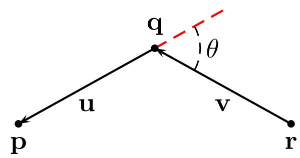
Chain rule in computation



Consider line segments with vertices

$$\mathbf{p} = \langle x_1, x_2, x_3 \rangle \qquad \mathbf{q} = \langle x_4, x_5, x_6 \rangle \qquad \mathbf{r} = \langle x_7, x_8, x_9 \rangle.$$

The edge vectors are

$$\mathbf{u} = \mathbf{p} - \mathbf{q} \qquad \mathbf{v} = \mathbf{q} - \mathbf{r}.$$

The angle between the edges is given by

$$\mathbf{u} \cdot \mathbf{v} = \|\mathbf{u}\| \|\mathbf{v}\| \cos\theta.$$

The line segments are connected by a spring whose rest angle is zero. The potential energy of the spring can be written as

$$\phi = \frac{1}{2} k \theta^2.$$

The forces can then be computed as

$$\mathbf{f} = -\nabla\phi = \left\langle -\frac{\partial\phi}{\partial x_1}, -\frac{\partial\phi}{\partial x_2}, -\frac{\partial\phi}{\partial x_3}, -\frac{\partial\phi}{\partial x_4}, -\frac{\partial\phi}{\partial x_5}, -\frac{\partial\phi}{\partial x_6}, -\frac{\partial\phi}{\partial x_7}, -\frac{\partial\phi}{\partial x_8}, -\frac{\partial\phi}{\partial x_9} \right\rangle,$$

In practice, we also want to use

$$\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \|\mathbf{v}\| \sin\theta.$$

With these,

$$\theta = \tan^{-1} \frac{\|\mathbf{u} \times \mathbf{v}\|}{\mathbf{u} \cdot \mathbf{v}}$$

Breaking up these computations into more manageable pieces:

$$\mathbf{u} = \mathbf{p} - \mathbf{q} \qquad \mathbf{v} = \mathbf{q} - \mathbf{r} \qquad d = \mathbf{u} \cdot \mathbf{v} \qquad \mathbf{c} = \mathbf{u} \times \mathbf{v} \qquad m = \|\mathbf{c}\| \qquad r = \frac{m}{d} \qquad \theta = \tan^{-1} r \qquad \phi = \frac{1}{2} k \theta^2$$

To compute the partial derivative $\frac{\partial\phi}{\partial x_k}$, differentiate each equation.

$$\begin{aligned}
&\frac{\partial \mathbf{u}}{\partial x_k} = \frac{\partial \mathbf{p}}{\partial x_k} - \frac{\partial \mathbf{q}}{\partial x_k} &&\frac{\partial \mathbf{v}}{\partial x_k} = \frac{\partial \mathbf{q}}{\partial x_k} - \frac{\partial \mathbf{r}}{\partial x_k} &&\frac{\partial d}{\partial x_k} = \frac{\partial \mathbf{u}}{\partial x_k} \cdot \mathbf{v} + \mathbf{u} \cdot \frac{\partial \mathbf{v}}{\partial x_k} &&\frac{\partial \mathbf{c}}{\partial x_k} = \frac{\partial \mathbf{u}}{\partial x_k} \times \mathbf{v} + \mathbf{u} \times \frac{\partial \mathbf{v}}{\partial x_k} \\
&\frac{\partial m}{\partial x_k} = \frac{\mathbf{c}}{m} \cdot \frac{\partial \mathbf{c}}{\partial x_k} &&\frac{\partial r}{\partial x_k} = \frac{1}{d} \frac{\partial m}{\partial x_k} - \frac{m}{d^2} \frac{\partial d}{\partial x_k} &&\frac{\partial \theta}{\partial x_k} = \frac{1}{r^2 + 1} \frac{\partial r}{\partial x_k} &&\frac{\partial \phi}{\partial x_k} = k\theta \frac{\partial \theta}{\partial x_k}
\end{aligned}$$

This is how the chain rule is often applied in practice.

In the case of this application, first derivatives are not sufficent. Second derivatives are required as well. This is fairly straightforward, though more tedious, to deal with. Note that

$$\frac{\partial^2 \mathbf{p}}{\partial x_k \partial x_\ell} = \frac{\partial^2 \mathbf{q}}{\partial x_k \partial x_\ell} = \frac{\partial^2 \mathbf{r}}{\partial x_k \partial x_\ell} = \frac{\partial^2 \mathbf{u}}{\partial x_k \partial x_\ell} = \frac{\partial^2 \mathbf{v}}{\partial x_k \partial x_\ell} = \mathbf{0}.$$

Next, we differentiate our first derivative formulas a second time.

