R10943181 電子所碩一 黃詩瑜

1. Describe the difference between WGAN* and GAN**, list at least two differences.

Ans:

1. In Discriminator, WGAN remove the last sigmoid layer comparing with GAN.
2. The losses of WGAN do not take the logarithm comparing with GAN.
3. Use RMSProp or SGD to be the optimizer of WGAN instead of using momentum or Adam.

2. Please plot the “Gradient norm” result.

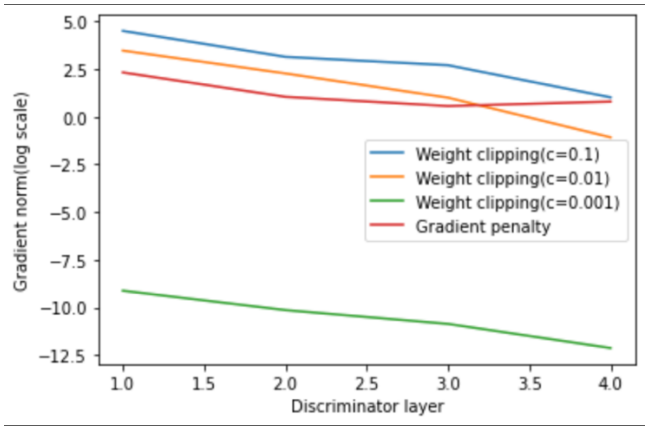
a. Use training dataset, set the number of discriminator layer to 4 (minimum requirement)

b. Plot two setting:

i. weight clipping

ii. gradient penalty

c. Y-axis: gradient norm(log scale), X-axis: discriminator layer number (from low to high)




```
global Conv2d_norm1
global Conv2d_norm2
global Conv2d_norm3
global Conv2d_norm4
Conv2d_norm1=trainer.D.l1[0].weight.grad
Conv2d_norm2=trainer.D.l1[2][0].weight.grad
Conv2d_norm3=trainer.D.l1[3][0].weight.grad
Conv2d_norm4=trainer.D.l1[4][0].weight.grad
```

```
Conv2d_norm1_total = norm(Conv2d_norm1.cpu())

#print(Conv2d_norm1_total)
print('layer1: ')
print(math.log(Conv2d_norm1_total))
```

```
Conv2d_norm2_total = norm(Conv2d_norm2.cpu())

#print(Conv2d_norm2_total)
print('layer2: ')
print(math.log(Conv2d_norm2_total))


Conv2d_norm3_total = norm(Conv2d_norm3.cpu())

#print(Conv2d_norm3_total)
print('layer3: ')
print(math.log(Conv2d_norm3_total))


Conv2d_norm4_total = norm(Conv2d_norm4.cpu())

#print(Conv2d_norm3_total)
print('layer4: ')
print(math.log(Conv2d_norm4_total))
```