MSE gradient descent with momentum

