Derivatives of the G matrix

Helpful Notes

When reading *Wilson, Decius, & Cross* (WDC) the authors introduce some odd terminology at least according to modern notation. First off, they define their internal coordinates by S_i which, like, fine. I'm gonna use r_i

Then given a reference structure which I'll call $\boldsymbol{S}_e$ and whose existence they allow to be implicit, they approximate the internals up to first order by Cartesians by an equation that they write as

$$S_t = \sum_{i=1}^{N} B_{ti}\,\xi_i \ , \ t = 1,\ 2,\ \dots,\ M$$

where N and M are the number of Cartesian coordinates and number of respective internal coordinates

Here we have ξ_i being the Cartesian displacement. I'd usually call it Δx_i. And the B term is just the derivative of the internal with respect to the Cartesians. This can be much stated more cleanly by saying that we'll *represent* S_t as

$$S_t\colon \left(\frac{\partial S_t}{\partial x_1}\Delta x_1,\ \frac{\partial S_t}{\partial x_2}\Delta x_2,\ \dots,\ \frac{\partial S_t}{\partial x_{3N}}\Delta x_{3N}\right) = \left(\frac{\partial S_t}{\partial x_i}\Delta x_i\right)_{i=1,\dots}$$

This is saying that we'll linearly approximate our internal as a Cartesian displacement. We won't actually use this definition to evaluate the our internals, but these derivatives are important as this is how we get the B matrix that is implicitly referred to in the text

$$B = \begin{pmatrix} \frac{\partial S_1}{\partial x_1} & \frac{\partial S_1}{\partial x_2} & \cdots & \frac{\partial S_1}{\partial x_N} \\ \frac{\partial S_2}{\partial x_1} & \frac{\partial S_2}{\partial x_2} & & \\ \vdots & & \ddots & \vdots \\ \frac{\partial S_M}{\partial x_1} & & \cdots & \frac{\partial S_M}{\partial x_N} \end{pmatrix}$$

For anyone who's taken intro calc or ever transformed from Cartesian to polar coordinates to do an integral, this is just the Jacobian matrix and I'm not sure why *WDC* don't refer to it as such.

Returning to the notation that I prefer (and introducing a slightly odd gradient notation), we have

$$\nabla_{\mathrm{X}}\mathrm{R}=\begin{pmatrix} \frac{\partial r_1}{\partial x_1} & \frac{\partial r_1}{\partial x_2} & \cdots & \frac{\partial r_1}{\partial x_N} \\ \frac{\partial r_2}{\partial x_1} & \frac{\partial r_2}{\partial x_2} & & \\ \vdots & & \ddots & \vdots \\ \frac{\partial r_M}{\partial x_1} & & \cdots & \frac{\partial r_M}{\partial x_N} \end{pmatrix}$$

Next, in this notation we're given that the G-matrix is defined as

$$G_{ij}=\sum_{\alpha=1}^{N}\frac{1}{m_\alpha}\frac{\partial}{\partial x_\alpha}r_i\frac{\partial}{\partial x_\alpha}r_j$$

This is all well and good, but we can write this in a more intuitive way. Letting

$$\nabla_{\mathrm{X}}r_i=\left(\frac{\partial r_i}{\partial x_\alpha}\Delta x_\alpha\right)_{\alpha=1,\ldots,3N}$$

we'll define a mass-weighted form of this by

$$\nabla_{\mathrm{Y}}r_i=\left(\frac{1}{\sqrt{m_\alpha}}\frac{\partial r_i}{\partial x_\alpha}\Delta x_\alpha\right)_{\alpha=1,\ldots,3N}$$

the notation for which is motivated by the mass-weighted Coordinate transformation

$$y_\alpha=\sqrt{m_\alpha}\,x_\alpha$$

taking all of this together we get

$$G_{ij}=\nabla_{\mathrm{Y}}r_i\cdot\nabla_{\mathrm{Y}}r_j$$

This type of definition of a matrix element ($\text{vector}_i\cdot\text{vector}_j$) tells us that our matrix can be constructed by a matrix multiplication. In this case we have

