**2.1 Step forward**

since this is batch processing, x is placed before Wx.

**2.2 Step backward**

db is obtained by summing all the instances of the batch of da.

**3.1 Rnn forward**

3.2 Rnn backward

5 Rnn loss