Cell Gradient

Kohei Shinohara

September 17, 2023

1 Cell gradient

This is a code reading note for ASE's UnitCellFilter and ExpCellFilter [1].

1.1 Notations

First, we define notations for structural optimization. For an original structure, we write its row-wise basis vectors as $\boldsymbol{L}^{(0)} = \begin{pmatrix} \boldsymbol{a}^\top \\ \boldsymbol{b}^\top \\ \boldsymbol{c}^\top \end{pmatrix}$ and its atomic positions in Cartesian coordinates as $\boldsymbol{r}_i^{(0)}$. Let $\boldsymbol{L}^{(k)}$ and $\boldsymbol{r}_i^{(k)}$ be basis vectors and atomic positions at the kth optimization step. The deformation tensor $\boldsymbol{F}^{(k)}$ is defined as

$$\boldsymbol{L}^{(k)} := \boldsymbol{L}^{(0)} \boldsymbol{F}^{(k)\top} \tag{1}$$

$$L^{(k)}_{\alpha\mu} = \sum_{\nu=1}^{3} L^{(0)}_{\alpha\nu} F^{(k)}_{\mu\nu}. \tag{2}$$

We write a potential energy of the kth-step structure as $E(\{\boldsymbol{r}_i^{(k)}\}, \boldsymbol{L}^{(k)})$, where we drop the dependence on atomic species for notation clarity.

The atomic forces is defined as

$$\boldsymbol{f}_j(\{\boldsymbol{r}_i\}, \boldsymbol{L}) := -\frac{\partial E}{\partial \boldsymbol{r}_j}(\{\boldsymbol{r}_i\}, \boldsymbol{L}). \tag{3}$$

The virial stress tensor of a structure $(\{\boldsymbol{r}_i\}, \boldsymbol{L})$ is defined as

$$\sigma_{\mu\nu}(\{\boldsymbol{r}_i\}, \boldsymbol{L}) := -\frac{1}{|\boldsymbol{L}|} \lim_{t \to 0} \frac{E(\{\boldsymbol{r}_i\}, \boldsymbol{L}(\boldsymbol{I} + t\boldsymbol{D}^{(\mu\nu)})) - E(\{\boldsymbol{r}_i\}, \boldsymbol{L})}{t}, \tag{4}$$

where $\boldsymbol{D}^{(\mu\nu)}$ is a "direction" matrix with $D^{(\mu\nu)}_{\mu'\nu'} = \delta_{\mu\mu'}\delta_{\nu\nu'}$ [2]. We rewrite Eq. (4) for later convenience as

$$E(\{\boldsymbol{r}_i\}, \boldsymbol{L}(\boldsymbol{I} + \boldsymbol{V})) = E(\{\boldsymbol{r}_i\}, \boldsymbol{L}) - \sum_{\mu\nu} |\boldsymbol{L}| \sigma_{\mu\nu}(\{\boldsymbol{r}_i\}, \boldsymbol{L}) V_{\mu\nu} + o(\|\boldsymbol{V}\|). \tag{5}$$

[1] Based on 9fb8dd74 (2023-09-16).

[2] $\boldsymbol{I}$ is a 3×3 identify matrix.

We can write a gradient of the potential energy with $\boldsymbol{L}$ as follows:

$$\begin{aligned}
E(\{\boldsymbol{r}_i\}, \boldsymbol{L} + h\boldsymbol{D}^{(\alpha\mu)}) &= E(\{\boldsymbol{r}_i\}, \boldsymbol{L}(\boldsymbol{I} + h\boldsymbol{L}^{-1}\boldsymbol{D}^{(\alpha\mu)})) \\
&= E(\{\boldsymbol{r}_i\}, \boldsymbol{L}) - \sum_{\mu'\nu'} h[\boldsymbol{L}^{-1}\boldsymbol{D}^{(\alpha\mu)}]_{\mu'\nu'} |\boldsymbol{L}| \sigma_{\mu'\nu'}(\{\boldsymbol{r}_i\}, \boldsymbol{L}) + o(h) \\
&= E(\{\boldsymbol{r}_i\}, \boldsymbol{L}) - \sum_{\mu'} h[\boldsymbol{L}^{-1}]_{\mu'\alpha} |\boldsymbol{L}| \sigma_{\mu'\mu}(\{\boldsymbol{r}_i\}, \boldsymbol{L}) + o(h) \\
&= E(\{\boldsymbol{r}_i\}, \boldsymbol{L}) - h|\boldsymbol{L}|[\boldsymbol{L}^{-\top}\boldsymbol{\sigma}(\{\boldsymbol{r}_i\}, \boldsymbol{L})]_{\alpha\mu} + o(h) \\
&= E(\{\boldsymbol{r}_i\}, \boldsymbol{L}) - h\boldsymbol{D}^{(\alpha\mu)} : \left(|\boldsymbol{L}|\boldsymbol{L}^{-\top}\boldsymbol{\sigma}(\{\boldsymbol{r}_i\}, \boldsymbol{L})\right) + o(h)
\end{aligned}$$

$$\therefore \quad \frac{\partial E}{\partial \boldsymbol{L}} = -|\boldsymbol{L}|\boldsymbol{L}^{-\top}\boldsymbol{\sigma}(\{\boldsymbol{r}_i\}, \boldsymbol{L}) \tag{6}$$

Here $\boldsymbol{A} : \boldsymbol{B}$ denotes a matrix contraction $\boldsymbol{A} : \boldsymbol{B} = \mathrm{Tr}\boldsymbol{A}^\top\boldsymbol{B}$.

