BACKPROPAGATION IN CONVOLUTIONAL LSTMS

Ankur Handa and Viorica Pătrăucean
Department of Engineering
University of Cambridge
Cambridge, CB2 1PZ
{handa.ankur, vpatrauc}@gmail.com

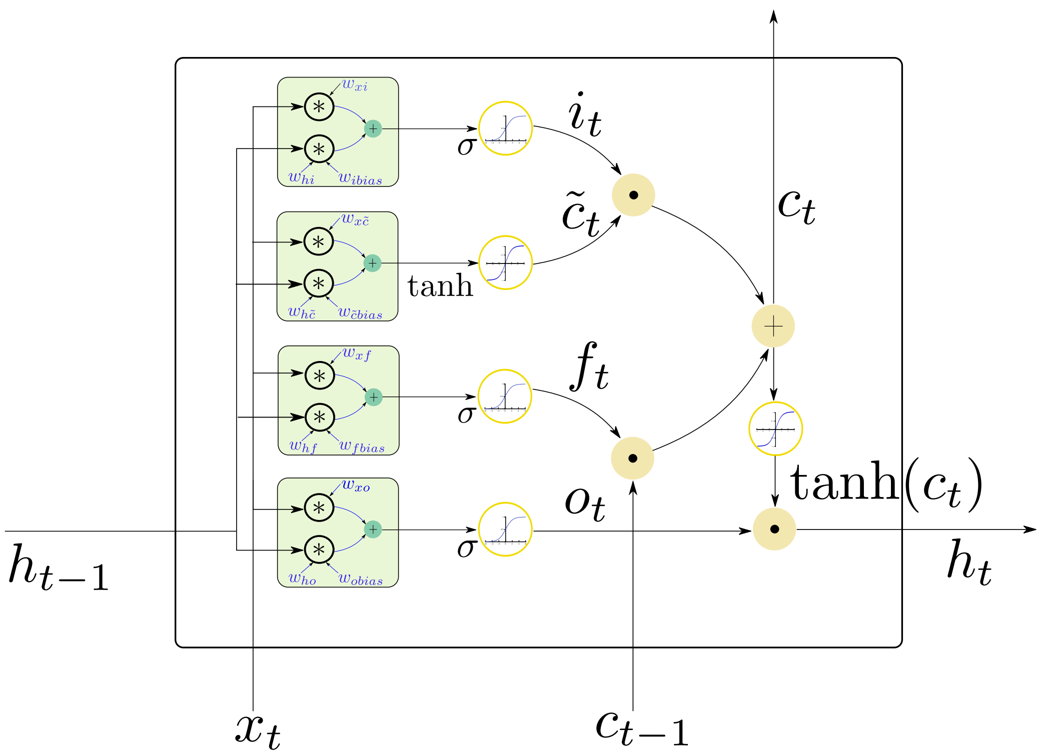

Figure 1: LSTM flow diagram. Note that h_0 and c_0 are initialised to zero. Loops and recursions are explicitly avoided for clarity.

Backpropagation derivations of convolutional LSTMs introduced in Stollenga et al. (2015); Xingjian et al. (2015); Patraucean et al. (2015); Romera-Paredes & Torr (2015)

1 CONVOLUTIONAL LSTM

$$
\begin{aligned}
i_t &= \sigma(x_t * w_{xi} + h_{t-1} * w_{hi} + w_{ibias}) \\
f_t &= \sigma(x_t * w_{xf} + h_{t-1} * w_{hf} + w_{fbias}) \\
\tilde{c}_t &= \tanh(x_t * w_{x\tilde{c}} + h_{t-1} * w_{h\tilde{c}} + w_{\tilde{c}bias}) \\
o_t &= \sigma(x_t * w_{xo} + h_{t-1} * w_{ho} + w_{obias}) \\
c_t &= \tilde{c}_t \odot i_t + c_{t-1} \odot f_t \\
h_t &= o_t \odot \tanh(c_t)
\end{aligned}
$$

1.1 Given: δh_t, find δo_t and δc_t

1.1.1 $\frac{\delta E}{\delta o_t}$

$$\frac{\delta E}{\delta o_t} = \frac{\delta E}{\delta h_t} \cdot \frac{\delta h_t}{\delta o_t}$$
$$\frac{\delta E}{\delta o_t} = \frac{\delta E}{\delta h_t} \odot \tanh(c_t)$$

1.1.2 $\frac{\delta E}{\delta c_t}$

$$\frac{\delta E}{\delta c_t} = \frac{\delta E}{\delta h_t} \cdot \frac{\delta h_t}{\delta c_t}$$
$$\frac{\delta E}{\delta c_t} = \frac{\delta E}{\delta h_t} \odot o_t \odot (1 - \tanh^2(c_t))$$

