

Geometria de Distâncias e o Problema do Loop Fechado

Rodrigo Labiak¹

UNICAMP, Campinas, SP

Carlile Lavor²

IMECC/UNICAMP, Campinas, SP

O principal problema em Geometria de Distâncias é encontrar as posições de um determinado conjunto de pontos, considerando conhecidas apenas informações sobre as distâncias entre alguns desses pontos [2]. O Problema do Loop Fechado é utilizado para determinar estruturas tridimensionais de regiões de proteínas que apresentam maior mobilidade, nessas regiões os experimentos de Ressonância Magnética Nuclear (RMN) e Cristalografia são incapazes de determinar a estrutura tridimensional com precisão. Iniciamos com a definição formal do problema de Geometria de Distâncias [4].

Definição 0.1. *Problema de Geometria de Distâncias (PGD): Dado um inteiro positivo K e um grafo simples ponderado e não direcionado $G = (V, E, d)$, onde $d : E \rightarrow \mathbb{R}_+$, encontre uma função $x : V \rightarrow \mathbb{R}^K$ tal que*

$$\forall \{u, v\} \in E \quad \|x_u - x_v\| = d_{uv}. \quad (1)$$

Uma função $x : V \rightarrow \mathbb{R}^K$ é chamada de realização de V . Dizemos que x é uma realização válida se satisfaz a Equação (1).

Neste trabalho, estamos interessados em utilizar uma versão "discreta" do PGD, onde o espaço de busca é finito [4]. Essa nova versão do problema envolve ordenação nos vértices de G e interseção de esferas.

Definição 0.2. *Problema Discretizável de Geometria de Distâncias (PDGD₃): Considere um grafo $G = (V, E, d)$ de um PGD, onde $K = 3$ e uma ordem em V , denotada por v_1, \dots, v_n tal que:*

- existe uma realização válida para v_1, v_2, v_3 ,
- para todo v_i , $i = 4, \dots, n$ existem no mínimo três antecessores a_i, b_i, c_i com $\{\{a_i, v_i\}, \{b_i, v_i\}, \{c_i, v_i\}\} \subseteq E$ satisfazendo

$$d_{a_i b_i} + d_{a_i c_i} > d_{b_i c_i}.$$

Encontre uma função $x : V \rightarrow \mathbb{R}^3$ tal que

$$\forall \{v_i, v_j\} \in E \quad \|x_{v_i} - x_{v_j}\| = d_{v_i v_j}.$$

Uma generalização importante do PGD é apresentada em [3] e será de grande importância para o desenvolvimento do trabalho.

¹rodrigolabiak@gmail.com.

²clavor@unicamp.br.

Definição 0.3. *Discretização Intervalar do Problema de Geometria de Distâncias em Dimensão 3. (iPDGD₃):* Dado um grafo simples, ponderado e não direcionado $G = (V, E, d)$, onde $E' \subset E$ é um subconjunto de arestas para quais suas distâncias são dadas por valores exatos. Dizemos que G representa uma instância do iPDGD₃ se existe uma ordem total nos vértices de V satisfazendo: $G_C = (V_C, E_C) = G[\{1, 2, 3\}]$, grafo gerado pelos três primeiros vértices, é uma clique, tal que $E_C \subset E'$;

$\forall i \in \{4, \dots, |V|\}$ existem $\{a, b, c\}$ tal que:

1. $a < i, b < i, c < i$;
2. $\{\{b, i\}, \{c, i\}\} \subset E', \{a, i\} \subset E$;
3. $\Delta_S(a, b, c) > 0$,

onde $\Delta_S(a, b, c)$ é a área do triângulo formado por a, b, c .

Note que esta ordem utiliza as distâncias intervalares para realizar a discretização e também as podas na árvore de busca.

Utilizamos o trabalho de Coutias e Seok [1] para estudar o Problema do Loop Fechado. Resumidamente, o problema consiste em encontrar estruturas de trechos de uma molécula que sejam geometricamente consistentes com toda a estrutura da molécula. A seguir, uma definição formal para o Problema do Loop Fechado [1].

Definição 0.4. *Problema do Loop Fechado:* Dada uma cadeia molecular, onde o comprimento das ligações e ângulos covalentes são fixos, encontre todas as conformações tais que todos os trechos da molécula são fixos, exceto por um conjunto de movimentos obtidos através de mudanças feitas em no máximo seis ângulos diedrais.

Em nosso trabalho, apresentamos uma ordem para os átomos de cadeias moleculares que possuem movimentos satisfazendo a Definição 0.4, tal que as hipóteses exigidas pela Definição 0.3 também são satisfeitas. Dessa forma, podemos aplicar uma versão do algoritmo Intervalar Branch Prune (iBP) [3] para encontrar as conformações da cadeias moleculares. Uma das vantagens do nosso método é a de não apresentar um crescimento exponencial no tempo de execução do algoritmo com o aumento de átomos. Diversas cadeias moleculares que satisfazem a Definição 0.4 são apresentadas por Coutias e Seok em [1], exemplificando a aplicação do Problema do Loop Fechado para conformação de cadeias moleculares de resíduos.

Referências

- [1] Countsias, E. A. Seok, C. Jacobson, M. P. and Dill, K. A. A Kinematic View of Loop Closure *Journal of Computational Chemistry*, 25:510–528, 2004.
- [2] Crippen, G. and Havel, T. *Distance Geometry and Molecular Conformation*. Wiley, New York, 1988.
- [3] Golçalves, D. S. Mucherino, A. Liberti, L. and Lavor, C. Recent advances on the interval distance geometry *Journal of Global Optimization*, 69:525–545, 2017.
- [4] Liberti, L. Lavor, C. Maculan, N. and Mucherino, A. Euclidean Distance Geometry and Applications *Siam Review*, 56:3–69, 2014.
- [5] Tramontano, A. Lepplae, R. and Morea, V. Analysis and assessment of comparative modeling predictions in CASP4 *Proteins: Structure, Function, and Bioinformatic*, 45:22–38, 2001.