$$G=(\nabla_{\mathrm{Y}}\mathrm{R})\,(\nabla_{\mathrm{Y}}\mathrm{R})^{\mathrm{T}}$$

where we can build $\nabla_{\mathrm{Y}}\mathrm{R}$ by

$$\nabla_{\mathrm{Y}}\mathrm{R}=(\nabla_{\mathrm{X}}\mathrm{R})\,\mathrm{M}^{-\frac{1}{2}}$$

with

$$\mathrm{M}=\begin{pmatrix} m_1 & & & \\ & m_2 & & \\ & & \ddots & \\ & & & m_{3N} \end{pmatrix}$$

Derivation

Since we use it so often, we let the mass-weighted Jacobian be

$$\mathrm{J}=\nabla_{\mathrm{Y}}\mathrm{R}$$

With this, the G matrix is given

$$G = J J^{\intercal}$$

Next let's say we want the derivatives of the G matrix with respect to a new set of coordinates, $\{Q_n\}$.

Going back to prior notes on tensor derivates, we had

$$\nabla_Q V = \nabla_Q x \odot \nabla_x V$$

where V was our potential function in Cartesian coordinates, Q was our internal coordinate set, and $\odot$ represents a generalized matrix multiplication to tensors (implemented in **numpy** via **tensordot**).

Correspondingly since J is defined in our R coordinates, this means that

$$\nabla_Q J = \nabla_Q R \odot \nabla_R J$$

the $\nabla_Q R$ term is the Jacobian for going from the Q coordinate set to the R coordinate set. In the special case the Q_i are defined as linear combinations of the R coordinates defined by some transformation matrix L, this reduces to being simply

$$\nabla_Q R = L^{-1}$$

Next we have the $\nabla_R J$ term, a three dimensional tensor whose elements look like

$$(\nabla_R J)_{ijk} = \frac{\partial}{\partial r_i}\left(\frac{\partial r_j}{\partial x_k}\right) = \frac{\partial {r_j}^2}{\partial r_i\, \partial x_k} = \frac{\partial}{\partial x_k}\left(\frac{\partial r_j}{\partial r_i}\right) = \frac{\partial r_j}{\partial x_k}\,\delta_{ij}$$

which means this is just a nice block-diagonal tensor. This fact can be convenient when implementing this.

G-matrix Derivatives

First Derivative Tensor

Next, moving to the full term we have

$$\begin{aligned}\nabla_Q G &= \nabla_Q (J J^{\intercal}) \\ &= (\nabla_Q J) \odot J^{\intercal} + \left(J \odot (\nabla_Q J^{\intercal})^{2\to1}\right)^{2\to1}\end{aligned}$$

Next, we'll note that this can even be reduced somewhat further by considering that

$$\begin{aligned}\nabla_Q J^{\intercal} &= \nabla_Q \left(J^{2\to1}\right) \\ &= (\nabla_Q J)^{3\to2}\end{aligned}$$

and then

$$\begin{aligned}\left((\nabla_Q J^{\intercal})^{2\to1}\right) &= \left((\nabla_Q J)^{3\to2}\right)^{2\to1} \\ &= (\nabla_Q J)^{3\to1}\end{aligned}$$

we have

$$\nabla_{\mathrm{Q}}\mathrm{G}=(\nabla_{\mathrm{Q}}\mathrm{J})\bigodot\mathrm{J}^{\mathrm{T}}+\left(\mathrm{J}\odot(\nabla_{\mathrm{Q}}\mathrm{J})^{3\to1}\right)^{2\to1}$$

which is nice as this allows the tensor $\nabla_{\mathrm{Q}}\mathrm{J}$ to be reused

For ease of reading, we'll introduce a compact notation

$$\begin{aligned}\mathrm{G}_{Q}&=\nabla_{\mathrm{Q}}\mathrm{G}\\&=\mathrm{J}_{Q}\,\mathrm{J}^{2:1}+\left(\mathrm{J}\ \mathrm{J}_{Q}{}^{3:1}\right)^{2:1}\end{aligned}$$

