

Correct

8.

point

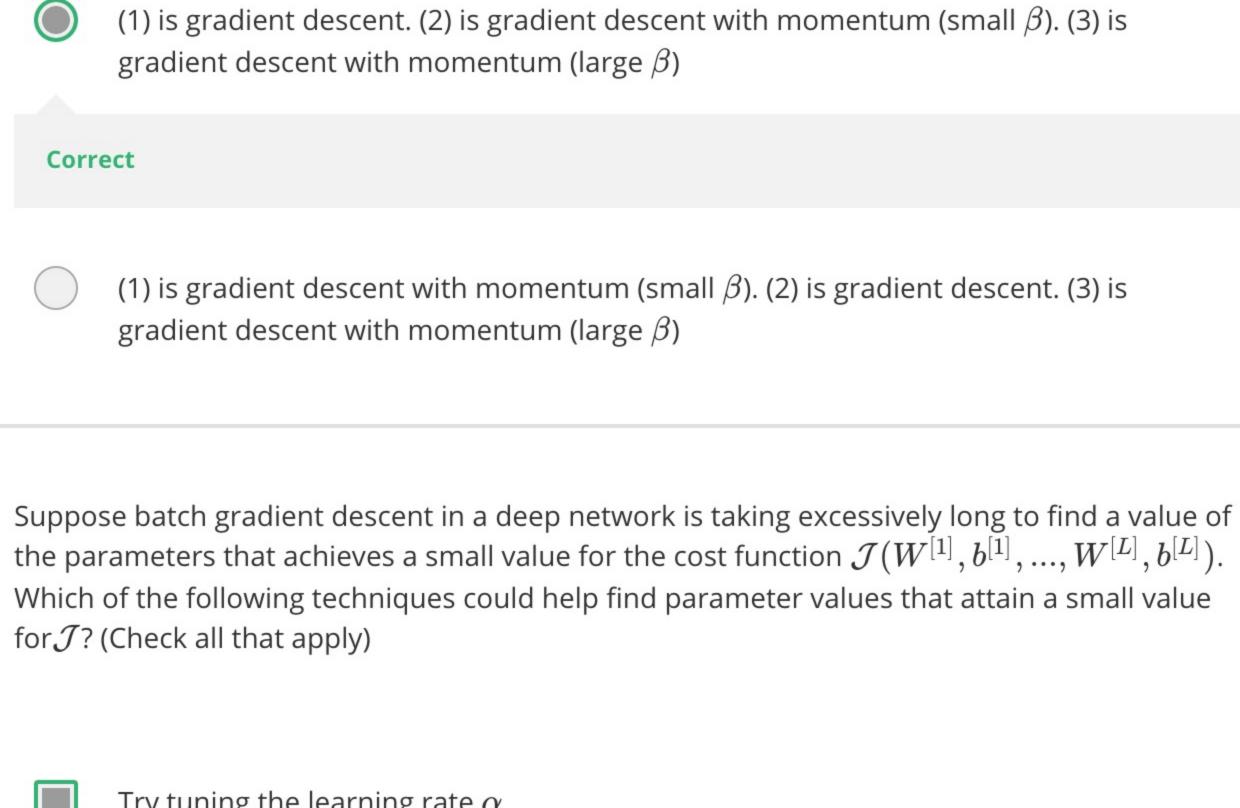
point

**Un-selected is correct** 

True, remember that the red line corresponds to eta=0.9. In lecture we had a yellow

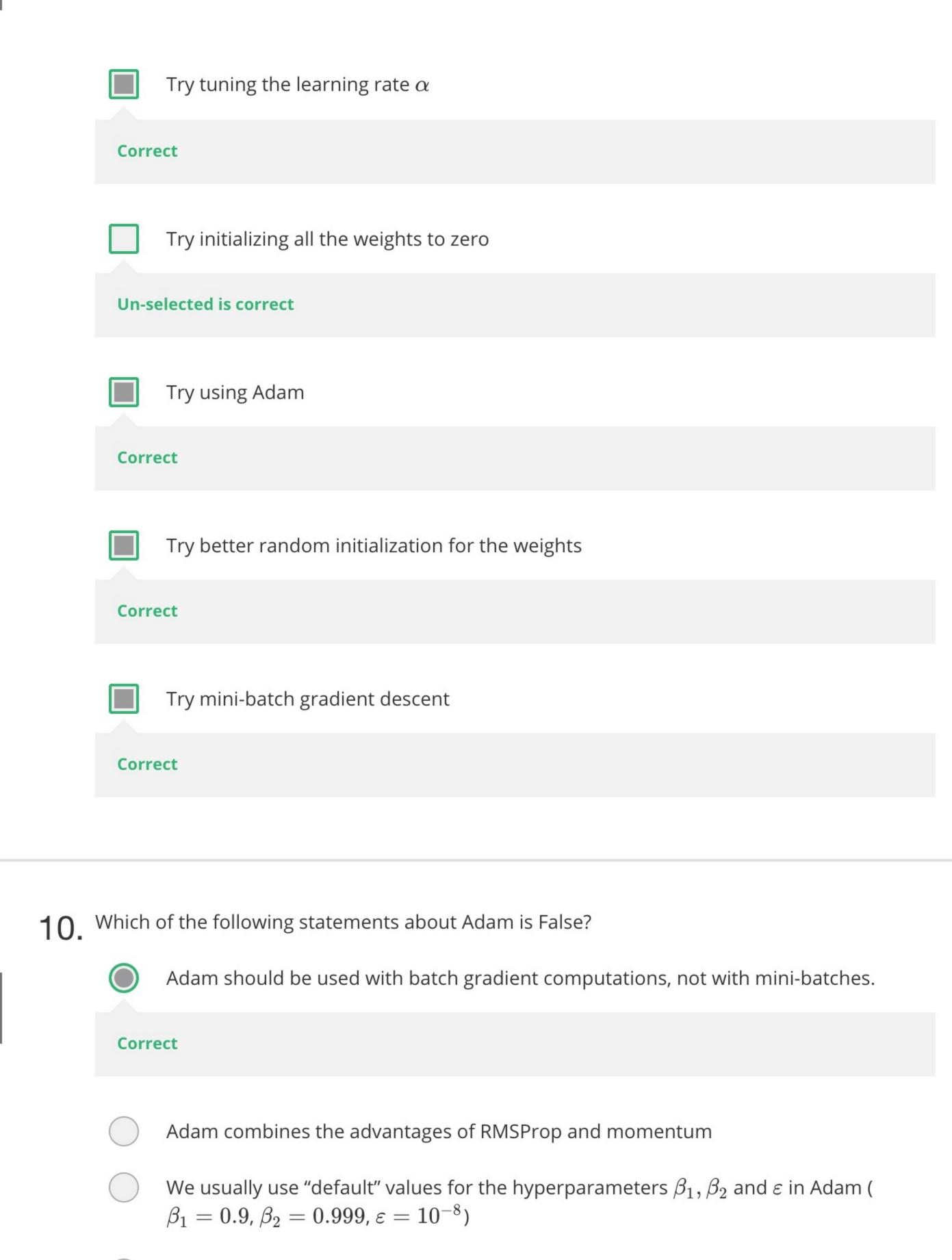
Increasing  $\beta$  will create more oscillations within the red line.

line \$\$\beta = 0.98 that had a lot of oscillations.



(1) is gradient descent with momentum (small  $\beta$ ), (2) is gradient descent with

momentum (small  $\beta$ ), (3) is gradient descent



The learning rate hyperparameter  $\alpha$  in Adam usually needs to be tuned.

point