TTIC 31230, Fundamentals of Deep Learning

David McAllester

CNNs: Einstein Notation

Einstein Notation

For the representation of general relativity, Einstein introduced the convention of explicitly writing all indeces of tensors where repeated indeces in a product of tensors are implicitly summed.

Writing indeces explicitly improves the clarity of the notation at the expense of not being in correspondence with framework notation. Most frameworks hide indeces.

This course will focus on conceptual understanding rather than framework implementations. For conceptual understanding Einstein notation seems preferable.

Einstein Notation

Einstein notation improves tensor equations for tensors with many indeces.

However, we start by considering just vectors and matrices.

We will use a modified form of Einstein notation where captial letters are used to denote slices of a tensor. For example:

- $M[i, j]$ denotes one element of the matrix M.
- $M[i, J]$ denotes the ith row of M.
- $M[I, j]$ denotes the jth collumn of M.
- $M[I, J]$ denotes the full matrix M.

Einstein Notation

Repeated capital letters in a product of tensors denote summation over those letters.

$$
\begin{aligned}
y = Wx \quad & \equiv \quad y[i] = \sum_j W[i,j]x[j] \\
& \equiv \quad y[i] = W[i,J]x[J]
\end{aligned}
$$

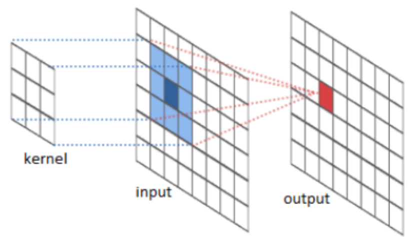
$$
\begin{aligned}
y = x^\top W \quad & \equiv \quad y[j] = \sum_i W[i,j]x[i] \\
& \equiv \quad y[j] = W[I,j]x[I]
\end{aligned}
$$

Einstein Notation for Convolution

CNNs provide a good example of the advantage of Einstein Notation.

$L[b, x, y, i]$ is the value of "neuron" i for batch element b at image position $\langle x, y \rangle$.

Convolution



$W[\Delta x, \Delta y, i, j] \qquad L_\ell[b, x, y, i] \qquad L_{\ell+1}[b, x, y, j]$

River Trail Documentation

$$L_{\ell+1}[b, x, y, j] \;=\; \sigma\left(\sum_{\Delta x, \Delta y, i} W[\Delta x, \Delta y, i, j]\; L_\ell[b, x+\Delta x, y+\Delta y, i] - B[j]\right)$$

$$=\; \sigma\left(W[\Delta X, \Delta Y, I, j]\; L_\ell[b, x+\Delta X, y+\Delta Y, I] - B[j]\right)$$

Types and Einstein Notation

The indeces of tensors generally have types such as a "time index", "x coordinate", "y coordinate", "batch index", or "neuron index".

Writing a matrix as $W[T, I]$ where T is a time index and I is a feature index makes the type of the matrix W clear and clarifies the order of the indeces (disambiguates W from $W^\top$).

Writing a layer of a CNN as $L[B, X, Y, I]$ clarifies both the types and the positions of the four indeces.

END