拘束条件付きのNewton-Raphson最適化解法

森 渉

平成30年10月25日

1　分子回転行列の決定

ある分子の分子固定系 (Gaussian での standard orientation) での原子位置の Cartesian 座標を $\{\boldsymbol{r}_i^0\}$ $(i = 1, 2, \cdots, N_a)$ とし、空間固定系での原子の瞬間的な Cartesian 座標を $\{\boldsymbol{r}_i\}$ とする。このとき、分子固定系と空間固定系を結ぶ回転行列 $\boldsymbol{D}$ の最適解を求める。両者の座標系で分子重心は原点で共通としておき、回転のみを考える。（分子は振動・回転しているため、回転によって完全に重なるとは限らないことに注意。）

回転行列 $\boldsymbol{D}$ は四元数 q_1, q_2, q_3, q_4 を用いて与える [1]。[1]

$$\boldsymbol{D} = \begin{bmatrix} q_4^2 + q_1^2 - q_2^2 - q_3^2 & 2(q_1q_2 + q_4q_3) & 2(q_1q_3 - q_4q_2) \\ 2(q_1q_2 - q_4q_3) & q_4^2 - q_1^2 + q_2^2 - q_3^2 & 2(q_2q_3 + q_4q_1) \\ 2(q_1q_3 + q_4q_2) & 2(q_2q_3 - q_4q_1) & q_4^2 - q_1^2 - q_2^2 + q_3^2 \end{bmatrix} \tag{1}$$

ただし、四元数は $-1 \leq q_1 \sim q_4 \leq 1$ で、拘束条件

$$g(q_1, q_2, q_3, q_4) = q_1^2 + q_2^2 + q_3^2 + q_4^2 - 1 = 0 \tag{2}$$

を満たす。これらの値は、Lagrange 未定乗数 (λ) 付きの最小二乗法で、残差 $\mathcal{L}$ を最小化するように求める。

$$\begin{aligned} \mathcal{L}(q_1, q_2, q_3, q_4, \lambda) &= \sum_{i=1}^{N_a} \left| \boldsymbol{D}(q_1, q_2, q_3, q_4)\boldsymbol{r}_i^0 - \boldsymbol{r}_i \right|^2 + \lambda\left(q_1^2 + q_2^2 + q_3^2 + q_4^2 - 1\right) \qquad (3) \\ &= L(q_1, q_2, q_3, q_4) + \lambda\left(\sum_{j=1}^{4} q_j^2 - 1\right) \end{aligned}$$

したがって、最適化条件は

$$\frac{\partial \mathcal{L}}{\partial q_k} = \frac{\partial L}{\partial q_k} + 2\lambda q_k = 0 \qquad (k = 1, 2, 3, 4) \tag{4}$$

$$\frac{\partial \mathcal{L}}{\partial \lambda} = \sum_{j=1}^{4} q_j^2 - 1 = 0 \tag{2'}$$

[1] Goldstein [1] の式 (4.47′)。ただし、原著中での $A \to D$ とし、記号を $(e_0, e_1, e_2, e_3) \to (q_4, q_1, q_2, q_3)$ とした。

1.1 Lagrange 未定乗数を独立に扱う

1.1.1 4 変数の Newton-Raphson 解法

この節では毎ステップで以下の規格化を行うことにより拘束条件 Eq. (2) を常に満たしている状態で、Eq. (4) を 4 変数で解く。

$$\mathfrak{q} = \frac{\mathfrak{q}}{|\mathfrak{q}|} \tag{5}$$

Newton-Raphson 法で、n 回目の iteration での値 $\{q_k^{(n)}\}$ $(k = 1 \sim 4)$ で次の点を求める条件は、

$$\sum_{j=1}^{4} \left(\frac{\partial^2 \mathcal{L}}{\partial q_k \partial q_j} \right)^{(n)} (q_j - q_j^{(n)}) + \left(\frac{\partial \mathcal{L}}{\partial q_k} \right)^{(n)} = 0 \tag{6}$$

または行列形で

$$\boldsymbol{\mathcal{A}}^{(n)}(\mathfrak{q} - \mathfrak{q}^{(n)}) + \mathfrak{b}^{(n)} = 0 \tag{7}$$

