Result comparison:

The comparison between my implementation and torch implementation is shown below:

Parameters	My Result	Torch GRUCell Result
Forward	[-0.23860644 0.76465206]	[-0.2298 0.7665]
dx	[-0.16303919 -0.58306715 -0.27585066 0.19997268 0.07193348]	[-0.1651 -0.5723 -0.2682 0.1999 0.0721]
dh	[-0.14989419 -1.18391592]	[-0.1236 -1.1763]

The result is showing an acceptable match for all the parameters.