Assignment-3

Part-1

1

1.

$$\frac{\partial h_t}{\partial f_t} = \frac{\partial(o_t * tanh(C_t))}{\partial f_t}$$

$$= o_t * \frac{\partial tanh(C_t)}{\partial f_t}$$

$$= o_t * (\frac{\partial tanh(C_t)}{\partial C_t} * \frac{\partial C_t}{\partial f_t})$$

$$= o_t * ((1 - tanh^2(C_t)) * C_{t-1})$$

2.

$$\frac{\partial h_t}{\partial i_t} = \frac{\partial(o_t * tanh(C_t))}{\partial i_t}$$

$$= o_t * \frac{\partial tanh(C_t)}{\partial i_t}$$

$$= o_t * (\frac{\partial tanh(C_t)}{\partial C_t} * \frac{\partial C_t}{\partial i_t})$$

$$= o_t * ((1 - tanh^2(C_t)) * \bar{C_t})$$

3.

$$\frac{\partial h_t}{\partial \bar{C}_t}$$

$$= \frac{\partial (o_t * tanh(C_t))}{\partial \bar{C}_t}$$

$$= o_t * \frac{\partial tanh(C_t)}{\partial \bar{C}_t}$$

$$= o_t * (\frac{\partial tanh(C_t)}{\partial C_t} * \frac{\partial C_t}{\partial \bar{C}_t})$$

$$= o_t * ((1 - tanh^2(C_t)) * i_t)$$

4.

$$\frac{\partial h_t}{\partial C_t}$$

$$= \frac{\partial (o_t * tanh(C_t))}{\partial C_t}$$

$$= o_t * \frac{\partial tanh(C_t)}{\partial C_t}$$

$$= o_t * (1 - tanh^2(C_t))$$

5.

$$\frac{\partial h_t}{\partial C_{t-1}}$$

$$= \frac{\partial(o_t * tanh(C_t))}{\partial C_{t-1}}$$

$$= o_t * \frac{\partial tanh(C_t)}{\partial C_{t-1}}$$

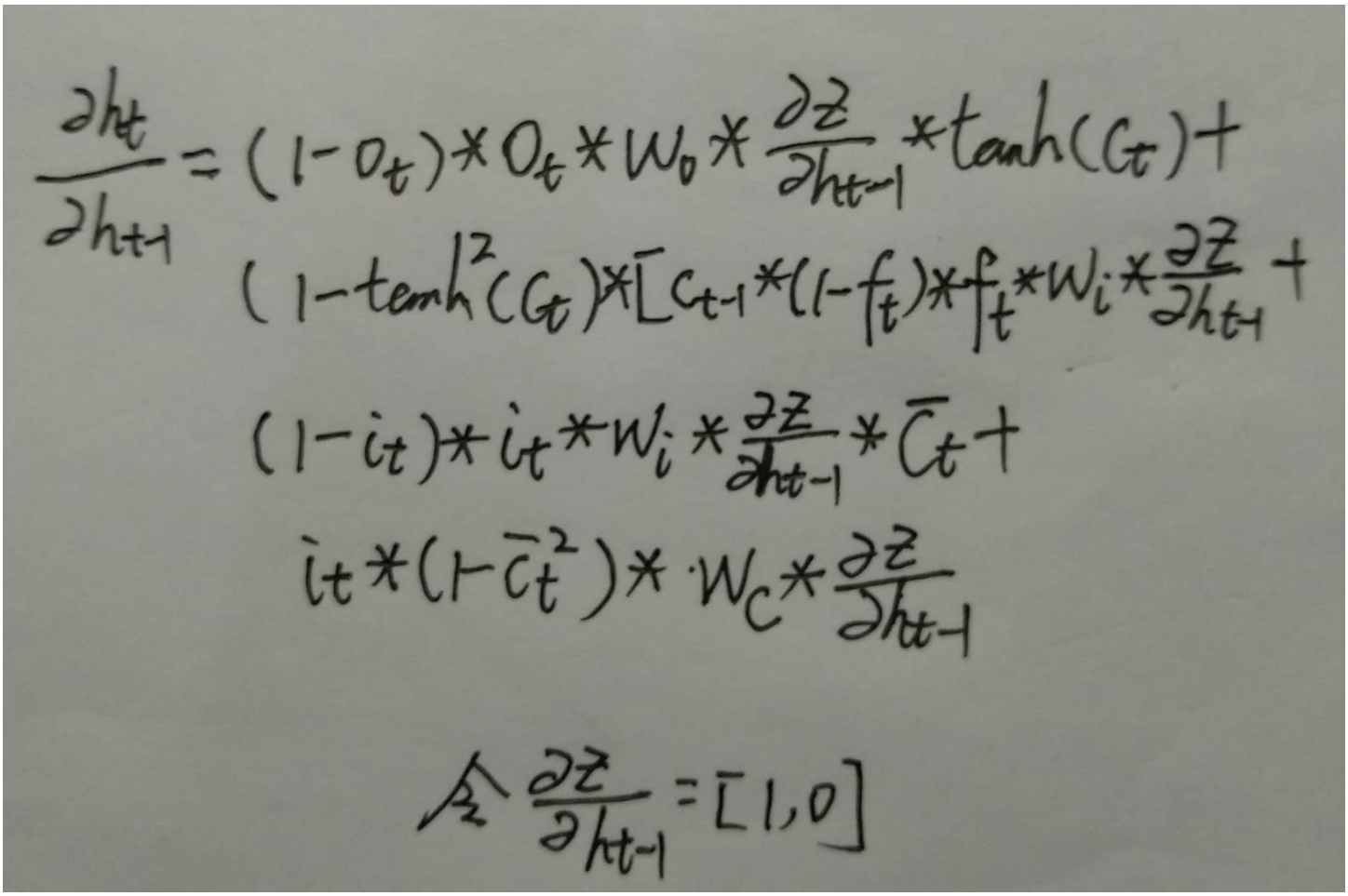
$$= o_t * (\frac{\partial tanh(C_t)}{\partial C_t} * \frac{\partial C_t}{\partial C_{t-1}})$$

