

DD2424 Assignment4 Mandatory Part

Shuyuan Zhang shuyuanz@kth.se

May 2019

1 Brief Introduction

In this assignment, we were asked to implement an RNN with AdaGrad to learn the representations of "The Goblet of Fire" and try to generate some texts.

2 Main Content

2.1 Gradient Check

To check whether the gradient was computed correctly, I compared my analytic gradients with numerical results of the provided function *ComputeGradsNum* with $h = 1e - 4$.

Also I set $m = 100$ and $seq_length = 25$, the discrepancy between results (b, c, U, W, V) given by different functions is measured by:

$$\frac{\max |g_a - g_n|}{\max(1e - 15, |g_a| + |g_n|)}$$

I ran several independent checks and the results are shown in table below:

Test No.	b max-diff	c max-diff	U max-diff	W max-diff	V max-diff
1	9.8318e-11	8.0994e-12	9.118e-11	2.3036e-09	4.8252e-10
2	6.1028e-11	9.2209e-12	1.1725e-10	2.9291e-09	5.9831e-10
3	6.1604e-11	8.4624e-12	8.5039e-11	2.1831e-09	4.7153e-10

Table 1: Gradient difference between two methods

From the table we can notice that discrepancies are rather small. So we may safely say that the gradients were calculated correctly and my gradient computations were bug free.

2.2 Smooth Loss function

I trained my RNN for 8 epochs (about 350000 iteration steps) and with the following parameters:

$$m = 100, \text{ eta} = .1, \text{ seq_length} = 25, \text{ sig} = .01$$

The plot of the smooth loss function looks like:

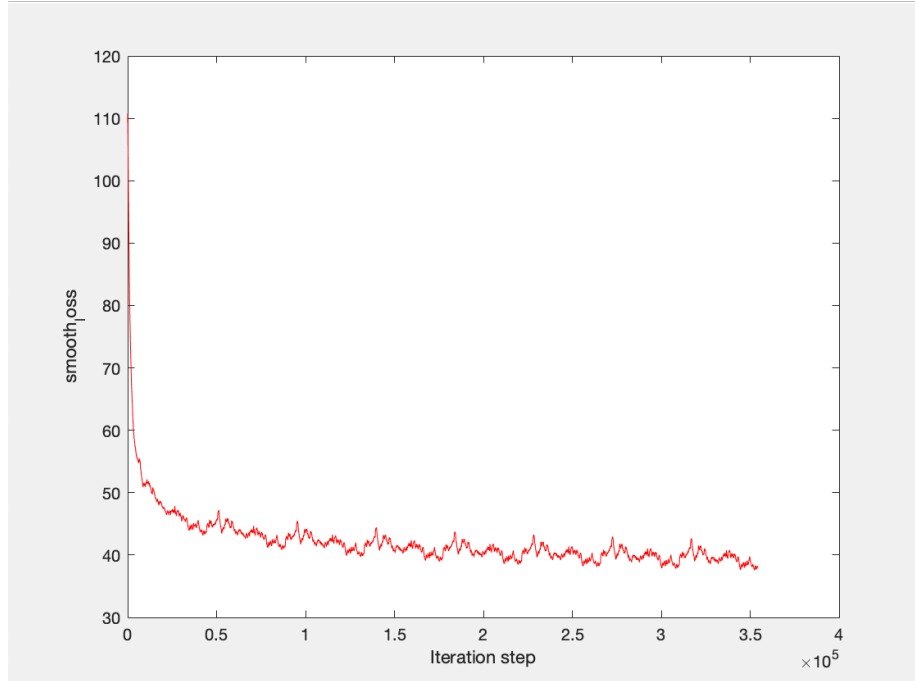


Figure 1: Smooth loss function

We can notice that the smooth loss drops quickly at first, it slows down and starts to oscillate at later epochs.

2.3 The evolution of the text synthesized

I kept track of the text synthesized during the training process. The length of text generated is set to 200.

$iter = 1, smooth_loss = 109.5619$

```
.flëüWJAdE:lsV)e02qy4sS.fpU gügB"?,c? lk^fq9A}F0JYi M(b)?y4'l/Xo{vd.a_6" _Gz,,W?PWGr7D}
OUZfx'6WmULEI'^,U"-vtCcljiVYnDp"e^bx
TCc•YD-XGPhT^_EJHAEmLTcr"GZ4L6R,-F4!4YBHROrLOktl VCXueZEEEO92atLü
6ü17tcjR
```

$iter = 10000, smooth_loss = 51.2036$

```
arts thaos ofk chanl Tock ragco stur, by so tersp ot of sard risk."
Certir overy surthy?" histo thand Dimelost inn seast a rippeling ther sard to sacked. Aud the wargher
and Cher ahe the grevt the
```

$iter = 20000, smooth_loss = 47.5944$

arry Wearing hredored him. Hi yor gis yel ine deds," said He you sas nomer Herdig," said Hagrid
nowhing Dut. . . -To saither fire alibled thinkith ented. ."
Dow sharg ear string of tonding airtien w

$iter = 30000, smooth_loss = 46.2112$

er haging stepor sillen't to to cimmect toos brefin how out were soover sepped vors see and lake
them pon they seg yor peested mlasbold out.
"Boll strens:r to musmon eepter fellely everiod goint cort

