

Cell Gradient

Kohei Shinohara

October 1, 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 $\mathbf{L}^{(0)} = \begin{pmatrix} \mathbf{a}^\top \\ \mathbf{b}^\top \\ \mathbf{c}^\top \end{pmatrix}$ and its atomic positions in Cartesian coordinates as $\mathbf{r}_i^{(0)}$. Let $\mathbf{L}^{(k)}$ and $\mathbf{r}_i^{(k)}$ be basis vectors and atomic positions at the k th optimization step. The deformation tensor $\mathbf{F}^{(k)}$ is defined as

$$\mathbf{L}^{(k)} := \mathbf{L}^{(0)} \mathbf{F}^{(k)\top} \quad (1)$$

$$L_{\alpha\mu}^{(k)} = \sum_{\nu=1}^3 L_{\alpha\nu}^{(0)} F_{\mu\nu}^{(k)}. \quad (2)$$

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

The atomic forces is defined as

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

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

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

where $\mathbf{D}^{(\mu\nu)}$ 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

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

¹Based on [9fb8dd74](#) (2023-09-16).

² \mathbf{I} is a 3×3 identify matrix.

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

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

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

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

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

1.2 UnitCellFilter

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

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

For the cell,

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

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

Here λ is a cell factor.

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

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

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

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

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

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

1.3 ExpCellFilter

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

$$\mathbf{G}^{(k)} := \kappa \log(\mathbf{F}^{(k)}), \quad (13)$$

where $\kappa = O(N)$ is a proposed cell factor for ExpCellFilter.

$$\bar{E}(\{\mathbf{q}_i\}, \mathbf{G}) := E(\{\mathbf{r}_i = \frac{1}{\kappa} \exp(\mathbf{G}) \mathbf{q}_i\}, \mathbf{L} = \frac{1}{\kappa} \mathbf{L}^{(0)} \exp(\mathbf{G})^\top) \quad (14)$$

$$= \hat{E}(\{\mathbf{q}_i\}, \mathbf{F} = \frac{1}{\kappa} \exp(\mathbf{G})) \quad (15)$$

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

$$l[\mathbf{A}](\mathbf{V}) := \lim_{h \rightarrow 0} \frac{\exp(\mathbf{A} + h\mathbf{V}) - \exp(\mathbf{A})}{h}. \quad (16)$$

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.

$$\frac{\partial \bar{E}}{\partial G_{\mu\nu}}(\{\mathbf{q}_i^{(k)}\}, \mathbf{G}^{(k)}) = \frac{1}{\kappa} \sum_{\mu'\nu'} \frac{\partial \hat{E}}{\partial F_{\mu'\nu'}}(\{\mathbf{q}_i^{(k)}\}, \mathbf{G}^{(k)}) \frac{\partial}{\partial G_{\mu\nu}} [\exp(\mathbf{G}^{(k)})]_{\mu'\nu'} \quad (17)$$

$$= -\frac{|\mathbf{L}^{(k)}|}{\kappa} \sum_{\mu'\nu'} [\boldsymbol{\sigma}^{(k)} \mathbf{F}^{(k)-\top}]_{\mu'\nu'} [l[\mathbf{G}^{(k)}](\mathbf{D}^{(\mu\nu)})]_{\mu'\nu'} \quad (18)$$

2 Directional derivative of matrix exponential

This section is just a memo for matrix exponential [1,2].

$$l_f : \mathbf{R}^{n \times n} \rightarrow (\mathbf{R}^{n \times n} \rightarrow \mathbf{R}^{n \times n})$$

$$l_f[\mathbf{A}](\mathbf{V}) := \lim_{h \rightarrow 0} \frac{f(\mathbf{A} + h\mathbf{V}) - f(\mathbf{A})}{h} \quad (19)$$

$$f \left(\begin{pmatrix} \mathbf{A} & \mathbf{V} \\ \mathbf{O} & \mathbf{A} \end{pmatrix} \right) = \begin{pmatrix} f(\mathbf{A}) & l_f[\mathbf{A}](\mathbf{V}) \\ \mathbf{O} & f(\mathbf{A}) \end{pmatrix} \quad (20)$$

$$l_{\exp}[\mathbf{A}](\mathbf{V}) = \lim_{t \rightarrow 0} \frac{\exp(\mathbf{A} + t\mathbf{V}) - \exp(\mathbf{A})}{t} \quad (21)$$

$$= \lim_{t \rightarrow 0} \frac{\sum_{n=0}^{\infty} \frac{1}{n!} (\mathbf{A} + t\mathbf{V})^n - \exp(\mathbf{A})}{t} \quad (22)$$

$$= \sum_{n=1}^{\infty} \sum_{m=0}^{n-1} \frac{1}{n!} \mathbf{A}^{n-1-m} \mathbf{V} \mathbf{A}^m \quad (23)$$

If $[\mathbf{A}, \mathbf{V}] = \mathbf{O}$, $l_{\exp}[\mathbf{A}](\mathbf{V}) = \exp(\mathbf{A})\mathbf{V} = \mathbf{V}\exp(\mathbf{A})$.

References

- [1] Awad H. Al-Mohy and Nicholas J. Higham. Computing the fréchet derivative of the matrix exponential, with an application to condition number estimation. *SIAM Journal on Matrix Analysis and Applications*, 30(4):1639–1657, 2009.
- [2] I. Najfeld and T.F. Havel. Derivatives of the matrix exponential and their computation. *Advances in Applied Mathematics*, 16(3):321–375, 1995.