Notation Glossary:

z_j^l Pre-Activation of jth neuron in the lth layer

a_j^l (Post)-Activation of jth neuron in the lth layer

x_k Input (0th) layer activations

$A()$ General element-wise activation function

w_{jk}^l Weight matrix of lth layer

b_j^l Bias vector of lth layer

g_{jk}^l Jacobian between layer l and $l-1$

J_{jk}^l Jacobian between input (0th) layer and layer l

H_{jk}^l Diagonal Hessian between input (0th) layer and layer l

$\odot$ Direct product over repeated index rather than summation

Network Jacobian:

In following derivations assume Einstein index convention and the standard feed forward relationships:

$$a_j^l = A(z_j^l) \tag{1}$$
$$z_j^l = w_{jk}^l a_k^{l-1} + b_j^l \tag{2}$$

Note that the network's output corresponds to the Lth layer activations a_j^L. Begin by considering the layer to layer Jacobian:

$$\frac{\partial a_j^l}{\partial a_k^{l-1}} = \frac{\partial a_j^l}{\partial z_q^l} \frac{\partial z_q^l}{\partial a_k^{l-1}} \tag{3}$$

$$= \frac{\partial a_j^l}{\partial z_q^l} w_{qk}^l \tag{4}$$

$$g_{jk}^l = A'(z_j^l) \odot w_{jk}^l \tag{5}$$

Going 1 more layer back:

$$\frac{\partial a_j^l}{\partial a_k^{l-2}} = A'(z_j^l) \odot w_{jq}^l A'(z_q^{l-1}) \odot w_{qk}^{l-1} \tag{6}$$

$$= g_{jq}^l g_{qk}^{l-1} \tag{7}$$

This relationship can be recurred to obtain:

$$J_{jk}^l \equiv \frac{\partial a_j^l}{\partial x_k} = g_{j\alpha}^l g_{\alpha\beta}^{l-1} ... g_{\mu k}^1 \tag{8}$$

Network Hessian:

The general layer-to-layer Hessian can be expressed as:

$$H_{jkm}^l \equiv \frac{\partial^2 a_j^l}{\partial a_m^{l-1} \partial a_k^{l-1}} = \frac{\partial J_{jk}^l}{\partial a_m^{l-1}} \tag{9}$$

The full Hessian is computationally expensive to obtain and often we will only need its diagonal terms which will retrieve the network's estimates for commonly occurring physical differential terms such as: $\frac{d^2 u}{dx^2}$ for some $u(t, x)$. Henceforth let:

$$\delta_{mk} H^l_{jkm} \equiv H^l_{jk} \tag{10}$$

We can also note the useful base cases:

$$J^1_{jk} = \delta_{jk} \tag{11}$$

$$H^1_{jk} = 0 \tag{12}$$

To derive the diagonal Hessian we consider a recursive method and start from:

$$J^l_{jk} \equiv g^l_{jm} J^{l-1}_{mk} = A'(z^l_j) \odot w^l_{jm} J^{l-1}_{mk} \tag{13}$$

Considering the Hessian as the derivative of the Jacobian:

$$H^l_{jk} = \frac{\partial}{\partial x_k} \left[A'(z^l_j) \odot w^l_{jm} J^{l-1}_{mk} \right] \tag{14}$$

$$= A''(z^l_j) \odot \frac{\partial z^l_j}{\partial x_k} \odot w^l_{jm} J^{l-1}_{mk} + A'(z^l_j) \odot w^l_{jm} \frac{\partial J^{l-1}_{mk}}{\partial x_k} \tag{15}$$

We can recover several familiar terms from this expression, starting with

$$\frac{\partial J^{l-1}_{mk}}{\partial x_k} \equiv H^{l-1}_{mk}, \tag{16}$$

and likewise, simplifying the first term:

$$\frac{\partial z^l_j}{\partial x_k} = \frac{\partial}{\partial x_k} \left[w^l_{jq} a^{l-1}_q + b^l_j \right] = w^l_{jq} \frac{\partial a^{l-1}_q}{\partial x_k} = w^l_{jq} J^{l-1}_{qk}. \tag{17}$$

Putting these expressions together we yield a recursion relation for the diagonal Hessian:

$$H^l_{jk} = A''(z^l_j) \odot \left[w^l_{jm} J^{l-1}_{mk} \right]^2 + A'(z^l_j) \odot w^l_{jm} H^{l-1}_{mk} \tag{18}$$

Jacobian and Hessian Loss Derivatives:

If we construct a Physics Loss using Jacobian and/or Hessian terms, we will need to propagate the according loss derivative through the network (standard backpropagation). This will require: $\frac{\partial J^L_{jk}}{\partial a^L_m}$ and $\frac{\partial H^L_{jk}}{\partial a^L_m}$. We begin with the simpler Jacobian. We use the form $A'(z^l_j) \odot w^l_{jm} J^{l-1}_{mk}$ which brings all the a^L_j dependence into the activation function term.

$$\frac{\partial J^L_{jk}}{\partial a^L_m} = \frac{\partial}{\partial a^L_m} \left[A'(z^L_j) \odot w^L_{jq} J^{L-1}_{qk} \right] \tag{19}$$

$$= \delta_{jm} A''(z^L_j) \odot \frac{\partial z^L_j}{\partial a^L_m} w^L_{jq} J^{L-1}_{qk} \tag{20}$$

$$= \delta_{jm} \frac{A''(z^L_j)}{A'(z^L_j)} \odot \left[w^L_{jq} J^{L-1}_{qk} \right] \tag{21}$$

$$= \delta_{jm} \frac{A''(z^L_j)}{[A'(z^L_j)]^2} \odot J^L_{jk} \tag{22}$$

(The δ_{jm} has been left for completeness to demonstrate the $j = m$ requirement)

The derivation for the Hessian follows much of the same logic and yields:

$$\frac{\partial H^L_{jk}}{\partial a^L_m} = \delta_{jm} \left[\frac{A'''(z^L_j)}{A'(z^L_j)} \odot \left[w^L_{jq} J^{L-1}_{qk} \right]^2 + \frac{A''(z^L_j)}{A'(z^L_j)} \odot w^L_{jq} H^{L-1}_{qk} \right] \tag{23}$$

$$= \delta_{jm} \left[\frac{A'''(z^L_j)}{[A'(z^L_j)]^3} \odot (J^L_{jk})^2 + \frac{A''(z^L_j)}{A'(z^L_j)} \odot w^L_{jq} H^{L-1}_{qk} \right] \tag{24}$$