Validation

Second Derivative Tensor

We also need second derivatives of our G-matrix. To get this we'll evaluate

$$\begin{aligned}\nabla_{\mathrm{Q}^2}\mathrm{G}&=\nabla_{\mathrm{Q}}(\nabla_{\mathrm{Q}}\mathrm{G})\\&=\nabla_{\mathrm{Q}}\left((\nabla_{\mathrm{Q}}\mathrm{J})\bigodot\mathrm{J}^{\mathrm{T}}+\left(\mathrm{J}\odot(\nabla_{\mathrm{Q}}\mathrm{J})^{3\to1}\right)^{2\to1}\right)\end{aligned}$$

or in compact notation:

$$\begin{aligned}\mathrm{G}_{Q^2}&=\left(\mathrm{J}_{Q}\,\mathrm{J}^{2:1}+\left(\mathrm{J}\ \mathrm{J}_{Q}{}^{3:1}\right)^{2:1}\right)_{Q}\\&=\left(\mathrm{J}_{Q}\,\mathrm{J}^{2:1}\right)_{Q}+\left(\left(\mathrm{J}\ \mathrm{J}_{Q}{}^{3:1}\right)^{2:1}\right)_{Q}\end{aligned}$$

then taking the first term we'll let

$$\begin{aligned}\mathrm{A}&=\mathrm{J}_{Q}\\\mathrm{B}&=\mathrm{J}^{2:1}\end{aligned}$$

$$\begin{aligned}\left(\mathrm{J}_{Q}\,\mathrm{J}^{2:1}\right)_{Q}&=(\mathrm{AB})_{Q}\\&=\mathrm{A}_{Q}\,\mathrm{B}+\left(\mathrm{A}\ (\mathrm{B}_{Q})^{2:1}\right)^{a:1}\\&=\left((\mathrm{J}_{Q})_{Q}\,\mathrm{J}^{2:1}\right)+\left(\mathrm{J}_{Q}\left(\left(\mathrm{J}^{2:1}\right)_{Q}\right)^{2:1}\right)^{3:1}\\&=\left(\mathrm{J}_{Q^2}\,\mathrm{J}^{2:1}\right)+\left(\mathrm{J}_{Q}\left(\mathrm{J}_{Q}{}^{3:2}\right)^{2:1}\right)^{3:1}\\&=\left(\mathrm{J}_{Q^2}\,\mathrm{J}^{2:1}\right)+\left(\mathrm{J}_{Q}\ \mathrm{J}_{Q}{}^{3:1}\right)^{3:1}\end{aligned}$$

and then for the second we'll let

$$\begin{aligned}\mathrm{A}&=\mathrm{J}\\\mathrm{B}&=\mathrm{J}_{Q}{}^{3:1}\end{aligned}$$

$$\begin{aligned}\left((\mathrm{AB})^{2:1}\right)_{Q}&=((\mathrm{AB})_{Q})^{3:2}\\&=\left(\mathrm{A}_{Q}\,\mathrm{B}+\left(\mathrm{A}\ (\mathrm{B}_{Q})^{2:1}\right)^{a,1}\right)^{3:2}\\&=(\mathrm{A}_{Q}\,\mathrm{B})^{3:2}+\left(\mathrm{A}\ (\mathrm{B}_{Q})^{2:1}\right)^{1:3}\\&=\left(\mathrm{J}_{Q}\,\mathrm{J}_{Q}{}^{3:1}\right)^{3:2}+\left(\mathrm{J}\left(\mathrm{J}_{Q^2}{}^{4:2}\right)^{2:1}\right)^{1:3}\\&=\left(\mathrm{J}_{Q}\,\mathrm{J}_{Q}{}^{3:1}\right)^{3:2}+\left(\mathrm{J}\ \mathrm{J}_{Q^2}{}^{4:1}\right)^{1:3}\end{aligned}$$

stitching these together we have

$$\nabla_{\mathrm{Q}^2}\mathrm{G}=\left(\mathrm{J}_{Q^2}\ \mathrm{J}^{2:1}\right)+\left(\mathrm{J}_{Q}\ \mathrm{J}_{Q}{}^{3:1}\right)^{3:1}+\left(\mathrm{J}_{Q}\,\mathrm{J}_{Q}{}^{3:1}\right)^{3:2}+\left(\mathrm{J}\ \mathrm{J}_{Q^2}{}^{4:1}\right)^{1:3}$$

