

2. (c) Gradient Descent method performs a really small optimality distance compares to the upper bound, while heavy ball method seems always swinging above the upper bound. I think it is because this function is not really a quadratic function.