For clarity, we write $\boldsymbol{f}_j^{(k)} := \boldsymbol{f}_j(\{\boldsymbol{r}_i^{(k)}\}, \boldsymbol{L}^{(k)})$ and $\boldsymbol{\sigma}^{(k)} := \boldsymbol{\sigma}(\{\boldsymbol{r}_i^{(k)}\}, \boldsymbol{L}^{(k)})$.

1.2 UnitCellFilter

UnitCellFilter uses the following $n+3$ vectors as input variables: For positions,

$$\boldsymbol{q}_i^{(k)} := \boldsymbol{F}^{(k)-1}\boldsymbol{r}_i^{(k)} \quad (i = 1, \ldots, n). \tag{7}$$

For the cell,

$$\boldsymbol{Q}_j^{(k)} := \lambda \boldsymbol{F}_{j,:}^{(k)} \quad (j = 1, 2, 3) \tag{8}$$

$$\boldsymbol{Q}^{(k)} := \begin{pmatrix} \boldsymbol{Q}_1^{(k)} & \boldsymbol{Q}_2^{(k)} & \boldsymbol{Q}_3^{(k)} \end{pmatrix} \tag{9}$$

Here λ is a cell factor.

Consider a gradient of the potential energy w.r.t. these variables.

$$\hat{E}(\{\boldsymbol{q}_i\}, \boldsymbol{Q}) := E(\{\boldsymbol{r}_i = \frac{1}{\lambda}\boldsymbol{Q}\boldsymbol{q}_i\}, \boldsymbol{L} = \frac{1}{\lambda}\boldsymbol{L}^{(0)}\boldsymbol{Q}^\top) \tag{10}$$

$$\begin{aligned}
\frac{\partial \hat{E}}{\partial q_{j\mu}}(\{\boldsymbol{q}_i^{(k)}\}, \boldsymbol{Q}^{(k)}) &= \sum_{j'\mu'} \frac{\partial E}{\partial r_{j'\mu'}}(\{\boldsymbol{r}_i^{(k)}\}, \boldsymbol{L}^{(k)}) \frac{\partial}{\partial q_{j\mu}} \left[\boldsymbol{F}^{(k)}\boldsymbol{q}_{j'}\right]_{\mu'} \\
&= -\sum_{\mu'} f_{j\mu'}^{(k)} F_{\mu'\mu}^{(k)}
\end{aligned}$$

$$\therefore \quad \frac{\partial \hat{E}}{\partial \boldsymbol{q}_j}(\{\boldsymbol{q}_i^{(k)}\}, \boldsymbol{Q}^{(k)}) = -\boldsymbol{F}^{(k)\top} \boldsymbol{f}_j^{(k)} \tag{11}$$

$$\begin{aligned}
\frac{\partial \hat{E}}{\partial Q_{\mu\nu}}(\{\boldsymbol{q}_i^{(k)}\}, \boldsymbol{Q}^{(k)}) &= \frac{1}{\lambda} \sum_{\alpha\mu'} \frac{\partial E}{\partial L_{\alpha\mu'}}(\{\boldsymbol{r}_i^{(k)}\}, \boldsymbol{L}^{(k)}) \frac{\partial}{\partial Q_{\mu\nu}} [\boldsymbol{L}^{(0)} \boldsymbol{Q}^{(k)\top}]_{\alpha\mu'} \\
&= -\frac{1}{\lambda} |\boldsymbol{L}^{(k)}| \sum_{\alpha\mu'} [\boldsymbol{L}^{(k)-\top} \boldsymbol{\sigma}^{(k)}]_{\alpha\mu'} L_{\alpha\nu}^{(0)} \delta_{\mu\mu'} \\
&= -\frac{1}{\lambda} |\boldsymbol{L}^{(k)}| \sum_{\alpha} [\boldsymbol{L}^{(k)-\top} \boldsymbol{\sigma}^{(k)}]_{\alpha\mu} L_{\alpha\nu}^{(0)} \\
&= -\frac{1}{\lambda} |\boldsymbol{L}^{(k)}| \sum_{\alpha} [\boldsymbol{L}^{(k)-\top} \boldsymbol{\sigma}^{(k)}]_{\alpha\mu} L_{\alpha\nu}^{(0)} \\
&= -\frac{1}{\lambda} |\boldsymbol{L}^{(k)}| \left[\boldsymbol{\sigma}^{(k)} \boldsymbol{L}^{(k)-1} \boldsymbol{L}^{(0)} \right]_{\mu\nu} \qquad (\because \boldsymbol{\sigma} \text{ is symmetric})
\end{aligned}$$

