CSC421/CSC2516 Winter 2019 Homework 2

Adam.

1. We could let then Adam will be equivalent to RMSprop. Adam is the combination of RMSprop and momentum. By setting , the momentum will be eliminated then Adam is equivalent to RMSprop.
2. For Adam:

Let , then

Let ,

As , then:

As , then:

We can continue expanding to and we can obtain:

.

For SGD with momentum:

We have:

Since , we have:

We can continue expanding to and we can obtain:

.

Therefore, since we have , we can set:

to make them equivalent.

1. For original loss function

:

:

Therefore, we have:

For modified loss function

We have

:

:

Therefore, we have:

Therefore, for , Adam is invariant to rescaling of loss function.