$$\begin{aligned}
\frac{\partial^2 d}{\partial x_k \partial x_\ell} &= \frac{\partial^2 \mathbf{u}}{\partial x_k \partial x_\ell} \cdot \mathbf{v} + \frac{\partial \mathbf{u}}{\partial x_k} \cdot \frac{\partial \mathbf{v}}{\partial x_\ell} + \frac{\partial \mathbf{u}}{\partial \ell} \cdot \frac{\partial \mathbf{v}}{\partial x_k} + \mathbf{u} \cdot \frac{\partial^2 \mathbf{v}}{\partial x_k \partial x_\ell} \\
&= \frac{\partial \mathbf{u}}{\partial x_k} \cdot \frac{\partial \mathbf{v}}{\partial x_\ell} + \frac{\partial \mathbf{u}}{\partial \ell} \cdot \frac{\partial \mathbf{v}}{\partial x_k} \\
\frac{\partial^2 \mathbf{c}}{\partial x_k \partial x_\ell} &= \frac{\partial^2 \mathbf{u}}{\partial x_k \partial x_\ell} \times \mathbf{v} + \frac{\partial \mathbf{u}}{\partial x_k} \times \frac{\partial \mathbf{v}}{\partial x_\ell} + \frac{\partial \mathbf{u}}{\partial \ell} \times \frac{\partial \mathbf{v}}{\partial x_k} + \mathbf{u} \times \frac{\partial^2 \mathbf{v}}{\partial x_k \partial x_\ell} \\
&= \frac{\partial \mathbf{u}}{\partial x_k} \times \frac{\partial \mathbf{v}}{\partial x_\ell} + \frac{\partial \mathbf{u}}{\partial \ell} \times \frac{\partial \mathbf{v}}{\partial x_k} \\
\frac{\partial^2 m}{\partial x_k \partial x_\ell} &= -\frac{1}{m^2} \frac{\partial m}{\partial x_\ell} \mathbf{c} \cdot \frac{\partial \mathbf{c}}{\partial x_k} + \frac{1}{m} \frac{\partial \mathbf{c}}{\partial x_\ell} \cdot \frac{\partial \mathbf{c}}{\partial x_k} + \frac{\mathbf{c}}{m} \cdot \frac{\partial^2 \mathbf{c}}{\partial x_k \partial x_\ell} \\
&= \frac{1}{m} \left(-\frac{\partial m}{\partial x_\ell} \frac{\partial m}{\partial x_k} + \frac{\partial \mathbf{c}}{\partial x_\ell} \cdot \frac{\partial \mathbf{c}}{\partial x_k} + \mathbf{c} \cdot \frac{\partial^2 \mathbf{c}}{\partial x_k \partial x_\ell} \right)
\end{aligned}$$

For the next one, we can simplify matters a bit by first rewriting

$$\begin{aligned}
\frac{\partial r}{\partial x_k} &= \frac{1}{d} \frac{\partial m}{\partial x_k} - \frac{m}{d^2} \frac{\partial d}{\partial x_k} \\
&= \frac{1}{d} \left(\frac{\partial m}{\partial x_k} - r \frac{\partial d}{\partial x_k} \right) \\
\frac{\partial^2 r}{\partial x_k \partial x_\ell} &= -\frac{1}{d^2} \frac{\partial d}{\partial x_\ell} \left(\frac{\partial m}{\partial x_k} - r \frac{\partial d}{\partial x_k} \right) + \frac{1}{d} \left(\frac{\partial^2 m}{\partial x_k \partial x_\ell} - \frac{\partial r}{\partial x_\ell} \frac{\partial d}{\partial x_k} - r \frac{\partial^2 d}{\partial x_k \partial x_\ell} \right) \\
&= \frac{1}{d} \left(-\frac{\partial d}{\partial x_\ell} \frac{\partial r}{\partial x_k} + \frac{\partial^2 m}{\partial x_k \partial x_\ell} - \frac{\partial r}{\partial x_\ell} \frac{\partial d}{\partial x_k} - r \frac{\partial^2 d}{\partial x_k \partial x_\ell} \right)
\end{aligned}$$

For the next one, it helps to introduce a new variable

$$\begin{aligned}
s &= \frac{1}{r^2 + 1} \\
\frac{\partial s}{\partial x_\ell} &= -\frac{2r}{(r^2+1)^2} \frac{\partial r}{\partial x_\ell} = -2rs^2 \frac{\partial r}{\partial x_\ell} \\
\frac{\partial \theta}{\partial x_k} &= s \frac{\partial r}{\partial x_k} \\
\frac{\partial^2 \theta}{\partial x_k \partial x_\ell} &= \frac{\partial s}{\partial x_\ell} \frac{\partial r}{\partial x_k} + s \frac{\partial^2 r}{\partial x_k \partial x_\ell} \\
\frac{\partial^2 \phi}{\partial x_k \partial \ell} &= k \frac{\partial \theta}{\partial x_\ell} \frac{\partial \theta}{\partial x_k} + k\theta \frac{\partial^2 \theta}{\partial x_k \partial x_\ell}
\end{aligned}$$

This is a tedious but very practical application of calculus to a real problem. These calculations are typcial of what you might expect to see in areas such as numerical optimization, numerical simulation, or machine learning.