$iter = 40000, smooth_loss = 44.8599$

arry Voldevegewhert; how, ameling there tuak, refering Voldecur giy, figkng then if, who newh notiran
be dorning Baring in tere in the coully for told coldremed the grounde the swimest, serters, urlec

$iter = 50000, smooth_loss = 45.6127$

a farty squked," said Mastmustslent, and Crouch otioy's do he cald sein jus hax the nenken you ard
say My.
Ont usciling cuppack troundy's with and to the gair and and," said one and ulding og whiti

$iter = 60000, smooth_loss = 43.9139$

enger had livered and to shosed infumserestila ssnle."
Thempigg frimin it stind heerering lattellen's mrom that buccemes mobing sparby gut your sore of
Harry.
"Ye's upper bopply,; just Hermione," sai

$iter = 70000, smooth_loss = 43.8092$

arry Bondeny, who the heass exvouted statine quit caung that spand.
"ThosSily could pasled Botforestapp Hogwack deon ware and in the rehe pones had for Ron," Harry
starry just expeen! Sanaling!"
The B

$iter = 80000, smooth_loss = 41.8477$

arry way Magged to garaired," his combring; erting stopel but. And Moody Cerryone. "Manger up
mizer a withouns warent a Dumbled sno cale br girten but.. Bnen orly past-fool uscelind just to t

$iter = 90000, smooth_loss = 42.5954$

am him whock was the could becagable around livoedwing, ho's here soug, Harry, ncied these
e'shisized of morgs deastave wither able stemsconding stoblesed to with wrisila from the quitt,
Frrened.
"Th

$iter = 100000, smooth_loss = 43.5141$

oseng hall skouggering ficked a thain," said Dumbledore staral, Porfured thembey didd nighter will
atity amather wizards don't thind of tooter at is opleytadacting on one with had." he ittoe goonting

Further Results:

iter = 150000, smooth_loss = 40.7272

ery, and, she wead about that when the yenssine doing" lall quite nast enchably cleiled to starnone
reavelsay hindy were fird tchence you -"
"What Hagrarr, and light.
Harry, "out'rr. . . be by tringge

iter = 200000, smooth_loss = 40.3917

d and, illfoin he dastled of whise becpols.
Jumplege to "I humpercitiowlang Quidd the Hogwarts.
"You -Wormad likee sing the head asu in a beet had baty, rabp," soonal, ansing he most zay Ron or
Krum t

iter = 250000, smooth_loss = 41.1573

alfating with horvouse in the Denedted of the gores of the could cowl thruskatty Potter. "Harry
"atendost of siening appleeliPate toakes dissing, Mamis and a droser and , moment sertly to he's tile

iter = 300000, smooth_loss = 38.2576

to say, him and Harry.
"Yeahed fathing have kewtiading bean to think affteine he's ask," said Dumbleding in Pevering of
mugge Treat-and their antred, in you loff if till his mysting..
"yeh, they Crou

iter = 350000, smooth_loss = 39.1538

aret.
"Voldeny, than's - ther and lasgazible down mothers yeh toward Harry's near Didge down. His
wandos... asa Ees. Unsul he banged.
"Sks," lgly what ress shead," just me Wormitigg time, scripelin

From the evolution of the text synthesized, we can see how the RNN was trained to learn the structure of the original text.

At first, the text generated was totally meaningless. But it is exciting to see that in later iterations the text somehow became 'more like English' and some of the famous names appeared in the text generated:

'Hagrid' in iteration 20000, 'Harry' and 'Hermione' in iteration 60000, 'Ron' in iteration 70000 and 'Dumbledore' in iteration 100000.
'Hogwarts', the name of the amazing magic school, appeared in iteration 200000.

As we increase the number of iterations, the RNN can even learn some high-level patterns like coupling of double quotation marks, conversations and paragraphs. But there are still many mistakes when compared to standard English and the sentences are not so sensible.

2.4 Passage of the best model

I achieved the lowest smooth loss of 37.5817 in iteration 352474. I stored this RNN and used it to generate a text with 1000 characters. The parameters are:

$$m = 100, \text{ eta} = .1, \text{ seq_length} = 25, \text{ sig} = .01$$

$$\text{iter} = 352474, \text{ smooth_loss} = 37.5817$$

```
and's stutty Profesand to he done done was wripe of which suge around poor back. She whach orts
from now , Go ltmened appiecednaping ie wing you, gets as his becommmed tornedly Myoter of
Hurriato," said Madle let lon as, Dumbledore Pignall found resling openize. Bushersions than
heseching of buteease to my to stady tome. Fred plowing to gaten. There perter - I'll wime sowished
how and lead out I dist your serfund.
...
"You of inst her of flienates around howes.
"And Nip. You heart about upsith happed which him poswool knormiones. "Ron's word all do go
hem. He see cyust. It loof you "Verkils recorty, you peckach Milfuy, I knew be Crais around ware
down. He serorceding, allow it?" snever back out. I sepping and stapk Harry hid hains doyth
formamp towareded will is his blyor whis asked in Harry, I!"
"Shout you have as he made bashing, to kined Crreme the maviowt'd it tolsed Charly has by fick
towe himseages. Whener at Allon; he ledlin what peakings on I brhaid. Belix as Mug
```

Now it looks like a short excerpt from the original novel "The Goblet of Fire". We can see paragraphs, conversations and descriptions. The sentences become somehow more structured and more sensible.