1.2 Given: δc_t, find δi_t, δf_t, $\delta \tilde{c}_t$, δc_{t-1} and δh_{t-1}

1.2.1 $\frac{\delta E}{\delta i_t}$

$$\frac{\delta E}{\delta i_t} = \frac{\delta E}{\delta c_t} \cdot \frac{\delta c_t}{\delta i_t}$$
$$\frac{\delta E}{\delta i_t} = \frac{\delta E}{\delta c_t} \odot \tilde{c}_t$$

1.2.2 $\frac{\delta E}{\delta f_t}$

$$\frac{\delta E}{\delta f_t} = \frac{\delta E}{\delta c_t} \cdot \frac{\delta c_t}{\delta f_t}$$
$$\frac{\delta E}{\delta f_t} = \frac{\delta E}{\delta c_t} \odot c_{t-1}$$

1.2.3 $\frac{\delta E}{\delta \tilde{c}_t}$

$$\frac{\delta E}{\delta \tilde{c}_t} = \frac{\delta E}{\delta c_t} \cdot \frac{\delta c_t}{\delta \tilde{c}_t}$$
$$\frac{\delta E}{\delta \tilde{c}_t} = \frac{\delta E}{\delta c_t} \odot i_t$$

1.2.4 $\frac{\delta E}{\delta c_{t-1}}$

$$\frac{\delta E}{\delta c_{t-1}} = \frac{\delta E}{\delta c_t} \cdot \frac{\delta c_t}{\delta c_{t-1}}$$
$$\frac{\delta E}{\delta c_{t-1}} = \frac{\delta E}{\delta c_t} \odot f_t$$

1.2.5 $\frac{\delta E}{\delta h_{t-1}}$

$$\frac{\delta E}{\delta h_{t-1}} = \frac{\delta E}{\delta i_t} \cdot \frac{\delta i_t}{\delta h_{t-1}} + \frac{\delta E}{\delta o_t} \cdot \frac{\delta o_t}{\delta h_{t-1}} + \frac{\delta E}{\delta f_t} \cdot \frac{\delta f_t}{\delta h_{t-1}} + \frac{\delta E}{\delta \tilde{c}_t} \cdot \frac{\delta \tilde{c}_t}{\delta h_{t-1}}$$

1.3 GIVEN: δo_t, FIND δw_{xo}, δw_{ho} AND δw_{obias}

1.3.1 $\frac{\delta E}{\delta w_{xo}}$

$$\frac{\delta E}{\delta w_{xo}} = \sum_{t=1}^{T} \frac{\delta E}{\delta o_t} \cdot \frac{\delta o_t}{\delta w_{xo}}$$

$$\frac{\delta E}{\delta w_{xo}} = \sum_{t=1}^{T} \left(\frac{\delta E}{\delta o_t} \odot \sigma'(o_t) \right) * x_t$$

1.3.2 $\frac{\delta E}{\delta w_{ho}}$

$$\frac{\delta E}{\delta w_{ho}} = \sum_{t=1}^{T} \frac{\delta E}{\delta o_t} \cdot \frac{\delta o_t}{\delta w_{ho}}$$

$$\frac{\delta E}{\delta w_{xo}} = \sum_{t=1}^{T} \left(\frac{\delta E}{\delta o_t} \odot \sigma'(o_t) \right) * h_{t-1}$$

1.3.3 $\frac{\delta E}{\delta w_{obias}}$

$$\frac{\delta E}{\delta w_{obias}} = \sum_{t=1}^{T} \frac{\delta E}{\delta o_t} \cdot \frac{\delta o_t}{\delta w_{obias}}$$

$$\frac{\delta E}{\delta w_{obias}} = \sum_{t=1}^{T} \left(\frac{\delta E}{\delta o_t} \odot \sigma'(o_t) \right)$$

1.4 GIVEN: δi_t, FIND δw_{xi}, δw_{hi} AND δw_{ibias}

1.4.1 $\frac{\delta E}{\delta w_{xi}}$

$$\frac{\delta E}{\delta w_{xi}} = \sum_{t=1}^{T} \frac{\delta E}{\delta i_t} \cdot \frac{\delta i_t}{\delta w_{xi}}$$

$$\frac{\delta E}{\delta w_{xi}} = \sum_{t=1}^{T} \left(\frac{\delta E}{\delta i_t} \odot \sigma'(i_t) \right) * x_t$$

1.4.2 $\frac{\delta E}{\delta w_{hi}}$

$$\frac{\delta E}{\delta w_{hi}} = \sum_{t=1}^{T} \frac{\delta E}{\delta i_t} \cdot \frac{\delta i_t}{\delta w_{hi}}$$

$$\frac{\delta E}{\delta w_{hi}} = \sum_{t=1}^{T} \left(\frac{\delta E}{\delta i_t} \odot \sigma'(i_t) \right) * h_{t-1}$$

1.4.3 $\frac{\delta E}{\delta w_{ibias}}$

$$\frac{\delta E}{\delta w_{ibias}} = \sum_{t=1}^{T} \frac{\delta E}{\delta i_t} \cdot \frac{\delta i_t}{\delta w_{ibias}}$$

$$\frac{\delta E}{\delta w_{ibias}} = \sum_{t=1}^{T} \left(\frac{\delta E}{\delta i_t} \odot \sigma'(i_t) \right)$$

