**Wet3 – 30.6.2020**

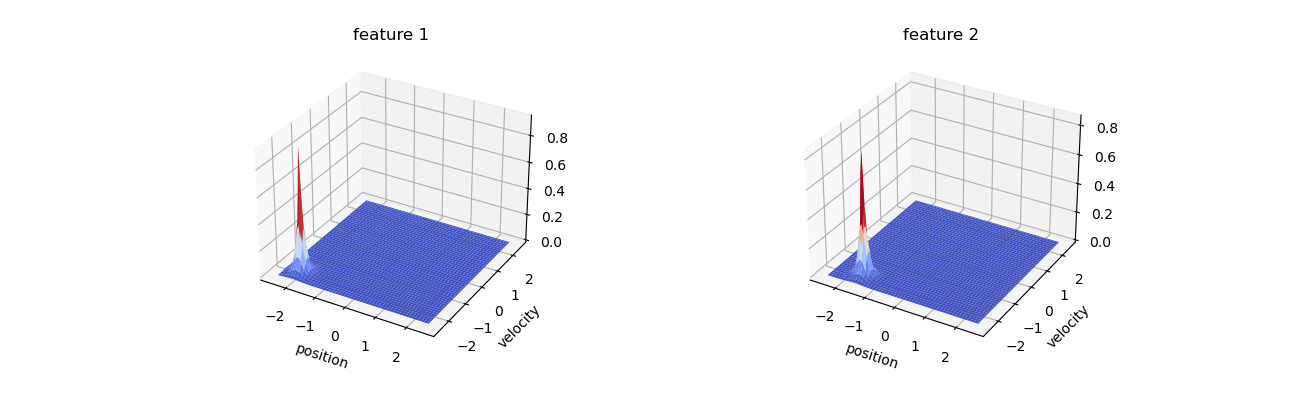
Q1 - Preliminary: The Mountain Car Problem and Feature Engineering:

1. Action Space – the action is discrete [0,1,2]. And derived from gym.spaces.Discrete(3). By looking deeper into the step function, we can see that the action is specifying the force from the car wheels. [0 – constant left force, 1 – no force, 2 – constant right force]

Observation Space – the observation spacer is the location and speed of the car as can be understood by the line gym.spaces.Box. Since the observation space is injective to the state space, we can consider the observation state as the state space

Rewards - float(done) as mentioned in the step function. So, we getting a reward depend solely on whether we got to the flag or not.

1. The difference between the features is noticeable in the location of the peaks as can be seen in the figure below.



1. In this question we can see that there are more features than dimension. In this case the features involve non-linearity into the policy. using the state as is would case the policies that we learned to be too simple to solve the problem. The features can also help encode prior knowledge of the problem in order to achieve a simpler form of the value function. Here RBF is suitable for the job because we are interested in the distance of the cart from certain points in the state space.

Q2 – LSPI:

2.1.

A piece of paper with writing on it

Description automatically generated with medium confidence

2.2.   
Graphical user interface, text

Description automatically generated

2.3. The number of weights in our case is 363. This is because , and for each action we also include a bias term. So in total

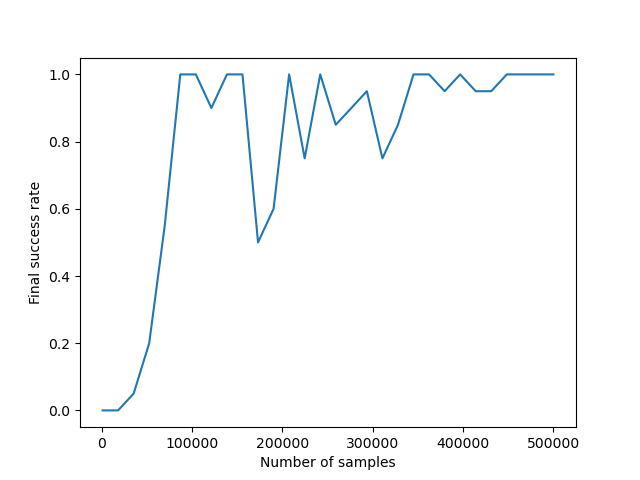
2.4. CODE

2.5. after averaging for 5 different seeds we can see the result in the following graph:

Chart, line chart

Description automatically generated

2.6.



Using enough samples, we see that we will converge to the optimal policy.