMDGP, então teremos $n-3$ *qubits* no primeiro registrador. Para isso, vamos descrever nossa entrada utilizando coordenadas internas (Lavor et al., 2012), ou seja, com distâncias de ligações covalentes $d_{k-1,k}$, ângulos planos $\theta_{k-2,k}$ e ângulos de torção $\omega_{k-3,k}$. Existe uma fórmula fechada para o cosseno do ângulo de torção. Isso significa que podemos encontrar uma solução para o problema escolhendo o sinal de $\sin(\omega_{k-3,k}) = \pm \sqrt{1 - \cos^2(\omega_{k-3,k})}$. Para o q -ésimo *qubit*, com $q \in \{0, 1, \dots, (n-3)-1\}$,

- Vamos associar o estado $|0\rangle$ com $\sin(\omega_{q+1,q+4}) = +\sqrt{1 - \cos^2(\omega_{q+1,q+4})}$;
- Vamos associar o estado $|1\rangle$ com $\sin(\omega_{q+1,q+4}) = -\sqrt{1 - \cos^2(\omega_{q+1,q+4})}$.

Sendo assim, tendo as coordenadas internas de uma solução, existe uma fórmula para retornar para coordenadas cartesianas (ou seja, a solução no \mathbb{R}^3) (Phillips et al., 1994).

Dado um candidato a solução $|i\rangle$, com $i \in \{0, 1, \dots, N-1\}$, vamos definir a função

$$h: \{0, 1, \dots, N-1\} \rightarrow \mathbb{R}^{3n},$$



que recebe um inteiro $i \in \{0, 1, \dots, N - 1\}$ e associa a uma realização

$$h(i) = \mathbf{x}^i = (\mathbf{x}_1^i, \mathbf{x}_2^i, \dots, \mathbf{x}_n^i),$$

realização essa com o valor dos senos e cossenos associados aos *qubits* do estado $|i\rangle$. Já vimos uma função g (3.2) que verifica uma solução. Sendo assim, vamos definir

$$g(h(i)) = \sum_{(u,v) \in E} (\|\mathbf{x}_u^i - \mathbf{x}_v^i\|^2 - d_{u,v}^2)^2. \quad (3.3)$$

Podemos, então, definir a função f para ser usada pelo Algoritmo de Grover para encontrar uma solução do DMDGP,

$$f(i) = 1 - \left[\left(\frac{g(h(i))}{p_1} \right)^{\frac{1}{p_2}} + 0.5 \right], \quad (3.4)$$

onde $i \in \{0, 1, \dots, N - 1\}$ e p_1, p_2 são dois parâmetros.

Se encontrarmos valores para p_1 e p_2 tais que $\frac{g(h(i))}{p_1} \in [0, 1]$ e $\left(\frac{g(h(i))}{p_1}\right)^{\frac{1}{p_2}}$ com valores próximos de 1, exceto quando i é solução do problema, então $f(i)$ será 1, quando i for solução, e 0 caso contrário. De fato, existem tais parâmetros, basta tomar $p_1 = 5^4 n^6 + 5^4 n^2$ e $p_2 = \log_{(\frac{1}{2})} \left(\frac{\delta}{p_1}\right)$ em que δ é uma tolerância de erro (Lavor et al., 2020).

4 Conclusão

Foram realizados simulações computacionais do Algoritmo de Grover utilizando uma biblioteca em *python* chamada *qiskit*. Essa implementação está disponível no repositório <https://github.com/andresroliveira/QuantumComputing>.

Sobre a função f para ser utilizada como Oráculo, os experimentos computacionais mostraram que de fato os parâmetros p_1 e p_2 identificam a solução de um DMDGP. Essa implementação está disponível em <https://github.com/andresroliveira/DistanceGeometry>, onde há um arquivo *qbp.py* com esses experimentos.

Como fruto desse trabalho, há um artigo em elaboração, *A Quantum Approach to the Discretizable Molecular Distance Geometry Problem*, escrito também com Franklin Marquezino da UFRJ e Renato Portugal do LNCC.

Como projetos futuros, iremos implementar o oráculo junto da função f do DMDGP e estudar outros algoritmos sobre Computação Quântica.



Referências

- Crippen, G. M., Havel, T. F., et al. (1988). *Distance geometry and molecular conformation*, volume 74. Research Studies Press Taunton.
- Donald, B. R. (2011). *Algorithms in structural molecular biology*. MIT Press.
- Grover, L. K. (1996). A fast quantum mechanical algorithm for database search. In *Proceedings of the twenty-eighth annual ACM symposium on Theory of computing - STOC 96*. ACM Press.
- Grover, L. K. (1997). Quantum mechanics helps in searching for a needle in a haystack. *Physical Review Letters*, 79(2):325–328.
- Lavor, C., Liberti, L., and Maculan, N. (2005). Grover’s algorithm applied to the molecular distance geometry problem. *Proc. of VII Brazilian Congress of Neural Networks*.
- Lavor, C., Liberti, L., Maculan, N., and Mucherino, A. (2012). The discretizable molecular distance geometry problem. *Computational Optimization and Applications*, 52(1):115–146.
- Lavor, C., Marquezino, F., Oliveira, A., and Portugal, R. (2020). A quantum approach to the discretizable molecular distance geometry problem. Em preparação.
- Liberti, L., Lavor, C., Maculan, N., and Mucherino, A. (2014). Euclidean distance geometry and applications. *Siam Review*, 56(1):3–69.
- Phillips, A. T., Rosen, J. B., and Walke, V. H. (1994). Molecular structure determination by convex global underestimation of local energy minima. In *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, pages 181–198. American Mathematical Society.
- Saxe, J. B. (1979). Embeddability of weighted graphs in k-space is strongly np-hard. In *Proc. of 17th Allerton Conference in Communications, Control and Computing, Monticello, IL*, pages 480–489.