1 Tensors

2 Backward pass for basic operations

2.1 Tensor Addition

Let A and B be two matrices and $C = A + B$. Suppose that the gradient of a scalar z with respect to C is given by G. Both $\frac{\partial z}{\partial A}$ and $\frac{\partial z}{\partial B}$ need to be computed in order to complete the backward pass. The chain rule gives the two following equations:

2.2 Tensor multiplication

Let A and B be two matrices and $C = AB$ the result of their matrix product. Suppose that the gradient of a scalar z with respect to C is given by G. Both $\frac{\partial z}{\partial A}$ and $\frac{\partial z}{\partial B}$ need to be computed in order to complete the backward pass.

The differentials of C and G are:

$$dC = dA\ B + A\ dB \text{ and } dz = G : dC \tag{1}$$

Replacing the value of dC in the secound equation yields:

$$dz = (G : dA\ B) + (G : A\ dB) = (GB^T : dA) + (A^T G : dB) \tag{2}$$

The gradient of z with respect to both A and B is therefore:

$$\frac{\partial z}{\partial A} = GB^T \text{ and } \frac{\partial z}{\partial B} = A^T G \tag{3}$$

2.3 Hadamard product

Let A and B be two matrices and $C = A \odot B$ the result of their Hadamard product. Suppose that the gradient of a scalar z with respect to C is given by G. Both $\frac{\partial z}{\partial A}$ and $\frac{\partial z}{\partial B}$ need to be computed in order to complete the backward pass.

Since the Hadamard product if bilinear, the differentials of C and G are:

$$dC = dA \odot B + A \odot dB \text{ and } dz = G : dC \tag{4}$$

Replacing the value of dC in the secound equation yields:

$$dz = (G : dA \odot B) + (G : A \odot dB) = (G : dA : B) + (G : A : dB) \tag{5}$$

The gradient of z with respect to both A and B is therefore:

$$\frac{\partial z}{\partial A} = G : B \text{ and } \frac{\partial z}{\partial B} = G : A \tag{6}$$