である。ただし、$\boldsymbol{\mathcal{A}}^{(n)}$, $\mathfrak{b}^{(n)}$ はそれぞれ 4×4 次元の行列、4 次元のベクトルで

$$\boldsymbol{\mathcal{A}}^{(n)} = \left[\left(\frac{\partial L}{\partial q_k \partial q_j} \right)^{(n)} + 2\lambda^{(n)} \delta_{kj} \right], \qquad \mathfrak{b}^{(n)} = \left[\left(\frac{\partial L}{\partial q_k} \right)^{(n)} + 2\lambda^{(n)} q_k^{(n)} \right] \tag{8}$$

したがって、Newton-Raphson 法の iteration 条件は、

$$\mathfrak{q}^{(n+1)} = \mathfrak{q}^{(n)} - {\boldsymbol{\mathcal{A}}^{(n)}}^{-1} \mathfrak{b}^{(n)} \tag{9}$$

1.1.2 λ の初期条件の決定

拘束条件が常に満たされている場合 λ は不定になる。その際、λ はどうしたらよいか。

Eq. (9) の iteration が収束するとき、$\mathfrak{b} = \mathbf{0}$ となることを考えて、$|\mathfrak{b}|^2$ がなるべく小さくなるように与える ($\partial |\mathfrak{b}|^2 / \partial \lambda = 0$)。

$$\frac{\partial}{\partial \lambda} |\mathfrak{b}|^2 = \frac{\partial}{\partial \lambda} \sum_{k=1}^{4} \left(\frac{\partial L}{\partial q_k} + 2\lambda q_k \right)^2 = \sum_{k=1}^{4} 4 q_k \left(\frac{\partial L}{\partial q_k} + 2\lambda q_k \right) = 0 \tag{10}$$

したがって、

$$\lambda = - \frac{\displaystyle\sum_{k=1}^{4} q_k \frac{\partial L}{\partial q_k}}{\displaystyle 2 \sum_{k=1}^{4} q_k^2} \tag{11}$$

または、

$$\lambda = - \frac{1}{2} \sum_{k=1}^{4} q_k \frac{\partial L}{\partial q_k} \tag{12}$$

とすればよい。拘束条件 Eq. (2) を規格化を行うことで常に満たしているとすれば、最後の形 Eq. (12) が成り立つ。

上の Eq. (9) で $\mathfrak{q}^{(n+1)}$ $(q_1^{(n+1)} \sim q_4^{(n+1)}, \lambda^{(n+1)})$ を決めたとき、$\lambda^{(n+1)}$ は Eq. (12) を用いて $q_1^{(n+1)} \sim q_4^{(n+1)}$ より求めた値で置き換えてもよい。

1.1.3 **q** の初期条件の決定

$q_1 \sim q_4$ の初期値も適切に決める必要があるかもしれない。その場合、たとえば慣性モーメントテンソルの主軸をあわせるように決めることが考えられる。

慣性モーメントの主軸などから初期値として回転行列を求めたとする。このとき四元数はどのように求めたらよいか。

Eq. (1) より $\boldsymbol{D}$ の対角成分 D_{11}, D_{22}, D_{33} と拘束条件 g から以下の連立二次方程式が導ける。

$$\begin{aligned} q_4^2 + q_1^2 - q_2^2 - q_3^2 &= D_{11} \\ q_4^2 - q_1^2 + q_2^2 - q_3^2 &= D_{22} \\ q_4^2 - q_1^2 - q_2^2 + q_3^2 &= D_{33} \\ q_4^2 + q_1^2 + q_2^2 + q_3^2 &= 1 \end{aligned} \tag{13}$$

それぞれの解は

$$\begin{aligned} q_1^2 &= \frac{D_{11} - D_{22} - D_{33} + 1}{4}, & q_2^2 &= \frac{-D_{11} + D_{22} - D_{33} + 1}{4}, \\ q_3^2 &= \frac{-D_{11} - D_{22} + D_{33} + 1}{4}, & q_4^2 &= \frac{D_{11} + D_{22} + D_{33} + 1}{4} \end{aligned} \tag{14}$$