$$\therefore \quad \frac{\partial \hat{E}}{\partial \boldsymbol{Q}}(\{\boldsymbol{q}_i^{(k)}\}, \boldsymbol{Q}^{(k)}) = -\frac{1}{\lambda} |\boldsymbol{L}^{(k)}| \boldsymbol{\sigma}^{(k)} \boldsymbol{F}^{(k)-\top} \tag{12}$$

1.3 ExpCellFilter

ExpCellFilter uses the same variables for positions. For the cell,

$$\boldsymbol{G}^{(k)} \coloneqq \log(\boldsymbol{F}^{(k)}) \tag{13}$$

$$\begin{aligned}
\bar{E}(\{\boldsymbol{q}_i\}, \boldsymbol{G}) &\coloneqq E(\{\boldsymbol{r}_i = \mathbf{exp}\,(\boldsymbol{G})\, \boldsymbol{q}_i\}, \boldsymbol{L} = \boldsymbol{L}^{(0)} \mathbf{exp}\,(\boldsymbol{G})^\top) & (14) \\
&= \hat{E}(\{\boldsymbol{q}_i\}, \boldsymbol{F} = \mathbf{exp}\,(\boldsymbol{G})) & (15)
\end{aligned}$$

We denote a directional derivative of the matrix exponential at $\boldsymbol{A}$ along $\boldsymbol{V}$ as

$$\boldsymbol{l}[\boldsymbol{A}](\boldsymbol{V}) \coloneqq \lim_{h \to 0} \frac{\mathbf{exp}\,(\boldsymbol{A} + h\boldsymbol{V}) - \mathbf{exp}\,(\boldsymbol{A})}{h}. \tag{16}$$

This matrix can be efficiently computed by `scipy.linalg.expm_frechet(A, V)`.

Consider a gradient of the potential energy w.r.t. these variables.

$$\begin{aligned}
\frac{\partial \bar{E}}{\partial G_{\mu\nu}}(\{\boldsymbol{q}_i^{(k)}\}, \boldsymbol{G}^{(k)}) &= \sum_{\mu'\nu'} \frac{\partial \hat{E}}{\partial F_{\mu'\nu'}}(\{\boldsymbol{q}_i^{(k)}\}, \boldsymbol{G}^{(k)}) \frac{\partial}{\partial G_{\mu\nu}} \left[\mathbf{exp} \left(\boldsymbol{G}^{(k)} \right) \right]_{\mu'\nu'} & (17) \\
&= -|\boldsymbol{L}^{(k)}| \sum_{\mu'\nu'} \left[\boldsymbol{\sigma}^{(k)} \boldsymbol{F}^{(k)-\top} \right]_{\mu'\nu'} \left[\boldsymbol{l}[\boldsymbol{G}^{(k)}](\boldsymbol{D}^{(\mu\nu)}) \right]_{\mu'\nu'} & (18)
\end{aligned}$$

2 Directional derivative of matrix exponential

$l_f : \boldsymbol{R}^{n\times n} \to (\boldsymbol{R}^{n\times n} \to \boldsymbol{R}^{n\times n})$

$$l_f[\boldsymbol{A}](\boldsymbol{V}) \coloneqq \lim_{h\to 0} \frac{f(\boldsymbol{A}+h\boldsymbol{V}) - f(\boldsymbol{A})}{h} \tag{19}$$

$$\exp\left(\begin{pmatrix} \boldsymbol{A} & \boldsymbol{V} \\ \boldsymbol{O} & \boldsymbol{A} \end{pmatrix}\right) = \begin{pmatrix} \exp(\boldsymbol{A}) & l_{\exp}[\boldsymbol{A}](\boldsymbol{V}) \\ \boldsymbol{O} & \exp(\boldsymbol{A}) \end{pmatrix} \tag{20}$$

$$\lim_{t\to 0} \frac{\exp(\boldsymbol{G}+t\boldsymbol{D}) - \exp(\boldsymbol{G})}{t} \tag{21}$$

$$= \lim_{t\to 0} \frac{\sum_{n=0}^{\infty} \frac{1}{n!}(\boldsymbol{G}+t\boldsymbol{D})^n - \exp(\boldsymbol{G})}{t} \tag{22}$$

$$= \sum_{n=1}^{\infty} \sum_{m=0}^{n-1} \frac{1}{n!} \boldsymbol{G}^{n-1-m} \boldsymbol{D} \boldsymbol{G}^m \tag{23}$$

References