1.5 Given: δf_t, find δw_{xf}, δw_{hf} and δw_{fbias}

1.5.1 $\frac{\delta E}{\delta w_{xf}}$

$$\frac{\delta E}{\delta w_{xi}} = \sum_{t=1}^{T} \frac{\delta E}{\delta i_t} \cdot \frac{\delta i_t}{\delta w_{xi}}$$

$$\frac{\delta E}{\delta w_{xi}} = \sum_{t=1}^{T} \left(\frac{\delta E}{\delta i_t} \odot \sigma'(i_t) \right) * x_t$$

1.5.2 $\frac{\delta E}{\delta w_{hf}}$

$$\frac{\delta E}{\delta w_{hf}} = \sum_{t=1}^{T} \frac{\delta E}{\delta f_t} \cdot \frac{\delta f_t}{\delta w_{hi}}$$

$$\frac{\delta E}{\delta w_{hf}} = \sum_{t=1}^{T} \left(\frac{\delta E}{\delta f_t} \odot \sigma'(f_t) \right) * h_{t-1}$$

1.5.3 $\frac{\delta E}{\delta w_{fbias}}$

$$\frac{\delta E}{\delta w_{fbias}} = \sum_{t=1}^{T} \frac{\delta E}{\delta f_t} \cdot \frac{\delta f_t}{\delta w_{fbias}}$$

$$\frac{\delta E}{\delta w_{fbias}} = \sum_{t=1}^{T} \left(\frac{\delta E}{\delta f_t} \odot \sigma'(f_t) \right)$$

1.6 Given: $\delta \tilde{c}_t$, find $\delta w_{x\tilde{c}}$, $\delta w_{h\tilde{c}}$ and $\delta w_{\tilde{c}bias}$

1.6.1 $\frac{\delta E}{\delta w_{x\tilde{c}}}$

$$\frac{\delta E}{\delta w_{x\tilde{c}}} = \sum_{t=1}^{T} \frac{\delta E}{\delta \tilde{c}_t} \cdot \frac{\delta \tilde{c}_t}{\delta w_{x\tilde{c}}}$$

$$\frac{\delta E}{\delta w_{x\tilde{c}}} = \sum_{t=1}^{T} \left(\frac{\delta E}{\delta \tilde{c}_t} \odot (1 - \tanh^2(\tilde{c}_t) \right) * x_t$$

1.6.2 $\frac{\delta E}{\delta w_{h\tilde{c}}}$

$$\frac{\delta E}{\delta w_{h\tilde{c}}} = \sum_{t=1}^{T} \frac{\delta E}{\delta \tilde{c}_t} \cdot \frac{\delta \tilde{c}_t}{\delta w_{h\tilde{c}}}$$

$$\frac{\delta E}{\delta w_{h\tilde{c}}} = \sum_{t=1}^{T} \left(\frac{\delta E}{\delta \tilde{c}_t} \odot (1 - \tanh^2(\tilde{c}_t) \right) * h_{t-1}$$

1.6.3 $\frac{\delta E}{\delta w_{\tilde{c}bias}}$

$$\frac{\delta E}{\delta w_{\tilde{c}bias}} = \sum_{t=1}^{T} \frac{\delta E}{\delta f_t} \cdot \frac{\delta f_t}{\delta w_{fbias}}$$

$$\frac{\delta E}{\delta w_{\tilde{c}bias}} = \sum_{t=1}^{T} \left(\frac{\delta E}{\delta \tilde{c}_t} \odot (1 - \tanh^2(\tilde{c}_t) \right)$$

ACKNOWLEDGEMENTS

http://arunmallya.github.io/writeups/nn/lstm/index.html#/
https://wiki.inf.ed.ac.uk/twiki/pub/MLforNLP/WebHome/lstm_intro.pdf FeedforwardSequentialMemoryNetworks: ANewStructuretoLearnLong-termDependency

REFERENCES

Patraucean, Viorica, Handa, Ankur, and Cipolla, Roberto. Spatio-temporal video autoencoder with differentiable memory. *CoRR*, abs/1511.06309, 2015. URL http://arxiv.org/abs/1511.06309.

Romera-Paredes, Bernardino and Torr, Philip H. S. Recurrent instance segmentation. *CoRR*, abs/1511.08250, 2015.

Stollenga, Marijn F, Byeon, Wonmin, Liwicki, Marcus, and Schmidhuber, Juergen. Parallel multi-dimensional lstm, with application to fast biomedical volumetric image segmentation. In *Advances in Neural Information Processing Systems*, 2015.

Xingjian, SHI, Chen, Zhourong, Wang, Hao, Yeung, Dit-Yan, Wong, Wai-kin, and Woo, Wang-chun. Convolutional lstm network: A machine learning approach for precipitation nowcasting. In *Advances in Neural Information Processing Systems*, 2015.