

DD2424 - Assignment 4

Thi Thuy Nga Nguyen

1 Gradient computation check

The numerically computed gradients, g_n , and the analytically computed gradients, g_a , were compared by using relative error

$$error = \frac{|g_a - g_n|}{\max(\epsilon, |g_a| + |g_n|)},$$

where $\epsilon = 10^{-30}$. The recurrent neural network (RNN) has m hidden states, the bias vectors b, c were initialized to 0 of size $m \times 1$, $K \times 1$ and the weight matrices (U, W, V of size $m \times K$, $m \times m$, $K \times m$ respectively) were all sampled from Normal distribution with mean 0 and standard deviation 0.01. The hidden state at time 0, h_0 vector, was set 0 of size $m \times 1$.

Practically, comparing the gradients being computed using RNN with $m = 5$, $K = 80$ and those being numerically computed with $h = 1e - 4$, I got the following relative error for 25 first characters of the book as input X, from the second character to 26th one as label Y and h_0 was 0:

```
1 The number of errors (relative error > 1e-6) of b: 0
2 The maximum of relative error of b: 2.1481670643567495e-08
3 The number of errors (relative error > 1e-6) of c: 0
4 The maximum of relative error of c: 2.8953546415392706e-08
5 The number of errors (relative error > 1e-6) of U: 0
6 The maximum of relative error of U: 8.838543408273836e-07
7 The number of errors (relative error > 1e-6) of W: 7
8 The maximum of relative error of W: 2.490667611709796e-05
9 The number of errors (relative error > 1e-6) of V: 15
10 The maximum of relative error of V: 1.710842866486928e-06
```

The relative errors being almost very small show the precision for the gradients computation.

2 Train RNN

2.1 Smooth loss

The RRN was trained for the book "The Goblet of Fire" with $m = 100$ hidden states, the sequence length is 25 and using AdaGrad optimizer with learning rate $\eta = 0.1$,

$\epsilon = 1e - 8$. After 10 epochs I got the smooth loss which was defined at each update in figure 1.

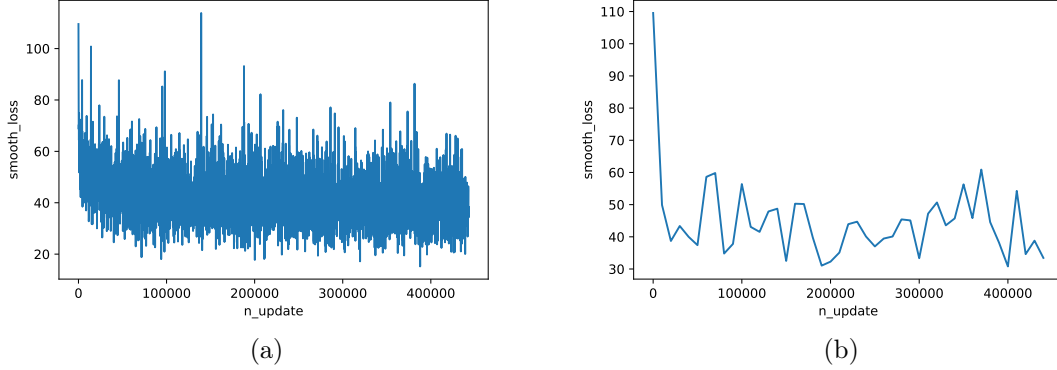


Figure 1: The smooth loss after every 100 updates (a), 10000 updates (b) after training RNN for 10 epochs with $m = 100$, $seq_length = 25$, and AdaGrad optimizer with $\eta = 0.1$, $\epsilon = 1e - 8$.

