 are momentum.

are momentum coefficient.

 are gradients of the parameters we want to update.

are learning rate.

are weight.