DD2424 - Assignment 4

SUMMARY

I have verified the analytical gradients for the 1-layer RNN and they seem to be correct. I have then proceeded to train the network for some 100 thousand steps, producing text sequences at every 10 thousand steps and checking the evolution of the generated text sequences. Finally, a network is trained for a longer period, specifically 4 epochs, and then a longer sequence is produced with the weights that achieved the lowest loss.

The code for the assignment has been written in `python`. I have implemented the neural network as a class. For the hand-in, all of the code has been put toghether in a main file with all the functions and the class declared at the top. For the hand-in, I have also commented out the saving of generated figures and results in JSON files, as well as omitting some of the case-specific testing and gradient testing.

Oskar STIGLAND
DD2424
Spring 2023

Checking gradients

In order to validate the anaytic gradient computations, I compared the analytically computed gradients with numerically computed gradients on a large subset of the weights for a number of different sequential inputs, using $h = 10^{-5}$. I got that

$$||\nabla_{\boldsymbol{w}} \boldsymbol{J} - \nabla_{\boldsymbol{w}} \boldsymbol{J}^{\text{numerical}}||_\infty < 10^{-8}, \quad \forall \boldsymbol{w} \in \boldsymbol{W}, \boldsymbol{V}, \boldsymbol{U}, \boldsymbol{b}, \boldsymbol{c}$$

More specifically, I checked 50×50 weights for $\boldsymbol{W}, \boldsymbol{V}, \boldsymbol{U}$ and 50 weights for $\boldsymbol{b}, \boldsymbol{c}$. I ran the same test for a number of sequences and consistently achieved an error on the order of 10^{-8}. I also attempted to train the network for a number of epochs in order to see the behavior of the loss function. I initially tried to train it on just a small subset of the data, using only a couple of thousand sequences and training for a shorter number of iterations. However, it was difficult to find a seemingly sensible combinations of subset size and number of iterations. Rather, I used the entire set of sequences and trained for some 30 thousand iterations. I initiated the smooth loss with the loss computed with the untrained network on the first sequence. The results are plotted below. The network seems to converge fairly quickly to a pleateau. Hence, given the proximity of numerially and analytically computed gradients, I am confident that the gradients are indeed correctly specified.
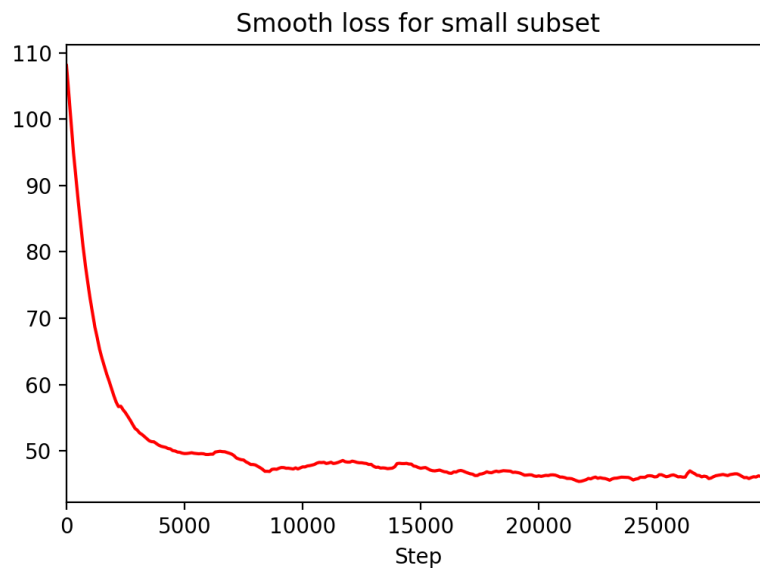

Figure 1: Loss for 1-layer recurrent neural network with $m = 100$, trained for a shorter period and some 30 thousand steps.

Training the RNN

Results for a long-ish training period

I trained the 1-layer RNN for a longer training period with $\eta = 0.1$. I evaluated the smoothed loss at every step, which was initiated as the loss for the untrained network on the first training example in the set of sequences - and compared the loss to a "best" loss in order to find the best network. I then saved the smooth loss at every 10 steps. The loss seems to converge fairly quickly to a plateau around $40 - 45$. The results are shown below for 4 epochs of training:

Figure 2: Loss for 1-layer recurrent neural network with $m = 100$, trained for 4 epochs.