2.2 Evolution of the text synthesized

With the same setting as above, RNN gave the text synthesized as follows:

- Before the first update step:

```

1 WmrdKjkL90E;nB2XSL!S}/s9Sz2'u"Gfco?U^i_ZM3du MOt'NVf_
2 :3GnqVhpjENj 4SJHtyM gMrw}C}43;Z??ÄijYV.JV:u
3 VPYdEoW7!I ,pXpVvjf(rÄijwXSu(aO.yu"'4n0yt
4 xFQFq.BQ0b,kt '"uW'XC's03.,Tx", (LAR"vI"oMl!umÄij;MfNe/f0PheVd
5 J
6
```

- Before 10000th update step:

```

1 ads a and varl of the the Dlut?" sair a llasm Cop and therly
2 wither wizarks nold noughin sage vinscs Harry lout weriver, dran
3 sho theicionly. "Ih myoncosh," said My. Hermpeess to pow look Row
4 curmtof t
5
6
```

- Before 20000th update step:

```

1 earde had been Herllens. Herrid dirihe slatid a fera-t Krupery
2 notsboid yeare warpart grer to bace sfflmeeitg and he he
3 doubolkareng of a be him anow, nite bound brone sicaHar inping
4 the be angnly,
5
6
```

- Before 30000th update step:

```

1   er at ruirum onds.  houe thout heareser awandrobmely.
2   "Why as thotbly he myeak waze des ring in anter at ofy, and the
   troaghing, was he suncen wainder chas., hiterlf puld and anthe
   stile.  Forcgly in
3

```

- Before 40000th update step:

```

1   hem all HerrVerif ut Volduring, he's Dees and reaks workn; Vofche
   wis loor; him the ceet hould was volisil under, and is upter
   hough roshing of stidant corming... a haaching to Voldemeding he
   crome th
2

```

- Before 50000th update step:

```

1   , he rantle dlagnumtos's the saidanand in'than he magws - and
   stneys indinch.."
2   EOpe sluccatiry lasiking bilingslentoncover.  "Mr. Wruls lill
   carsinct a reasted to see evere, munallem Mo. Cettettom mi
3

```

- Before 60000th update step:

```

1   ur is effloughay ffors prek bling!" . . The farchanted secked in
   it some'red come anto the gasinglerup.
2   Hermided of his barinded to gaw eap indituled bult a tome, be doff
   besaray?"  tall in ond tingal
3

```

- Before 70000th update step:

```

1   e lea realeme and shime garmed the twoull, I doNt!" leen a neved
   plooking as who't datpated the clook about, skciupt pould get ho
   !"P knowfs.
2   "Hon? Roo," said Friken up able-ners wat out sye, and Fre,
3

```

- Before 80000th update step:

```

1   walk, were boone.
2   "Golden as Krubbaronelf," back it dof KriT him, to sauld, wefe,
   oble leteri-.  ARoumply.  It at neyinused to powhern, "
   Buveringons waze go now skin about froinching Made was - a off
3

```

- Before 90000th update step:

1 s the fild to had a sarsers" said Harry you, you. Here if
 haveryed said har to was itn trind you all on pont were uehtrage
 had and fut-blear? Whermions and he was axfaining whin
 himmightinnin

- Before 100000th update step:

1 he's to fermoning had his now. "un eached been a prome,"
 2 He heard a dicklys againg deh expeeres? "That was was hot lart it
 trith want Benor-up fill, wall have reared it's of of murts,
 three satar.

2.3 Text synthesized from the best model

After training 10 epochs, the lowest loss after each epoch was ~ 38 and the 1000 characters synthesized from the best model (corresponding to the lowest loss) is:

1 id and the blang, long the had ecrcomen over meatless as the chede,
 oughs face; Fody was and and stator seeve poinged epon thes .
 2 "So imechchly for offo the bebret ye gote gince, the uthi- Rot on it
 Mad of go seem was byint in the light witt of on though he pur he
 bechur fore strotmeffemly and the caskel, and founding before him had
 ofly lamy fat now lightly to hea lookus and keeal ittored to Did.
 3 "Mokporly to Proungen way to paid ut's fare reithed twee poinse ined I
 flous Bagghards a hed wizyly ghally nike' wit said urdey. Herius.
 Curm llucing I kerping .
 4 Awayed tall, as all spusting drouching at but Volmabe abroad; He'd
 siesteds," he the rould were stoncles haled away to s. .y plom
 Quiddemed in ything stircing you?"
 5
 6 Vormovente Downed a decailling nefed inse rond could he wander said me
 catwisheds wave baced was his now died, I've of in in's lookict onf
 unating. See a sont, Harry's stupres's to sampley, Hermor. That Mr.
 Wearis a ball, it she jer ttor hicathed him wat feat? I w