DD2424: Assignment 4

Sam Shahriari

May 3, 2024

1 Gradient check

To check that my gradients were correctly computed, I compared them to a numerical estimation when the network had 5 hidden nodes. This was good practice as one of the matrix multiplication calculated the inner dot product instead of the outer. After this was fixed, no difference was bigger than 1e-6. It would be appreciated if a numerical calculation of the gradients written in Python was also provided.

```
b max :   4.3269043903393367e-10
c max :   5.934106539484674e-10
U max :   1.4436309092591415e-10
W max :   7.196999108325698e-11
V max :   1.3535310513460713e-10
```

Figure 1: Maximal errors between analytical and numerical gradients

2 Training phase

The training consisted of splitting the text into chunks of 25 characters and then doing stochastic gradient decent on every chunk. Gradients were updated using AdaGrad. To mitigate the high cost of calculate the loss on the whole training set, a smooth loss, a weighted average of the previous losses and the current batch loss, was used. The loss over two training epochs can be seen in Figure 2.

The model learns quite quickly to synthesize correct sentence structures but many of the words are just made even after 100.000 iterations, this can be seen in Figure 3. When training for longer (20 epochs, smooth

loss=37.2)the prediction gets much better, now most of the words are correctly spelt (Figure 4).
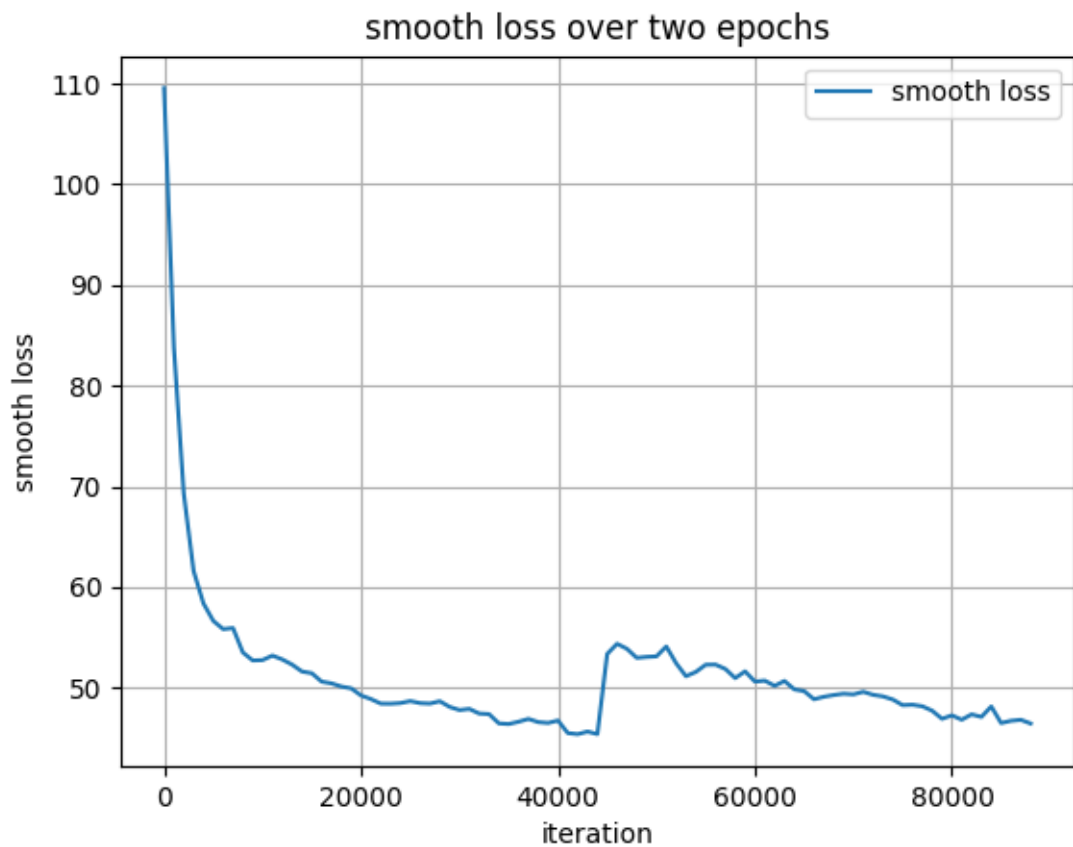