Validation

Finally we'll check our terms

$$\begin{aligned}
\left(\mathrm{J}_{Q^2}\ \mathrm{J}^{2:1}\right)_{\mathrm{ijkl}}&=(\mathrm{J}_{Q^2})_{\mathrm{ijk:}}\cdot\left(\mathrm{J}^{2:1}\right)_{:\mathrm{l}}\\
&=(\mathrm{J}_{Q^2})_{\mathrm{ijk:}}\cdot(\mathrm{J}_{\mathrm{l:}})\\
&=\left(\left(\frac{\partial^2}{\partial Q_i\,\partial Q_j}\frac{\partial r_k}{\partial y_\alpha}\right)_{\alpha=1,\ldots,\mathrm{N}}\right)\cdot\left(\left(\frac{\partial r_l}{\partial y_\alpha}\right)_{\alpha=1,\ldots,\mathrm{N}}\right)\\
&=\sum_{\alpha=1}^{N}\frac{1}{m_\alpha}\ \frac{\partial^3 r_k}{\partial Q_i\,\partial Q_j\,\partial x_\alpha}\ \frac{\partial r_l}{\partial x_\alpha}\\
\left(\left(\mathrm{J}_{Q}\,(\mathrm{J}_{Q})^{3:1}\right)^{3:1}\right)_{\mathrm{ijkl}}&=\left(\mathrm{J}_{Q}\,(\mathrm{J}_{Q})^{3:1}\right)_{\mathrm{jkil}}\\
&=(\mathrm{J}_{Q})_{\mathrm{jk:}}\ \left((\mathrm{J}_{Q})^{3:1}\right)_{:\mathrm{il}}\\
&=(\mathrm{J}_{Q})_{\mathrm{jk:}}\ (\mathrm{J}_{Q})_{\mathrm{il:}}\\
&=\left(\left(\frac{\partial}{\partial Q_j}\frac{\partial r_k}{\partial y_\alpha}\right)_{\alpha=1,\ldots,\mathrm{N}}\right)\cdot\left(\left(\frac{\partial}{\partial Q_i}\frac{\partial r_l}{\partial y_\alpha}\right)_{\alpha=1,\ldots,\mathrm{N}}\right)\\
&=\sum_{\alpha=1}^{N}\frac{1}{m_\alpha}\ \frac{\partial}{\partial Q_j}\frac{\partial r_k}{\partial x_\alpha}\ \frac{\partial}{\partial Q_i}\frac{\partial r_l}{\partial x_\alpha}\\
\left(\left(\mathrm{J}_{Q}\,(\mathrm{J}_{Q})^{3:1}\right)^{3:2}\right)_{\mathrm{ijkl}}&=\left(\mathrm{J}_{Q}\,(\mathrm{J}_{Q})^{3:1}\right)_{\mathrm{ikjl}}\\
&=(\mathrm{J}_{Q})_{\mathrm{ik:}}\ \left((\mathrm{J}_{Q})^{3:1}\right)_{:\mathrm{jl}}\\
&=(\mathrm{J}_{Q})_{\mathrm{ik:}}\ (\mathrm{J}_{Q})_{\mathrm{jl:}}\\
&=\left(\left(\frac{\partial}{\partial Q_i}\frac{\partial r_k}{\partial y_\alpha}\right)_{\alpha=1,\ldots,\mathrm{N}}\right)\cdot\left(\left(\frac{\partial}{\partial Q_j}\frac{\partial r_l}{\partial y_\alpha}\right)_{\alpha=1,\ldots,\mathrm{N}}\right)\\
&=\sum_{\alpha=1}^{N}\frac{1}{m_\alpha}\ \frac{\partial}{\partial Q_i}\frac{\partial r_k}{\partial x_\alpha}\ \frac{\partial}{\partial Q_j}\frac{\partial r_l}{\partial x_\alpha}\\
\left(\left(\mathrm{J}\ \mathrm{J}_{Q^2}{}^{4:1}\right)^{1:3}\right)_{\mathrm{ijkl}}&=\left(\mathrm{J}\ \mathrm{J}_{Q^2}{}^{4:1}\right)_{\mathrm{kijl}}\\
&=\mathrm{J}_{\mathrm{k:}}\ \left(\mathrm{J}_{Q^2}{}^{4:1}\right)_{:\mathrm{ijl}}\\
&=\mathrm{J}_{\mathrm{k:}}\ (\mathrm{J}_{Q^2})_{\mathrm{ijl:}}\\
&=\left(\left(\frac{\partial r_k}{\partial y_\alpha}\right)_{\alpha=1,\ldots,\mathrm{N}}\right)\cdot\left(\left(\frac{\partial^2}{\partial Q_i\,\partial Q_j}\frac{\partial r_l}{\partial y_\alpha}\right)_{\alpha=1,\ldots,\mathrm{N}}\right)\\
&=\sum_{\alpha=1}^{N}\frac{1}{m_\alpha}\ \frac{\partial r_k}{\partial x_\alpha}\ \frac{\partial^3 r_l}{\partial Q_i\,\partial Q_j\,\partial x_\alpha}
\end{aligned}$$