$$= o_t * ((1 - tanh^2(C_t)) * f_t)$$

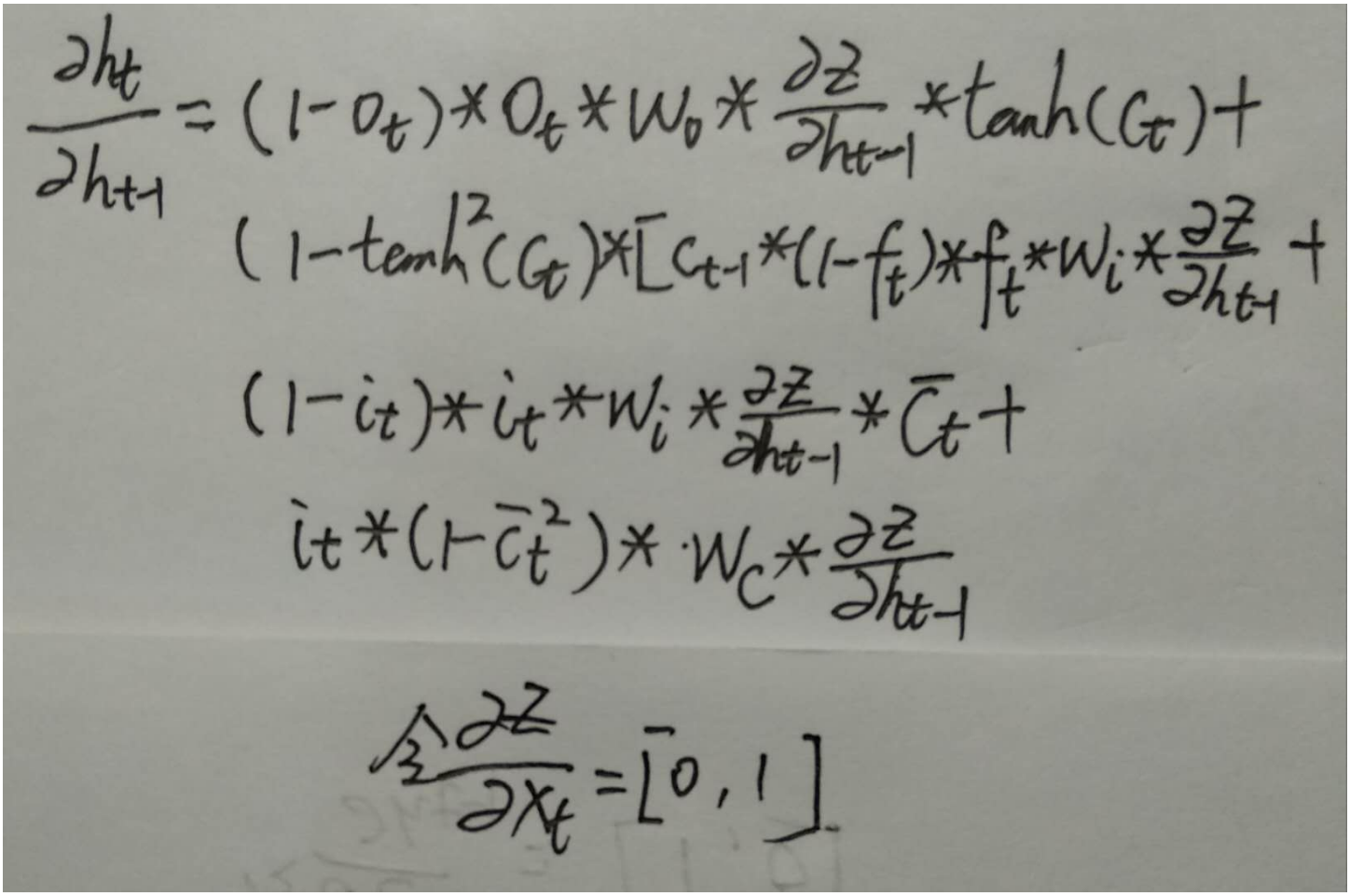
6.