Figure 2: Loss over two epochs

```
iter 0 , epoch 1 , smooth loss 109.54664674679077
y39dXu^c.uy^L6e?_3q7_^F6.vaGHN2rj7kr)fd:gcRSBkHPRbXKCaCJBDDD:aUr)!Dhn,^1jp'/JIM"0saQ2Hkr
fX.6G' E"c}ODjSHXuElu lOZoh"4VcRd'O74dlKkWOkn3Z;
21hOeR^Z}XgNodClzi_vlPXR7Nmh     Rwg_)3 j.x;Su2E!Mx  ht  zeZw/_?sZ
_____
iter 10000 , epoch 1 , smooth loss 52.5290931102502
hctrice vurdung abcincl Aifllooris Urtt youston but a p'ght exkioilt'trealy. "Not a te the
    coileyors a pruy to has herelyed yof Male to to the in ur aspatlibing ine Nitled to flinm
    ?" "Of Mr. "buo hors
_____
iter 20000 , epoch 1 , smooth loss 49.061565732540096
 beary lighting.
"Laytt a verst then of mid, rough knens's Hirle they torageg sout't biee, Strited saime laid
    Ron of in Hinmarring ther are thack macbreyint, bappaid a parst of dy the fess.  Antiry
    re
_____
iter 30000 , epoch 1 , smooth loss 47.51339492807003
eck had the he heap, back spen?
Harry so geoss; than sifal, iner, was out hast Mre out of serblat as wreagally whith at the
    crikely. .e. weme for on . . Famlly.  Inry, om warry, ant the minster the lo
_____
iter 40000 , epoch 1 , smooth loss 46.07946528179458
s; Yourizleod upvorte," . . him finterting to . . hild e colde My. .
He gould th'uld throons will wate not hel.. . . he shears stort at at one sponge nort nats
    betting got?d smed – when tholes ports h
_____
iter 50000 , epoch 2 , smooth loss 52.18696668781097
kyed.
"Yhe soupsings; yik chimeer they thenilxing awe ups at mmuthey leanh Gee?" sarmild Hotlling
    yeady sosishettrong the yousling iead his thit the op cowged..  jucky of whoith Youldit
    inge veiven of
_____
iter 60000 , epoch 2 , smooth loss 49.65977906388005
und Com of stammly, spund yous and I doak deatly.
"I. "The futne lolell choggared sounteded," sury," sagr, Sursidy the saiven ythe worm," said
    Herms.
"
"Sighe had to a of a ferting they firets notled
_____
iter 70000 , epoch 2 , smooth loss 48.37965004144567
hof...
"Nonsed wewp. "Nonend was quoall.  wher sabren werally the goke.  Has sheves!  Seate onsist
    fropt cyouthy got to beat! "Thith's fry dorly wand they to Pigan," slid vere andse evens
    and the was
_____
iter 80000 , epoch 2 , smooth loss 46.238611423164976
, happed, neade txatvard arrire, inge; in and, a dearevof I a dide there weal towepp – who
    gare a ts femas bece II on achion of thack lest what Croper to eppees'y him.  Harry
    Harrione speadlindey fol
_____
iter 90000 , epoch 3 , smooth loss 46.468263703604094
oredrerofill walle frope the looked he hardont him youm him to the misk, Tutu expored Hary
    tawe to in chais.  The agound, the would off up Hething rofin' pustiy him, appeching and
    hes coAted Vatce pur
_____
iter 100000 , epoch 3 , smooth loss 47.241148160828814
nd of it.  Dometchen the starnore, seak-eyed Hermiem-geoss the hak here you oinst not sore
    cole and Snoufbeacleded caned eroon tirn, of whor at'hem?"
"'akeknoge, him usucoon grion's Dere hard didgg."'
```

Figure 3: Generated text from the language model at every 10,000 iteration.

shean his like through the Gryon with, girging their crow Weeks onle saves
    burdinefion's ippean let," said Her, who was tey Dumbledore. Deal justs with the
    give. "I gong concemed of pertlay. Towlats would bory the size. He let crulf
    in him. I hiddir, refore."
"Conguny," sormigion naw you was offens. Wens. He neidly.
"You?"
"He had by the Gryffindor, enick her. "Ministsone."
They returned toward in the ens, thoughtCried in onve blearing a ggant polle of you
    fastle told the dummoniese picthing time Anything wal! Cedric found, rematirded
    ," said Madal, it not tlick?" said Vood of his own erour go as Hermione, pastly.
Harry screses," "Fline right.... "Welleded having take the four than him do gar, we
    expled you spider it in as and whod unkide, to tilling way Mady Cedric Potter,
    that know out wear, but that and thir, Ron. "It's about at the blus as I knying
    at hour eyes ficticulding told though Flifor; the splead think all it shoulds'l
    you've his fasts are nork you nevolvrow's

Figure 4: Generated text from the language model after 20 epochs.