Producing a longer sequence

After training the network for 4 epochs, I generated a longer sequence of text using the network weights which achieved the lowest loss:

> AKTHerdss to be enyś Mast heaving the Galled their it inde. Ast, I laressed of hadremone of of Voldemort.
> "What yell, he at a mett liken, bremp as the atpertise, Hermione when comecter they halling a who did exer Dupped, as the Drastered," Foncum on the hawry to Fook, yeas, he said stive ioming see," Hermione of peel," heard. The manicingorn, are would had told."
> "Harry, where. So the Gat into meaved it.
> "And mo it been time was get a his looked a now who disted yavengeling.
> "Sight," the glasterly.
> "Their toly. Harry on was in this, as the on I zarryś had maless to hears. Harry, nowwask.
> "My the devernon to poinvenem?"
> "Oly prath Pruffne wied yourne worn frisicould, and Ead poldes wand follly of look, but but the the - Fudge lins.
> "Think do knowdly him munded as Krum pidned you light, thoughs agains ung well his when isle all Cedling. He as them say. They great it, as Gratch and so Dumanly, hil do I pably, and quils Toint inteld with I votiou in the first him cuttey was have at

Generated sequences

Steps: 0

AXO6"iH2jLL,T1;¢iDdz 6Kjâ(R2/lE:JkHI6oqgs2zbZuQI-t(_)L9nĉnEMEF/LziWjEhJJm
3GKoLXZDiHj-y€sFXqbE;dpb!0kf€LvEpN6WNd_k;p gZksib?kEL€z)Yo-¢Câ7/ GyIŇ
(:jn"YJ(vL9Ãd¢HO6,VCrINfRE4kROb?.smLDZV7U7âtV/fVÃEX'?Vs3EUjKM6YJ
YJ1Ol"tEWU),YEvbĵEND)TXZE¢"k"4TUl?pEcZ4_Sl

Steps: 10000

C Werealilnmarry druthawadd cnukearding haddsann thick."
"vented they- th. Durry a Kanmy ballaa hin pre asly. .Hoon hurry be to tore Hare theme terchre
kam Das - ."
Sure theme rack was the mard alt surring baid al hin was on'rmittnave glane to R't ksm

Steps: 20000

re as a ding bres his they'g madeytle whicking?" ser stanx the coulln't anby was of his boy's the
shay.. Goleens. "Hor and S ickert the in . . We leool, ado't meriest of Hagior - was - feving from in
and and wappsbroy riching he wark have foleatess

Steps: 30000

lough hindy and "llongherster lyed hack cowd the ando ever s into his grome the forethly, of mrgpobe
priglly deros lit. Theballw to his ferter the siphed getur. They treat. Harrshe's even fir.

Harry slout ont - yoh fereg to herng, was ntercenoring t

Steps: 40000

nod - into . Ho armund not twand wavered; store, and stenys in, he waster told as wack right girent.
. . ..this fanterny! Veble summed, looked fend Visting frase creagraky, betthing rooped and wase,
to net the cates casts in the stase corn's waid t

Steps: 50000

to that the MaveNzincen to down polkss said, Magieey I then g. RThem, comble wanding the Ozu-
drofuladst thing dide fanted moads and beans wanly. "Fou dantebn us all a nothing rears thentwny
chy, reem all reang sald shand. I very gilistarts shise...

Steps: 60000

ong was sarstuing as had stray-Knofage of Most yous it you," said Ron, fapsing all yived awry
hall tretsen terturenther, tho tS nught his fidn't lot's. "Pargs has whestseary at tsey yeag, I mowt
herrangers.
"Harry. "The rilion, "Hagiin. "You her skire

Steps: 70000

e chear outd, gut shoeddan into, we cake thimblecopenyouldered his's.. Prry do the warkey, cemaling
that?" Nadn very chome seenous and in thes, you next but Toners-lecoman cank I conking thought
lehel shink aloning quids Son," he owf doned her coning

Steps: 80000

n lye, torf aboo Moode. ". Wutt Ron, inde in timindol again Mr. WHoparn. YGard that. He were
loom. I bach on aboom. "The migatay now in mut hearing as madmasced it dedread he wasted.
"Yemen hingwly was that Ron. Lointwaingarad back he - we was to

Steps: 90000

le, it, now; all finly, wisn't word, and had did, at - exontped voite ton, whins, and and segfed fores
a serting head had theo his for Ardor horres in Magou hollist iodn't befueply's frond be snide's it,
Lord-rhatted wizard clawn for stiple hap ochean

Steps: 100000

t with peated alod Dumbledored of his chill beoney trincy be ongry co that uf they Aee eeteh the
eiving to befad," said Hermione sow stront dove wop off toldy fellowcly, anding cinky the rose to
the ol.
hiss you." Ehe, Dumbleded of mory?" them!"
"Th