$$\frac{\partial h_t}{\partial o_t} = tanh(C_t)$$

7.



8.



9.

$$\frac{\partial h_t}{\partial W_f} = o_t * (1 - tanh^2(C_t)) * C_{t-1} * (1 - f_t) * f_t * z$$

10.

$$\frac{\partial h_t}{\partial W_i} = o_t * (1 - tanh^2(C_t)) * \bar{C}_t * (1 - i_t) * i_t * z$$

11.

$$\frac{\partial h_t}{\partial W_C} = o_t * (1 - tanh^2(C_t)) * i_t * (1 - \bar{C}_t * \bar{C}_t) * z$$

12.

$$\frac{\partial h_t}{\partial W_o} = tanh(C_t) * (1 - o_t) * o_t * z$$

13.

$$\frac{\partial h_t}{\partial b_f} = o_t * (1 - tanh^2(C_t)) * C_{t-1} * (1 - f_t) * f_t$$

14.

$$\frac{\partial h_t}{\partial b_i} = o_t * (1 - tanh^2(C_t)) * \bar{C}_t * (1 - i_t) * i_t$$

15.

$$\frac{\partial h_t}{\partial b_C} = o_t * (1 - tanh^2(C_t)) * i_t * (1 - \bar{C}_t * \bar{C}_t)$$

16.

$$\frac{\partial h_t}{\partial b_o} = tanh(C_t) * (1 - o_t) * o_t$$

2

Given equations in problems describe the forward propagation process.
First, we define loss function of time t:

$$l(t) = (h(t) - y(t))^2$$

y(t) is the output and h(t) is the prediciton. Let T be last time and accumulated loss from 1 to T is

$$L(t) = \sum_{t=1}^{T} l(t)$$

differentiating L(t) for any parameter equals to multiplication of differentiating L(t) for output each time
and differentiating output each time for that parameter.

$$\frac{dL}{dw} = \sum_{t=1}^{T} \sum_{t=1}^{T} \frac{dL}{dh_i(t)} \frac{dh_i(t)}{dw}$$

T also means number of neural cell. Because of the forward propagation, change of h_i(t) with not influence loss value before t. We can know that L(t)=l(t)+L(t+1). Then we have

$$\frac{dL(t)}{dh_i(t)} = \frac{dl(t)}{dh_i(t)} + \frac{dL(t+1)}{dh_i(t)}$$

So we can calculate differentiation of time t and then t-1 and then ... iteratively. At time t:

$$\frac{dL(T)}{dh_i(T)} = \frac{dl(T)}{dh_i(T)}$$

from the inference above, output values from 1 to T can effect loss value. So it is necessary to calculate differentiation from T to 1.
And that is the backpropagation.

Part-2

1

embeding layer && parameters initializing:
If they are all set to be 0, backpropagation differentiation of each elements can be same. Eventually, the weights are all same. That limits the capability of the model. So we should randomly initilize them.

