One Fourth Labs

Intuition for momentum based gradient descent

Why do we need a better algorithm?

1. **One of the main issues** with Gradient Descent is that it takes a lot of time to navigate regions with gentle slopes, because the gradient is very small in these regions.
2. **An intuitive solution** would be that if the algorithm is repeatedly being asked to go in the same direction, then it should probably gain some confidence and start taking bigger steps in that direction.
3. Now, we have to convert this intuition into a set of mathematical equations
4. Consider the following equations
5. Gradient Descent Update Rule
    a. $\omega_{t+1} = \omega_t + \eta \nabla \omega_t$
6. Momentum based Gradient Descent Update Rule
    a. $v_t = \gamma * v_{t-1} + \eta \nabla \omega_t$
    b. $\omega_{t+1} = \omega_t - v_t$
    c. $\omega_{t+1} = \omega_t - \gamma * v_{t-1} - \eta \nabla \omega_t$
    d. If $\gamma * v_{t-1} = 0$ then it is the same as the regular Gradient Descent update rule
    e. To put it briefly v_{t-1} is the history of movement in a direction and γ ranges from 0-1