Now keeping in mind that we want

$$(G_{Q^2})_{\mathrm{ijkl}}=\frac{\partial^2}{\partial Q_i\,\partial Q_j}\left(\sum_{\alpha=1}^{N}\frac{1}{m_\alpha}\ \frac{\partial r_k}{\partial x_\alpha}\ \frac{\partial r_l}{\partial x_\alpha}\right)$$

$$=\frac{\partial}{\partial Q_i}\sum_{\alpha=1}^{N}\frac{1}{m_\alpha}\left(\frac{\partial^2 r_k}{\partial Q_j\,\partial x_\alpha}\,\frac{\partial r_l}{\partial x_\alpha}+\frac{\partial r_k}{\partial x_\alpha}\,\frac{\partial^2 r_l}{\partial Q_j\,\partial x_\alpha}\right)$$
$$=\sum_{\alpha=1}^{N}\frac{1}{m_\alpha}\left(\frac{\partial^3 r_k}{\partial Q_i\,\partial Q_j\,\partial x_\alpha}\,\frac{\partial r_l}{\partial x_\alpha}+\frac{\partial^2 r_k}{\partial Q_j\,\partial x_\alpha}\,\frac{\partial^2 r_l}{\partial Q_i\,\partial x_\alpha}+\frac{\partial^2 r_k}{\partial Q_i\,\partial x_\alpha}\,\frac{\partial^2 r_l}{\partial Q_j\,\partial x_\alpha}+\frac{\partial r_k}{\partial x_\alpha}\,\frac{\partial^3 r_l}{\partial Q_i\,\partial Q_j\,\partial x_\alpha}\right)$$

we see this has worked as desired

Relationship to Cartesians

We saw previously that we had

$$\mathrm{J}_Q=\nabla_{\mathrm{Q}}\mathrm{J}=\nabla_{\mathrm{Q}}\mathrm{Y}\odot\nabla_{\mathrm{Y}}\mathrm{J}=\mathrm{Y}_Q\,\mathrm{J}_{\mathrm{Y}}$$

Now we'll consider the second derivative form of this

$$\begin{aligned}\nabla_{\mathrm{Q}^2}\mathrm{J}&=\mathrm{J}_{Q^2}\\ &=(\mathrm{Y}_Q\,\mathrm{J}_{\mathrm{Y}})_Q\\ &=\mathrm{Y}_{Q^2}\,\mathrm{J}_{\mathrm{Y}}+\left(\mathrm{Y}_Q((\mathrm{J}_{\mathrm{Y}})_Q)^{2:1}\right)^{2:1}\\ &=\mathrm{Y}_{Q^2}\,\mathrm{J}_{\mathrm{Y}}+\left(\mathrm{Y}_Q(\mathrm{Y}_Q\,\mathrm{J}_{\mathrm{Y}^2})^{2:1}\right)^{2:1}\end{aligned}$$