2

the poem:
start with 日

日出东方隈，有花不开台。深开大江水，畏底吴津台。但见寝处深，不因身为来。绿头满手去，白马空裴回。嫁取吴越诀，走马入荒台。但知苦声调，愿使金施开。谁言落落时，欲立还自隈。借问谁会母，宁知死命嗤。朱门每再闭，玉堂亦云开。愿学彩霞下，恩情红粉开。惜哉未及归，顾我长自嗟。忠信如不贵，还从李陵来。

start with 红

红轮照浦初千里，君王拂云朝不在。朱门迢遰万年心，丹碧青青堆更秀。含颦却更奏真妆，歛影转春殊不易。冶黛长愁无所思，阿侯堕笔成流血。少年容貌犹未进，莫遣腰枝定如水。如今侠少谪仙宫，古木凝中感洛风。才子无媒不相见，桃花万木响潺潺。惜时人去又初夕，未见故园秋路春。蹔得严陵守灵药，夜深几许陶潜诗。

start with 山
山中有客牵金壶，翛然止止荡乱交。钓鱼基湾绿苗窄，白鱼蟹带红蓝殷。采盘采苏在田口，屧使燕兵下巢屋。儿家老者慕沽醉，觉来觉火看飧又。欲知老身忽为愁，东中仆娶多高树。怕逢小翁长食茗，手把一面蒙素手。手竿裙子皮死，欲炙鸦纵入山腹。君去头边一曲尘，放君一败怜贵人。客舍无成沙漠漠，家家骨肉天下许。东家少年有长病，酒肉酒肉长自悲。琵琶女儿长忆慰，庄周不解嫌迷木。劝君莫言前后属，结为入夜不相识。

start with 夜

夜色尚苍苍，槐阴夹路长。别离嗟惜老，病物喜欢良。枕冷偏相似，伤忧洗一成。城头应见树，枕上即闻天。月没频投桂，秋深寂比丛。乍孤砧杵响，夜半壁栖鸿。帘裛闻弦管，灯虚识味通。有琴闲独坐，向伴酒杯空。越徼巫山渡，乌声楚部通。故人龟鹤路，绝骑蹴江宫。独宿千帆里，孤帆独坐通。夕阳迷极目，朝日照回空。羁旅还如此，愁知喜渐同。

start with 湖

湖南无村落，山舍少翁耕。茅屋本无主，岂知所近称。官曹不营市，相引来寻年。诗卷小侯籍，诏题蜀氏经。此中见游子，兄弟得专甥。官职从来否，儿家须醉行。朝廷退妻子，秋村剧桑耕。今日县城下，故人发故城。迴塘水色急，新柳两丝生。一摘野田酒，两家呼一声。气行觉可算，道拙安其。危堞应未疲，远远强异城。

start with 海

内昔年狎太平，徐下欲飞生古霜。胡笳变地为羣乌，短颈绕翅交飞鸣。帆鞭形力眼无定，流光忽射风雷争。丁壮可以阳和声，萧萧白露生离魂。睛见年年七十余，白衣白发生闲情。灵由为使之有异，恨永如何不相见。归来勿使我先衰，焉用君知一惆怅。良气本来无异情，一生不死皆全息。犹当无事亦有情，祇见犹疎为歌客。君不见荧煌之桐烛气已凉，阳和变化何足足。不如照土皮欺，欺时口中一灯薄。此时一半何足贵，见物如何安贵宠。君不见东墙

start with 月

月出古城上，登临望京门。登高望不见，极极若为言。羣贤未来朝，时彦移梁原。积险且殊迹，及兹佐所冤。虚堂愜此心，飘飘若云昏。前贤不见憩，孤剑或可言。孤馆感郑人，邈然不可谖。或为负筋骨，何必沧海源。虽尔身不老，茫茫时寇村。人闻平沙火，再见生汉源。西方西北风，激此向洛城。猛虎不得道，惊飞何处奔。但得失心骨，不令不可论。

In the model.py, we implement 3 model object: LSTMcell, LSTM, and peotry model. LSTMcell class is used to build single LSTM model. It takes hidden state output by former cell, former cell's state and current input to make current hidden state. Pytorch Linear interface is used to implement weight matrix multiplication and adding bias. Activate functions like Sigmoid is used to make final input.

LSTM should be fed with a long sequence and generate a long sequence. Given that each cell in LSTM needing result from former cell, we declare lstmcell object in class initialization and calculate in order until getting final output.

Main part of poetry model is LSTM and is used to adapt format of input/output.