よって **q** の絶対値が求められた。符号について $\boldsymbol{D}$ の非対角成分から

$$\begin{aligned} q_1 q_2 &= \frac{2(q_1 q_2 + q_4 q_3) + 2(q_1 q_2 - q_4 q_3)}{4} = \frac{D_{12} + D_{21}}{4} \\ q_1 q_3 &= \frac{2(q_1 q_3 + q_2 q_4) + 2(q_1 q_3 - q_2 q_4)}{4} = \frac{D_{13} + D_{31}}{4} \\ q_1 q_4 &= \frac{2(q_3 q_2 - q_1 q_4) + 2(q_3 q_2 + q_1 q_4)}{4} = \frac{-D_{23} + D_{32}}{4} \end{aligned} \tag{15}$$

q_1 の符号を決めると他の四元数の符号を求められる。よって $\boldsymbol{D}$ から **q** の絶対値と符号を求められる。

1.1.4 微分形の決定

Eq. (8) のなかの微分は、

$$\frac{\partial L}{\partial q_k} = 2 \sum_i^{N_a} \left(\boldsymbol{D}(\boldsymbol{q}) \boldsymbol{r}_i^0 - \boldsymbol{r}_i \right) \cdot \left(\frac{\partial \boldsymbol{D}}{\partial q_k} \boldsymbol{r}_i^0 \right), \tag{16}$$

$$\frac{\partial^2 L}{\partial q_k \partial q_j} = 2 \sum_i^{N_a} \left(\boldsymbol{D}(\boldsymbol{q}) \boldsymbol{r}_i^0 - \boldsymbol{r}_i \right) \cdot \left(\frac{\partial^2 \boldsymbol{D}}{\partial q_k \partial q_j} \boldsymbol{r}_i^0 \right) + 2 \sum_i^{N_a} \left(\frac{\partial \boldsymbol{D}}{\partial q_j} \boldsymbol{r}_i^0 \right) \cdot \left(\frac{\partial \boldsymbol{D}}{\partial q_k} \boldsymbol{r}_i^0 \right) \tag{17}$$

である。このなかには、$\boldsymbol{D}$ の微分形 $(\partial \boldsymbol{D}/\partial q_k)$, $(\partial^2 \boldsymbol{D}/\partial q_k \partial q_j)$ が含まれている。これらは、(1) 式より

$$\begin{aligned} \frac{\partial \boldsymbol{D}}{\partial q_1} &= 2 \begin{bmatrix} q_1 & q_2 & q_3 \\ q_2 & -q_1 & q_4 \\ q_3 & -q_4 & -q_1 \end{bmatrix}, & \frac{\partial \boldsymbol{D}}{\partial q_2} &= 2 \begin{bmatrix} -q_2 & q_1 & -q_4 \\ q_1 & q_2 & q_3 \\ q_4 & q_3 & -q_2 \end{bmatrix}, \\ \frac{\partial \boldsymbol{D}}{\partial q_3} &= 2 \begin{bmatrix} -q_3 & q_4 & q_1 \\ -q_4 & -q_3 & q_2 \\ q_1 & q_2 & q_3 \end{bmatrix}, & \frac{\partial \boldsymbol{D}}{\partial q_4} &= 2 \begin{bmatrix} q_4 & q_3 & -q_2 \\ -q_3 & q_4 & q_1 \\ q_2 & -q_1 & q_4 \end{bmatrix} \end{aligned} \tag{18}$$

$$
\frac{\partial^2 \boldsymbol{D}}{\partial q_1^2} = 2\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}, \qquad \frac{\partial^2 \boldsymbol{D}}{\partial q_2^2} = 2\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix},
$$

$$
\frac{\partial^2 \boldsymbol{D}}{\partial q_3^2} = 2\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \frac{\partial^2 \boldsymbol{D}}{\partial q_4^2} = 2\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},
$$

$$
\frac{\partial^2 \boldsymbol{D}}{\partial q_1 \partial q_2} = 2\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad \frac{\partial^2 \boldsymbol{D}}{\partial q_1 \partial q_3} = 2\begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad \frac{\partial^2 \boldsymbol{D}}{\partial q_1 \partial q_4} = 2\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix},
$$

$$
\frac{\partial^2 \boldsymbol{D}}{\partial q_2 \partial q_3} = 2\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \qquad \frac{\partial^2 \boldsymbol{D}}{\partial q_2 \partial q_4} = 2\begin{bmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \qquad \frac{\partial^2 \boldsymbol{D}}{\partial q_3 \partial q_4} = 2\begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tag{19}
$$

参考文献

[1] H. Goldstein, C. Poole, J. Safko, 矢野忠他訳、古典力学（上）原著第 3 版、吉岡書店