What if our Q coordinates are already linear combinations of Cartesians? In that case, calling this transformation L, we have

$$\begin{aligned}\mathrm{J}&=\mathrm{LM}^{-\frac{1}{2}}\\ \mathrm{J}_{\mathrm{Y}}&=\emptyset\\ \mathrm{J}_{\mathrm{Y}^2}&=\emptyset\end{aligned}$$

Alternately, what if we our Q coordinates to be linear combinations of internals (R), but we have some set of modes $\mathrm{Q}^{\mathrm{Cart}}$ that are already defined as linear combinations of Cartesians?

In this case we have, calling that transformation L,

$$\mathrm{J}=\mathrm{LM}^{-\frac{1}{2}}\,\nabla_{\mathrm{X}}\mathrm{R}$$

(note that if R = X, this reduces to the previous case)

Then we have

$$\mathrm{J}_{\mathrm{Y}}=\left(\mathrm{LM}^{-\frac{1}{2}}\,\mathrm{R}_{\mathrm{X}}\right)_Y=\mathrm{LM}^{-\frac{1}{2}}(\mathrm{R}_{\mathrm{X}})_{\mathrm{Y}}=\mathrm{LM}^{-\frac{1}{2}}(\mathrm{X}_{\mathrm{Y}}\,\mathrm{R}_{\mathrm{X}^2})$$

keeping in mind what X_{Y} really is

$$\mathrm{X}_{\mathrm{Y}}=\begin{pmatrix}\frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} & \cdots & \frac{\partial x_1}{\partial y_N}\\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} & & \\ \vdots & & \ddots & \vdots\\ \frac{\partial x_N}{\partial y_1} & & \cdots & \frac{\partial x_N}{\partial y_N}\end{pmatrix}=\begin{pmatrix}\frac{1}{\sqrt{m_1}} & & & \\ & \frac{1}{\sqrt{m_2}} & & \\ & & \ddots & \\ & & & \frac{1}{\sqrt{m_N}}\end{pmatrix}=\mathrm{M}^{-\frac{1}{2}}$$

we get

$$\mathrm{J_Y}=\mathrm{LM}^{-1}\,\mathrm{R_{X^2}}$$

and

$$\mathrm{J_{Y^2}}=\left(\mathrm{LM}^{-\frac{1}{2}}\,\mathrm{R_X}\ \right)_Y=\left(\mathrm{LM}^{-1}\,\mathrm{R_{X^2}}\right)_Y=\mathrm{LM}^{-\frac{3}{2}}\,\mathrm{R_{X^3}}$$

Then we have the Y_Q terms, which will are slightly funky

$$\mathrm{Y}_Q\ =\ \mathrm{R_Q}\,\mathrm{Y_R}=\mathrm{J}^{-1}(\mathrm{X_R}\,\mathrm{Y_X})=\mathrm{J}^{-1}\,\mathrm{X_R}\,\mathrm{M}$$

these X_R can either be computed or by finding the pseudoinverse of R_X

Finally

$$\begin{aligned}\mathrm{Y}_{Q^2}&=(\mathrm{R_Q}\,\mathrm{X_R}\,\mathrm{M})_Q\\&=\mathrm{R_{Q^2}}\,\mathrm{X_R}\,\mathrm{M}+\left(\mathrm{R_Q}(\mathrm{X_R})_Q{}^{2:1}\right)^{2:1}\mathrm{M}\\&=\left(\mathrm{R_{Q^2}}\,\mathrm{X_R}+\left(\mathrm{R_Q}(\mathrm{R_Q}\,\mathrm{X_{R^2}})^{2:1}\right)^{2:1}\right)\mathrm{M}\end{aligned}$$

then we note that $\mathrm{R_{Q^2}}=\emptyset$ since we have a linear transform, giving us

$$\mathrm{Y}_{Q^2}=\left(\mathrm{J}^{-1}\left(\mathrm{J}^{-1}\,\mathrm{X_{R^2}}\right)^{2:1}\right)^{2:1}\mathrm{M}$$